

IZA DP No. 5412

**Social Support Substitution and the Earnings Rebound:
Evidence from a Regression Discontinuity in Disability
Insurance Reform**

Lex Borghans
Anne C. Gielen
Erzo F.P. Luttmer

December 2010 (revised: July 2012)

Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform

Lex Borghans

Maastricht University, ROA and IZA

Anne C. Gielen

IZA and Erasmus University Rotterdam

Erzo F.P. Luttmer

Dartmouth College, NBER and IZA

Discussion Paper No. 5412
December 2010
(revised: July 2012)

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

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

ABSTRACT

Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform^{*}

In this paper, we exploit a cohort discontinuity in the stringency of the 1993 Dutch disability reforms to obtain causal estimates of the effects of decreased generosity of disability insurance (DI) on behavior of existing DI recipients. We find evidence of substantial “social support substitution”: individuals on average offset a euro of lost DI benefits by collecting 31 cents more from other social assistance programs. This benefit-substitution effect declines somewhat over time, but is still a significant 20% eight years later. Individuals also exhibit a strong rebound in earnings: labor earnings increase by 62 cents on average per euro of lost DI benefits. This is novel evidence of substantial remaining earnings capacity in a sample of long-term claimants of DI. On average, individuals make up for almost the entire DI benefit reduction through increases in other forms of social assistance and in labor earnings.

JEL Classification: H53, J22, I38

Keywords: crowd out, spillover effects, benefit substitution, social insurance, income assistance, welfare, regression discontinuity, administrative data

Corresponding author:

Anne C. Gielen
IZA
P.O. Box 7240
53072 Bonn
Germany
E-mail: gielen@iza.org

^{*} We would like to thank Gerard van den Berg, Mark Duggan, Amy Finkelstein, Paul Frijters, Pierre Gielen, Ed Glaeser, David Johnston, Pierre Koning, Peter Kuhn, Ellen Meara, Jon Skinner, Doug Staiger, and Frank Vella, for helpful conversations. We thank seminar participants at IZA, Tilburg University, Maastricht University, University of Milan, the NBER Public Economics meetings, SciencesPo, University of New South Wales, Queensland University, the CPB, Erasmus University of Rotterdam, the University of Chicago, and the Tinbergen Institute, as well as participants at the IZA/IFAU, LEW, EALE, IZA/CEPR, and IZA/SOLE conferences for insightful comments. We would like to especially thank Bas ter Weel, who greatly contributed to the conception of this paper. All errors are our own.

1. Introduction

Because estimates of labor supply responses are of tremendous policy relevance, the literature on the effects of social assistance on behavior has rightly focused on labor supply responses to changes in the eligibility criteria or generosity of social assistance programs. These estimates, however, do not capture two additional policy-relevant dimensions of the response to changes in social assistance programs. First, they do not capture potential spillover effects to other social assistance programs that arise when individuals substitute between programs. Such social support substitution may decrease the reduced-form labor supply response to changes in generosity of a particular program when individuals take up other programs instead of adjusting their labor supply. Similarly, social support substitution reduces the welfare impact of reductions in generosity of any given program on recipients of that program. Evidence on the extent of social support substitution is also important for policy makers because it allows them to make more accurate predictions of the budgetary impact of a reform to a social assistance program by taking into account the spillover effects of the reform on participation in other programs. Second, existing estimates of supply responses generally do not distinguish between the responses by existing claimants and responses by (potential) new enrollees. Evidence of labor supply responses among existing long-term DI recipients is of great importance for policy reforms because effects that operate on the existing stock of recipients have the potential for a much greater immediate impact than reforms that operate on the comparatively small inflow into a DI program.

In this paper, we estimate the extent of social support substitution in response to reforms in disability insurance in the Netherlands as well as the reduced-form labor supply response to this reform among existing DI recipients. Together, this allows us to estimate the impact of the reform on the total income of affected DI recipients. The reforms to the Dutch DI system, which also insures against partial loss of earnings capacity, entailed medical re-examinations of existing recipients. A common consequence of the re-examination was a change in the benefit level. Two features make this reform particularly suitable for studying substitution between different social assistance programs as well as labor supply responses among current recipients. First, we have administrative panel data on the universe of Dutch disability insurance claimants,

including information on their future labor market earnings and their future income from all other government cash social assistance programs. These data allow us to track for a period of nearly a decade what happens to (former) disability insurance claimants in the wake of the reform. Second, the reform contains a cohort discontinuity: the reform was significantly more stringent and led to an average benefit reduction of an additional 10% for the cohort that turned 45 after August 1st, 1993. Because we have each individual's month of birth, we exploit this discontinuity by comparing later labor market earnings and social assistance income for the cohort just below this age cutoff to outcomes for the cohort just above the age cutoff. We scale this difference in outcomes by the discontinuity in disability benefit levels around the age cut-off. This yields two key ratios: (i) the benefit-substitution ratio, which is the average causal effect of the more stringent DI rules on income from other social assistance as a fraction of average lost DI income, and (ii) the earnings crowd-out ratio, which is the average causal effect of the more stringent DI rules on earnings as a fraction of average lost DI income.

We find that, in the short term (about 2 years after reform), the more stringent DI rules increase the probability of receiving any income from other social assistance programs by 5 percentage points (on a base of 14 percent), and the income from these other social assistance programs replaces 31% of lost DI income. In other words, we find a substantial amount of social support substitution with a short-term benefit-substitution ratio of 0.31. The more stringent rules increase the probability of having any labor market earnings by 3 percentage points (on a base of 35 percent) and also increase labor earnings on average. The additional earnings replace 62% of foregone DI income, i.e., the earnings crowd-out ratio is 0.62. Recipients classified as fully disabled have a crowd-out ratio of 0.52, which is novel evidence of a substantial labor supply response among long-term fully disabled individuals. Combining the effects of social assistance substitution and earnings crowd out, we find that individuals are able to replace basically all of their foregone DI income on average. Specifically, we cannot reject that individuals fully offset the cut in their DI benefit by increased income from other social assistance programs and labor earnings.

Over time (up to 8 years after the re-examinations for our cohorts), the benefit-substitution ratio falls somewhat and the earnings crowd-out increases slightly, but these

trends are not significant in either ratio. Even in 2005, or about 8 years after the implementation of the reform, the benefit-substitution ratio is still a statistically and economically significant 0.20 and the earnings crowd-out ratio stands at 0.71. Spillovers between social assistance programs may operate not only through former DI recipients' own choices, but could potentially also operate through the decisions of their spouses. However, we do not find any statistically significant evidence of responses by spouses in terms of labor supply or social assistance receipt. The point estimates indicate that spouses increase their labor supply but do not change their social assistance receipt. This implies that if we measure earnings crowd out and benefit substitution at the household level rather than at the individual level, we find a slightly (about 18 percentage point) higher earnings crowd-out ratio and a similar benefit-substitution ratio.

While the precise magnitudes of our findings are obviously specific to this particular Dutch disability insurance reform, we believe our paper offers important lessons that are widely applicable. First, our evidence demonstrates that social support substitution occurs at an economically meaningful scale for prime-age disability insurance recipients. Hence, a carefully designed reform of a social assistance program needs to take into account its effects on other social assistance programs. Second, our findings show that even long-term disability insurance recipients can still exhibit a meaningful rebound in their labor supply. Third, to measure the full impact of social insurance reforms on labor supply and the reliance on other forms of social insurance, it is important to also consider effects over the longer term and to take possible behavioral responses of spouses into account.

Our findings on the existence of spillover effects between different social assistance programs confirm earlier results from other contexts.¹ With respect to child-related benefits, Garrett and Glied (2000) show that the increase in child Supplemental Security Income (SSI) eligibility in the early 1990s led to a greater increase in SSI enrollment in states with less generous benefits for Aid to Families with Dependent

¹ The one exception is a paper by Autor and Duggan (2008), in which they exploit a ruling that suddenly expanded the eligibility for Veterans' Disability Compensation (DC) for a subgroup of Vietnam Veterans. They find that the increased take-up of Veteran's Disability Compensation due to this ruling raised the receipt of Social Security Disability Insurance (SSDI) benefits. As Autor and Duggan note, this result may be explained by the fact that one needs leave the labor force to qualify for SSDI, and leaving the labor force is less costly for people who already receive DC. Thus, this institutional feature may explain the complementarity between two social assistance programs in this case.

Children (AFDC) while Schmidt and Sevak (2004) show that states that reduced the generosity of their AFDC program experienced increases in SSI enrollments. Both studies suggest that families use SSI and AFDC as substitutes. Kubik (2003) shows that the substitution of SSI for AFDC is larger in states with negative fiscal shocks, suggesting that states actively encouraged this substitution (because the state-share in SSI payments is generally lower than the state-share in AFDC payments). Duggan and Kearney (2007) examine individual-level panel data to find that households in which a child becomes eligible for SSI subsequently receive less income from AFDC, WIC, and food stamps.² With respect to early retirement, Duggan, Singleton, and Song (2007), Li and Maestas (2008), and Coe and Haverstick (2010) exploit differences by cohort in the generosity of Social Security retirement benefits to show that the reduction in the generosity of Social Security retirement benefits led to increases in applications for or receipt of Social Security disability benefits. Koning and Van Vuuren (2010) use administrative data to describe program enrolment after dismissal, and find that DI substitutes for UI but fail to find evidence that UI substitutes for DI. Karlström, Palme, and Svensson (2008) use a difference-in-differences design to examine the effect of the abolition of DI as a path to early retirement for 60-64 year olds in Sweden. They find that, in the 2-3 years following the reform, this group responded by taking up other forms of social assistance rather than by increasing their labor supply. Finally, Staubli (2011) also uses a difference-in-differences approach to show that a disability insurance reform that affected 55-56 year-old males in Austria had spillover effects on their take up of unemployment insurance and sick leave, in addition to affecting their labor supply. Our paper contributes to this literature by estimating substitution between social support programs for prime-age individuals, and doing this in a setting that allows us to very cleanly identify the degree of spillovers between programs. In addition, we extend the literature by examining substitution effects over longer horizons (up to eight years after the re-examinations for our cohorts took place).

Our estimates also contribute to an extensive literature on the labor supply effects of disability insurance (see Bound and Burkhauser 1999 for an overview) by showing well-identified effects on the labor supply of existing long-term DI participants. Parsons

² WIC provides nutritional assistance to low-income families with young children and pregnant women.

(1980) shows cross-sectional evidence that suggests that the rise in DI generosity has contributed to the decline in male labor force participation. Gruber (2000) exploits a natural experiment in Canada and finds a sizeable labor force participation response among older workers to the generosity of DI benefits. Much of the more recent work in the U.S. on labor supply effects of DI compares accepted to rejected DI applicants. Bound (1989) and Von Wachter, Song, and Manchester (2011) directly compare accepted and rejected applicants, and those estimates are probably upperbounds of labor supply effects because there are likely unmeasured determinants of the rejection decision that are correlated with labor supply. To get around this issue, other studies use plausibly exogenous variation in rejection rates. Gruber and Kubic (1997) use variation across states and time in rejection rates, Chen and Van der Klaauw (2008) use an age discontinuity in rejection rates for a particular subgroup, De Jong, Lindeboom, and Van der Klaauw (2010) use variation in screening severity induced by an experimental intervention, French and Song (2011) use variation in rejection rates due to the essentially random assignment of administrative law judges to DI cases, and Maestas, Mullen, and Strand (2011) use variation in rejection rates due to the essentially random assignment of DI examiners to DI cases. These studies all find clear evidence of labor supply responses to disability insurance. Autor and Duggan (2003) exploit the interaction of state disability replacement rates with national changes in program stringency to find credible evidence that increased DI generosity reduced labor force participation of high school dropouts. Finally, Maestas and Song (2011) use the automatic conversion of DI into Social Security benefits to show that there is a labor supply effect of DI on older disability insurance recipients. Our study contributes to this literature by showing the labor supply responses not only occur at the point of the initial eligibility determination, but that there are strong labor supply responses to changes in DI generosity even among prime-age long-term disability insurance recipients, including individuals classified as fully disabled.

The remainder of the paper is organized as follows. In Section 2, we describe the reform in the Dutch disability act that we use for identification in this paper. Section 3 describes the data, and Section 4 presents the results. Section 5 concludes.

2. The 1993 Dutch Disability Insurance Reform

After a waiting period of one year, individuals in the Netherlands are entitled to disability benefits if an illness or infirmity prevents them from earning the amount they used to earn before the onset of the disability.³ The replacement rate offered by DI depends on the “degree of disability,” which is defined by the percentage difference between the prior earnings and the remaining potential earnings capacity of the DI applicant. In other words, the Dutch DI system also covers partial disability. Health insurance is not linked to DI in the Netherlands, so this is not a consideration for DI participation.

In order to explain the 1993 DI reform, we first describe how the Dutch disability insurance system determined eligibility and replacement rates prior to the reform. Prior to the reform, the potential earnings capacity was determined by the following procedure. First, a medical doctor examined the applicant and compiled a list of work activities that, according to the doctor’s judgment, the applicant could still perform.⁴ Second, using a dictionary of occupations that specified for each occupation the required education level and work activities, a list of occupations that an applicant could still perform was compiled, but occupations that were more than two “education levels” (on a 7-level scale) below the education level required for the applicant’s previous occupation were not considered. Finally, if the list contained at least 5 suitable occupations with at least 10 active workers (though not necessarily vacancies) in the applicant’s region⁵, then the mean wage of the 5 highest paying occupations on the list was taken as the applicant’s potential earnings capacity. The loss of earnings due to the disability, measured by the difference between the prior labor earnings and the potential earnings capacity, determined the degree of disability. If it was not possible to specify 5 suitable occupations with at least 10 workers, the degree of disability was set at 100%. The

³ Also see Bovenberg (2000), who provides useful institutional background information on the Dutch disability act. See García-Gómez, Von Gaudecker, and Lindeboom (2011) for further background information and descriptive evidence on DI enrollment trends and patterns.

⁴ The list includes 27 physical activities (such as “lifting,” “kneeling,” and “ability to deal with temperature fluctuations”) and a list of 10 psychological abilities (such as “ability to work under time pressure,” “ability to perform monotonous work,” and “ability to deal with conflict”).

⁵ The Netherlands was divided up in 5 regions and in 16 “start regions.” Alternative jobs had to be found in the “start regions” first. Only if none were available, the labor market expert could look for jobs in the neighboring regions (within one of the main 5 regions).

measured degrees of disability were grouped into 8 categories varying from 0-15% to 80-100%, and these categories determined the replacement rate (see Table 1). The replacement rate was applied to the individual's indexed previous earnings, where the previous earnings are subject to cap (about €36,000/year in 1999). Individuals on DI have an earnings exemption equal to their capped indexed previous earnings times the degree they are not disabled (which is set at 100% minus the lowerbound of their degree disabled category). Earnings above the exemption lead to a one-for-one reduction in DI benefits in the short term, and to a reclassification in the degree disabled in the longer term (typically after about 3 years).

The DI reforms of 1993 tightened this procedure in two respects.⁶ First, the determination of disability had to be based on objective medical information (rather than just the doctor's judgment). In other words, the applicant needed to have a clearly observable functional work limitation, and a direct relationship between the functional work limitation and the medical diagnosis had to be plausible. Disabilities due to mental health problems became more difficult to prove than physical health problems. Second, the criteria for the list of suitable alternative occupations were relaxed: (i) occupations more than two "education levels" below the applicant's education level were included from now on, (ii) the list only needed to contain 3 suitable alternative occupations (rather than 5), and (iii) the geographic region in which these occupations had to exist with at least 10 active workers was expanded roughly threefold.⁷ With these relaxed criteria, it became more likely to find higher-paying alternative occupations and less likely that the list would not contain at least the minimum number (now 3) of occupations needed to avoid declaring the applicant as fully disabled. By changing the criteria for what constituted suitable alternative employment, the reform aimed to lower the generosity of disability benefits and to reduce the number of claimants.⁸

⁶ The formal name of the 1993 UI reforms is "Terugdringing Beroep op Arbeidsongeschiktheidsverzekeringen (TBA)," which roughly translates as "Reducing claims on disability insurance."

⁷ Now all available jobs within the main region where the individual was residing (out of 5 main regions) could be used to calculate the potential earnings capacity, rather than just jobs in one of the 16 "start regions."

⁸ Another important change of the 1993 DI reform was the introduction of an age- and duration-dependent benefit for new applicants. To those already receiving disability benefits as of August 1993, i.e. the group that we are studying here, these changes did not apply and the benefit level remained a function of the indexed previous earnings.

The new procedure for determining benefits was applied to new DI applicants as well as to existing DI claimants who were 50 or younger at the time the reform went into effect (August 1st, 1993). Because re-examinations of existing claimants are time consuming, these re-examinations were scheduled to take place by cohort over a period of several years. Disability claimants who were age 34 or younger on the 1st of August 1993 were re-examined in 1994, the 35-40 year-old cohort in 1995, the 41-44 year-old cohort in 1996/1997, and the 45-50 year-old cohort were to be re-examined in 1997-2001. However, shortly before the re-examinations for this latter cohort started, political pressure led the government to decide to that the 45-50 age cohort would be re-examined based on the previous and more generous procedure for determining replacement rates rather than the new and more stringent procedure. While re-examinations tended to lower benefits, this was not necessarily the case for every claimant. Some disability claimants saw their benefits rise, for example because their medical condition had deteriorated. However, because the new procedure is more stringent in all respects, the benefit determined under the new procedure is weakly lower than what it would have been for that particular individual if the old procedure had still been used.

3. Data

3.1 Data sources

This paper relies on administrative data that Statistics Netherlands has assembled from several sources. Information from these various sources is merged at the individual level by using a so-called RIN-number (which is a coded version of the Dutch equivalent of the U.S. Social Security number).⁹

First, we have administrative data on all disability benefits recipients aged 15-64 in the Netherlands for the period 1995-2005. The data were collected by the organizations responsible for administering disability benefits. The information from these administrative records include the start and end dates of a disability spell, the degree of disability (in categories), disability benefit payments, previous earnings, industry information, and the reason for the termination of the disability spell, but does

⁹ These data can be accessed via a remote-access computer after a confidentiality statement has been signed.

not contain reliable or consistent information about the medical condition that gave rise to the disability spell.

We obtain the demographic characteristics of the disability claimants from the municipal registries (“GBA”), which contain all residents of the Netherlands. This database includes information on each person’s month and year of birth, marital status, number of children, national origin, and place of residence, as well as the identification numbers (RIN-codes) of their partners. The RIN-codes of the partners allow us to include income sources of the partner. We collect these demographic characteristics as of January 1996, which is the start of our sample period.

Finally, we obtain information on labor market earnings and sources of social assistance income other than DI by merging five administrative datasets: earnings of all employees, self-employment earnings, unemployment benefits (“WW”), general assistance (“Bijstand”), and receipt of any other form of social assistance (from about 30 relatively minor programs). Data on social assistance come from the organizations that administer these programs. Information about the earnings from paid labor and self-employment are gathered by Statistics Netherlands using information from the tax authorities and social insurance records. All these files are available from 1999 onwards, which is why 1999 is the start year for our empirical analysis. Unemployment insurance covers any income loss due to unemployment for a duration of up to 5 years, where the duration depends on one’s work history. General assistance is unlimited in duration and does not require dependents (unlike the U.S. welfare program), but is means tested. Apart from the programs mentioned here, there are no additional cash social assistance programs in the Netherlands that are relevant for individuals in the age range of our sample. Exact variable definitions are provided in Appendix B.

3.2 Sample definition

In our baseline analysis, we restrict the sample to all individuals who (i) received disability benefits on August 1st, 1993, (ii) who were between the ages of 42.5 and 47.5 at that date, and (iii) who were still on DI as of January 1st, 1996. The first restriction is necessary because the discontinuity in benefit rules only applies to existing claimants on the date the reform went into effect. The second restriction limits the sample to those

who are close to age 45, where the discontinuity in benefit rules occurs. We selected this bandwidth based on the criterion by Imbens and Kalyanaraman (2009).¹⁰ The last restriction is driven by data availability. While our data on disability starts in 1995, the information in the 1995 file does not contain previous earnings, so we use the files from 1996 onwards instead. Thus, we can only observe individuals who were on disability at the time of the passage of the reform legislation if they remained on disability until January of 1996 or later. We believe it is highly unlikely that differential attrition occurred around the age discontinuity prior to January 1st of 1996 because the re-examinations for the individuals in our sample did not start until later in 1996 and the government decided only in 1997 that the age 45+ cohort would not be subject to the new, more stringent criteria.¹¹ A plot of the density of disability claimants by cohort is relatively smooth and therefore gives no indication of differential attrition (see Appendix Figure A1). The plot further indicates that heaping-induced bias (Barreca, Lindo, and Waddell, 2011) is not a concern and that there is no discontinuity in the density around the cutoff age of 45 (p-value of 0.126 in the McCrary, 2008, density test). Our data extends until 2005 because in 2006 there was a fundamental reform to DI.

We exclude all individuals that appeared on more than one disability record in our data in a given month (about 3 percent of the sample). We exclude these observations because it is not clear whether they reflect administrative/coding errors or whether they truly concern individuals who are entitled to two (or more) different disability insurance benefits because they were employed in two (or more) jobs before they became disabled. In the latter case it is hard to understand why we observe that in many of these cases there has been a health improvement (i.e., a reduction in the degree of disability) during a particular period for one of the benefit claims, but not for the other. We have checked that no discontinuity occurs at age 45 in the likelihood that an individual has more than one disability record, and are therefore not concerned that the omission or inclusion of

¹⁰ The Imbens-Kalyanaraman criterion yields different optimal bandwidths for different outcome variables. Rather than changing the sample for each outcome variable, we selected a bandwidth in the middle of the optimal bandwidths suggested by the Imbens-Kalyanaraman criterion, and applied this bandwidth to all our specifications. Below, we show that our key results are robust to using different bandwidths.

¹¹ Please note that we refer to cohorts by their age at the time the reform went into effect (i.e., as of August 1st, 1993).

the 3 percent of observations with multiple records would substantively affect our results. After these sample restrictions, our baseline sample contains 84,185 observations.

3.3 Summary statistics

Table 2 presents summary statistics for our key variables. Panel A shows the characteristics of our sample measured as of January 1996, i.e. before the re-examinations took place. About a third of disability claimants are female and about two thirds are married. The average DI spell started in 1985. So when the reform was implemented in 1996-1998 for the cohorts in our sample, the average claimant in our data had been on DI for more than a decade. Finally, about two thirds of the sample is classified as fully disabled (earnings capacity reduction of 80% or more) and is therefore eligible for a replacement rate of 70%. The fraction fully disabled is markedly higher among females than among males. Only about 4% of the sample is considered to have lost between 55% and 80% of their earnings capacity. The remaining 30% of the sample is considered to have lost between 15% and 55% of their earnings capacity and is eligible for replacement rates between 14% and 35%.

Panel B presents the means of our key outcome variables. While we have these variables for all years from 1999 through 2005, we only present the values for 1999 and 2005 in the interest of space. In 1999, so about 1 to 3 years after the re-examinations took place for the age cohorts in our sample, 92% of those on DI at the start of 1996 were still on DI, where being on DI in 1999 is defined as having received positive income from DI in 1999. In short, the re-examinations cannot have had a dramatic effect on DI participation, though our next section will show evidence of a clear discontinuity in exit rates around the age cutoff. About 36% percent of our sample was working, defined as having positive earnings (including from self-employment) in 1999, which is consistent with DI also covering partial disability in the Netherlands. Of those who had left DI, 53% were employed, whereas 33% of those on DI were employed, so a considerable number combined DI receipt with work. The fraction of men working (45%) is more than twice as high as the fraction of females with positive labor earnings (18%). Sixteen percent of our sample also had social assistance income (other than from the original DI spell) in 1999. Another four percent are not observed in any of our administrative files.

Most of these individuals did not have any formal labor or social assistance income in 1999 but about a third of them died or emigrated during our sample period.¹²

The average income in our sample is about €17,000, of which roughly two-thirds comes from DI benefits with the remaining third coming mostly from labor earnings. Income from other social assistance programs accounts only for about 6% of total income.

In 2005, so about 7 to 9 years after the re-examinations, 81% of those on DI at the start of 1996 were still on DI. Between 1999 and 2005, the fraction employed fell from 36% to 28% and the fraction with income from social assistance other than from the original DI spell increased from 16% to 25%. These trends are consistent with the general decline in labor force participation in the Netherlands as people approach retirement. In 2005, still about two-thirds of total income in our sample comes from DI benefits.

4. Results

4.1 Magnitude of the reform

To what extent did the more stringent re-examinations reduce the generosity of the DI program for the under-45 cohort? The answer to this question allows us to interpret the magnitude of the effects of the reform on earnings and on receipt of other forms of social assistance. Figure 1 shows three measures by which to gauge the magnitude of the reform: the effect on benefit amounts, the effect on replacement rates, and the effect on participation in the DI program.

Panel A plots annual disability benefit amounts in 1999, including zeros for those who have exited, by cohort. Visually, there is a clear discontinuity at the cutoff age. We estimate the size of this discontinuity by running an OLS regression of the outcome variable on a treatment dummy that equals 1 for cohorts subject to the more stringent re-examinations (so age less than 45), a linear term in age, and an interaction of the treatment dummy with (age-45). All ages are specified as ages as of 8/1/93, so they

¹² About 1.1% of our sample had died by 1999 and 0.3% had emigrated. These observations are included in the main analysis and their income and participation variables are all set to zero in 1999. Results are extremely similar if we exclude these observations altogether. The more stringent reforms had no significant impact on the probability of death or the probability of emigration.

effectively measure cohorts. We run this specification for all RD estimates reported in the figures in the paper. The fitted regression line is indicated in the plot, and the RD estimate on the treatment dummy is -1.076, indicating that the more stringent re-examinations for the younger cohort reduced their annual DI benefits by €1076, or about 10%. All reduced-form RD estimates can be seen as local average treatment effects – local in the sense that they only apply at the discontinuity (so for the cohort being exactly 45 as of 8/1/93) and average in the sense that it is the average effect for all those at the discontinuity. The key identifying assumption behind all our RD estimates is that the only discontinuous change at the age cutoff is the stringency of the DI re-examinations. While we cannot test this assumption, we know of no other policy changes that would create a discontinuity at this cutoff. Further, when we run our reduced-form RD specification using all our demographic characteristics as dependent variables, we only find two significant coefficients, in line with what would expect under the null hypothesis of no effect when running 39 placebo regressions.¹³

Panel B shows that the replacement rate, including zeros for those who exited, is 5.9 percentage points lower for the affected cohort at the discontinuity.¹⁴ The average replacement rate for those who just escaped the more stringent re-examinations is 0.55, so the 5.9 percentage point drop represents an 11 percent decline. Panel C shows that the fraction of the sample that is still on the original DI spell in 1999 falls discontinuously by 3.8 percentage points at the age cutoff. Overall, Figure 1 shows that the more stringent re-examinations roughly translate into a 10% benefit reduction. The effects of the reform on labor supply and other benefit receipt should be viewed in light of this magnitude.

The reforms led to somewhat larger reductions in benefits and replacement rates for men (reductions by 12%) than for women (reductions by 7%), but induced 6.1 percent

¹³ The demographic characteristics are measured as of January 1996, i.e., before the implementation of the reform for our sample, and therefore should not contain a discontinuity. Reassuringly, only two of the 39 estimates are significant at the 5-percent level. This is roughly what one would expect by pure chance given that, even if there is no true effect, one out of every 20 regressions on average shows a coefficient that is significant at the 5-percent level. In addition, there is one coefficient that is significant at the 10-percent level. The full results are presented in Appendix Table A1. It would be instructive to do similar checks on the identification strategy with our key outcome variables: labor income and income from other social assistance programs. Unfortunately, we do not have data on these variables prior to 1999.

¹⁴ The data do not contain the post-reform replacement rate for those who exited from DI. Based on discussions with the DI administration, our impression is that most exits occurred for those who were no longer eligible for DI, but we cannot rule out that some of those who exited were still eligible for a positive replacement rate.

of female recipients to exit but only 2.7 percent of male recipients. It is not clear what exactly drives these differences, but many factors (types of jobs, types of disabilities, outside options) obviously differ by gender, and it is therefore plausible that the reform had a differential impact by gender. In light of this differential impact, we will split out our key results by gender.

Because we do not have income measures from before the reform, we cannot estimate effects of the reform on income from labor and from other social assistance programs separately for those staying on DI and for those leaving DI. However, we do have the pre-reform replacement rate, which allows us to examine heterogeneity in the effects of the reform on replacement rates.

The first column of Table 3 shows the change in replacement rates at age 45.0 for those subject to the less stringent re-examination.¹⁵ About 72% of this group saw no change in their replacement rate, 12% experienced an increase in their replacement rate, and 16% faced a decrease in their replacement rate. The second column shows the change in the replacement rates at age 45.0 for those who underwent the more stringent re-examination. A much larger fraction (29%) in is latter group experienced a reduction in the replacement rate, and a much lower fraction (5%) saw an increase in the replacement rate. Still, even in the group subject to the more stringent re-examination, about two thirds experienced no change in the replacement rate. The third column shows the treatment effect of the more stringent re-examination on the change in the replacement rate, which is simply the difference between the first two columns. This column shows a downward shift in probability mass throughout the distribution of changes in replacement rates, showing that the re-examination made DI less generous for each counterfactual change in replacement rates.

¹⁵ We run our standard reduced-form RD specification using dummies for each possible change in the replacement rate as the dependent variables. We estimate changes in the replacement rate at age 45.0 for those subject to the less stringent re-examination as the intercept of the right segment of regression line at age 45.0. Changes in the replacement rate for those subject to the more stringent re-examination are estimated by the intercept of the left segment of regression line at age 45.0. More detailed estimates are provided in Appendix Table A2, which presents the joint distribution of the replacement rate in 1996 (pre reform) and 1999 (post reform) at age 45.0 for those who were subject to the less stringent re-examination rules and well as the impact of the reform on these joint probabilities (estimated by the RD effect). The estimates in Table 3 summarize these joint probabilities by presenting the sums of the diagonal entries of the Appendix Table.

As an additional partial check on our identifying assumption that no factors besides the DI re-examinations had a discontinuous impact at age 45, Figure 2 reports DI exit rates separately for 1995, 1996/97, 1998, and 1999. Exit is defined as the end date of the original DI spell, as recorded in the administrative data file, occurring during the year in question. We calculate these rates as fractions of DI claimants in our sample on January 1st, 1996.¹⁶ Since no re-examinations took place in 1995 for DI claimants in the age 40+ cohorts, a discontinuity at age 45 in the 1995 exit rate would invalidate our identifying assumption. Reassuringly, the 1995 exit rate shows no sign of a discontinuity at age 45. In 1996 en 1997, the age 40-44 cohort was re-examined as well as part of the age 45 cohort. Exactly in these years, the discontinuity at age 45 is very pronounced. In 1998, the remainder of the age 45 cohort and some of the age 46 cohort were re-examined, which explains the statistically significant discontinuity in the opposite direction. This discontinuity, however, is much smaller in size because the age 45+ cohort was re-examined under the old and less stringent standards. Hence, if we calculate the total exit rate over the 1996-1998 period, we find a discontinuous increase in exit for the group subject to the more stringent re-examinations. In 1999, all the re-examinations for the age 44-45 cohort were completed and we find no discontinuity in exit rates at the age cutoff.

4.2 Reduced-form impacts on labor market and social assistance outcomes

To what extent did individuals whose DI was reduced by the reform end up in other social assistance programs and to what extent did they find paid work? The answer to this question is critical for judging the effectiveness of the reform. In the former case, the reform merely shuffles individuals across programs and budgetary savings only occur to the extent that benefits in other programs are lower than DI benefits. In the latter case, increased earnings are an indication of moral hazard among existing disability recipients. In this subsection, we examine labor market and social assistance outcomes in 1999, which is the first year for which we have the required data, and which is about two years after the re-examinations took place. In subsection 4.6, we will examine the effects over a longer horizon.

¹⁶ However, the exit rate for 1995 is calculated as a fraction of DI recipients as of January 1st, 1995.

We start by analyzing the reduced-form effects of the DI reform on receipt of other forms of social assistance. The first panel of Figure 3 plots income from other social assistance (including income from new DI spells, but excluding income from the original DI spell) by cohort. The figure shows a clear upward jump in income from other social assistance for the cohort that underwent the more stringent re-examinations. In fact, the RD regression estimates that the more stringent re-examinations increased other social assistance income by €314 per year. The second panel shows that the fraction receiving any social assistance income from a source other than the original DI spell discontinuously increases by 4.7 percentage points at the age cutoff for the more stringent re-examinations. Both increases are highly significant and represent an increase of about a third in the amount and in the participation rate. In other words, we find clear evidence of substitution of other forms of social assistance for DI benefits.

Do people who leave DI fully account for the increased income from other social assistance or is some of the increase also caused by those who remain on DI receiving higher amounts? To answer this question, we re-run the RD regression on the subsample of those who left DI. The regression shows that income from other social assistance jumps by €601/year at the age discontinuity; the regression line just to the right of the age cutoff (where the re-examination was less stringent) lies at €1962 of other social assistance income per year, whereas it lies at €2564/year just to the left of the age cutoff (where the re-examination was more stringent). Some of those who exited are “inframarginal” leavers – they would have left DI even under the less stringent re-examination. These inframarginal leavers were by definition not affected by the increase in stringency of the re-examination at the age discontinuity, and we therefore assume that other social assistance income is equal to €1962/year on both sides of the age cutoff for inframarginal leavers. The participation rate just at the right of the age continuity is 93.2% (see Figure 1c), so 6.8% are inframarginal leavers. Figure 1c also showed that 3.8% are marginal leavers, so 35.8% ($=3.8/(6.8+3.8)$) of all leavers are marginal leavers and 64.2% are inframarginal leavers. This implies that income from other social assistance for marginal leavers must have been €3644 per year so that the weighted average of all leavers to the left of the discontinuity is the amount we found above, namely €2564 (so $0.358*3644 + 0.642*1962 = 2564$). If marginal leavers receive €3644

per year from other social assistance, the maximum amount by which their income from other social assistance could have possibly increased due to the more stringent re-examination is €3644. Given that they comprise only 3.8% of the of the entire sample (so both stayers and leavers) at the age cutoff, they can at most account for an increase of $0.038 * 3644 = €138$ in the RD estimate of €314/year that we found for the entire sample. Thus, we conclude that those leaving DI altogether can at most be responsible for 44% ($=138/314$) of the overall jump in income from other social assistance, and that at least 56% must be due to increases in other social assistance income by those remaining on DI.

Next, we present the reduced-form effects of the DI reform on labor market outcomes. The first panel of Figure 4 plots labor earnings, including self-employment income, in 1999 by cohort. The figure shows a discontinuity in earnings at the cutoff age but the discontinuity is not as visually compelling as in the earlier figures due to the higher variance in earnings. However, the RD regression estimates that earnings are €624 per year higher at the cutoff age for those who were subject to the more stringent re-examinations, and this estimate is highly significant. The €624 increase represents an 11 percent increase in annual earnings. This figure establishes our qualitative finding that disability income crowds out labor income. It also contributes to the literature on the labor supply disincentive effects of disability insurance by showing novel evidence of labor supply responses among prime-age DI recipients who are long-term recipients of DI (with durations of at least 2 years at the time of the reform, but on average 10 years). We will discuss the economic magnitude of the labor supply response in the next subsection.

Because we do not have earnings for prior years, we cannot precisely determine the extent to which the average increase in earnings stems from non-workers finding employment (extensive margin) and from workers increasing their earnings (intensive margin). However, at least some of the increase comes from the extensive margin because the second panel of Figure 4 shows a clear discontinuity in the fraction of individuals with strictly positive income from wages or self-employment. The RD regression estimates that the more stringent re-examinations caused the fraction working to increase by 2.9 percentage points. To explain the observed increase in earnings in the absence of an intensive-margin labor supply response, average earnings for those who started working would need to be €21,500 ($=624/0.029$) per year, which is higher than the

observed average earnings for those with positive earnings (€17,000/year). It therefore seems likely that some of the response also occurred along the intensive margin. We can also do a bounding calculation similar to the one we did above for income from other social benefits. We find that those who left DI altogether can at most be responsible for 61% of the overall jump in earnings caused by the more stringent re-examination, and that at least 39% must be due to earnings increases among those remaining on DI.

4.3 Benefit Substitution and Earnings Crowd Out

Figures 3 and 4 establish that people substitute between DI income and other forms of social assistance, and that DI benefits crowd out labor income. We now turn to the economic magnitudes of earnings crowd out and substitution of social assistance. In the first column of Table 4, we scale our reduced-form estimates by the amount by which disability benefits from the original spell decrease at the age discontinuity whereas in the second column we scale by the discontinuity in the replacement rate. We implement this scaling by running IV regressions following the standard “fuzzy RD” specification.¹⁷ We include a rich set of demographic control variables to increase the precision of the estimates. As should be the case with a valid RD design, the control variables do not substantially affect the magnitudes of our estimates (See Appendix Table A3 for the corresponding estimates without controls). Given that the re-examination was in all respects more stringent for those below the cutoff age, the monotonicity assumption required for the fuzzy RD design should be satisfied; being subject to a re-examination following the more stringent new protocol rather than the old protocol weakly decreases the benefit amount for any given individual and weakly decreases the replacement rate for any given individual. We do not interpret these IV estimates as causal impacts of the level of DI benefits per se or as causal estimates of the DI replacement rate per se because at the age discontinuity both the level of benefits and the earnings exemption change. Rather, we see the IV estimates as a way to relate the magnitudes of the behavior effects of the reform to alternative measures of the size of the reform. In other words, we view this solely as a scaling exercise.

¹⁷ Excellent discussions of the theoretical underpinnings and the practical application of RD methods can be found in Hahn, Todd, and Van der Klaauw (2001), Imbens and Lemieux (2008), and Lee and Lemieux (2010).

Panel A of Table 4 examines to what extent the reduced generosity of disability benefits induces individuals to shift to other forms of social assistance (including new DI spells). The first row of column 1 of panel A shows that per euro reduction in disability benefits, individuals receive €0.31 more from other social assistance programs in 1999. Thus, the benefit-substitution ratio is 0.31. A government not taking this substitution into account would overestimate the reduction in government expenditure from tightening the DI eligibility rules by 44%. The second row shows that per €1000 per year decrease in DI benefits caused by the more stringent rules, the probability that an individual receives income from another social assistance program increases by 4.5 percentage points. An alternative way of scaling the degree of substitution between social assistance programs is provided in the second column, which shows that for a 10-percentage point reduction in disability replacement rates, income from other social assistance programs increases by €535 per year (an increase of more than 50%) and the probability of participation in other social assistance programs increases by 8.0 percentage points. The estimates of panel A establish that benefit substitution is not only statistically significant but also important in economic terms.

Benefit substitution can occur mechanically when individuals automatically receive more income from other social assistance programs as their DI benefits decrease. While this might explain some of the substitution, it cannot account for the entire reaction because we also observe individuals enrolling in other forms of social assistance, for which enrollment is not automatic. Benefit substitution can also be a result of the individual actively looking for alternative sources of benefits and trying to qualify for them. Finally, benefit substitution can occur when caseworkers steer individuals towards alternative sources of support. We have no direct evidence on the relative importance of these three channels, and suspect that all three may have contributed to some degree to the observed amount of benefit substitution.

The estimate in the first row of column 1 of panel B indicates that per euro of benefits decrease caused by the reform, the reform induced individuals to increase earnings by €0.62 in 1999. In other words, we find an earnings crowd-out ratio of 0.62: a

euro of DI benefits crowds out 62 cents of labor earnings.¹⁸ Alternatively, one can scale the change in earnings by the change in total benefits (including the original DI benefits) due to the more stringent re-examination. Given that for each euro in decreased DI benefits, other benefits went up by €0.31, total benefits only decreased by €0.69. Thus, per euro decrease in total benefits, earnings went up by €0.90 ($=0.62/0.69$). The second row of Panel B examines the extensive margin response and shows that, per €1000 of disability benefits decrease caused by the reform, the probability of being employed in 1999 increases by 2.9 percentage points. The second column presents the analogous estimates, but now scaled by the change in replacement rates caused by the more stringent re-examinations. We find that for a 10-percentage point decrease in replacement rates, earnings increase by €1085 per year (or about 19 percent) and the probability of employment increases by 5.1 percentage points. All four estimates in panel B are highly statistically significant and establish that the degree to which DI benefits crowd out labor market earnings and participation is economically meaningful.

In the late 1990s, the Netherlands experienced an economic boom, which likely made it relatively easy for individuals to increase their labor supply and reduced the incentive to look for other forms of social assistance. As a result, our estimate for earnings crowd out may be higher than it would be during average economic times, and the estimate of benefit substitution may be lower than it would be during average economic times. During the early 2000s, the Netherlands experienced a recession, and as we will see in Section 4.6 below, our estimates of earnings crowd out and benefit substitution remained very similar during that period. A further reason why the earnings crowd out estimate may be relatively high is that the Netherlands had a variety of policies to help individuals re-integrate into the labor market.

DI recipients have an earnings exemption that equals their indexed previous earnings times the degree to which they deemed able to work (i.e., one minus the degree disabled). Any earnings beyond the exemption are effectively taxed 100% on the margin through reduced DI benefits. Thus, if the re-examination led to a reduction in the degree disabled, this both reduced the DI benefit (which is an income effect) and it increased the

¹⁸ All amounts are gross of tax and social insurance contributions. Both benefits and earnings are subject to taxation and mandatory social insurance contributions, so we can directly compare changes in income to changes in benefits.

earnings exemption (which is a substitution effect). Therefore, like most of the previous literature on the labor supply response to DI, we cannot determine the extent to which the response is driven by the substitution effect and by the income effect. However, given the large magnitude of the earnings reaction (especially if compared to the change in total benefits), we suspect that incentive effects stemming from the change in the earnings exemption played a large role.

Panel C presents the combined effect of benefit substitution and labor crowd out. The estimate in the first row and column indicates that individuals increased income from other social assistance and work by €0.92 per euro of DI benefits lost. In other words, on average individuals almost fully offset the decrease in DI benefits by increased income from other sources, and we cannot reject the hypothesis that the offset was complete (p-value 0.494). Even if the offset is complete, individuals now have to work more and will be worse off to the extent they receive disutility from supplying labor.¹⁹ The second row shows the effect on a dummy for working or receiving income from a social assistance program other than the original DI spell. We find that per €1000 decrease in DI, an individual is 5.7 percentage points more likely to obtain income from a new source. The fact that this estimate is less than the sum of the estimates in row 2 of panels A and B indicates that some individuals both started working and started drawing income from other forms of social assistance. In particular, per €1000 decrease in DI, individuals became 1.7 (=2.9+4.5-5.7) percentage points more likely to have both income from other social assistance programs and labor income in 1999.

The estimates in Table 4 are based on a bandwidth of +/- 2.5 years around the cutoff age, which is the bandwidth suggested by applying the Imbens-Kalyanaraman criterion (2009) to our data. Appendix Table A4 explores the sensitivity of the benefit-substitution ratio and the earnings crowd-out ratio to the choice of bandwidth. For any bandwidth between +/- 1 year and +/- 5 years, both ratios are statistically significant at the 1-percent level. The size of the benefit-substitution ratio is relatively insensitive to

¹⁹ Even if we cannot reject that individuals are on average able to fully offset the DI benefit cut, it is conceivable that certain subgroups are not able to do so. To examine this possibility, we ran reduced-form quantile RD regressions on log total income (including DI income). We find statistically significant declines in log total income between the 15th and 40th percentiles of the income distribution. See Appendix Figure A2 for details.

the choice of bandwidth but the earnings crowd-out ratio is substantially larger for smaller bandwidths – it rises to 0.78 for a bandwidth of +/- 1 year.

4.4 Effects by Income Source, Gender, and Degree of Disability

The first column of panel A of Table 5 splits out the results of the first column of Table 4 by source of income. Of the increase in benefits, 67% comes from increased UI benefits, 10% from increased General Assistance, 3% from re-entry into DI, and 20% from all other types of social support benefits. We find that the increases in wage earnings account for 80% ($=0.492/0.618$) of the earnings response and changes in self-employment income only for 20%. Given that evasion of wage earnings is hard and limited, this breakdown indicates that the earnings response is unlikely to be largely driven by people not changing their actual labor supply but simply starting to report their earnings. Panel B examines receipt of any amount by income sources, and its results are very similar to the results of panel A.

Columns 2 and 3 of Table 5 split out the results of the first column by gender. These columns suggest that social support substitution is more predominant among women. In particular, the point estimate of the benefit-substitution ratio is much larger for women than for men (0.48 vs. 0.26) but this difference is not statistically significant (p-value 0.149). The difference is statistically significant, however, if we look at the participation response for other forms of social assistance. Per €1000 decrease in DI benefits, women increase their participation in other social assistance programs by 7.4 percentage points, which is nearly twice the 3.8 percentage point increase by men. In response to a given DI benefit cut, women are also significantly more likely than men to start working whereas the point estimate of the earnings response is actually slightly larger for men than for women (though not significantly so). The fact that labor force participation is only 18% for women but 45% for men may explain why women experience a larger response on the extensive margin but that total earnings increase slightly more for men because the scope for an intensive-margin response is larger among men. There is no significant difference in the degree to which men and women are able to offset the decrease in DI benefits by other sources of income. As noted earlier, it is hard to determine what exactly drives differences in the effects of the DI reform on men

and women, but we suspect differences in initial DI benefit levels, differences in types of disabilities, and differences in opportunities in market and household production are likely explanations. For example, a primary diagnosis of a psychiatric condition is much more prevalent among women than men in our sample (39.6% vs. 28.2%) whereas a primary diagnosis of a musculoskeletal disorder is much more prevalent among men than women (39.2% vs. 28.6%).

Table 6 analyzes benefit substitution and earnings crowd out by the degree disabled into which individuals were classified as of January 1st 1996 (so before the re-examinations took place).²⁰ Panel A shows that social support substitution is much more prevalent among individuals classified as fully disabled than among those classified as partially disabled. This finding applies both for social assistance benefit amounts and for social assistance participation, and holds in the entire sample as well as the subsamples by gender. These differences are not only statistically significant but also large in magnitude. The benefit-substitution ratio for the fully disabled is 0.50, which is four times as large as the ratio of 0.12 for partially disabled recipients. About 90% of this difference in benefit substitution ratios is accounted for by the larger increases of UI benefits among the fully disabled relative to partially disabled individuals.

In contrast, we find high rates of labor crowd out both for those classified as partially disabled and for those classified as fully disabled. Panel B shows that the point estimate of crowd out is somewhat higher for the partially disabled than for the fully disabled (0.68 vs. 0.52), but this difference is not statistically significant. The high degree to which the fully disabled are able to replace foregone disability income with labor income is striking, though it should be kept in mind that degree of disability depended on the availability of suitable jobs in the applicant's region, and that an applicant could also be classified as fully disabled if not enough of such jobs were found.

Panel C shows that both the partially and fully disabled are able to offset basically all of their lost DI income by other sources of income. The point estimates indicate that the fully disabled actually offset somewhat more of the lost DI benefits than the partially disabled, but this difference is not statistically significant.

²⁰ We also investigated whether the benefit-substitution ratio and the labor crowd-out ratio varied by marital status, previous earnings, duration of the DI spell, and national origin. We found no significant differences along these dimensions. See appendix Table A5 for details.

Our data has information on primary medical diagnoses, but much of this data was retrospectively added. As a result, this data is significantly less likely to be missing for those who remained on DI, and there is a strong discontinuity at the age cutoff in the indicator for the medical diagnosis being missing. This implies that we cannot stratify our estimates by medical diagnosis. Because the medical part of the re-examination was the same on either side of the age cutoff, there should be no causal effect of the more stringent re-examination on the medical diagnosis itself.²¹ Therefore any discontinuity at the age cutoff in the prevalence of a given medical diagnosis among those who remained on DI must be due to differential exit by that medical diagnosis. It turns out that we lack statistical precision on the inferred distribution of differential exit by medical diagnosis, but the point estimates indicate that the more stringent reforms led to disproportionately high exit rates among those with diagnoses of musculoskeletal, psychiatric, and neurological conditions and disproportionately low exit rates among those whose diagnosis was labeled “general”. Full results are in Appendix Table A6.

4.5 Responses of Partners of DI recipients

In Table 7, we provide estimates of benefit substitution and earnings crowd out at the household level. These estimates differ from our baseline estimates of Table 4 in that the current estimates account for possible responses of partners of (former) DI recipients. We find that our point estimates of benefit substitution in the entire sample are virtually identical whether or not we take the partners’ response into account. For men the benefit substitution ratio becomes somewhat larger and for women it becomes smaller when we take the partners’ response into account, but neither difference is statistically significant. The increase for earnings crowd out, while at 18 percentage points not insubstantial in economic terms, is statistically insignificant. Earnings responses of partners were previously studied by Cullen and Gruber (2000) who estimate that increased UI benefits paid to unemployed males are largely offset by decreased labor market earnings of their wives. While partner responses could potentially be important, and therefore are important to consider, we find only a limited role for them in our setting. Including the

²¹ As explained in Section 2, the re-examination was more stringent for the younger cohort only because the procedure that translated medical diagnoses into replacement rates was less generous for them.

partner responses, however, decreases the precision of our estimates, which is why we exclude them from our other analyses.

4.6 Responses over time

Responses to reductions in DI benefits could vary over time, for example because it can take time to find the right match in the labor market or because certain forms of social assistance have time limits. Hence, focusing only on 1999, the first year that re-examinations are completed for individuals near each side of the age discontinuity, yields an incomplete picture of the consequences of the reform. We therefore repeated our main analyses for all years until 2005, which is the last year in our dataset.²²

Panel A of Figure 5 presents estimates of the reduced-form RD regression of DI benefit amounts for each of the years from 1999 to 2005. This panel shows that the effect of the reform on DI benefit amounts is remarkably constant over time. Panel B shows estimates of the benefit-substitution ratio and the earnings crowd-out ratio over time. In other words, the figure plots the coefficients from the same fuzzy RD IV regression that we presented in the first row of Panels A and B of Table 4, but now for all years until 2005. We find that both the earnings crowd-out ratio and benefit-substitution ratio are positive and statistically significant in each year. The degree to which individuals replace lost DI benefits with other forms of social assistance decreases over time, from 31% in 1999 to 20% in 2005. This decrease, however, is not statistically significant. The decline of the benefit-substitution ratio is driven by the decreased reliance on UI benefits over time. Whereas UI benefits accounted for 20 percentage points of the benefit-substitution ratio in 1999, they only account for 5 percentage points in 2005. This decline is consistent with the fact that unemployment assistance is only available for a limited duration. This decline is partly offset by increased reliance on new DI spells. Income from new DI spells account for about 1 percentage point of the 1999 benefit-substitution ratio but for 6 percentage points in 2005. Reliance on General Assistance and other forms of social assistance remains roughly constant over time. The figure shows a slight

²² Additional DI reforms took place in 2002 and 2004. The first reform only affected new entrants while the second reform led to a re-examination of people on DI who were younger than 50 on July 1st, 2004. All individuals in our sample were older than 50 at that time. These reforms therefore do not affect the individuals in our sample. There was a major overhaul of the DI system in 2006. This overhaul also affected the individuals in our sample, which is why we end our sample period in 2005.

increase over time in the earnings crowd-out ratio, which rises from 62% in 1999 to 71% in 2005, but this increase is not statistically significant and the fraction of earnings that comes from self-employment remains roughly constant over time. Finally, Panel C shows the benefit-substitution ration and the crowd-out ratio scaled by the change in DI benefits in 1999 in order to isolate the movements that are solely due to changes in earnings and social assistance receipt (other than the original DI spell). This panel is very similar to the previous panel, which is not surprising given that we found before that the impact of the more stringent re-examination on DI benefit amounts was basically constant over time.

5. Conclusion

In this paper, we investigate the consequences of a reduction in the generosity of one social support program when that program is part of a larger system of social assistance programs. Especially in the case of social assistance to people in their prime age, it was unknown to what extent reduced generosity of one program induces them to increase labor supply and to what extent it leads them to rely more on other social assistance programs instead. Examining the labor supply response of existing beneficiaries (as opposed to labor supply responses to qualify for a program) is important for policy because a response by the large stock of existing beneficiaries can quickly affect DI participation rates whereas a response by new enrollees will only slowly affect overall DI participation. Showing a labor supply response of long-term DI beneficiaries, including individuals who are classified as fully disabled, also establishes that long-term participation in DI does not severely degrade one's labor market skills. Benefit substitution is of obvious policy relevance in many countries. While existing studies have investigated spillover effects among programs for children or for people close to retirement, this paper examines benefit-substitution and earnings crowd-out effects for people on DI in their late 40s. Finally, our paper recognizes that spillovers from a reform to one program can be partly driven by responses by the partners of people affected by the reform and that the spillovers may vary with the amount of time passed since the reform.

The combination of access to extensive administrative panel data and the presence of a cohort discontinuity in a reform law allows us to produce causal estimates of the effect of the 1993 Dutch disability insurance reform on the participation in other social assistance programs. We find economically meaningful and statistically significant evidence of social support substitution. About 2 years after the implementation of the DI reform for our sample members, income drawn from other social assistance programs increases by 31 cents for each euro of reduced DI benefits. Thus, ignoring this benefit-substitution effect of 31% would lead one to overestimate the cost savings of the DI reform by nearly one half. At 50%, the benefit substitution effect is especially pronounced for the fully disabled whereas it is just 12% for partially disabled DI recipients. While the benefit-substitution ratio decreases over time, the benefit-substitution ratio still stands at 20% about 8 years after the implementation of the reform for our sample.

We also find a remarkable earnings rebound given that all members of our sample were at least partially disabled and on average had been receiving DI for over a decade when the reform was implemented for our cohorts. On average, individuals were able to make up 62% of their foregone DI benefits through increases in earnings, and this figure is similar for partially and fully disabled individuals. Between increased income from labor and other social assistance programs, individuals almost fully offset the decrease in DI benefits. Of course, these estimates are based on a relatively minor (10% on average) cut in DI benefits, and may not apply for larger cuts. Also, because these estimates reveal average responses they can mask more severe impacts on total income for certain subgroups of DI recipients.

Benefit-substitution and earnings crowd-out estimates would obviously be different in different settings, but the direction in which the estimates would change is not clear. Our benefit-substitution figure may be higher than it would be in other countries because the Netherlands has a relatively generous system of alternative social assistance programs. On the other hand, the reform we analyzed concerned a relatively minor reduction in DI generosity. Thus, many of those affected by the reform may not have qualified for means-tested alternative forms of social assistance, or alternative forms of

social assistance may still have been less attractive than DI (despite the reduction in DI generosity).

While our specific coefficient estimates only directly apply to this particular Dutch DI reform, we believe our paper offers three general lessons that are widely applicable. First, our paper provides strong evidence that spillover effects between social assistance programs can be substantial, also for prime-aged individuals. Thus, any analysis of a reform of a social assistance program would be wise to consider the possibility of benefit substitution. Second, we show that among long-term disability recipients there may still be a substantial capacity to change labor income in response to relatively moderate changes in DI generosity. In other words, labor supply among DI recipients is not just determined by limitations from the disability, but also by economic incentives. Finally, our work emphasizes that it can potentially be important to take into account the responses of the partners of the individuals directly affected by the reform and to consider the amount of benefit substitution and earnings crowd out over the longer term.

Because the discontinuity in the stringency of disability reform applies to existing recipients, we examine social support substitution and labor supply responses among those already receiving disability insurance at the time the reform went into effect. Our setting does not allow us to estimate spillover effects and labor supply responses stemming from people who would have flowed into DI under the less stringent rules but not under the more stringent rules. We view estimates of such spillover effects and labor supply responses as complementary to our estimates.

Literature

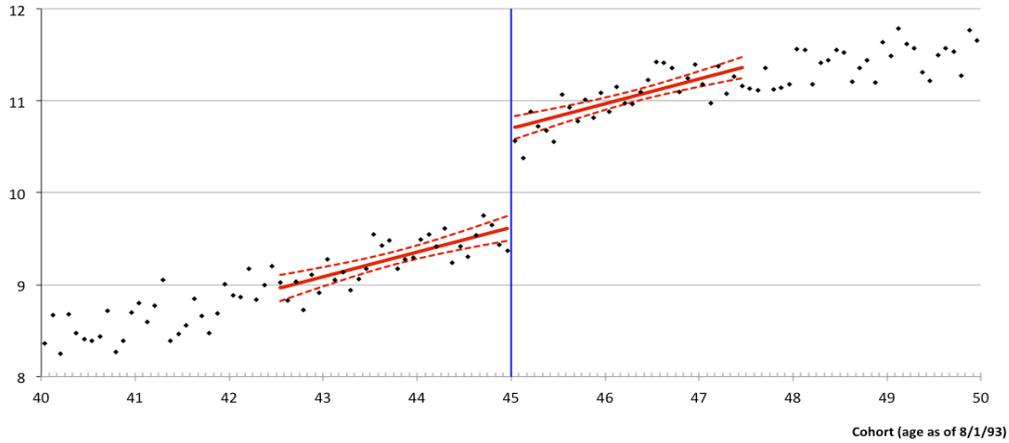
- Autor, David H., and Duggan, Mark G. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment," *Quarterly Journal of Economics* 118(1): 157 - 206.
- Autor, David H., and Duggan, Mark G. 2008. "The Effect of Transfer Income on Labor Force Participation and Enrollment in Federal Benefits Programs: Evidence from the Veterans Disability Compensation Program," unpublished manuscript.
- Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell. 2011. "Heaping-Induced Bias in Regression-Discontinuity Designs," NBER Working Paper No. 17408.
- Bound, John. 1989. "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economic Review* 79(3): 482-503.
- Bound, John, and Richard V. Burkhauser. 1999. "Economic Analysis of Transfer Programs Targeted on People with Disabilities," in Orley C. Ashenfelter and David Card (eds.), *Handbook of Labor Economics*, Vol. 3, Elsevier, Ch. 51, pp. 3417 – 3528.
- Bovenberg, A. Lans. 2000. "Reforming Social Insurance in the Netherlands," *International Tax and Public Finance* 7(3): 345-368.
- Chen, Susan, and Wilbert van der Klaauw. 2008. "The Work Disincentive Effects of the Disability Insurance Program in the 1990s," *Journal of Econometrics* 142(2): 757-784.
- Coe, Norma B., and Kelly Haverstick. 2010. "Measuring the Spillover to Disability Insurance Due to the Rise in the Full Retirement Age," CRR WP 2010-21, Boston College.
- Cullen, Julie Berry, and Jonathan Gruber. 2000. "Does Unemployment Insurance Crowd out Spousal Labor Supply?" *Journal of Labor Economics* 18(3): 546-572.
- De Jong, Philip R., Maarten Lindeboom, and Bas van der Klaauw. 2011. "Screening Disability Insurance Applications," *Journal of the European Economic Association* 9(1): 106-129.
- Duggan, Mark G., Perry Singleton, and Jae Song. 2007. "Aching to Retire? The Rise in the Full Retirement Age and its Impact on the Disability Rolls," *Journal of Public Economics* 91(7-8): 1327-1350.

- Duggan, Mark G., and Melissa Schettini Kearney. 2007. "The Impact of Child SSI Enrollment on Household Outcomes," *Journal of Policy Analysis and Management* 26(4): 861-886.
- French, Eric and Jae Song. 2011. "The Effect of Disability Insurance Receipt on Labor Supply." Federal Reserve Bank of Chicago Working Paper Series, WP 2009-05.
- Garrett, Bowen, and Sherry Glied. 2000. "Does State AFDC Generosity Affect Child SSI Participation?" *Journal of Policy Analysis and Management* 19(2): 275–295.
- García-Gómez, Pilar, Hans-Martin von Gaudecker, and Maarten Lindeboom. 2011. "Health, Disability and Work: Patterns for the Working Age Population," *International Tax And Public Finance* 18(2): 146-165.
- Gruber, Jonathan. 2000. "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy* 108(6): 1162-1183.
- Gruber, Jonathan, and Jeffrey D. Kubik. 1997. "Disability Insurance Rejection Rates and the Labor Supply of Older Workers," *Journal of Public Economics* 64(1): 1-23.
- Hahn, Jinyong, Petra E. Todd, and Wilbert van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica* 69(1): 201-209.
- Imbens, Guido W., and Karthik Kalyanaraman. 2009. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," NBER working paper no. 14726.
- Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics* 142(2): 615-635.
- Li, Xiaoyan, and Nicole Maestas. 2008. "Does the Rise in the Full Retirement Age Encourage Disability Benefits Applications? Evidence from the Health and Retirement Study," MRRC Working Paper #198.
- Karlström, Anders, Mårten Palme, and Ingemar Svensson. 2008. "The Employment Effect of Stricter Rules for Eligibility for DI: Evidence from a Natural Experiment in Sweden," *Journal of Public Economics* 92(10-11): 2071-2082.
- Koning, Pierre W. C., and Daniel J. van Vuuren. 2010. "Disability Insurance and Unemployment Insurance as Substitute Pathways," *Applied Economics* 42(5): 575-588.

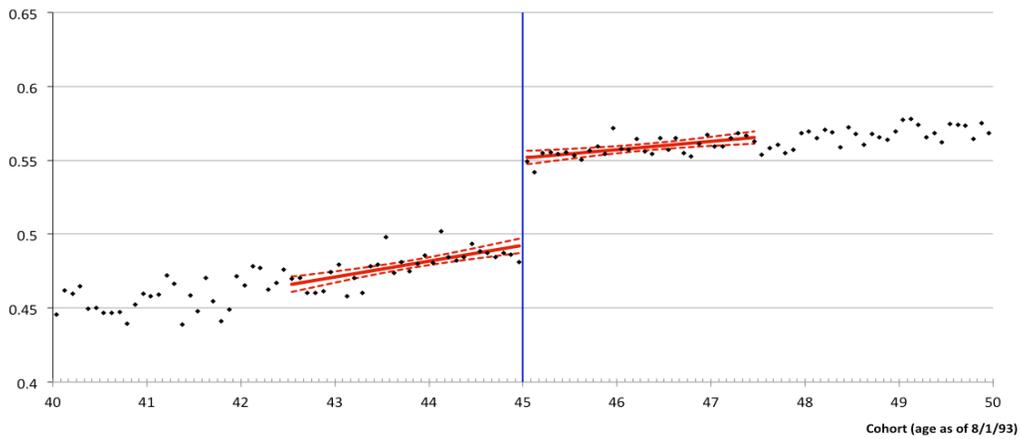
- Kubik, Jeffrey D. 2003. "Fiscal Federalism and Welfare Policy: The Role of States in the Growth of Child SSI," *National Tax Journal* 56(1): 61-79.
- Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics," *Journal of Economic Literature* 48(2): 281-355.
- McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics* 142(2): 698-714.
- Maestas, Nicole, Kathleen Mullen, and Alexander Strand. 2011. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt," Unpublished manuscript, RAND.
- Maestas, Nicole, and Jae Song. 2011. "The Labor Supply Effects of Disability Insurance: Evidence from Automatic Conversion Using Administrative Data," MRRC Working Paper #247.
- Parsons, Donald O. 1980. "The Decline in Male Labor Force Participation," *Journal of Political Economy* 88(1): 117-134.
- Schmidt, Lucie, and Purvi Sevak. 2004. "AFDC, SSI, and Welfare Reform Aggressiveness: Caseload Reductions vs. Caseload Shifting," *Journal of Human Resources* 39(3): 792-812.
- Staubli, Stefan. 2011. "The Impact of Stricter Criteria for Disability Insurance on Labor Force Participation," *Journal of Public Economics* 95 (9-10): 1223–1235.
- UWV (The agency that administers social insurance for employees in the Netherlands). 2006. *Kroniek van de Sociale Verzekeringen 2006 - Wetgeving en Volume-Ontwikkeling in Historisch Perspectief* (in Dutch), UWV, Amsterdam.
- von Wachter, Till, Jae Song, and Joyce Manchester. 2011. "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program," *American Economic Review* 101(7): 3308-3329.

Figure 1: Magnitude of the reform

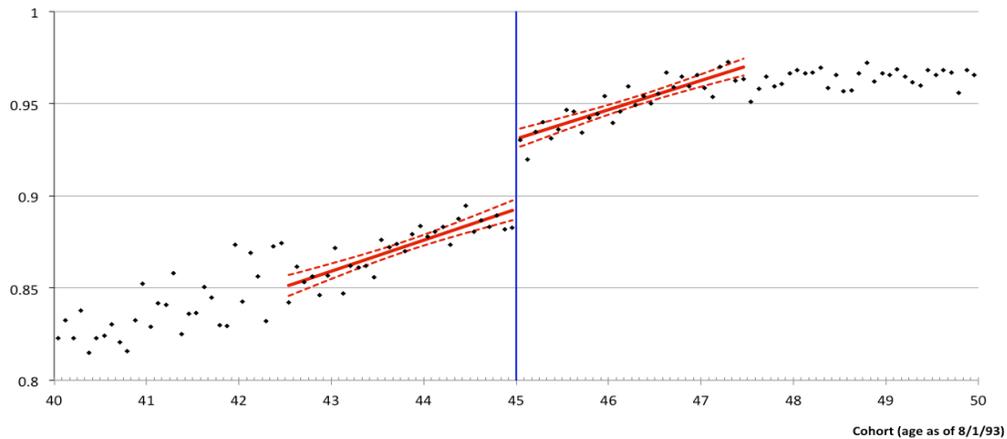
Panel A: Effect on DI Benefit Amounts (1000 €/yr). Estimate of the discontinuity: $-1.076 (0.096)^{***}$



Panel B: Effect on the DI Replacement Rate. Estimate of the discontinuity: $-0.059 (0.003)^{***}$



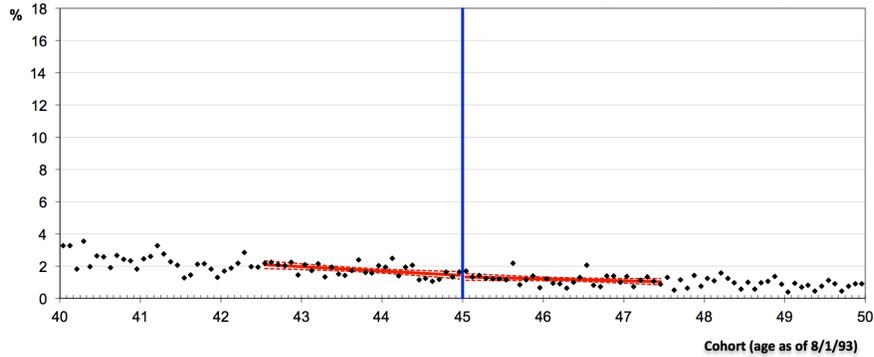
Panel C: Effect on Participation in DI in 1999. Estimate of the discontinuity: $-0.038 (0.004)^{***}$



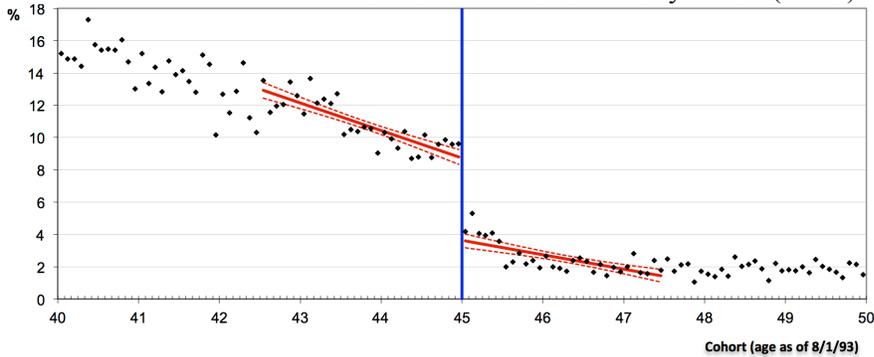
Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each figure is based on 84,185 observations. The dotted lines represent the 95% confidence intervals. Regression estimates come from reduced-form RD regressions without demographic control variables.

Figure 2: Exit by Year

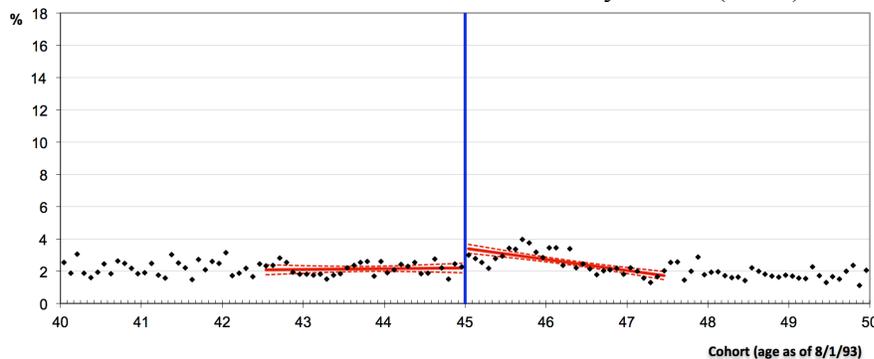
Panel A: Exit in 1995. Estimate of the discontinuity: 0.001 (0.002)



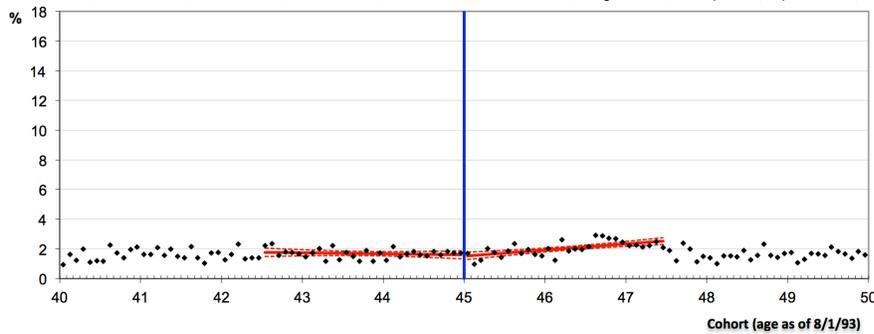
Panel B: Exit in 1996/1997. Estimate of the discontinuity: 0.051 (0.003)***



Panel C: Exit in 1998. Estimate of the discontinuity: -0.012 (0.002)***



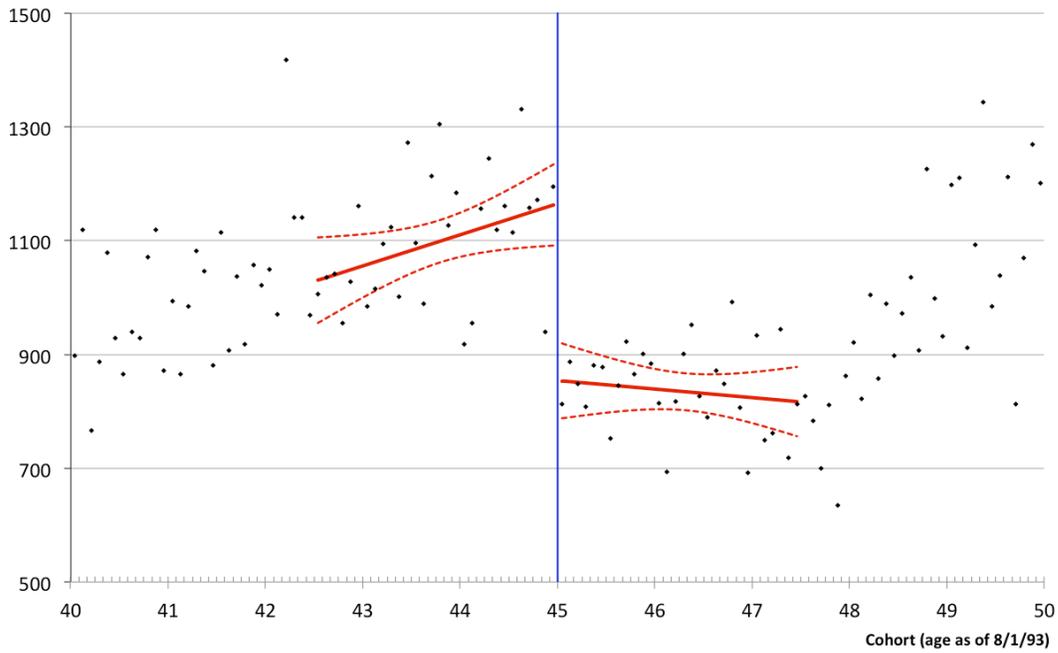
Panel D: Exit in 1999. Estimate of the discontinuity: 0.001 (0.002)



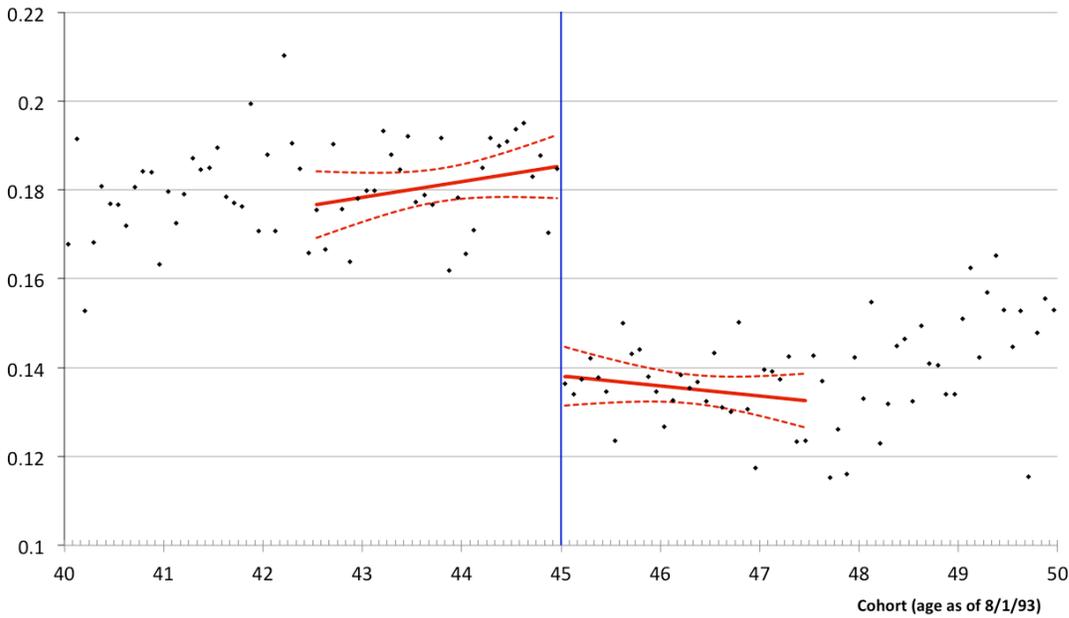
Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. The dotted lines represent the 95% confidence intervals. The exit rate is defined as a fraction of our sample in January 1996, except for panel A where it is a fraction of the sample in January 1995. Regression estimates come from reduced-form RD regressions without demographic control variables.

Figure 3: Effects of DI Reform on Social Assistance Other Than the Original DI Spell

Panel A: Social Assistance Income in 1999. Estimate of the discontinuity: 314 (51)***



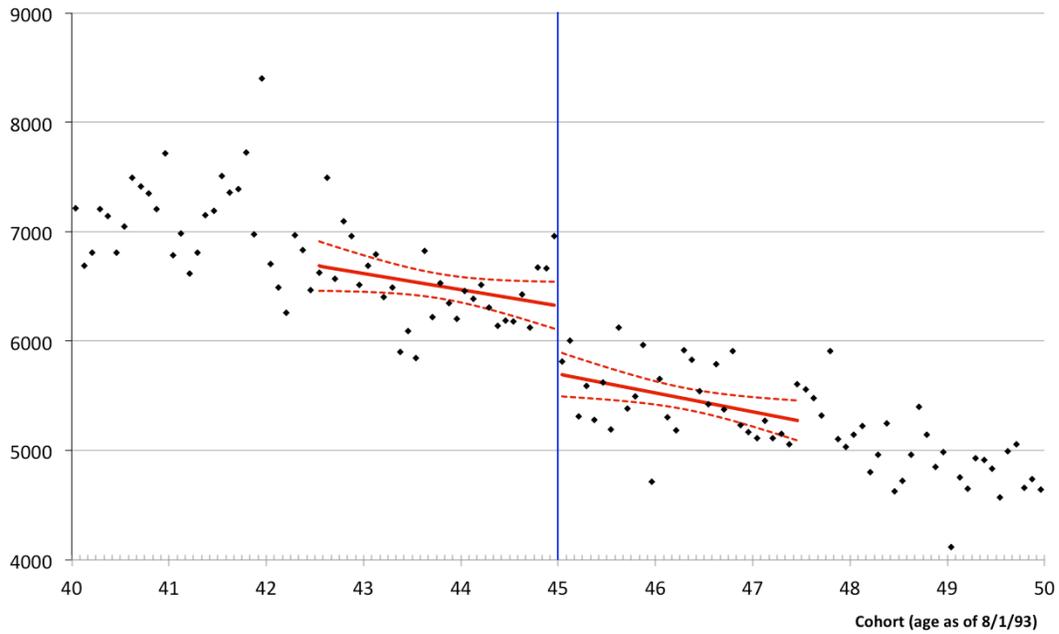
Panel B: Social Assistance Participation in 1999. Estimate of the discontinuity: 0.047 (0.005)***



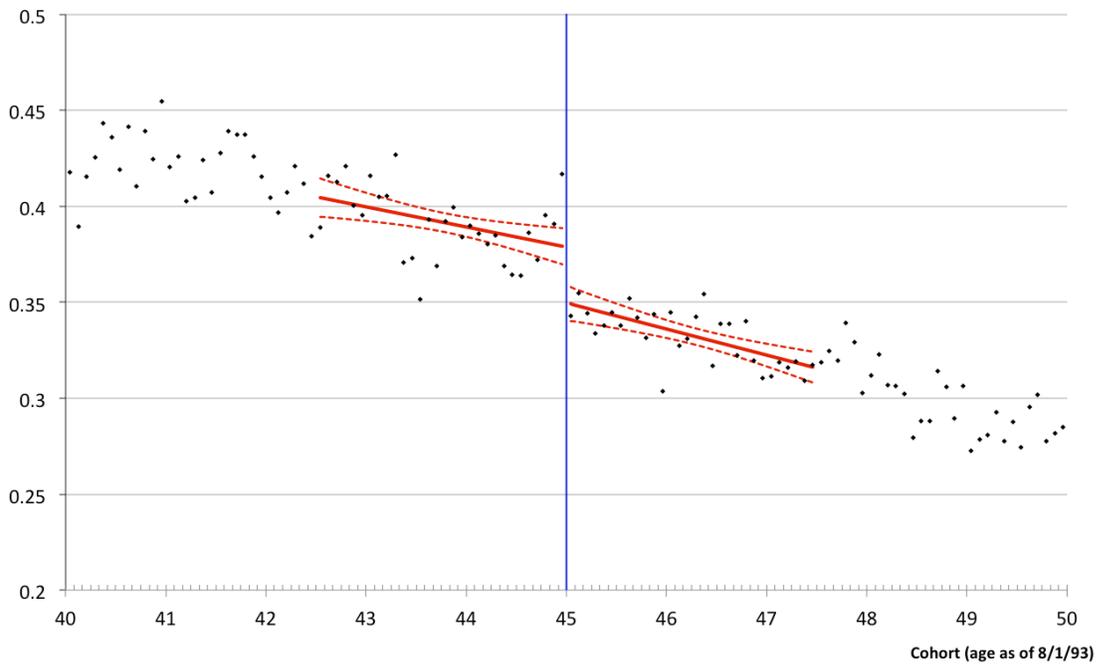
Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Figures are based on 84,185 observations. The dotted lines represent the 95% confidence intervals. Benefits from the original DI spell are not included in social assistance income, and participation rates exclude the original DI spell. Regression estimates come from regressions without demographic control variables. Standard errors are in parentheses.

Figure 4: Effects of DI Reform on Labor Market Outcomes

Panel A: Effect on Earnings in 1999. Estimate of the discontinuity: 624 (154)***



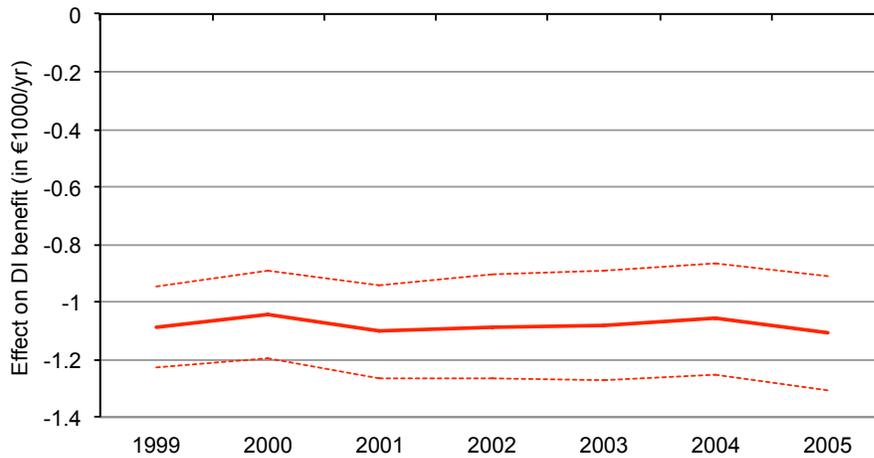
Panel B: Effect on Employment in 1999. Estimate of the discontinuity: 0.029 (0.007)***



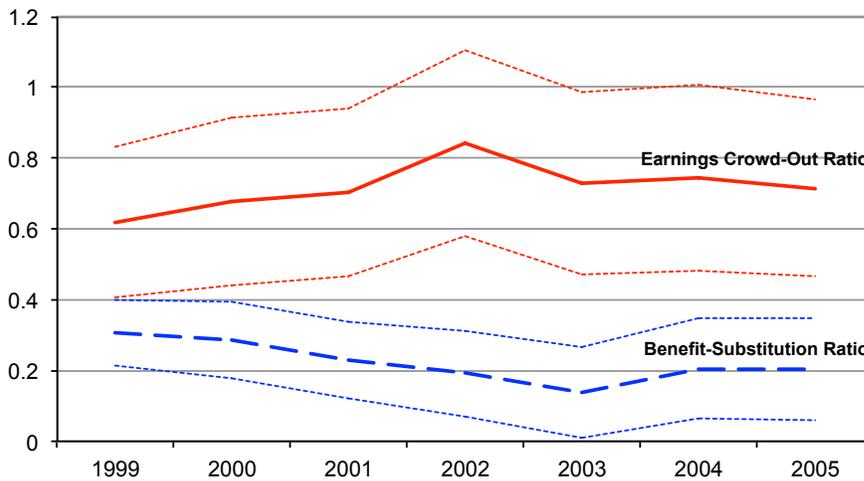
Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Figures are based on 84,185 observations. The dotted lines represent the 95% confidence intervals. Employment is defined as having positive earnings from employment or self-employment. Regression estimates come from reduced-form RD regressions without demographic control variables.

Figure 5: Earnings Crowd Out and Benefit Shifting over Time

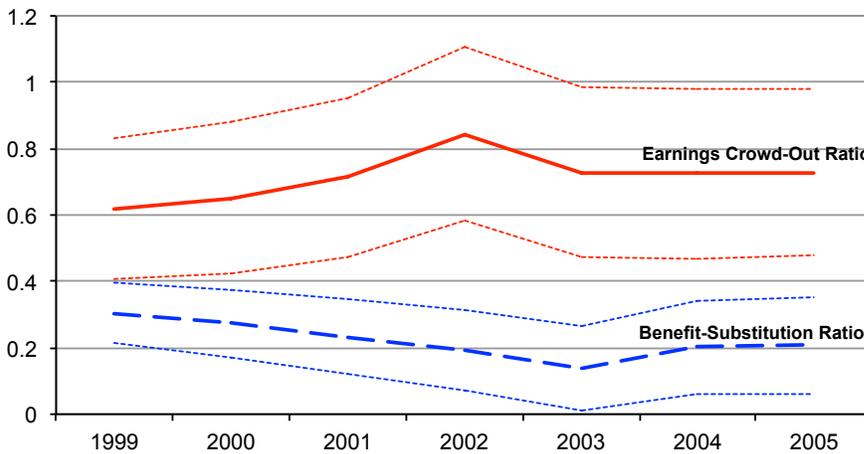
Panel A: Effect of More Stringent Re-examination on DI Benefits in €1000/year (First Stage)



Panel B: Earnings Crowd Out and Benefit Shifting Using Year-Specific First Stage



Panel C: Earnings Crowd Out and Benefit Shifting Using 1999 First Stage



Note: The dotted lines represent the 95% confidence intervals.

Table 1: Relation between Degree of Disability and Replacement Rates

Degree of disability:	Replacement rate (% of last earned wage):
80 – 100 %	70 %
65 – 80 %	50.75 %
55 – 65 %	42 %
45 – 55 %	35 %
35 – 45 %	28 %
25 – 35 %	21 %
15 – 25 %	14 %
Less than 15 %	0 %

Source: UWV (2006). UWV is the abbreviation of the agency that administers all social insurance for employees in the Netherlands. See text for a description of how the degree of disability is determined. Disability insurance benefit levels are determined as a percentage of the last earned wage and adjusted for inflation over time.

Table 2: Descriptive Statistics

	Full sample	Males	Females
Panel A: Sample characteristics, measured prior to re-examination			
Female (0=no; 1=yes)	0.34	0.00	1.00
Married (0=no; 1=yes)	0.66	0.69	0.61
Age on August 1 st , 1993	45.18	45.19	45.18
Start date of DI spell (year)	1985.1	1984.8	1985.7
<i>Degree of disability (% of earnings capacity lost):</i>			
15-25	7.67	8.85	5.36
25-35	9.53	11.96	4.74
35-45	6.91	8.75	3.28
45-55	5.78	6.32	4.74
55-65	2.01	2.22	1.60
65-80	1.97	2.41	1.11
80-100	66.1	59.50	79.17
Panel B: Outcomes after re-examination			
<i>Labor market status in 1999 (%):</i>			
Still on DI (on the original spell)	91.75	91.68	91.89
Employed	35.75	44.78	18.03
Social assistance (other than original DI spell)	15.50	14.78	16.92
Zero income (dummy for no formal income)	3.91	3.66	4.39
<i>Labor market status in 2005 (%):</i>			
Still on DI (on the original spell)	81.01	80.04	82.91
Employed	28.84	36.36	14.09
Social assistance (other than original DI spell)	24.94	26.41	22.06
Zero income (dummy for no formal income)	8.46	8.97	7.46
<i>Income by source in 1999, €/year (including zeros):</i>			
DI from original DI spell	10,296	11,135	8,649
Earnings	5,916	7,753	2,309
Social assistance (other than original DI spell)	949	862	1,120
<i>Income by source in 1999, €/year (if non-zero):</i>			
DI from original DI spell	11,731	12,732	9,785
Earnings	17,045	17,814	13,282
Social assistance (other than original DI spell)	6,169	5,900	6,631
<i>Income by source in 2005, €/year (including zeros):</i>			
DI from original DI spell	11,421	12,343	9,611
Earnings	5,452	7,136	2,145
Social assistance (other than original DI spell)	1,940	1,854	2,107
<i>Income by source in 2005, €/year (if non-zero):</i>			
DI from original DI spell	14,491	15,887	11,862
Earnings	20,136	20,889	16,308
Social assistance (other than original DI spell)	7,793	7,038	9,569
N	84,185	55,772	28,413

Note: Since we have information available from 1996 onwards, both marital status and degree of disability are recorded in January 1996 (before the re-examinations).

Table 3: Re-examinations and the Change in the Replacement Rate between 1996 and 1999

Change in the replacement rate	(1)	(2)	(3)	
	Predicted probability at age 45.0 for the less stringent re-examination	Predicted probability at age 45.0 for the more stringent re-examination	Treatment effect of the more stringent examination on the probability of the specified change in replacement rate	
7 "steps" less generous	6.10	7.81	1.70	(0.35)
6 "steps" less generous	0.47	1.52	1.05	(0.11)
5 "steps" less generous	0.84	2.43	1.59	(0.13)
4 "steps" less generous	1.08	2.26	1.18	(0.15)
3 "steps" less generous	1.12	2.31	1.19	(0.16)
2 "steps" less generous	2.23	3.76	1.52	(0.21)
1 "step" less generous	3.90	9.01	5.11	(0.28)
Same generosity	72.17	65.52	-6.65	(0.59)
1 "step" more generous	3.52	1.40	-2.12	(0.18)
2 "steps" more generous	1.46	0.67	-0.79	(0.12)
3 "steps" more generous	1.57	0.90	-0.66	(0.14)
4 "steps" more generous	1.87	0.91	-0.96	(0.14)
5 "steps" more generous	2.13	0.85	-1.28	(0.15)
6 "steps" more generous	1.54	0.66	-0.89	(0.12)

Note: Standard errors are in parentheses. Each row is estimated using our standard reduced-form RD regression without demographic controls, where the outcome variable is a dummy for the change in the replacement rate between 1996 and 1999 that corresponds to row header. There are eight possible replacement rates: 0%, 14%, 21%, 28%, 35%, 42%, 50.75%, 70%, where we assign 0% to those who exit from DI before the post-examination replacement rate is recorded. Column 1 shows the intercept at age 45.0 from the regression line to the right of the discontinuity (i.e., for those who underwent the less stringent re-examination), column 2 shows the intercept at age 45.0 from the regression line to the left of the discontinuity (i.e., for those who underwent the more stringent re-examination), and column 3 shows the treatment effect (i.e., the difference between columns 1 and 2). N=84,185.

Table 4: Earnings Crowd Out and Benefit Shifting

	Effect scaled by decrease in amount of original DI (in 1000 €/year)	Effect scaled by decrease in the replacement rate (fraction)
Panel A: Other social assistance in 1999		
Income from other social assistance	0.305 (0.047) ^{***}	5.353 (0.801) ^{***}
Participation dummy	0.045 (0.005) ^{***}	0.797 (0.082) ^{***}
Panel B: Labor market outcomes in 1999		
Earnings	0.618 (0.108) ^{***}	10.848 (1.924) ^{***}
Employment dummy	0.029 (0.005) ^{***}	0.511 (0.084) ^{***}
Panel C: Total		
Income except from original DI spell	0.923 (0.113) ^{***}	16.201 (1.983) ^{***}
Dummy for work or other social assistance	0.057 (0.006) ^{***}	0.992 (0.092) ^{***}

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each entry in the table comes from a separate IV regression based on the fuzzy RD design. The dependent variable is listed in the rows. Other social assistance only includes disability income from re-entry in to disability (so it excludes disability income from the original spell). The variable that is instrumented (endogenous explanatory variable) is listed in the columns. The instrument itself is the treatment dummy (age as of 8/1/93 less than 45). Earnings and income are measured in thousands of euros per year. The replacement rate is expressed as a fraction. Each regression is based on 84,185 observations. The following controls are used in the regressions: age in months as of 8/1/93, (age-45) interacted with the treatment dummy, 6 dummies for degree of disability in 1996, a cubic polynomial in pre-DI earnings, 9 national origin dummies, a dummy for being married in 1996, 39 regional dummies, a cubic polynomial in duration in DI at the start of the reform, a full set of interactions between the dummies for the degree of disability and the cubic polynomial in pre-DI earnings, a gender dummy, and a full set of interactions between all previously listed controls and gender. In total, each regression has 163 control variables.

Table 5: Earnings Crowd Out and Benefit Shifting by Source and by Gender

	Effect of reform per 1000 €/year decrease in amount of original DI			P-value gender dif.
	Full sample	Males	Females	
Panel A: Amounts in 1999				
Total other social assistance, of which:	0.305 (0.047)***	0.261 (0.047)***	0.482 (0.146)***	0.149
Unemployment insurance	0.203 (0.020)***	0.187 (0.021)***	0.266 (0.051)***	
General assistance	0.030 (0.007)***	0.032 (0.007)***	0.024 (0.024)	
Re-entry into DI	0.008 (0.008)	-0.001 (0.009)	0.046 (0.025)*	
All other benefits	0.063 (0.041)	0.043 (0.040)	0.146 (0.129)	
Total earnings, of which:	0.618 (0.108)**	0.632 (0.124)**	0.564 (0.208)***	0.781
Wage earnings	0.492 (0.098)***	0.497 (0.114)***	0.471 (0.181)***	
Self-employment earnings	0.126 (0.061)**	0.135 (0.070)*	0.093 (0.113)	
Total income from other income sources	0.923 (0.113)**	0.892 (0.127)**	1.046 (0.250)***	0.582
Panel B: Participation in 1999				
Any income from other social assistance	0.045 (0.005)***	0.038 (0.005)***	0.074 (0.016)***	0.037
Any Unemployment insurance	0.034 (0.003)***	0.030 (0.003)***	0.050 (0.009)***	
Any general assistance	0.009 (0.002)***	0.010 (0.002)***	0.007 (0.006)	
Re-entry into DI	0.001 (0.001)**	0.001 (0.001)*	0.003 (0.002)	
Any other benefits	0.015 (0.004)***	0.010 (0.004)**	0.035 (0.013)***	
Any work	0.029 (0.005)***	0.023 (0.005)***	0.053 (0.013)***	0.026
Any wage income	0.024 (0.005)***	0.018 (0.005)***	0.050 (0.012)***	
Any self-employment income	0.006 (0.003)**	0.007 (0.003)**	-0.001 (0.005)	
Any other income source	0.057 (0.006)***	0.044 (0.006)***	0.105 (0.019)***	0.002

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each entry in the table comes from a separate IV regression based on the fuzzy RD design. The dependent variable is listed in the rows. Income and earnings are measured in thousands of euros per year. "Other income sources" excludes DI income from the original spell. General assistance provides an income floor for everyone and does not require having dependents. Any other benefits are benefits from a large number of smaller (about 30) benefit programs. The variable that is instrumented (endogenous explanatory variable) is the amount of DI, so all coefficients can be interpreted as effect size per €1000/year decrease in DI. The instrument itself is the treatment dummy (age less than 45 as of 8/1/93). The regressions are based on 84,185, 55,772, and 28,413 observations for the full sample, males, and females, respectively. See the note to Table 4 for the demographic controls included in the regression.

Table 6: Earnings Crowd Out and Benefit Shifting by Degree of Disability and Gender

	Effect of reform per 1000 €/yr decrease in amount of original DI			P-value gender dif.
	Full sample	Males	Females	
Panel A: Other social assistance in 1999				
Income from other social assistance				
Partially disabled in 1996	0.122 (0.048)**	0.124 (0.049)**	0.112 (0.143)	0.939
Fully disabled in 1996	0.501 (0.087)***	0.415 (0.084)***	0.803 (0.271)***	0.171
p-value on difference by disability	<0.001	0.003	0.024	
Participation dummy				
Partially disabled in 1996	0.015 (0.006)***	0.016 (0.006)***	0.012 (0.015)	0.840
Fully disabled in 1996	0.078 (0.010)***	0.063 (0.010)***	0.131 (0.035)***	0.062
p-value on difference by disability	<0.001	<0.001	0.002	
Panel B: Labor market outcomes in 1999				
Earnings				
Partially disabled in 1996	0.682 (0.166)***	0.732 (0.186)***	0.435 (0.357)	0.460
Fully disabled in 1996	0.520 (0.128)***	0.506 (0.148)***	0.572 (0.250)**	0.819
p-value on difference by disability	0.441	0.341	0.753	
Employment dummy				
Partially disabled in 1996	0.023 (0.006)***	0.020 (0.006)***	0.037 (0.017)**	0.357
Fully disabled in 1996	0.034 (0.008)***	0.026 (0.009)***	0.062 (0.019)***	0.087
p-value on difference by disability	0.246	0.555	0.333	
Panel C: Total				
Income except from original DI spell				
Partially disabled in 1996	0.804 (0.163)***	0.856 (0.182)***	0.547 (0.360)	0.444
Fully disabled in 1996	1.021 (0.153)***	0.921 (0.166)***	1.376 (0.392)***	0.285
p-value on difference by disability	0.332	0.793	0.119	
Dummy for work or other soc. asst.				
Partially disabled in 1996	0.030 (0.005)***	0.028 (0.005)***	0.036 (0.017)**	0.656
Fully disabled in 1996	0.084 (0.011)***	0.062 (0.011)***	0.163 (0.041)***	0.017
p-value on difference by disability	<0.001	0.005	0.004	
N				
Partially disabled in 1996	28,509	22,590	5,919	
Fully disabled in 1996	55,676	33,182	22,494	

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each entry in the table comes from a separate IV regression based on the fuzzy RD design. The dependent variable is listed in the rows. Income and earnings are measured in thousands of euros per year. The variable that is instrumented (endogenous explanatory variable) is the amount of DI, so all coefficients can be interpreted as effect size per €1000/year decrease in DI. The instrument itself is the treatment dummy (age less than 45 as of 8/1/93). Degree of disability is as determined by the disability administration (see text for the description of the procedure for the determination of degree of disability). See the note to Table 4 for the demographic controls included in the regression.

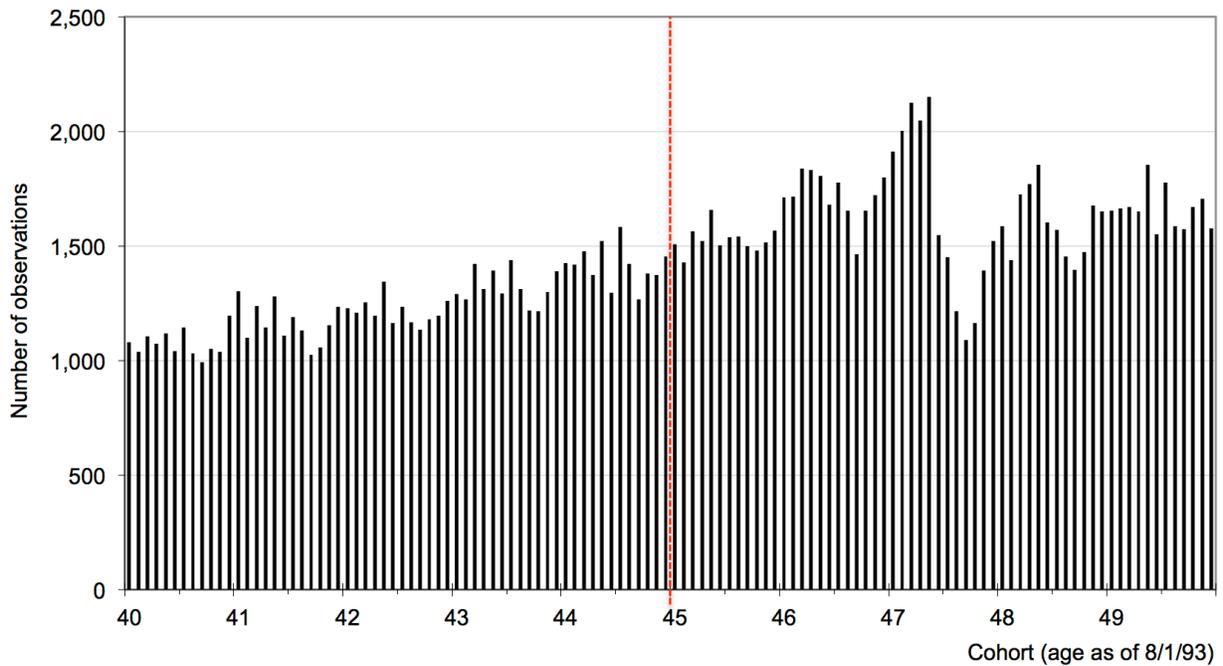
Table 7: Earnings Crowd Out and Benefit Shifting Including Partner Responses

	Effect of reform per 1000 €/yr decrease in amount of original DI			P-value gender dif.
	Full sample	Males	Females	
Panel A: Labor market outcomes in 1999				
Earnings	0.795 (0.209)***	0.718 (0.182)***	1.102 (0.749)	0.619
Panel B: Other social assistance in 1999				
Income from other social assistance	0.302 (0.083)***	0.307 (0.065)***	0.284 (0.318)	0.945
Panel C: Total				
Income except from original DI spell	1.097 (0.204)***	1.025 (0.185)***	1.386 (0.705)**	0.620

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each entry in the table comes from a separate IV regression based on the fuzzy RD design. The dependent variable is listed in the rows. Income and earnings are measured in thousands of euros per year. The variable that is instrumented (endogenous explanatory variable) is the amount of DI, so all coefficients can be interpreted as effect size per €1000/year decrease in DI. The instrument itself is the treatment dummy (age less than 45 as of 8/1/93). The regressions are based on 84,185 observations for the full sample, and on 55,772 and 28,413 observations for males and females, respectively. See the note to Table 4 for the demographic controls included in the regression.

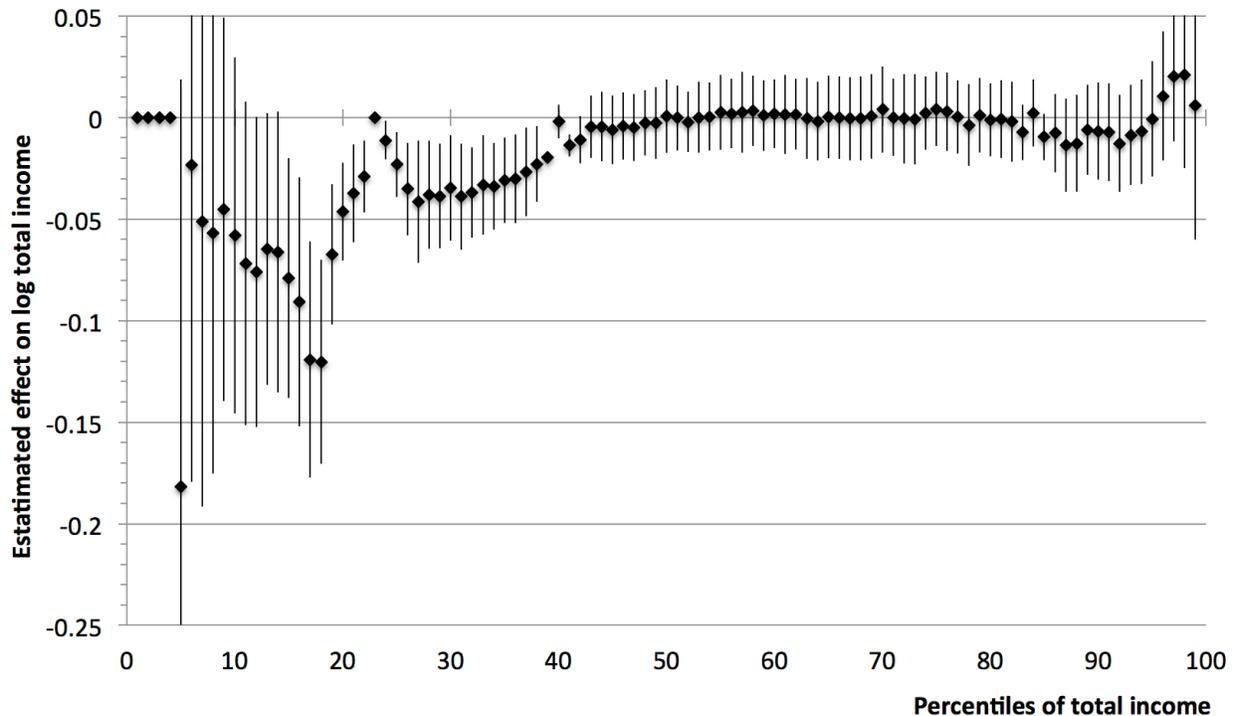
Appendix A: Additional Results (not for publication)

Figure A1: Number of Observations by Cohort



Note: The McCrary density test is insignificant (p-value 0.126). The large trough and spike around ages 47 and 48 are the effect of WWII (the “hunger winter of 1944”) and the subsequent baby boom.

Figure A2: Quantile Regressions of Effect of More Stringent Re-examination on Log Total Income



Note: Each dot corresponds the estimated effect from an RD quantile regression with log total income as the outcome variable. Lines correspond to 95-percent confidence intervals. Log total income is bottom coded at €1000/year to avoid very large percentage swings at very low incomes. Weighted by income (not logged), the average estimated effect of the more stringent re-examination is a 0.9% decline in total income.

Appendix Table A1: Placebo Regressions

	N	Treatment dummy: age<45	P-value	Sample mean
Panel A: Gender				
Female	84185	0.000 (0.007)	0.958	0.338
Panel B: Degree DI in 1996				
Degree DI in 1996 (scale of 1-8)	84185	0.037 (0.031)	0.231	6.554
Degree DI is 15-25% in 1996	84185	-0.001 (0.004)	0.797	0.077
Degree DI is 25-35% in 1996	84185	-0.009 (0.004)**	0.039	0.095
Degree DI is 35-45% in 1996	84185	0.003 (0.004)	0.327	0.069
Degree DI is 45-55% in 1996	84185	-0.001 (0.003)	0.666	0.058
Degree DI is 55-65% in 1996	84185	-0.000 (0.002)	0.827	0.020
Degree DI is 65-80% in 1996	84185	0.003 (0.002)	0.160	0.020
Degree DI is 80-100% in 1996	84185	0.005 (0.007)	0.444	0.661
Panel C: Province				
Province: Friesland	84185	-0.002 (0.003)	0.523	0.040
Province: Drenthe	84185	0.001 (0.003)	0.774	0.034
Province: Overijssel	84185	0.001 (0.004)	0.704	0.073
Province: Flevoland	84185	0.000 (0.002)	0.826	0.018
Province: Gelderland	84185	0.002 (0.004)	0.633	0.111
Province: Utrecht	84185	0.003 (0.003)	0.322	0.066
Province: Noord-Holland	84185	-0.008 (0.006)	0.164	0.192
Province: Zuid-Holland	84185	-0.006 (0.005)	0.258	0.168
Province: Zeeland	84185	-0.000 (0.002)	0.838	0.016
Province: Noord-Brabant	84185	0.002 (0.005)	0.731	0.151
Province: Limburg	84185	0.007 (0.004)*	0.074	0.090
Panel D: Duration in DI in 1993				
Duration on DI (months; as of 8/1993)	84185	-0.601 (1.037)	0.562	96.72
Duration: 5 years or more	84185	-0.007 (0.007)	0.284	0.592
Panel E: Marital status in 1996				
Married	84185	0.007 (0.007)	0.260	0.664
Panel F: Earnings before DI				
Previous earnings (euro/yr)	84185	13.0 (123.8)	0.916	16928
Quintile 1	84185	-0.003 (0.006)	0.601	0.200
Quintile 2	84185	0.005 (0.006)	0.413	0.200
Quintile 3	84185	-0.004 (0.006)	0.528	0.200
Quintile 4	84185	0.005 (0.006)	0.351	0.200
Quintile 5	84185	-0.003 (0.006)	0.550	0.200
Panel G: Origin				
Native Dutch	84185	0.003 (0.005)	0.608	0.834
Morocco	84185	-0.001 (0.001)	0.246	0.005
Turkey	84185	0.003 (0.002)	0.143	0.028
Surinam	84185	-0.005 (0.002)***	0.007	0.018
Antilles and Aruba	84185	-0.000 (0.001)	0.654	0.003
Other Non-Western Country	84185	-0.000 (0.001)	0.917	0.008
Netherlands, but born elsewhere	84185	-0.000 (0.003)	0.919	0.059
Other Western Country	84185	0.002 (0.002)	0.298	0.019
East-Europe	84185	0.001 (0.001)	0.330	0.009
Dutch East Indies	84185	-0.002 (0.002)	0.189	0.017

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each coefficient is estimated using our standard reduced-form RD regression without demographic controls. Previous earnings are capped at the maximum amount covered by DI (about €36,000/year in 1999).

Table A2: Replacement Rates and Re-examinations**Panel A: Joint distribution of the 1996 and 1999 replacement rates at age 45.0 under the less stringent re-examination**

Replacement rate in 1996	Replacement rate in 1999								Total
	0%	14%	21%	28%	35%	42%	50.75%	70%	
14%	1.65	2.43	0.74	0.23	0.13	0.04	0.02	1.54	6.79
21%	1.43	0.72	3.64	1.07	0.33	0.10	0.11	2.11	9.49
28%	0.59	0.15	0.60	3.10	0.65	0.18	0.08	1.72	7.06
35%	0.55	0.05	0.18	0.40	2.93	0.36	0.21	1.26	5.94
42%	0.22	0.00	0.04	0.07	0.15	0.92	0.14	0.51	2.07
50.75%	0.11	0.00	0.02	0.03	0.07	0.10	0.83	0.55	1.71
70%	6.14	0.35	0.62	0.52	0.38	0.28	0.32	58.33	66.93
Total	10.68	3.70	5.85	5.42	4.64	1.98	1.70	66.02	100.00

Panel B: Treatment effect of the more stringent re-examination on the joint distribution of 1996 and 1999 replacement rates

Replacement rate in 1996	Replacement rate in 1999								Total
	0%	14%	21%	28%	35%	42%	50.75%	70%	
14%	1.36	0.21	-0.52	-0.12	-0.11	-0.02	-0.01	-0.89	-0.10
	(0.21)	(0.28)	(0.08)	(0.05)	(0.03)	(0.03)	(0.02)	(0.12)	(0.37)
21%	0.33	1.07	-0.11	-0.59	-0.25	0.01	-0.04	-1.27	-0.85
	(0.18)	(0.11)	(0.32)	(0.10)	(0.05)	(0.04)	(0.04)	(0.15)	(0.41)
28%	0.31	0.42	1.38	-0.37	-0.44	-0.10	0.04	-0.89	0.35
	(0.12)	(0.06)	(0.11)	(0.28)	(0.07)	(0.04)	(0.04)	(0.13)	(0.36)
35%	0.14	0.16	0.36	0.53	-0.47	-0.19	-0.06	-0.61	-0.14
	(0.10)	(0.04)	(0.06)	(0.08)	(0.26)	(0.06)	(0.05)	(0.13)	(0.33)
42%	0.02	0.09	0.10	0.08	0.13	-0.16	-0.05	-0.25	-0.04
	(0.06)	(0.02)	(0.03)	(0.03)	(0.05)	(0.15)	(0.04)	(0.08)	(0.20)
50.75%	0.11	0.08	0.10	0.14	0.03	0.08	0.07	-0.33	0.27
	(0.05)	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.16)	(0.09)	(0.20)
70%	1.75	0.98	1.46	0.83	0.49	0.30	0.51	-5.82	0.51
	(0.35)	(0.09)	(0.11)	(0.10)	(0.09)	(0.07)	(0.08)	(0.69)	(0.66)
Total	4.01	3.01	2.78	0.50	-0.61	-0.09	0.46	-10.06	0.00
	(0.45)	(0.32)	(0.36)	(0.32)	(0.29)	(0.19)	(0.19)	(0.68)	(0.00)

Note: Standard errors are in parentheses. Panel A shows the predicted joint probability of replacement rates in 1996 and 1999 at exactly age 45. Each entry is estimated as the intercept at age 45.0 of the regression line to the right of the discontinuity (i.e., for those who underwent the less stringent re-examination). Panel B shows the treatment effect of the more stringent re-examination on each joint probability of replacement rates. Each entry is estimated using our standard reduced-form RD regression without demographic controls, where outcome variable is a dummy for combination of replacement rates in 1996 and 1999 that corresponds to that cell. Those who exit DI before the post-examination replacement rate is recorded are assigned a replacement rate of 0. N=84,185.

Appendix Table A3: Baseline Estimates Without Controls Variables

	Effect of reform per 1000 €/year decrease in amount of original DI					
	Full Sample		Males		Females	
Panel A: Other social assistance in 1999						
Income from other social assistance	0.292	(0.053) ^{***}	0.246	(0.051) ^{***}	0.511	(0.194) ^{***}
Participation dummy	0.044	(0.006) ^{***}	0.037	(0.006) ^{***}	0.080	(0.024) ^{***}
Panel B: Labor market outcomes in 1999						
Earnings	0.580	(0.135) ^{***}	0.659	(0.144) ^{***}	0.254	(0.285) ^{***}
Employment dummy	0.027	(0.006) ^{***}	0.026	(0.006) ^{***}	0.036	(0.017) ^{**}
Panel C: Total						
Income except from original DI spell	0.872	(0.142) ^{***}	0.906	(0.147) ^{***}	0.765	(0.343) ^{**}
Dummy for work or other soc. asst.	0.054	(0.007) ^{***}	0.045	(0.006) ^{***}	0.094	(0.027) ^{***}

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. The estimates in this table come from regressions that are identical to those used in Tables 4 and 5 except that the regressions for this table do not include demographic controls (except, of course, age as of 8/1/93 and the interaction of (age-45) with the treatment dummy).

Appendix Table A4: Sensitivity of Baseline Estimates to Bandwidth

		Effect scaled by decrease in amount of original DI (in 1000 €/yr)				
		First stage: DI amount in 1999 (in 1000 €/year)	Income sources in 1999		Participation in 1999	
N			Earnings	Income from other social assistance	Work	Other social assistance
+/- 5 years	160207	-1.047 (0.049) ^{***}	0.330 (0.088) ^{***}	0.412 (0.037) ^{***}	0.020 (0.003) ^{***}	0.050 (0.004) ^{***}
+/- 4 years	129557	-1.074 (0.055) ^{***}	0.417 (0.097) ^{***}	0.346 (0.038) ^{***}	0.023 (0.004) ^{***}	0.044 (0.004) ^{***}
+/- 3 years	98668	-1.112 (0.065) ^{***}	0.596 (0.097) ^{***}	0.266 (0.042) ^{***}	0.028 (0.004) ^{***}	0.041 (0.005) ^{***}
Baseline (+/- 2.5 years)	84185	-1.086 (0.071) ^{***}	0.618 (0.108) ^{***}	0.305 (0.047) ^{***}	0.029 (0.005) ^{***}	0.045 (0.005) ^{***}
+/- 2 years	66818	-1.101 (0.080) ^{***}	0.825 (0.119) ^{***}	0.305 (0.052) ^{***}	0.034 (0.005) ^{***}	0.044 (0.006) ^{***}
+/- 1 year	32772	-1.010 (0.116) ^{***}	0.778 (0.183) ^{***}	0.395 (0.083) ^{***}	0.036 (0.008) ^{***}	0.055 (0.009) ^{***}

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. The table shows our key estimates of the effect of the more stringent re-examination on the DI amount, the benefit-substitution ratio, and the earnings crowd-out ratio for different choices of bandwidth. Our baseline choice of bandwidth (+/- 2.5 years) was guided by the Imbens-Kalyanaraman (2009) criterion.

Appendix Table A5: Heterogeneity of Effect

		Effect scaled by decrease in amount of original DI (in 1000 €/year)	
		Earnings	Income from other social assistance
	N		
Panel A: By marital status			
Married	55913	0.551 (0.110) ^{***}	0.253 (0.042) ^{***}
Single	28272	0.879 (0.326) ^{***}	0.496 (0.183) ^{***}
p-value	on		0.196
difference		0.341	
Panel B: By previous earnings			
Below median	42095	0.563 (0.156) ^{***}	0.369 (0.070) ^{***}
Above median	42090	0.676 (0.144) ^{***}	0.255 (0.063) ^{***}
p-value	on		0.223
difference		0.597	
Panel C: By origin			
Native Dutch	70205	0.658 (0.128) ^{***}	0.286 (0.055) ^{***}
Other origin	13980	0.495 (0.186) ^{***}	0.360 (0.092) ^{***}
p-value	on		0.489
difference		0.470	
Panel D: By duration on DI as of 8/1/1993			
Less than 5 years	34378	0.897 (0.272) ^{***}	0.236 (0.126) [*]
More than 5 years	49807	0.512 (0.109) ^{***}	0.326 (0.044) ^{***}
p-value	on		0.477
difference		0.189	
Panel E: By degree of disability in 1996			
Partially disabled	28509	0.682 (0.166) ^{***}	0.122 (0.048) ^{**}
Fully disabled	55676	0.520 (0.128) ^{***}	0.501 (0.087) ^{***}
p-value	on	0.441	<0.001
difference			

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. Each entry comes from our standard fuzzy RD regression with the outcome variable indicated in the column header and the sample indicated in the row header. Previous earnings are the earnings to which the DI replacement rate is applied.

Appendix Table A6: Inferred Distribution of Medical Conditions Among Those Induced to Exit DI

Medical condition	Stayers with condition as a percentage of stayers		Stayers with condition as percentage of total population		Induced leavers with condition as percentage of total population	Induced leavers with condition as percentage of induced leavers	Odds ratio of prevalence of condition among induced leavers relative to untreated stayers	
	At left border of age 45 (treated) (1)	At right border of age 45 (untreated) (2)	Difference, (1)-(2): (3)	At left border of age 45 (treated) (4)	At right border of age 45 (untreated) (5)	Difference, (5)-(4): (6)	(7)/(2): (8)	
Musculoskeletal	37.4	38.0	-0.62 (0.69)	33.4	35.4	2.0	52.5	1.38
Psychiatric	31.8	32.8	-0.98 (0.67)	28.4	30.5	2.1	55.7	1.70
General	12.6	10.6	1.96 (0.46) ^{***}	11.3	9.9	-1.3	-35.5	-3.33
Neurological	4.1	4.7	-0.60 (0.32) [*]	3.6	4.3	0.7	18.7	4.00
Cardiovascular	4.3	4.4	-0.07 (0.32)	3.9	4.1	0.2	6.0	1.36
Digestive system	2.1	2.3	-0.13 (0.22)	1.9	2.1	0.2	5.2	2.31
Respiratory system	2.2	2.1	0.09 (0.22)	2.0	2.0	0.0	0.0	-0.02
Urological	1.5	1.3	0.14 (0.18)	1.3	1.2	-0.1	-1.9	-1.43
Visual impairment	1.0	0.9	0.04 (0.14)	0.9	0.9	0.0	0.1	0.10
Endocrinology	0.9	0.9	0.01 (0.15)	0.8	0.8	0.0	0.6	0.65
Hearing impairment	0.8	0.8	0.01 (0.14)	0.7	0.8	0.0	0.6	0.67
Dermatological	0.9	0.7	0.15 (0.14)	0.8	0.7	-0.1	-2.8	-3.80
Hematological	0.4	0.3	0.02 (0.09)	0.3	0.3	0.0	-0.2	-0.49
Pregnancy related	0.0	0.1	-0.04 (0.04)	0.0	0.1	0.0	0.9	11.64
Total	100.0	100.0	0.00	89.4	93.2	3.8	100.0	

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. We run a reduced-form RD regression with our standard controls (see note to Table 4) on the sample of those who remain on DI, where outcome variable is a dummy for the individual having as main diagnosis the condition indicated in the row header. Column 1 shows the intercept at the cutoff age (exactly 45 on 8/1/93) from the regression line to the left of the discontinuity (i.e., for those who underwent the more stringent re-examination), column 2 shows the intercept at the cutoff age from the regression line to the right of the discontinuity (i.e., for those who underwent the less stringent re-examination), and column 3 shows the “treatment effect” (i.e., the difference between columns 1 and 2) among the selected sample of those who remain on DI. Under the (reasonable) assumption that there is no difference between the more stringent and less stringent re-examination on the main diagnosis for a given individual, the “treatment effect” is due to differential exit by medical condition. The composition of medical conditions of those induced to exit by the more stringent re-examination is listed in columns (6) and (7). N=74,028 for the regressions that generate the estimates in columns (1) through (3).

Appendix B: Variable Definitions (not for publication)

Variables come from the datasets described below. The data sets are merged based on an individual identifier that is a scrambled version of the Dutch equivalent of a Social Security Number (“burgerservicenummer”).

Earnings and other sources of income

Information on income by source comes from various administrative datasets. Employment earnings (from the jobs database, “SSB-Banen”) and self-employment earnings (from the self-employed database, “SSB-Zelfst”) are obtained from yearly tax files, and hence comprise annual gross earnings measured in euros. Gross benefit payments are registered monthly by the respective benefit administration offices. From these monthly payments we construct the income flow for each source of benefits in euros per year. Benefits are gross; they are subject to income taxation and social insurance contributions. All amounts are nominal but inflation was low (around 3%/yr) in the time period of our data.

Variable	Dataset	Definition	Timing & Units
<i>I. Labor market earnings</i>			
Employment	SSB-Banen	<u>Employment participation</u> : having positive annual income from paid employment in a given year	Measured by year; 0/1 dummy
		<u>Employment earnings</u> : gross annual earnings from paid employment (0 if not employed)	Measured by year; euros/year
Self-employment	SSB-Zelfst	<u>Self-employment participation</u> : having positive annual income (profits) from self-employment in a given year	Measured by year; 0/1 dummy
		<u>Self-employment earnings</u> : gross annual earnings (profits) from self-employment (0 if not self-employed)	Measured by year; euros/year
<i>II. Social support benefits</i>			
Unemployment insurance	WW	<u>Unemployment insurance participation</u> : having positive annual income from UI in a given year	Measured by year; 0/1 dummy
		<u>Unemployment income</u> : gross annual income from UI (0 if not unemployed)	Measured by year; euros/year
General assistance	SSB-ABW	<u>General assistance participation</u> : having positive annual income from GA in a given year	Measured by year; 0/1 dummy

		<u>General assistance income</u> : gross annual income from GA (0 if not enrolled in GA)	Measured by year; euros/year
Re-entry into disability insurance	AO	<u>Re-entry in DI</u> : having positive annual income from DI in a given year coming from a new DI-spell	Measured by year; 0/1 dummy
		<u>New DI income</u> : gross annual income from DI (0 if not enrolled in DI) coming from a new DI-spell	Measured by year; euros/year
Other sources of benefits	QUITK	<u>Participation in other benefit programs</u> : having positive annual income from other programs in a given year	Measured by year; 0/1 dummy
		<u>Income from other benefit programs</u> : gross annual income from other programs (0 if not enrolled in scheme)	Measured by year; euros/year

DI administrative data

The disability offices in The Netherlands register many details for each DI claimant. From 1996 onwards, there is monthly information about all DI claimants in the Netherlands. Information includes start and end dates of DI-spells, degree of disability, amount benefit paid, and previous earnings, which is used as a reference to calculate the benefit level.

Variable	Dataset	Definition	Timing & Units
Disability	AO	<u>Participation in DI</u> : The original DI spell (i.e. the one that the individual was on when the reform was enacted on 8/1/93) continues into at least part of the current year	Measured by year; 0/1 dummy
		<u>Income from DI</u> : gross annual income from disability insurance benefits stemming from the original DI spell (0 if not enrolled in DI)	Measured by year; euros/year
Degree Disabled	AO	The degree of disability before the re-examination started. Categorical variable, based on percentage income loss due to DI (0-15%, 15-25%, 25-35%, 35-45%, 45-55%, 55-65%, 65-80%, 80-100%).	Measured as of January 1996; 1-8 scale, increasing in the degree disabled categories.
Previous Earnings	AO	Gross annual earnings (including vacation allowance, yearly bonus, extra pay for shift work, etc.) in the year before entry in DI.	Taken from the 1996 database, but refers to whichever

This amount was capped at a certain maximum level of earnings (€35,754 per year in 1999). was the last year of work prior to DI; euros/year.

Duration in DI	AO	Duration in DI at the time of the reform (August 1 st , 1993)	Measured on August 1 st , 1993; measured in months
----------------	----	--	---

Demographic characteristics

Each city (or aggregation of small villages) in The Netherlands keeps a registry (“GBA”) of all its native and foreign inhabitants. When an individual moves within the Netherlands, the person has to register in the new city (and unregister in the old city). Whenever there is a demographic change (e.g., child born, marriage), this is registered by the city administration. In this paper, we use the following demographics from the municipal registries: date of birth, gender, place of residence, origin, and marital status.

Variable	Dataset	Definition	Timing & Units
Age	GBA	Age in months at the time of the reform, calculated from year and month of birth	In months, measured on August 1 st , 1993
Married (dummy)	GBA	Dummy calculated from marital status information.	Measured as of January 1996; Dummy equals 1 if married
Gender	GBA	Dummy if person is female	Measured as of January 1996; Dummy equals 1 if female
Origin	GBA	Origin based on someone’s country of birth and that of his/her parents. Native Dutch are those (i) born in The Netherlands just as their parents, and those (ii) born in another country whose parents were both born in The Netherlands. Non-native Dutch are all others for whom at least one parent was not born in The Netherlands. Within the group of non-native Dutch we distinguish between countries of origin. For those born abroad, the country of birth is taken as the country of origin. When born in the Netherlands, the country of birth of the mother is taken as the country of origin. When both the individual and the mother are born in the Netherlands, the country of origin is based on the country	Measured as of January 1996; 10 dummy (0/1) variables

Region	GBA	<p>of birth of the father. We create 10 original dummies for the following categories: Native Dutch, Morocco, Turkey, Surinam, Dutch Antilles and Aruba, Other Non-Western Country, Dutch but born elsewhere, Other Western Country, East-European Country, and Dutch East Indies.</p> <p>Based on the place of residence (i.e., the city where someone is registered), we create 40 regional dummies that correspond to the COROP regions as defined by the Coordination Commission Regional Research Programme (see http://en.wikipedia.org/wiki/COROP for more information).</p>	<p>Measured as of January 1996; 40 dummy (0/1) variables</p>
--------	-----	--	--
