PRELIMINARY AND INCOMPLETE

Social Support Substitution and Labor Supply Shifting: Evidence from a Regression Discontinuity in Disability Insurance Reform

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

May 23, 2012

Abstract

This paper examines individuals' ability to compensate for a decrease in the generosity of one social assistance program by increasing their reliance on other social assistance programs and on labor market earnings. We exploit an age 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 individual behavior. We find substantial evidence of "social support substitution": individuals offset each Euro of lost DI benefits by collecting 31 cents more on average from other social assistance programs. This benefit-substitution effect declines somewhat over time, but is still a significant 21% eight years after the reform. Individuals also react to the benefit reduction though an increase of labor earnings of 62 cents on average per Euro of lost DI benefits. Thus, on average, individuals make up for almost the entire DI benefit decrease though increases labor earnings and other forms of social assistance.

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

JEL codes: H53, J22, I38

* Borghans: Maastricht University, ROA, and IZA (e-mail: lex.borghans@maastrichtuniversity.nl); Gielen: IZA and Erasmus University Rotterdam (e-mail: Gielen@ese.eur.nl); Luttmer: Dartmouth College, NBER, and IZA (e-mail: Erzo.FP.Luttmer@Dartmouth.edu). We would like to thank Gerard van den Berg, Mark Duggan, Pierre Gielen, Amy Finkelstein, Paul Frijters, Ed Glaeser, David Johnston, Pierre Koning, Peter Kuhn, Jon Skinner, Doug Staiger, Frank Vella, and seminar participants at IZA, Tilburg University, Maastricht University, University of Milan, the NBER Public Economics meetings, SciencesPo, University of New South Wales, and Queensland University, the CPB, Erasmus University of Rotterdam, and participants at the IZA/IFAU and LEW conference for helpful conversations and 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

The literature on the effects of social assistance on behavior has focused on labor supply responses to changes in the generosity or eligibility criteria of particular social assistance programs. Estimates of labor supply responses are of course tremendously important but they miss 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.

Social support substitution is pertinent in the U.S, where welfare reform has led many individuals to take up Supplemental Security Income rather than welfare (Schmidt and Sevak, 2004). Similarly, Autor and Duggan (2003) have shown that Disability Insurance applications have become more sensitive to local labor market shocks over the last few decades, indicating that Disability Insurance is increasingly functioning as a form of unemployment insurance. The scope for substitution among social assistance programs in response to changes in their relative generosity or stringency of eligibility is even larger in countries with a more extensive social safety net than that of the U.S.

In this paper, we estimate the extent of social support substitution in response to reforms to disability insurance in the Netherlands as well as the reduced-form labor supply response to this reform. Together, this allows us to estimate the impact of the reform on total income of affected DI recipients. The reforms entailed medical re-examinations of existing recipients and a common outcome of the re-examination was to reduce the benefit level or to terminate eligibility of a current recipient. Two features make this reform particularly suitable for studying substitution between different social assistance programs as well as labor supply responses. First, we have administrative panel data on the universe of Dutch disability insurance claimants, including information

on their future labor market earnings and income from all other government 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 an age discontinuity: the reform was significantly more stringent and led to an average benefit reduction of an additional 10% for the cohort that would turn 45 after August 1st, 1993. Since 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 these 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. The resulting estimates 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 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 forms of 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. Fully disabled recipients have a crowd-out ratio of 0.52, indicating that even they display a substantial labor supply response. 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. We cannot reject that individuals were able to fully offset the cut in their DI benefit by increased income from labor and other social assistance programs.

Over time (up to 8 years after the reform), 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 reform, the benefit-

substitution ratio is still a statistically and economically significant 0.21 and the earnings crowd-out ratio stands at 0.73. 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 16 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 is not just a theoretical possibility but that it 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 have a meaningful capacity to shift 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 Children (AFDC), suggesting that families use SSI and AFDC as substitutes. Kubik

[1] The one exception is the paper by Autor and Duggan (2008), who 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.

(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) and Li and Maestas (2008) use 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 an increased reliance on Social Security disability benefits. 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 has spillover effects on their take up of unemployment insurance and sick leave. Our paper contributes to this literature by estimating substitution between social support programs for prime-age individuals, 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 reform).

Our estimates also contribute to an extensive literature on the labor supply effects of disability insurance (see Bound and Burkhauser 1999 for an overview). Parsons (1980) shows cross-sectional evidence that suggests that the rise 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 et al. (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

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

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, French and Song (2011) use variation in rejection rates due the essentially random assignment of administrative law judges to DI cases, and Maestas et al. (2011) use variation in rejection rates due the essentially random assignment of DI examiners to DI cases. These studies all find clear evidence of labor supply responses to disability insurance. 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 fully disabled individuals.

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

To place the Dutch disability insurance reform in context, it is useful to know that the Netherlands has historically had high DI recipiency rates. Together with other countries such as Sweden (10.8%) and Denmark (7.2%), the Netherlands is in the top ranks among OECD countries in terms of the share of the working-age population receiving DI (8.3%).[3] The corresponding figure for the U.S. is 5.9%. 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.[4] The replacement rate offered by DI depends on the "degree of disability," which is defined by the gap between the prior earnings and the remaining potential earnings capacity of the DI applicant.

[3] These numbers were reported in 2007 and are obtained from the OECD project "*Sickness, Disability and Work*" (www.oecd.org/els/disability).
[4] Also see Bovenberg (2000), who provides useful institutional background information on the Dutch disability act.

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.[5] 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 applicant's level of education were not considered. Third, to prevent the list of suitable occupations from containing only very rare occupations, the list was further limited to occupations with at least 10 active workers in the applicant's region.[6] Because it was regarded as the responsibility of the worker to obtain a position, the occupations on the list did not need to have actual vacancies. Finally, if the list contained at least 5 suitable occupations with at least 10 active workers, 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 possibly to specify 5 suitable occupations with at least 10 workers, the degree of disability was set at 100%. The measured degrees of disability were grouped in 8 categories varying from 0-15% to 80-100%, and these categories determined the replacement rate (see Table 1).

The DI reforms of 1993 tightened this procedure in two respects.[7] 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

[5] 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").

[6] 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).

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

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.[8] With the relaxed criteria for the list with suitable alternative occupations, 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. A recent study by García-Gómez et al. (2011) provides some descriptive evidence that the recent reforms in Dutch DI (as of 2003) were effective in reducing the number of DI applicants.

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 of the reform (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 age group 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 for each individual than

[8] 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."

what it would have been if the old procedure had still been used. Hence, the monotonicity assumption, which is required for our IV regressions, should be satisfied.

In this paper, we exploit the fact that disability claimants aged 44 or younger as of August 1st, 1993 were re-examined according to the new and stricter criteria, while for those aged 45 and above the old criteria applied. This allows us to investigate the effects of the reform by comparing outcomes for those slightly younger than 45 on the date of reform to those of individuals aged slightly older than 45 on that date.

Another important change of the 1993 DI reform was the introduction of an age- and duration-dependent benefit for new applicants.[9] For those already receiving a 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) last earned wage.

Figure 1 shows the number of DI recipients as a percentage of the labor force for the period of 1950 to 2009. Until 1993, the percentage claimants increased steadily. There is a clear dip in the percentage DI recipients for about 3 years following the 1993 reform, after which the number again starts to rise until around 2002, when a new round of reforms was introduced to further reduce the use of disability benefits.[10]

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).[11]

[9] De Jong, Lindeboom, and Van der Klaauw (2011) analyze the effects on new applicants of an experiment in the Netherlands in which caseworkers in 2 (out of 26) regions were instructed to screen DI applicants more stringently. They find that stricter screening reduced DI applications, increased resumption of work (after a period of sickness), and did not affect UI take up.

[10] Initially the reforms of 2002 only concerned the inflow into DI. In 2004, a new round of re-examinations started for all people on DI younger than 50 in 2004. The complete sample in this paper was older than 50 at that time.

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

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), industry information, and the reason for the termination of the disability spell, but does 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, and the identification numbers (RIN-codes) of household members.

Finally, we obtain information on labor market earnings and other sources of social assistance income 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 relative minor programs). Data about 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 of consequences of the reform on labor market outcomes and receipt of other forms of social assistance. 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.

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 on January 1st, 1996. The first restriction is necessary because the discontinuity in benefit rules only applies to existing claimants on the date of the reform. 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 Imbens-Kalyanaraman (2009) test.[12] The last restriction is driven by data availability. While our data on disability starts in 1995, the information in the 1995 file has severe limitations, 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 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 those aged 45 or older would not be subject to the new, more stringent criteria.[13] Moreover, the McCrary (2008) density test shows no evidence of a discontinuity in the density of disability claimants around the cutoff age of 45 (p-value = 0.126).

Our data extends until 2005, when the oldest members of our sample approach age 60. This means that we do not examine the interaction of DI with retirement decisions. Autor and Duggan (2008), and Karlström, Palme, and Svensson (2008) show that for individuals nearing retirement, retirement benefits and other forms of social insurance are substitutes.

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.

[12] 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.

[13] Please note that the age as presented in this paper always refers to the age (in years and months) at the time of the reform (i.e., as of August 1st, 1993).

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 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 as measured at the date of the DI reform, 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 at the time of the reform, the average claimant in our sample had been on DI for about eight years. Finally, about two thirds of the sample are considered to be fully disabled (having lost more than 80% of their earnings capacity) and are therefore eligible for a replacement rate of 70%. The fraction fully disabled is markedly higher among females than 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, 90% of those on DI at the start of 1996 is 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. Over a third of our sample is working, defined as having positive earnings (including from self-employment) in 1999, which is consistent with DI also covering partial disability in the Netherlands. 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 though some of them may have been deceased or may have emigrated.

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 account only for about 6% of total income.

In 2005, so about 7 to 9 years after the re-examinations, 80% of those on DI at the start of 1996 are still on DI. Between 1999 and 2005, the fraction employed has fallen from 36% to 28% and the fraction with income from social assistance other than from the original DI spell has increased from 16% to 25%. These trends are consistent with the general decline in labor force participation in the Netherlands as people approach retirement.

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 2 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. There is a clear discontinuity at the cutoff age, indicating that the more stringent re-examinations for the younger cohort reduced their annual DI benefits by €1076, or about 10%. 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

[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

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 2 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 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. Panel A of Table 3 shows 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 initial, less stringent re-examination rules. Each entry is estimated as the value at exactly age 45 of the line estimated by a regression of a dummy for falling in the cell in question on age for the sample subject to the less stringent re-examination. Panel A shows that replacement rates were relatively stable for those subject to the less stringent re-examination. On average, they fell by less than 3 percentage points. Panel B shows the impact of the more stringent re-examination on the joint distribution of replacement rates in 1996 and 1999. Each entry is estimated as the treatment effect of a RD regression of a dummy for being in the cell in question. The panel shows that the more stringent reform reduced the probability of receiving a 70% replacement rate by about 10 percentage points. About 4 percentage points of this

---

longer eligible for DI, but we cannot rule out that some of those who exited were still eligible for a positive replacement rate.

probability mass moved to those who exited DI altogether and the remaining 6 percentage points showed up as an increase in the probability of receiving a replacement rate of 14% or 21%. In short, the reform not only reduced the generosity of DI for those who exited, but it also reduced the generosity for those remaining on DI. In an unreported RD regression with the change in the replacement rate as the dependent variable, we find that the more stringent re-examination on average decreased the replacement rate by 5.1 percentage points among those remaining on DI.

Table 4 provides an alternative way of describing the impact of the more stringent re-examination on replacement rates. The first column shows the change in replacement rates at age 45.0 for those subject to the less stringent re-examination; these entries are the sums of the diagonal entries in Panel A of Table 3. 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 had no change in the replacement rate. The third column shows the treatment effect of the more stringent re-examination on 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. As a partial check on our identifying assumption that no factors besides the DI re-examinations had a discontinuous impact at age 45, Figure 3 reports DI termination rates separately for 1996/97, 1998, and 1999. Termination is defined as the date for the end of the original DI spell, as recorded in the administrative data file, occurring during the year in question.[15] We calculate these rates as fractions of DI claimants in our sample on January 1st, 1996. In 1996 en 1997, all individuals aged 40-

[15] This means that exit is not equal to one minus the DI participation rate recorded in Figure 2C because for DI participation we also required that the person received a positive amount of DI benefits in the year in question. Some former DI recipients are still in the administrative records because they are repaying excess DI received in the past.

44 were re-examined as well as some of the 45 year olds. Exactly in these years, the discontinuity at age 45 is very pronounced. In 1998, the remainder of the 45 year olds and some of the 46 year olds 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 termination rate over the 1996-1998 period, we find a discontinuous increase in the termination rates for the group subject to the more stringent re-examinations. In 1999, all the re-examinations of 44 or 45 year olds were completed and we find no discontinuity in termination rates at the age cutoff.[16]

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 shuffles individuals across programs and budgetary savings only occur to the extent that benefits in other programs are lower than DI benefits. In the later case, not only can the government reduce DI expenditures, but it will also collect additional tax revenue. 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.

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 4 plots income from 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 reform increased other social assistance income by a €314 per year. The second panel shows that the fraction receiving social assistance income from a source other than the original DI spell discontinuously increases by 4.7

[16] 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.

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.

Next, we present the reduced-form effects of the DI reform on labor market outcomes. The first panel of Figure 5 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. The RD regression, however, estimates that at the cutoff age, earnings are €624 per year higher 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 earnings. This figure establishes our qualitative finding that disability income crowds out labor income. We will discuss the economic magnitude of our crowd-out estimate in the next subsection.

Because we do not have earnings for prior years, we cannot precisely determine to that extent 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 5 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 again 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/yr). It therefore seems likely that some of the response also occurred along the intensive margin. Our finding that the reduction in the generosity of DI increases labor supply contributes to the extensive literature on the labor supply disincentive effects of disability insurance by showing evidence of labor supply responses among prime-age DI recipients who are long-term recipients of DI (with duration of at least 2 years at the time of the reform, but on average 10 years).

4.3 Benefit Substitution and Earnings Crowd Out

Figures 4 and 5 established 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 5, we scale our outcome measures by the amount by which disability benefits from the original spell decrease at the age discontinuity whereas in the second column we scale our outcome measures 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 the case with a valid RD design, the control variables do not substantially affect the magnitudes of our estimates (See Appendix Table A1). 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 everyone and weakly decreases the replacement rate for everyone. 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 effective incentive for work change. Rather, we see the IV estimates as a way to relate the magnitudes of the behavior effects of the reform to alterative measures of the size of the reform. In other words, we view this primarily as a scaling exercise.

Panel A of Table 5 examines to what extent the reduced generosity of disability benefits causes individuals to shift to other forms of social assistance (including new DI spells). The first row of column 1 of panel B shows that for each Euro in decreased disability benefits, individuals receive €0.31 more from other social assistance programs in 1999. Thus, the benefit-substitution ratio is 0.31.[18] A government not taking this substitution into account would overestimate the reduction in government expenditure

[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).

[18] Most of the substitution happens towards unemployment insurance, a sizeable fraction occurs through general assistance, while only a small fraction is due to re-entry into DI.

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.

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 induces 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. The second row of Panel A 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. 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 effective taxed 100% on the margin through reduced DI benefits. Thus, if the re-examination led to a reduction in the degree of disability, this both reduced the DI benefit (which is an income effect) and it increased the 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 a substitution effect and by an income effect. 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.

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, 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). 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 some other source. The fact that this estimate is less than the sum of the estimates in row 2 of panels A and B indicates that the reform induced some individuals to both work and draw 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 5 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 A3 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 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 Gender and Degree of Disability

Table 6 splits out the results of the first column of Table 5 by gender. The table suggests 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 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 of DI benefits by other sources of income. As noted earlier, it hard to determine what 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.

Table 7 analyzes benefit substitution and earnings crowd out by degree of disability.[19] The degree of disability is measured as of January 1st 1996, so before the re-examinations took place. Panel A shows that social support substitution is much more prevalent among fully disabled DI recipients than among partially disabled ones. 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 fully disabled is 0.50, which is four times as large as the ratio of 0.12 for partially disabled recipients. Because about two thirds of the partially disabled have some earnings but only about 15 percent of the fully disabled do, the fully disabled may be more likely to be able to qualify for means-tested alternative sources of social assistance, such as general assistance.

In contrast, we find high rates of labor crowd out both for partially disabled and fully disabled recipients. 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

[19] 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 A3 for details.

region and that an applicant could also be classified as fully disabled if not enough of such jobs were found.

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

Our data has information on 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.[20] 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 disproportionally high exit rates among those with diagnoses of musculoskeletal, psychiatric, and neurological conditions and disproportionally low exit rates among those whose diagnosis was labeled "general". Full results are in Appendix Table A4.

4.5 Responses of Partners of DI recipients

In Table 8, we provide estimates of benefit substitution and earnings crowd out at the household level. These estimates differ from our baseline estimates of Table 5 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.

[20] 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.

The increase for earnings crowd out, while at 20% 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 partner responses, however, decreases the precision of our estimates, which is why we exclude them from the rest of our analysis.

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 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. Incidentally, looking beyond 2005 is less informative because at that point some members of our sample are starting to reach an age where early retirement is quite common in the Netherlands. While additional DI reforms took place in 2002 and 2004, these reforms had no impact on the people in our sample. 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 our estimates.

Figure 6 shows our estimates of the benefit-substitution ratio and the earnings crowd-out ratio for each of the years from 1999 to 2005. 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 5, but now for all years until 2005. The top panel shows the results for the entire sample while the bottom two panels split out the results by gender. We find that both the earnings crowd-out ratio and benefit-substitution ratio are positive and statistically significant for the full sample 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 21% in 2005. This decrease, however, is not statistically significant. The decline of the benefit-substitution ratio is consistent with the fact that unemployment assistance is only available for a limited duration. General assistance, however, has no time limits, which explains why the benefit-substitution ratio can remain positive even in the long term. The figure shows a slight increase over time in the earnings crowd-out ratio, which rises from 62% in 1999 to 73% in 2005, but this increase is not statistically significant. The separate figures for men and women roughly display the same patterns as the overall figure.

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 extend 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. This question 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. Moreover, it is important to recognize 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 an age-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 21% about 8 years after the implementation of the reform for our sample.

We also find a remarkable labor supply response given that all members of our sample were at least partially disabled. On average, individuals are able to make up 62% of their foregone DI benefits, 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 may mask much 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 is in which the estimates would change is not clear. Our benefit-substitution figure may 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 of 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 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 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 circumstances. 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 of the reform. Our setting does not allow us to estimate spillover effects stemming from people who would have claimed disability insurance under less stringent rules, but who take up another form of social assistance because of the increased stringency of the DI benefit rules. We view such spillover effects as complementary to our estimates of benefit substitution. We see the estimation of these complementary spillover effects as an important area for further research.

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.

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.

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, 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.

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.

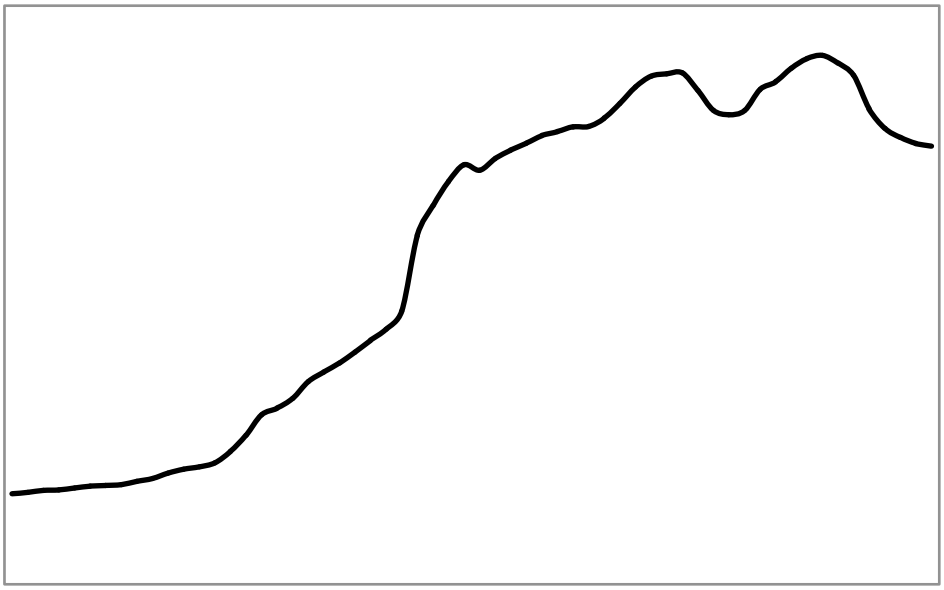
Staubli, Stefan. 2011. "The Impact of Stricter Criteria for Disability Insurance on Labor Force Participation," *Journal of Public Economics* 95 (9-10): 1223–1235.

SZW (Dutch Ministry of Social Affairs and Employment). 2002. "De Nederlandse WAO in Internationaal Perspectief" (in Dutch), Working paper 241, Ministry of SZW, Den Haag.

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.

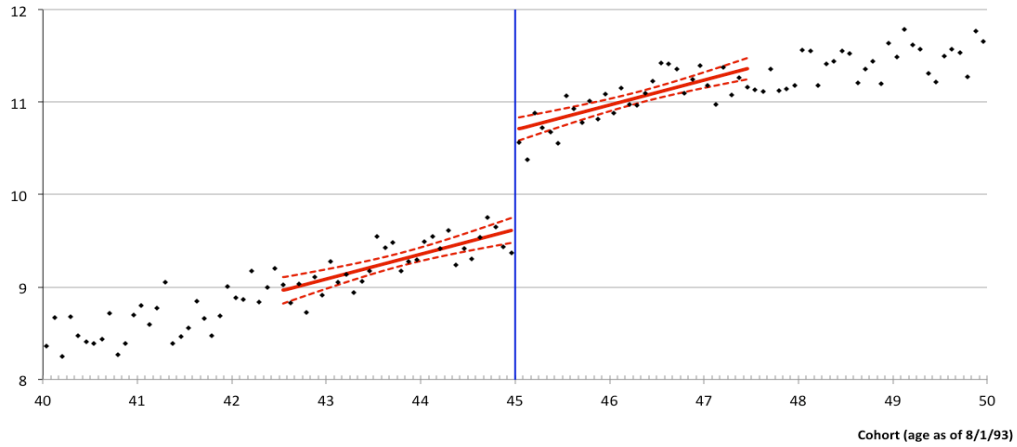
Figure 1: Percent of the labor force receiving DI benefits, 1950-2009.



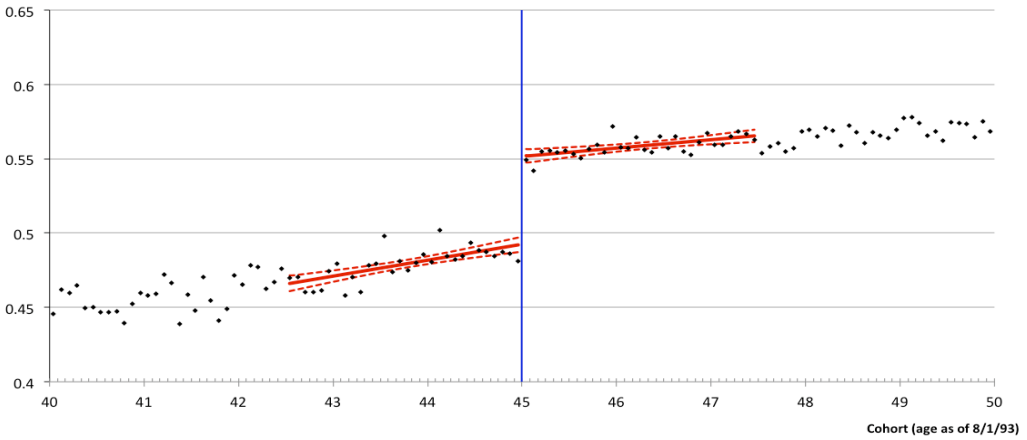
Source: Statline Statistics Netherlands

Figure 2: Magnitude of the reform

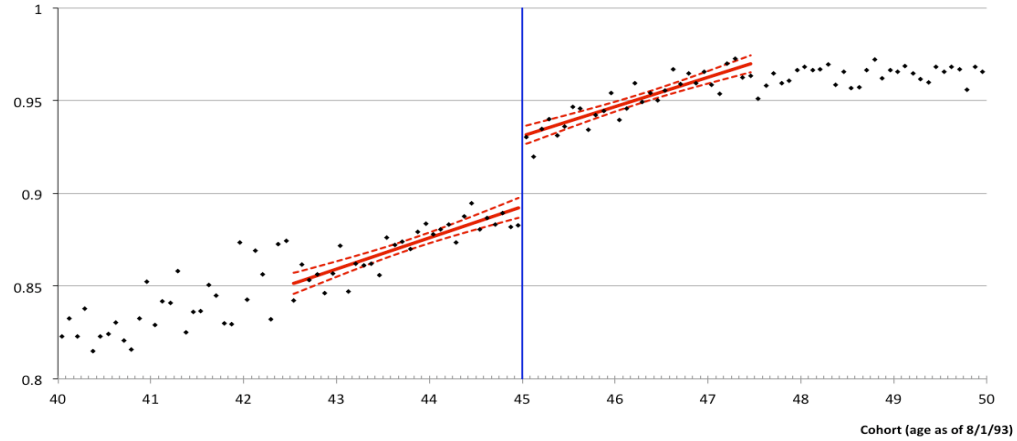
Panel A: Effect on DI Benefit Amounts (1000 €/yr). Estimate of the discontinuity: -1.076 (0.096)***



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



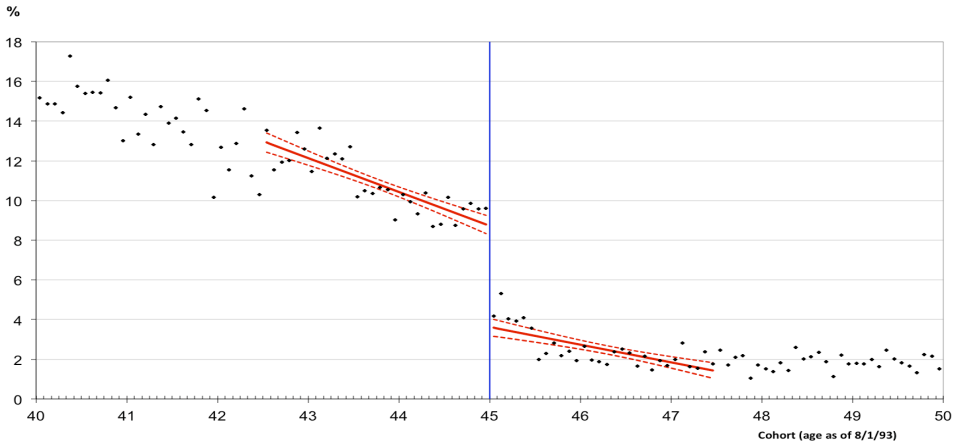
Panel C: Effect on Participation in DI in 1999. Estimate of the discontinuity: -0.038 (0.004)***



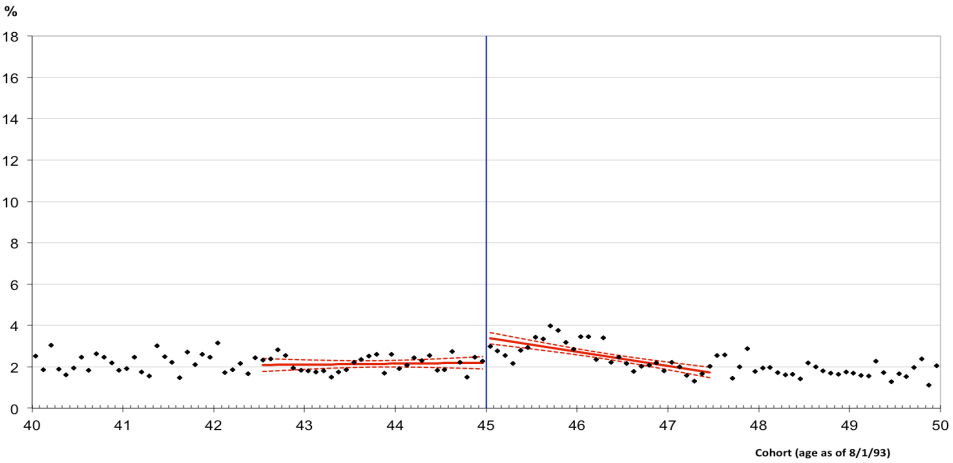
Note: Each figure is based on 84,185 observations. The dotted lines represent the 95% confidence intervals. Standard errors are in parentheses. Regression estimates come from regressions without demographic control variables.

Figure 3: Termination of administrative record in 1996/7, 1998 and 1999

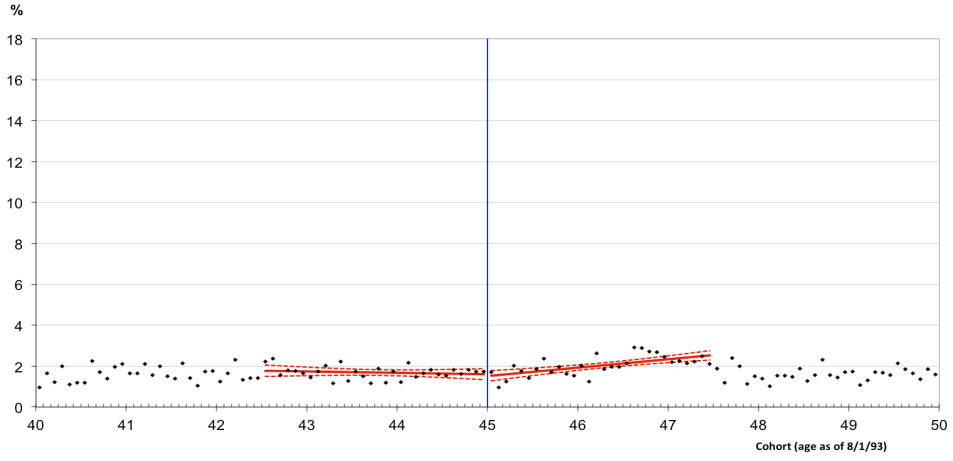
Panel A: Termination in 1996/1997. Estimate of the discontinuity: 0.051 (0.003)***



Panel B: Termination in 1998. Estimate of the discontinuity: -0.012 (0.002)***



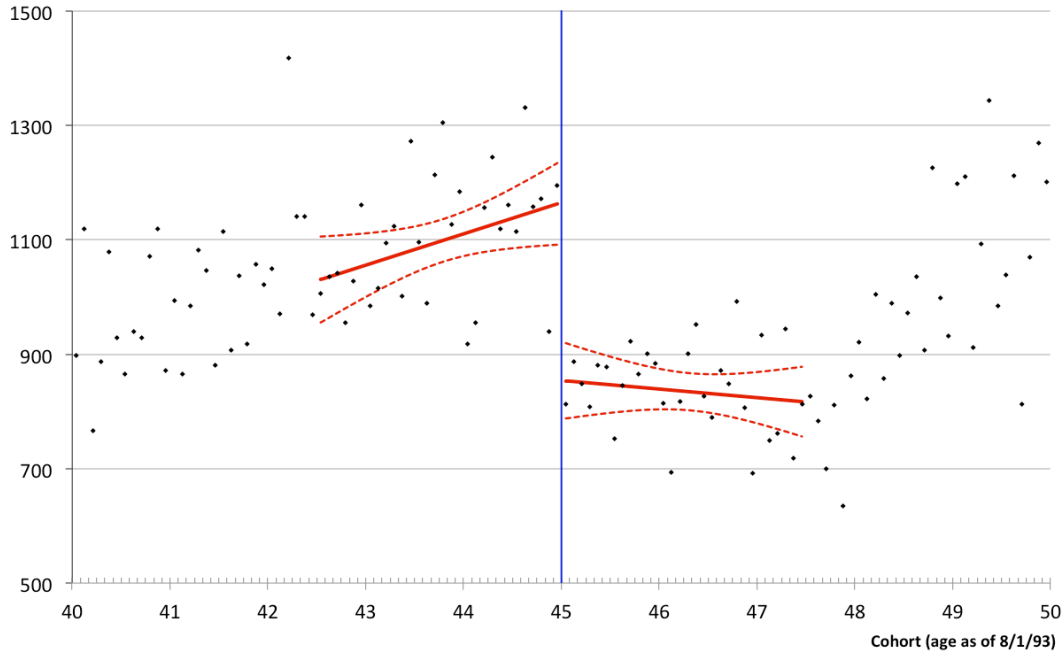
Panel C: Termination in 1999. Estimate of the discontinuity: 0.001 (0.002)



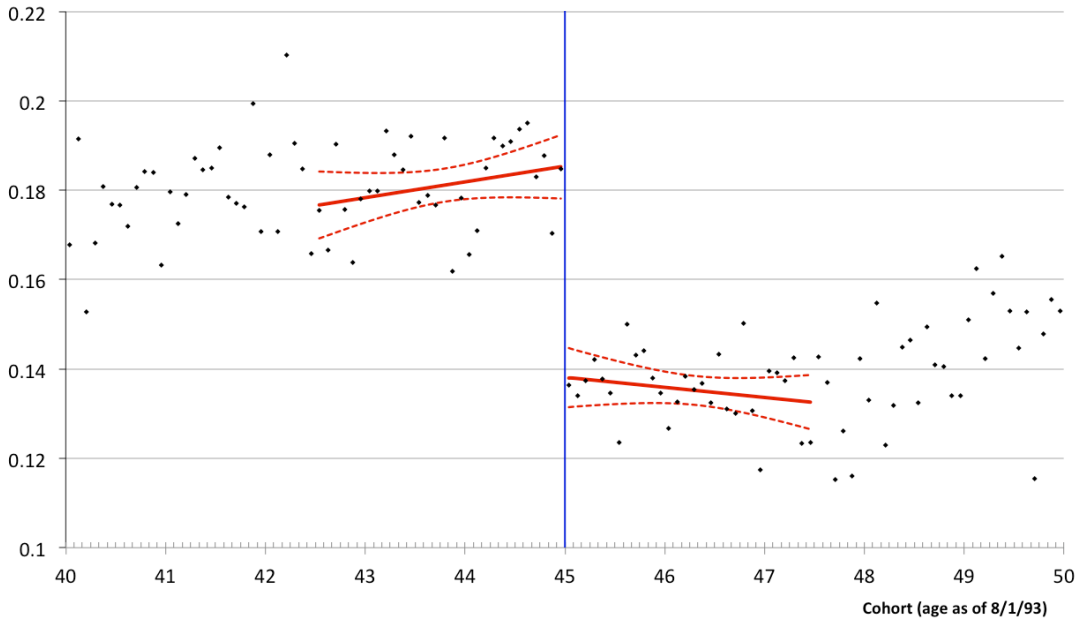
Note: The figures are based on 84,185 observations. The dotted lines represent the 95% confidence intervals. The exit rate is defined as a fraction of our sample in 1996. Regression estimates come from regressions without demographic control variables. Standard errors are in parentheses.

Figure 4: Effects of DI Reform on Other Social Assistance

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



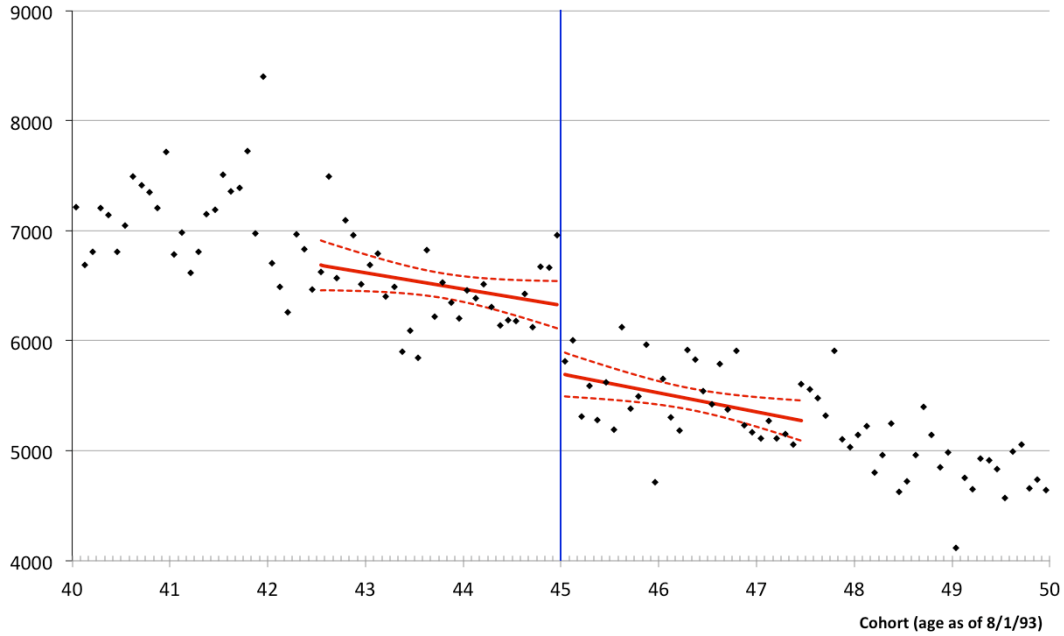
Panel B: Social Assistance Participation in 1999. Estimate of the discontinuity: 0.047 (0.005)***



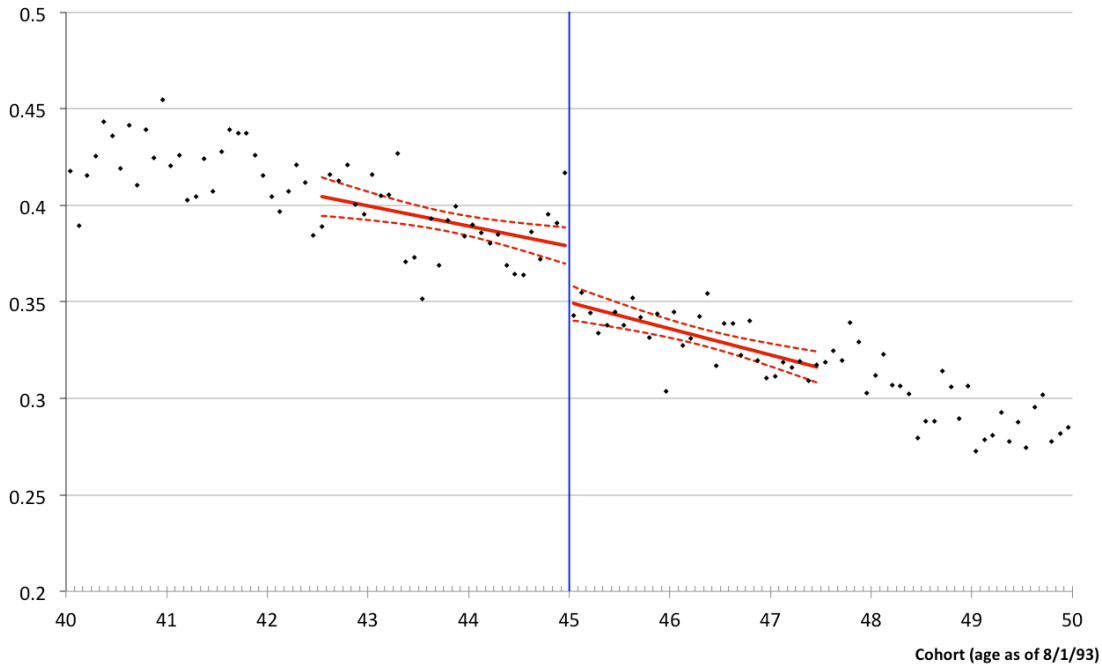
Note: Figures are based on 84,185 observations. The dotted lines represent the 95% confidence intervals. Benefits from the original DI spell are not included in social assistance income, and participation rates exclude the original DI spell. Regression estimates come from regressions without demographic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

Figure 5: Effects of DI Reform on Labor Market Outcomes

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

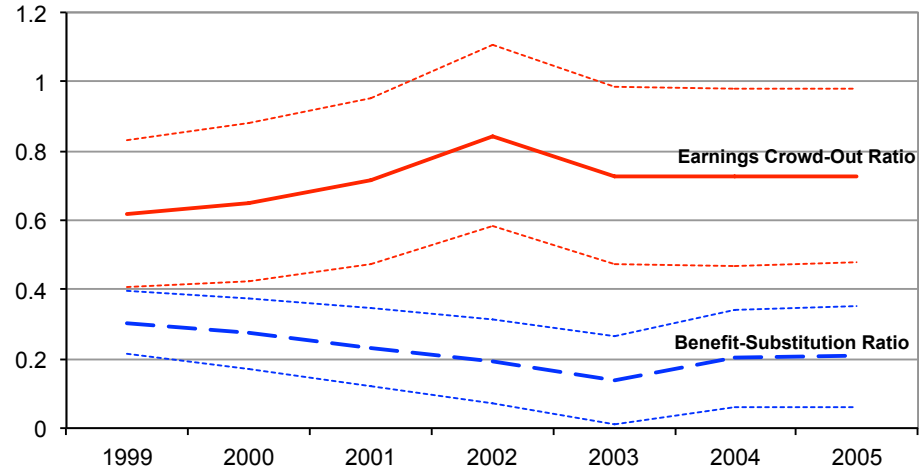


Panel B: Effect on Employment in 1999. Estimate of the discontinuity: 0.029 (0.007)***

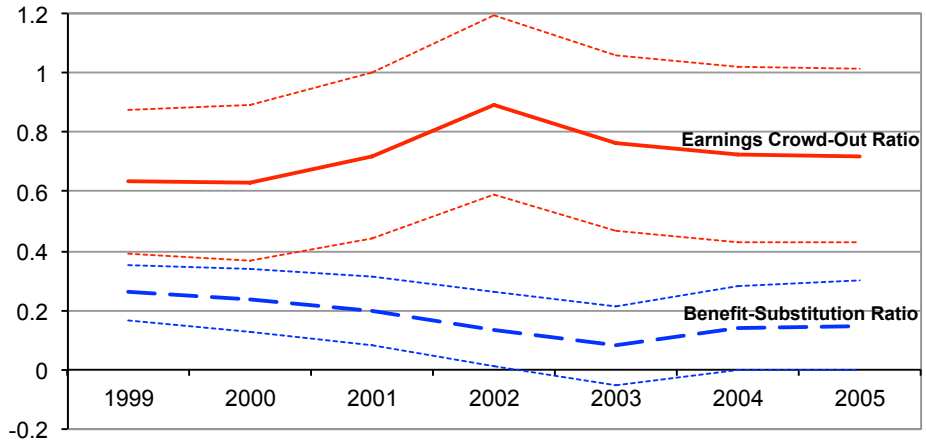


Note: Figures are based on 84,185 observations. The dotted lines represent the 95% confidence intervals. Employment is defined as having positive earnings from employment or self-employment. Regression estimates come from regressions without demographic control variables. Standard errors are in parentheses.

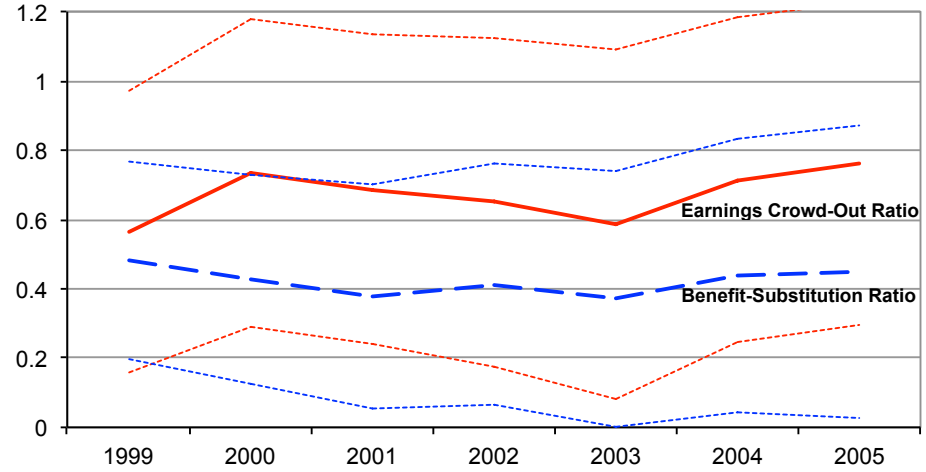
Figure 6: Earnings Crowd Out and Benefit Shifting over Time
Panel A: Full Sample



Panel B: Males



Panel C: Females



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 1st, 1993	45.18	45.19	45.18
Start date of DI spell	1985.13	1984.82	1985.73
Duration on DI (months; as of August 1993):	96.72	100.36	89.57
Degree of disability (% of earnings capacity lost):			
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			
Labor market status in 1999 (%):			
Still on DI (on the original spell)	89.87	89.68	90.24
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
Labor market status in 2005 (%):			
Still on DI (on the original spell)	79.85	78.83	81.86
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
Income by source in 1999, €/yr (including zeros):			
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
Income by source in 1999, €/yr (if non-zero):			
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
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 1995 (before the re-examinations).

Table 3: Replacement Rates and Re-examinations

Panel A: Joint distribution of 1996 and 1999 replacement rates at age 45.0 under the less stringent re-examination

	Replacement rate in 1999								
Replacement rate in 1996	0%	14%	21%	28%	35%	42%	50.75%	70%	Total
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 1999								
Replacement rate in 1996	0%	14%	21%	28%	35%	42%	50.75%	70%	Total
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: 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 but without demographic controls, where outcome variable is a dummy for combination of replacement rates in 1996 and 1999 that corresponds to that cell. Standard errors in parentheses. Those who exit DI before the post-examination replacement rate is recorded are assigned a replacement rate of 0. N=84,185.

Table 4: Re-examinations and the Change in the Replacement Rate between 1996 and 1999

Change in the replacement rate	(1) Predicted probability at age 45.0 for the less stringent re-examination	(2) Predicted probability at age 45.0 for the more stringent re-examination	(3) 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: Each row is estimated using our standard reduced-form RD regression but without demographic controls, where outcome variable is a dummy for the change in the replacement rates 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). Standard errors in parentheses. N=84,185.

Table 5: Earnings Crowd Out and Benefit Shifting

	Effect scaled by decrease in amount of original DI (in 1000 €/yr)	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 spells). The variable that is instrumented (endogenous explanatory variable) is listed in the columns. The instrument itself is the discontinuity at the cutoff age (45 as of 8/1/93). 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 analyses: age (in months), dummies for degree of disability in 1996, pre-DI earnings (and its squared and tripled value), national origin dummies, a dummy for being married in 1996, regional dummies, duration in DI at the start of the reform (and its squared and tripled value), interactions between the dummies for the degree of disability and pre-DI earnings, a gender dummy, and interactions between all controls and gender.

Table 6: Earnings Crowd Out and Benefit Shifting by Gender

	Effect of reform per 1000 €/yr decrease in amount of original DI				P-value gender difference
	Males		Females		
Panel A: Other social assistance in 1999					
Income from other social assistance	0.261	(0.047)***	0.482	(0.146)***	0.149
Participation dummy	0.038	(0.005)***	0.074	(0.016)***	0.037
Panel B: Labor market outcomes in 1999					
Earnings	0.632	(0.124)***	0.564	(0.208)***	0.781
Employment dummy	0.023	(0.005)***	0.053	(0.013)***	0.026
Panel C: Total					
Income except from original DI spell	0.892	(0.127)***	1.046	(0.250)***	0.582
Dummy for work or other social assistance	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. The variable that is instrumented (endogenous explanatory variable) is the amount of DI, so all coefficients can be interpreted as effect size per €1000/yr decrease in DI. The instrument itself is the discontinuity at the cutoff age (45 as of 8/1/93). The regressions are based on 55,772 and 28,413 observations for males and females, respectively. See the note to Table 5 for the demographic controls included in the regression.

Table 7: Earnings Crowd Out and Benefit Shifting by Degree of Disability

	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)***	0.015 (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 (0.006)***	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/yr decrease in DI. The instrument itself is the discontinuity at the cutoff age (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 5 for the demographic controls included in the regression.

Table 8: 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/yr decrease in DI. The instrument itself is the discontinuity at the cutoff age (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 5 for the demographic controls included in the regression.

APPENDIX

Appendix Table A1: Baseline Estimates Without Controls Variables

	Effect of reform per 1000 €/yr decrease in amount of original DI					
	Full Sample		Males		Females	
Panel A: Other social assistance in 1999						
Income from other social assistance	0.291	(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 5 and 6 except that the regressions for this table do not include demographic controls.

Appendix Table A2: Sensitivity of Baseline Estimates to Bandwidth

		Effect scaled by decrease in amount of original DI (in 1000 €/yr)			
	N	Earnings		Income from other social assistance	
+/- 5 years	160207	0.330	(0.088)***	0.412	(0.037)***
+/- 4 years	129557	0.417	(0.097)***	0.346	(0.038)***
+/- 3 years	98668	0.596	(0.097)***	0.266	(0.042)***
Baseline (+/- 2.5 years)	84185	0.618	(0.108)***	0.305	(0.047)***
+/- 2 years	66818	0.825	(0.119)***	0.305	(0.052)***
+/- 1 year	32772	0.778	(0.183)***	0.395	(0.083)***

Note: Standard errors are in parentheses. Significance levels: * 10 percent; ** 5 percent; *** 1 percent. The table shows our key estimates of 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 A3: Heterogeneity of Effect

	Effect scaled by decrease in amount of original DI (in 1000 €/yr)		
	N	Earnings	Income from other social assistance
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)***
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)***
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)***
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)***
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)***

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. None of the differences by subsample are statistically significant except the difference in the benefit-substitution ratio by degree of disability.

Appendix Table A4: Inferred Distribution of Medical Conditions Among Those Induced to Exit DI

	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
Medical condition	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)	(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 our standard reduced-form RD regression 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).