Netspar

Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

The impact of the disability insurance reform on work resumption and benefit substitution in the Netherlands

Tunga Kantarci, Jan-Maarten van Sonsbeek, Yi Zhang

DP 11/2019-039

The impact of the disability insurance reform on work resumption and benefit substitution in the Netherlands

Tunga Kantarcı, Jan-Maarten van Sonsbeek, and Yi Zhang

Tilburg University and Netspar

Netherlands Bureau for Economic Policy Analysis and Netspar

Tilburg University

18 November 2019

Preliminary version

Abstract

We evaluate the disability insurance reform introduced in 2006 in the Netherlands. Using unique administrative data on individuals who fall sick before and after the reform, we analyze how the reform affected labor participation and the use of benefits from alternative benefit programs. The results of the difference-in-difference regressions suggest that the stricter benefit regime of the reform limits the access to disability insurance substantially, and increases labor participation and unemployment benefit receipt to a limited extent. However, it increases wages earned and unemployment benefits received to sizable extents. The impact of the reform is persistent over time for disability benefit and labor participation and income. The reform is least effective for unemployed and older individuals who struggle most to resume working.

1 Introduction

In many western countries the number of disability benefit recipients and the share of the disability insurance program in the total public expenditure have grown in the past decades. During the past 50 years, benefit receipt in the Social Security Disability Insurance (SSDI) rose from less than 1 to 5 percent in the United States, and that in various disability benefit programs rose from 1 percent to 7 percentin the United Kingdom (Autor et al., 2019). During the period from 1990 to 2005, the total expenditure on disability benefits accounted for approximately 2.5 percent of the gross domestic product in the OECD countries, on average (OECD, 2010).

Through disability reforms, governments introduced different types of measures to reduce disability benefit claiming and increase labor participation among the sick individuals. Existing studies focus on analysing the effects of two main measures. The first is tightening the entrance criteria (Autor and Duggan, 2003; Karlström et al., 2008; De Jong et al., 2011; Staubli, 2011; Campolieti and Riddell, 2012; Borghans et al., 2014; Burkhauser et al., 2014; Moore, 2015; Autor et al., 2016), and the second is reducing the benefit level (Gruber, 2000; Campolieti, 2004; Maestas and Song, 2011; Kostøl and Mogstad, 2014; Deshpande, 2016; Mullen and

Staubli, 2016; Fevang et al., 2017; Koning and van Sonsbeek, 2017; Ruh and Staubli, 2018). A main finding in this literature is that restricting entitlement and reducing the benefit level increase labor participation and inflow into alternative benefit programs but that these effects are heterogeneous across subgroups of sick individuals.[1]

In the early 2000s the Netherlands had become one of the countries with the highest fraction of disabled workers in the working population, as the total number of disability benefit recipients reached almost one million whereas the working population was around 7 million. Successive governments implemented a series of radical reforms to limit access to the disability insurance scheme. In 2006, the Work and Income According to labor Capacity Act (WIA) came into effect as the final element of these reforms, as a successor to the Disability Insurance Act (WAO). The WIA extended the sickness benefit scheme that precedes the disability benefit scheme from one to two years, introduced stricter criteria to enter the disability benefit scheme, but also introduced financial incentives for work resumption at the more advanced stage of the disability benefit scheme. Among the three measures introduced, extension of the sickness period makes the 2006 reform distinctive compared to disability reforms in other countries. The strong incentive for the employer to facilitate work resumption is that the employer is obliged to compensate the employee for wage loss during the two years period of the sickness scheme. However, mandating employers to pay wages for a longer period may also make them become reluctant to hire workers with health problems (Koning and Lindeboom, 2015). These employer incentives during the sickness period, however, precede the incentives built into the disability scheme and therefore could affect individual behavior during the disability period. The WIA reform therefore offers a unique incentive for work resumption that is not seen in disability reforms in other countries.

Research analyzing the impact of the stricter rules of the WIA regime in comparison to the rules of the older WAO regime is limited to Van Sonsbeek and Gradus (2013). Based on data on disability applications, they find that the stricter eligibility criteria of the WIA have led to a sharp fall in the number of new benefit claims, in addition to what has already been achieved through previous reforms, generating large budgetary savings for the government. The lack of research on the impact of the WIA reform is due to lack of data on sick individuals who subsequently could file disability applications. The sickness benefit was reformed in 1994, 1996 and 2004 to mandate the employer to pay during the sickness period 70 percent of the earnings before sickness. Since no sickness benefit is paid by the government, but wage is paid by the employer, there is no registration of sickness absence by the government since these reforms. New reintegration regulations for employers were introduced in the sickness scheme in 2002 ("Gatekeeper protocol"). Only after this year the government started to register sickness cases to monitor whether employers comply with the new regulations. This has so far hampered evaluations of the disability reform directed at the sickness absence period. The existing research on the impact of the WIA reform is therefore limited in several respects. First, it is not analyzed to which extent the decrease in disability benefit use among sick individuals led to an increase in labor participation or use of benefits from alternative benefit programs such as

[1]Instead of exploiting exogenous variation in benefit rules due to policy reforms, Chen and van der Klaauw (2008), Maestas et al. (2013), French and Song (2014), and Gelber et al. (2017) exploit variation in benefit eligibility rules or benefit levels imbed in disability benefit programs to analyze the causal effect of benefit incentives on benefit receipt and labor supply. A smaller literature investigates how incentive changes on the employer side influence disability benefit receipt. For example, De Jong and Lindeboom (2004) show that mandating firms to use preventive and reintegration measures to reduce sickness absenteeism does not decrease absence rates. Koning (2009) and Groot and Koning (2016) analyze the introduction and abolishment of experience rating for firms' disability insurance premium, and find that experience rating effectively decreases disability benefit receipt. Another strand of the literature pays attention to the non-economic outcomes of disability reforms (Dahl and Gielen, 2018; García-Gómez and Gielen, 2018). García-Gómez and Gielen, for example, find that despite the gains in public finances, stricter eligibility criteria reduce life expectancy among women.

the unemployment benefit. Second, it is not known whether the effects of the WIA reform are structural or fade in the long run, for example because people who do not first enter the WIA later become more sick and still become incapacitated for work. Third, little is known about how the effects of the WIA reform vary across subgroups of sick individuals.

In this study we exploit unique administrative data from the Employee Insurance Agency (UWV) on individuals who fell sick in the third and fourth quarters of 2003 and the first quarter of 2004. The two groups of individuals who fell sick in the last two quarters of 2003 are insured under the old WAO scheme but are subject to different eligibility criteria, while the third group of individuals who fell sick in the beginning of 2004 is insured under the WIA scheme and is subject to additional and new eligibility criteria. To investigate the effects of the changes in the rules of the disability benefit scheme, we compare the labor market and benefit claiming behavior of the three groups of individuals before they fall sick and after they become eligible to apply for disability benefits. The three groups of sick individuals are comparable in background characteristics, and economic shocks are likely to affect the behavior of the three groups in similar ways since eligibility for different disability schemes are determined by falling sick within very close proximity in time. This allows to attribute the differences in the labor market and benefit claiming behavior across the three groups to the differences in the rules of the disability benefit schemes that apply to these groups. We take a difference-in-difference approach to identify the causal effects of the disability reform.

The regression results show that the WIA reform substantially decreased disability benefit receipt and the amount of disability benefits received. Individuals respond by increasing their labor participation to a limited extent, but their earnings substantially. They also increase their use of unemployment benefit, and substantially the amount they receive from unemployment benefit. Use of general assistance and other benefits decreases but the effects are small. The impact of the reform on disability benefit receipt and work resumption is persistent over time, while the spillover effect on alternative social security programs dies out in about seven years after individuals become eligible to apply for disability benefits. Unemployed and older individuals appear as worst affected by the reform. Compared to their counterparts insured under the WAO regime, they are less able to compensate the decrease in disability benefit receipt or income with higher labor participation or wages. whereas they more often rely on rely on unemployment benefit.

The remainder of this paper is organized as follows. Section 2 describes the institutional context. Section 3 describes the data. Section 4 presents descriptive statistics and exploratory graphical analysis. Section 5 describes the identification strategy. Section 6 presents the baseline results and analyzes the long-term and heterogeneous effects of the reforms. Section 7 discusses policy implications and concludes.

2 Institutional setting

The Disability Insurance Act (WAO) came into effect in 1967 to insure against loss of earnings due to long-term disability. The act was amended several times but since the main amendments in 1993 it preserved its main features until it was replaced by the Work and Income Act (WIA) in 2006.

The WAO consists of two schemes. The individual who earns wage or receives unemployment benefit is first admitted to the sickness scheme if he is unable to perform his work because of illness or injury irrespective of its cause. The employer is responsible to pay 70 percent of the former wage during the one year duration of the scheme. Most employers, however, pay the full amount of the former wage. When the sickness scheme expires, the individual is admitted to the disability scheme if his disability grade is at least 15 percent. The disability grade is

determined by dividing the estimated wage loss due to disability by the former wage, where estimated wage loss is given by the difference between the former wage and the potential wage that the sick individual can still earn. An ergonomist determines the potential wage by taking the average of the highest wages the sick individual could still earn in three suitable occupations. The individual is first entitled to the "Wage-loss benefit" that replaces 70 percent of the former wage multiplied by the disability grade. The duration of the benefit depends on the age of the individual and is limited to a maximum of 6 years. When the Wage-loss benefit expires, the disabled individual is entitled to the "Follow-up benefit" that is lower than the Wage-loss benefit and pays the minimum wage and an additional amount that depends on the former wage and the age at which the individual has become entitled to the benefit. The benefit is paid as long as the individual is disabled but expires when the individual becomes entitled to the state pension.

Due to easy access to the WAO, the annual inflow rate into the disability scheme increased to reach 1.5 percent of the insured working population in 2001. Subsequently the disability scheme was reformed. Before the WAO was abolished entirely, however, a transitional scheme was introduced on 1 October 2004 for people who have fallen sick during the period from 1 October 2003 until 1 January 2004. In the transitional scheme the features of the sickness and disability benefit schemes of the WAO have been preserved except that the criteria to enter the disability scheme have been made stricter. In particular, the transitional scheme has adapted a broader definition of what work can still be done by the applicant. This makes it easier to find potential jobs the individual can still perform. As a result, the wage loss due to disability can be estimated to be smaller making it difficult to reach the minimum disability grade to become eligible for disability benefit, or to reach a higher disability grade that leads to a higher Wage-loss benefit.

The WIA came into effect on 1 January 2006 for people who have fallen sick from 1 January 2004 onwards. It introduced major changes in both the sickness and disability schemes to facilitate work resumption and succeeded to limit the yearly inflow rate into the disability scheme to 0.5 percent of the insured working population during the first six years since its introduction (Koning and Lindeboom, 2015). For the sickness scheme, the duration of the scheme is extended from one to two years.[2] The strong incentive for the employer to facilitate work resumption is that the employer is obliged to compensate the employee for wage loss during the two years period of the scheme. The compensation must amount to 70 percent of the former wage. Most employers, however, pay the full amount of the former wage during the first year of sickness, and many pay more than 70 percent of the former wage during the second year of sickness.

The disability scheme of the WIA inherited the stricter eligibility criteria of the disability scheme of the transitional WAO that uses the broader definition of what work can still be done by the applicant. In addition, it introduced a number of major changes. First, the minimum grade of disability required to enter the scheme was increased from 15 to 35 percent. Therefore, workers with limited disability are expected to resume working with adaptations, or to apply for unemployment benefit.

Second, the scheme introduced a distinction between full and partial disability, and accordingly two specialized disability schemes. If the wage loss is more than 80 percent and there is no potential for any degree of recovery, the worker is admitted to the Full Invalidity Benefit Regulation (IVA), and is entitled to a benefit that replaces 75 percent of the former wage. Ad-

[2]Strictly speaking, the extension of the sickness benefit from one to two years is not part of the WIA, but is part of a separate law, called Verlenging Loondoorbetalingsverplichting bij Ziekte (VLZ). However, both laws are part of the same package of reforms and their starting dates were synchronised so that both laws are effective from 1 January 2004 onwards.

mission to the scheme is limited to a selective group of impairments that are expected to be permanent so that moral hazard is unlikely at least among this small group of workers.

If the wage loss is more than 35 percent and less than 80 percent, or if the wage loss is more than 80 percent but there is still a potential for recovery, the worker is insured under the Return to Work Regulation (WGA). The eligible worker is first entitled to the "Wage-related benefit". Like the Wage-loss benefit of the WAO, the Wage-related benefit is related to the former wage. It replaces 70 percent of the former wage multiplied by the disability grade if the individual utilizes his remaining work capacity to its full potential. The benefit has an unemployment benefit component that compensates the individual if he is not able to utilize his remaining work capacity. The duration of the benefit depends on the employment history, and is limited to a maximum of 38 months.

When the Wage-related benefit expires, the disabled individual is entitled to one of two types of benefits depending on whether he utilizes more than 50 percent of his remaining earning capacity. If the individual utilizes at least 50 percent of his remaining earning capacity, he is entitled to the "Wage-supplement benefit" which replaces 70 percent of the former wage multiplied by the disability grade. If the individual utilizes less than 50 percent of his remaining earning capacity, he is entitled to the "Follow-up benefit" which replaces 70 percent of the minimum wage multiplied by the disability grade. These mean that both the Wage-supplement and the Follow-up benefits make flat rate payments and hence disregard how much the individual is working below or above the threshold utilization rate of remaining work capacity. Both benefits are paid as long as the individual is disabled but expire when the individual becomes entitled to the state pension. At a given disability grade, the difference between the Wage-related benefit and the Follow-up benefit is as large as 70 percent of the difference between the former wage and the minimum wage, giving the partially disabled workers with higher former wages a stronger incentive to utilize at least 50 percent of their remaining work capacity when the wage-related benefit expires.

Finally, the WIA extended the "experience rating" period. The experience rating reform introduces differentiation in the premium firms pay to the disability insurance program. The premium amount is based on the costs of the disability benefits of the employees from the past. Hence, firms with high disability costs are punished with a higher premium. In the WAO, experience rating applied to employer contributions to disability insurance for a period of 5 years for all disabled workers. In the WIA, however, the experience rating period is extended to 10 years and applied to employer contributions for disabled workers participating in the WGA, but not in the IVA scheme.

In the WAO or the WIA, if the individual has no employer during his participation in the sickness scheme, he is eligible for the "Sickness benefit" (ZW). The Sickness benefit replaces 70 percent of the former wage. During participation in the disability scheme, the individual is eligible for the unemployment benefit. The amount of the unemployment benefit is a certain fraction of the remaining earning capacity. In the WAO, the individual is required to file an application to claim the unemployment benefit. In the WIA, however, the unemployment benefit is integrated into the disability benefit, and therefore no application is required. In fact, the duration of the Wage-related benefit is determined by the duration of the unemployment benefit.

If the benefit received from a benefit scheme (sickness, disability, or unemployment scheme), or the wage earned during the second year of sickness (in the WIA), is lower than the applicable social minimum, it is supplemented up to the social minimum according to the Supplementary Benefits Act (Toeslagenwet). The total of the benefit and the social minimum supplement cannot exceed the former wage. If the individual is living with a partner, the supplement is granted if the total income of the individual and the partner is below the social minimum. If

the individual is living alone, the amount of the supplement depends on whether the individual has children.

3 Data

We use unique administrative data from the Employee Insurance Agency (UWV) on three cohorts of sick people who face different criteria to enter the disability scheme and different incentives to resume working if participating in the disability scheme. In particular, the data contains information on all individuals who fell sick in the third quarter of 2003, fourth quarter of 2003, and the first quarter of 2004 and therefore became eligible to participate in the WAO, transitional WAO, and the WIA schemes, respectively. For these people we observe the beginning and ending dates of sickness, sector of employment, gender, date of birth, and country of birth. These people either earn wage or receive unemployment benefit at the time they fall sick since people of other labor market groups are not eligible to enter the sickness scheme. For people in employment, we observe whether they hold a regular contract, temporary contract, or a contract through a temporary work agency. We merge the administrative data on sickness with administrative data on labor participation, wage, and benefits, all available on a monthly basis from Statistics Netherlands. The benefits are from various benefit schemes which include the disability insurance, unemployment insurance, general assistance for low-wage earners, and other benefits from a large number of smaller benefit programs. The data from Statistics Netherlands extend from January 1999 to December 2015 and allow to study the differences in benefit claiming and labor market behavior of the three cohorts of sick individuals over a long period of time.

The initial sample of sick people consist of 51,319,668 observations for 251,567 individuals. We impose a number of restrictions on the initial sample. First, in the data, the sickness cases that last shorter than 180 days are under-reported among the individuals who fell sick in the earlier months of the period from July 2003 until March 2004. The earlier the date individuals fell sick, the more severe under-reporting is in the data. Factors that might have contributed to this are not clear. To ensure that the three cohorts of sick individuals are comparable in the number of days spent in sickness, we drop the sickness cases that last shorter than 180 days. This restriction leads to sample of 19,690,488 observations for 96,522 individuals. Second, we drop the individuals who start receiving the disability benefit before they become eligible to apply for the benefit. The majority of this group are existing disability benefit recipients due to an earlier sickness application, and assessing the impact of the reform is much harder among the members of this group as they participate in multiple schemes or participate in the same scheme multiple times. This restriction leads to a sample of 16,028,076 observations for 78,569 individuals. Third, we drop individuals if they are participants of the disability schemes for the self-employed (WAZ) or young people (WAJONG) since the institutional rules and incentives for work resumption are very different for them. This restriction leads to a sample of 15,864,876 observations for 77,769 individuals which constitute the study sample.

As described above, sick individuals become eligible to participate in one of three disability schemes depending on the date they fall sick. This allows to construct control and treatment groups and compare their responses to the disability reform in a quasi-experimental research design. In particular, we categorize the sick individuals into three groups based on the date they fall sick: the WAO group, the transitional WAO group, and the WIA group, which consist of individuals who fell sick in the third quarter of 2003, fourth quarter of 2003, and first quarter of 2004, respectively. We consider the WAO group as the "control group", and the transitional WAO or the WIA group as the "treatment group".

Based on the available data on labor and benefits, two sets of five outcome variables are

defined and used to compare behavioral responses of control and treatment groups to the disability reform. The first set considers labor participation and benefit receipt, and the second set considers income from labor and benefit programs. As described in Section 2, unemployment benefit is integrated into the disability benefit in the WIA while this is not the case in the WAO. With respect to disability benefit, an outcome variable of benefit receipt indicates receipt of disability benefit and possibly also unemployment benefit at the same time. This means that while the institutional setting requires that the individual who receives disability benefit also receives unemployment benefit at the same time if he is insured under the WIA, in our definition of benefit receipt we allow that the individual who receives disability benefit can also receive unemployment benefit if he is insured under the WAO and receives at least the disability benefit. An outcome variable of income indicates income from disability benefit and possibly also from unemployment benefit. Again, while the disability benefit includes the unemployment benefit if the individual is insured under the WIA, in our definition of income we add the unemployment benefit to the disability benefit if the individual is insured under the WAO and receives the unemployment benefit in addition to the disability benefit. With respect to unemployment benefit, an outcome variable of benefit receipt indicates receipt of unemployment benefit but not the disability benefit at the same time. An outcome variable of income indicates unemployment benefit exclusive of disability benefit. With respect to general assistance and other benefits, receiving benefits from these schemes and income from these schemes are defined as outcome variables. With respect to employment, labor participation and wage income are defined as outcome variables.

4 Descriptive analysis

Time trends in outcome variables

Figures 1a and 1b show the time profiles of labor participation and benefit receipt and income for control and treatment groups over a period of 16 years from January 1999 to July 2015. A time profile of a given group is generated as follows. First, within a group and in a given calendar month, the mean of the outcome variable (dummy variable that indicates labor participation or receipt of a benefit, or income from work or benefits) is calculated. The set of means calculated for each month of the 16 year period are then used to plot the time profile. In the plots vertical lines are added at the first instance individuals could become entitled to the sickness and disability benefits in the WAO, transitional WAO or WIA scheme.

The top left plot in Figure 1a shows the time profile of probability of receiving disability benefit and possibly also unemployment benefit at the same time. A large fraction of about 20 percent of the sick individuals in the WAO and the transitional WAO cohorts claim disability benefit, and possibly also unemployment benefit, at the onset of the disability period. This fraction shows no notable change until the end of the study period. The time trends of the WIA and WAO groups are similar, except in two respects. First, the inflow into the disability scheme for the WIA group is not as immediate as it is for the WAO groups when individuals become eligible to apply for disability benefit. It might be that stricter rules of the WIA make it difficult to claim the disability benefit at the first attempt. Second, the time trends for the WAO groups show a decrease during the first year after they show a peak at the onset of the disability period. It might be that the WAO groups have better chances of recovery shortly after entering the disability scheme. The WAO groups are relatively healthier than the WIA group since in the WIA the minimum disability grade to enter the disability scheme is higher but also disability assessment is stricter.

The probability of working shows a strong time trend that is common to both the control

and treatment groups. It increases until the date individuals fall sick. This pattern does not reflect behavioral responses but it reflects the fact that individuals can enter the sickness scheme, and get reported as sick in the administrative data, only if they are working or receiving the unemployment benefit at the time they fall sick (Section 3). Before this time, these individuals can have another status outside the labor force. The probability of working decreases dramatically during the first year of sickness, remains fairly stable for about three years, but decreases further throughout the remaining months of the study period.

Unemployment benefit use is strongly related to the use of sickness and disability benefits. For the WAO groups, during the sickness period, unemployment benefit use decreases sharply because unemployed people who fall sick change their unemployment benefit for sickness benefit. Unemployment benefit use rebounds thereafter because many of these people recover during the sickness period. It peaks as individuals become eligible to apply for disability benefit because when the sickness period ends, those who apply but get rejected to enter the disability scheme turn to the unemployment benefit. Unemployment benefit use decreases during the disability period. This is because unemployment benefit is temporary and ends after a maximum of 38 months. For the WIA group a similar time trend is observed except that unemployment benefit use increases further during the second year of sickness before it peaks in the beginning of the disability period since members of this group have more time for recovery during the longer sickness period, and change their sickness benefit for unemployment benefit.

The probability of receiving general assistance increases steadily from the time individuals fall sick. A similar but less pronounced increase is observed for the probability of receiving other benefits from various small benefit programs from the time individuals fall sick. These time patterns appear to be related to that of working. As individuals work less due to sickness, their earnings decrease enough to become entitled to these benefits.

In the right panels of Figures 1a and 1b, we show the time profiles of the means of income from work and benefit programs. In a mean calculation, we allow for zero income to analyze how the reform affected the income received due to changes in the amount of wages earned and benefits received at the intensive margin but also due to participation in the labor market and use of benefit programs at the extensive margin. The time trends of the means of different types of income resemble those of participation in the labor market and use of benefit programs shown in the left panels of the figures.

Differences between the control and treatment groups

Our main aim is to compare the behavioral responses of control and treatment groups after these groups become eligible to apply for disability benefits. Here we conduct exploratory graphical analysis of the differences between the control and treatment groups in the outcome variables during the pre-reform and reform periods to provide a first evidence on the impact of the reform. Based on Figures 1a and 1b, comparing the time profiles of a given outcome variable across the control and treatment groups is, however, difficult because time profiles of outcome variables often exhibit strong trends and across different groups they are close to each other at the same time, as in the plot for wage income in Figure 1a.

To make an explicit comparison of the behavioral responses across control and treatment groups, in Figures 2a and 2b we plot the time profiles of the outcome variables where we net out the effect of calendar time. This is done as follows. We obtain residuals from a regression where labor participation or benefit receipt, or income from labor or a benefit, is the outcome, and a full set of calendar month dummies for the entire study period and time-invariant individual fixed effects are controls. Within each group (control or treatment) and at given calendar months, we compute the average of the residuals. We use the computed set of mean residuals across all

calendar months of the study period to plot the time profile of the residualized outcome variable for each group. If the outcome variable is, for example, the residualized participation in labor or benefit receipt, a given point on the time profile of a given group can be interpreted as the propensity of that group to participate where calendar time plays no role. In the plots we do not present any pattern during the sickness period as our primary interest lies in comparing behavioral responses across groups before individuals fall sick and after they become eligible to apply for disability benefits.

The main observation is that, in all plots, the differences between the WIA and WAO groups are statistically insignificant before individuals fall sick, while the differences are large and significant when individuals are eligible for disability benefits. The differences between the transitional WAO and WAO groups are often insignificant before individuals fall sick or after they become eligible to apply for disability benefits. There are a number of other observations. In the top left plot of Figure 2a, for a period of about two years since individuals become eligible to apply for the disability benefit, the probability of receiving disability benefits is smaller for the transitional WAO group than for the WAO group. This is expected because the stricter eligibility criteria of the transitional WAO scheme makes it more difficult to reach the minimum disability grade to become eligible for disability benefits (Section 2). However, after this period, the gap between the two groups narrows and becomes statistically insignificant. This can be explained by the re-examinations of the WAO participants younger than 45 years of age that have taken place from 2004 until 2008 based on the eligibility criteria of the transitional WAO scheme (Mandicó et al., 2018). Another finding is that the WIA group is much less likely to apply for disability benefit compared to the WAO groups especially during the first years after all groups become eligible to apply for disability benefits. The gap between the WIA and the WAO groups remains about 5 percentage points until the end of the study period. Notably, the gap between the WIA and the transitional WAO group remains about the same suggesting that the impact of the eligibility criteria brought by the WIA compared to the criteria brought by the transitional WAO has been much stronger and lasted much longer.

The propensity to work in the WIA group is higher than, and statistically different from, those in the WAO groups during the months individuals are eligible for disability benefits. This suggests that the stricter WIA regime induces individuals to resume working. The size of the gap between the WIA and WAO groups is not particularly large, however. It might be that labor supply responses to the stricter WIA regime often occur at the intensive margin, and to a lesser extent at the extensive margin. Another possibility is that the WIA group seeks income from other benefit programs to replace lost income. This expectation is partly confirmed. The plots for other benefits show that the WIA group is more likely to claim unemployment benefit and less likely to fall back on general assistance and other benefits after the sickness period. On the other hand, the WIA group becomes as likely as the WAO groups to claim these benefits through the end of the study period. It can be that the duration of the unemployment benefit expires, and the higher wage earned by the WIA group makes this group less likely to fall below the income threshold to become eligible for the means-tested or subsistence level benefits. Another contributing factor can be that the criteria to become eligible for social minimum supplement (Section 2) have been relaxed for the WIA beneficiaries from 1 January 2006.

The time profiles of residualized income outcomes, presented in the right panels of Figures 2a and 2b, show that the WIA group receives smaller amounts of disability benefit and general assistance, and earns more wage during the disability period. The differences for other income outcomes are much smaller or insignificant. The differences between the groups are often small and insignificant during the period before individuals fall sick in all outcomes of income.

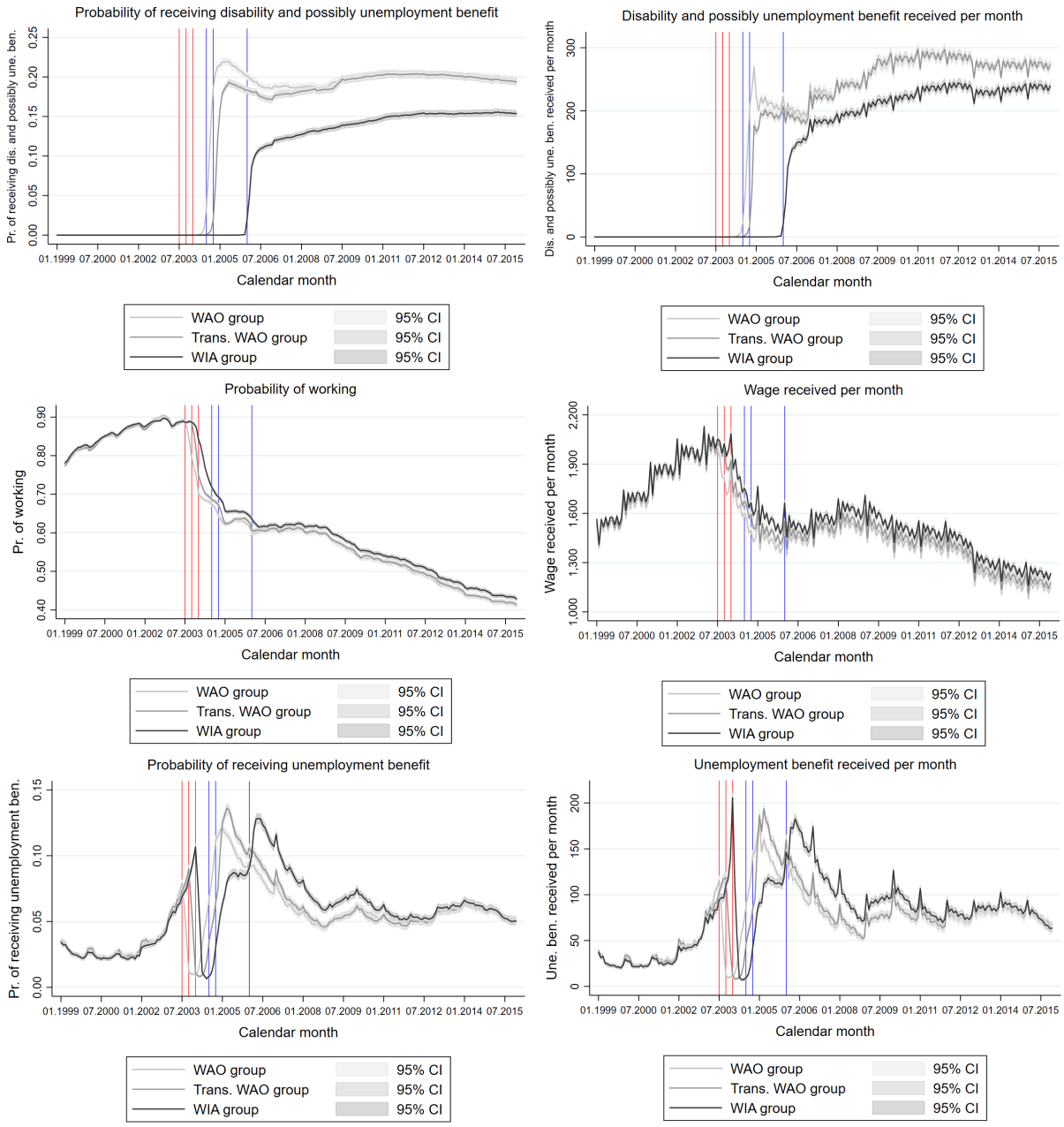

Figure 1a: Labor participation, benefit receipt (disability and unemployment), labor income and benefit income for control and treatment over calendar months. In a given plot, a point on a given time profile represents the mean of the outcome variable (dummy variable that indicates labor participation or benefit receipt, or income from work or benefits) within a group (control or treatment) in a given calendar month. Around the mean is a 95 percent confidence interval. Disability benefit might be supplemented with unemployment benefit. Unemployment benefit is defined so that receiving disability benefit at the same time is not allowed. Each plot is based on the study sample of 15,864,876 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

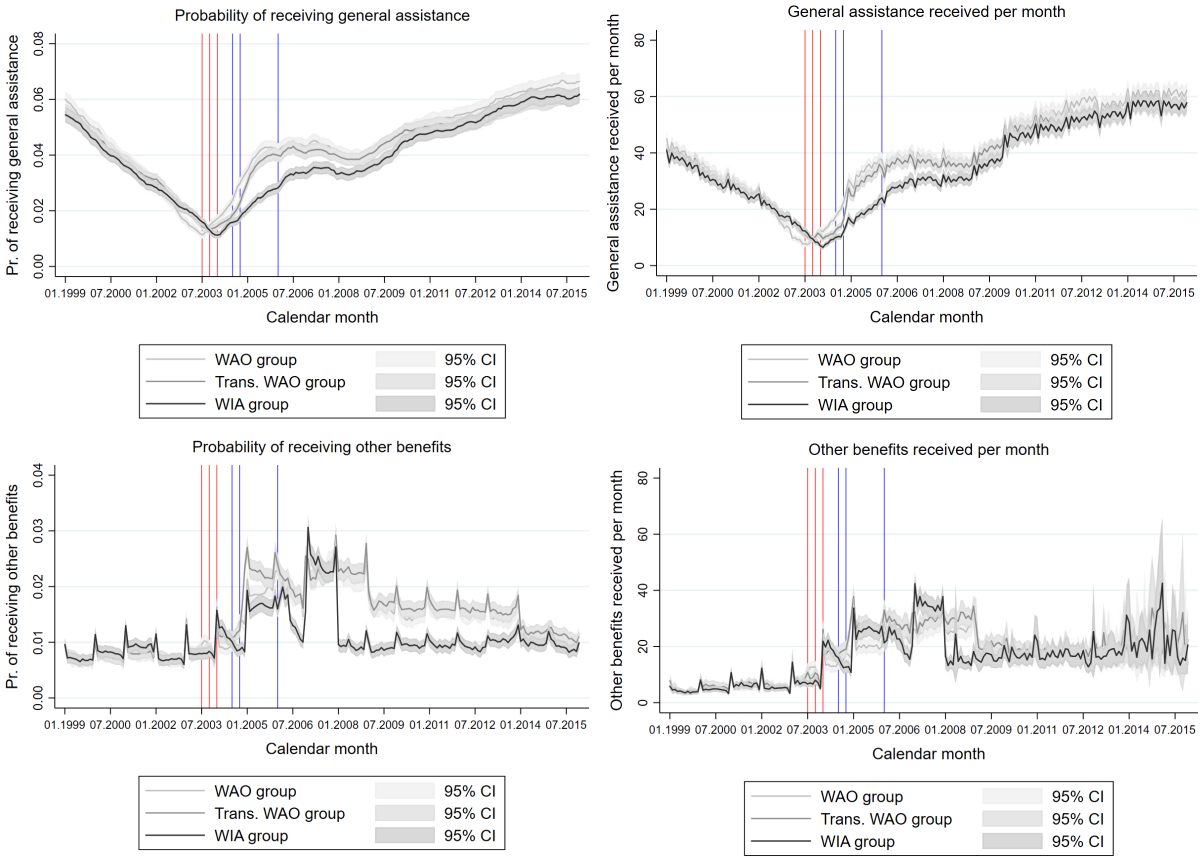

Figure 1b: Benefit (general assistance and other benefits) receipt and income for control and treatment over calendar months. In a given plot, a point on a given time profile represents the mean of the outcome variable (dummy variable that indicates receipt of a benefit, or income from benefits) within a group (control or treatment) in a given calendar month. Around the mean is a 95 percent confidence interval. Each plot is based on the study sample of 15,864,876 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

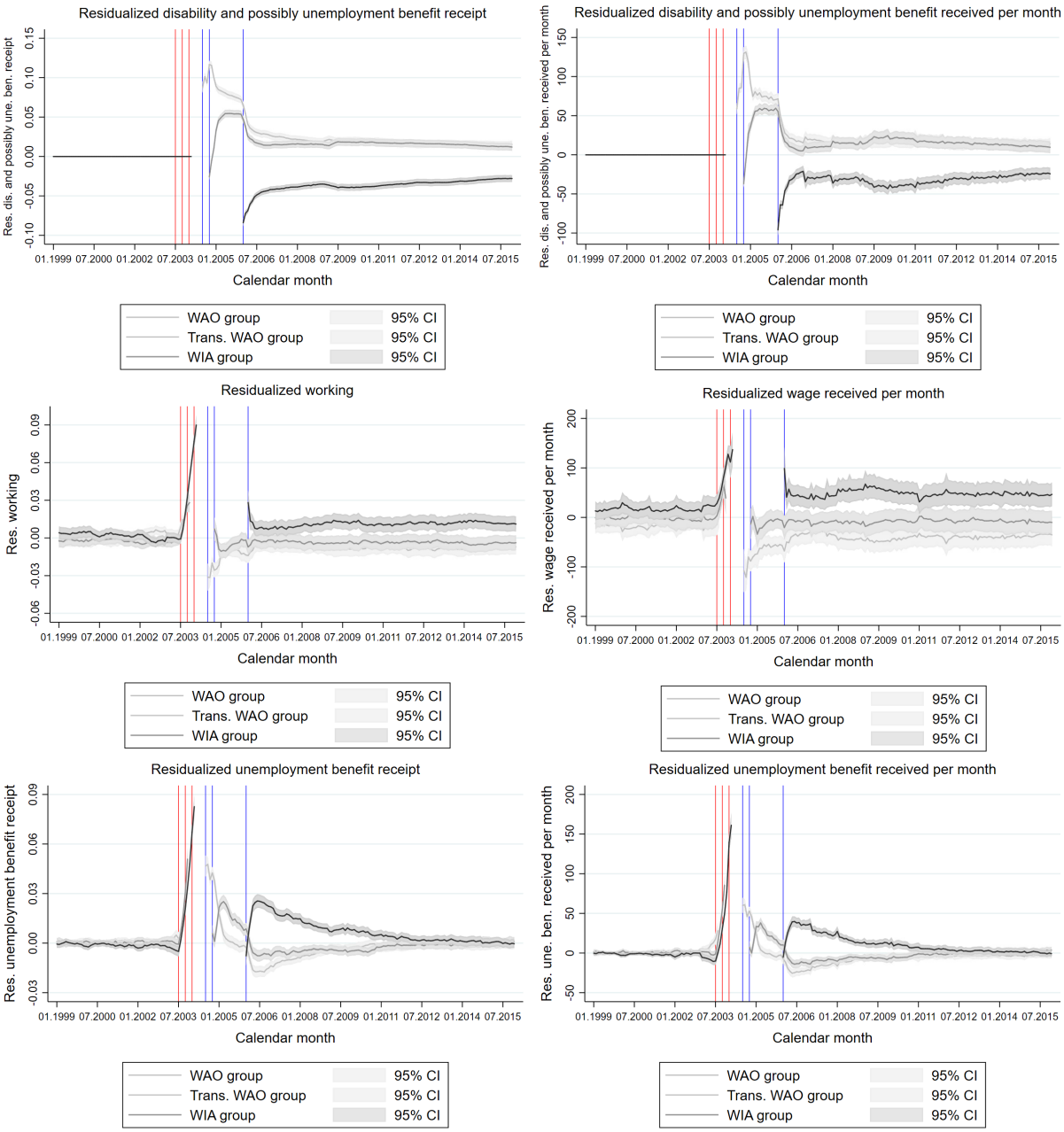

Figure 2a: Residualized labor participation, benefit receipt (disability and unemployment), labor income and benefit income for control and treatment over calendar months. In a given plot, a point on a given time profile represents the mean of residuals within a group (control or treatment) in a given calendar month. Around the mean is a 95 percent confidence interval. Residuals are from a regression where labor participation, benefit receipt, labor income, or benefit income is the outcome, and a full set of calendar month dummies for the entire study period and time-invariant individual fixed effects are controls (Section 4). Disability benefit might be supplemented with unemployment benefit. Unemployment benefit is defined so that receiving disability benefit at the same time is not allowed. Time profiles are not considered during the sickness period. Each plot is based on 4,509,329 and 10,117,027 observations during the periods that precede and succeed the sickness period, respectively, for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

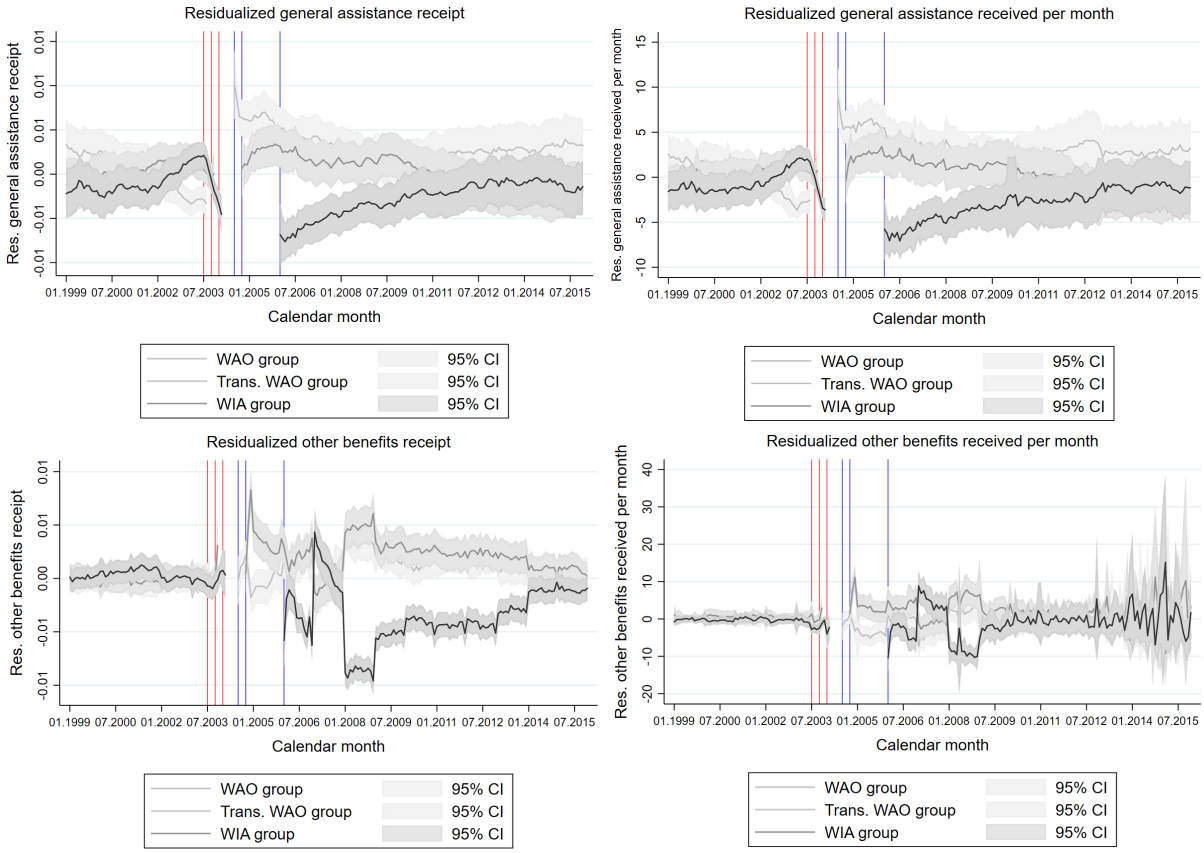

Figure 2b: Residualized benefit (general assistance and other benefits) receipt and income for control and treatment over calendar months. In a given plot, a point on a given time profile represents the mean of residuals within a group (control or treatment) in a given calendar month. Around the mean is a 95 percent confidence interval. Residuals are from a regression where benefit receipt or income is the outcome, and a full set of calendar month dummies for the entire study period and time-invariant individual fixed effects are controls (Section 4). Time profiles are not considered during the sickness period. Each plot is based on 4,509,329 and 10,117,027 observations during the periods that precede and succeed the sickness period, respectively, for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

5 Identification strategy

We use a difference-in-differences (DID) approach to identify the causal effects of the transitional WAO and the WIA reforms on labor participation and benefit receipt and on income from labor and benefit programs. The first difference is across groups. Those who fall sick in the last quarter of 2003 or the first quarter of 2004 are subject to different eligibility criteria to access the disability scheme, and they face different incentives to work or claim benefits, compared to individuals who fall sick in the third quarter of 2003 and are subject to a less restrictive benefit regime. The second difference is over (event) time. After individuals become eligible to apply for the disability benefit, the incentives to work or claim benefits change compared to the period before individuals fall sick.

We implement the DID comparison using the following regression:

$$y_{it} = \gamma_1 \left(Treat_i^{Trans.\ WAO} \times Post_t\right) + \gamma_2 \left(Treat_i^{WIA} \times Post_t\right) + \delta\, Post_t + \lambda_s + \alpha_i + \varepsilon_{it}. \quad (1)$$

i indexes individuals. t indexes the months of event time. Event time indicates the period of time before individuals fall sick, and subsequently the period of time when individuals are eligible for disability benefits. t indexes the months of the former period with values from –55 to 0, and those of the latter period with values from 1 to 120. y_{it} is the outcome variable of interest. λ_s is a monthly calendar time effect. s indexes the calendar months of the period before individuals fall sick, and subsequently the calendar months of the period when individuals are eligible for disability benefits. s is individual specific since individuals fall sick in any month from July 2003 to March 2004 in the available sickness data. For example, for the individual who falls sick in July 2003 and participates in the WAO scheme, where the sickness period lasts one year, s indexes the months from January 1999 to February 2004, and subsequently the months from July 2004 to July 2015. January 1999 is chosen as the base month for comparison. α_i is an individual-specific, time-invariant intercept term. ε_{it} represents the individual-specific, time-varying shocks that are not observed.

$Treat_i^{Trans.\,WAO}$ and $Treat_i^{WIA}$ are dummy variables that indicate the treatment groups, in particular the transitional WAO and the WIA groups, respectively.[3] The periods of time before individuals fall sick and after they become eligible to apply for disability benefit are the pre-treatment and post-treatment periods, respectively. $Post_t$ is a dummy variable that indicates the post-treatment period. We interact $Treat_i^{Trans.\,WAO}$ and $Treat_i^{WIA}$ with $Post_t$ to capture the mean difference in the outcome variable between the treatment and control groups during the post-treatment period compared to the mean difference between the two groups during the pre-treatment period. In this comparison, the latter difference aims to account for differences between the groups due to factors other than the policy reform. γ_1 and γ_2 are the coefficients of main interest and reflect the effects of the transitional WAO and WIA reforms. The assumptions needed to obtain unbiased estimates of these coefficients are discussed below. Standard errors are adjusted for clustering at the individual level.

It is important to note that, for the WIA reform in particular, we interpret that the differences in behavior between the treatment and control groups during the disability period are not only due to the stricter eligibility criteria for disability benefits (disability reform) but also due to the extension of the sickness period from one to two years (sickness reform). For example, the extra year spent in the sickness scheme provides the individuals in the treatment group with more time for recovery, and might reduce their propensity to apply for disability benefits. This means that we do not aim to separately identify the effects of the different incentives of the sickness and disability reforms on benefit and labor behavior during the disability period. The idiosyncratic impact of the disability reform could be identified, for example, by selecting

[3] These group dummies have no time variation and are omitted in the fixed effects regression.

subgroups of individuals from the treatment and control groups with similar health characteristics at the end of the sickness period, and comparing their subsequent behavior during the disability period. This is not done in this study. In fact, in the current DID comparison of interest, there is no reason to expect the treatment and control groups to share similar health or work characteristics at the end of the sickness period. Instead, we analyze the total effect of the WIA reform, comprising of the sickness and disability reforms, on behavior during the disability period, relative to the period before individuals fall sick.

To study the effects of the transitional WAO and WIA reforms in the long run, we consider the following regression:

$$y_{it} = \sum_{l=1}^{10} \gamma_{1l} \ \left(Treat_i^{Trans.\,WAO} \times d_{lt}\right) + \sum_{l=1}^{10} \gamma_{2l} \ \left(Treat_i^{WIA} \times d_{lt}\right) + \sum_{l=1}^{10} \delta_l d_{lt} + \lambda_s + \alpha_i + \varepsilon_{it}. \tag{2}$$

Compared to Equation (1), instead of the $Post_t$ dummy, which indicates the entire post-treatment period, we consider year dummies. That is, we divide the post-treatment period into ten equal periods of one year each. d_{lt} is a dummy variable that indicates a one-year period of the post-treatment period. For example, d_{2t} indicates the second year from the time the individual becomes eligible for disability benefit. The pre-reform period is chosen as the base period for comparison. The interaction terms of treatment and year dummies capture the mean difference in the outcome variable between the treatment and control groups in a given year compared to the mean difference between the two groups in the pre-reform period.

In the described DID setup, treatment and control groups are compared over "event time". Therefore, a distinction is made between event time and "calendar time". That is, according to our institutional setting, members of the treatment and control groups fall sick at different points in time, but also the duration of the sickness scheme they participate in is different, which both lead to different calendar dates to become eligible to apply for disability benefit. However, our aim is to compare the behavior of treatment and control groups across the periods these groups are eligible for disability benefits and before they fall sick. This means that our DID comparison requires comparing members of the treatment and control groups at different points in time with respect to the calendar months these groups fall sick but also the calendar months they become eligible to apply for disability benefits. Therefore, "event time" is defined to refer to the time periods before individuals fall sick and after they become eligible to apply for disability benefits. "Calendar time" is defined to define calendar time effects that aim at capturing calendar time trends in outcome variables. In Equation (2), the year dummies of event time are correlated with the month dummies of calendar time by construction. The concern is that this correlation structure could affect the point estimates or the standard errors of the main effects of interest given by the coefficients of the interaction terms. In the section on sensitivity analysis we show that this concern should be small.

Are the treatment effects likely to be confounded by unobserved shocks?

In Equation (1), to obtain unbiased estimates of the effects of the transitional WAO and WAO reforms, it must hold that the treatment and post interactions are not correlated with unobservable shocks that might be related to the policy changes. In the regression, α_i controls for time-invariant differences across individuals that might confound the estimated effects of the policy reforms. Individual-specific, time-varying shocks captured by the error of the regression would not necessarily confound the estimated effects of the reforms. According to the institutional setting, falling sick in one of three adjacent quarters of a nine-month period determines eligibility for one of three disability schemes. As eligibility for different schemes is

determined within a short period of time, concerns about the endogeneity of the treatment and post interactions should be small. For example, an exogenous economic shock should affect the three cohorts of sick individuals in similar ways so that the differences in outcomes across these cohorts can still be attributed to the differences in treatments under different benefit regimes. During the nine-month period from July 2003 to March 2004, however, no particular changes in macroeconomic conditions are observed, or no other major social security reforms are introduced.

Are sick individuals likely to select themselves to disability schemes?

Figure 3 presents the distribution of individuals by the month they fall sick. The distribution has modest peaks in September and January. As described in Section 2, the instance of falling sick during one of three adjacent quarters of the nine-month period from July 2003 to March 2004 determines eligibility for one of three disability schemes. Therefore, a concern for the identification of the impact of the reform is that individuals may react to select themselves into the old, transitional, or new disability schemes from the moment the transitional WAO and WIA reforms are announced. For example, individuals may select themselves into the lenient WAO scheme before it gets abolished at the end of September. Or, individuals may react in anticipation of the larger WIA reform having seen the transitional WAO reform. However, self-selection into different disability schemes is unlikely. The government presented a general policy program outlining, among other targets, its plan to reform the WAO scheme on 15 September 2003. In particular, it is announced that the sickness period will be extended from one to two years, and a stricter disability insurance law will be introduced for the individuals who fall sick from 1 January 2004. The details of the law were announced on 18 August 2004. This means that, in Figure 3, individuals who fall sick in the peak month of September are unlikely to have selected themselves into the lenient WAO scheme since they have little time to react to the planned reform announced in mid September. The transitional WAO reform, that introduced stricter criteria to become eligible for disability benefit after a sickness period of one year, was announced much later, on 12 March 2004. This means that sick reporting in reaction or anticipation to the transitional reform is impossible. We expect that the third peak month of January is unlikely to reflect self-selection since the WIA reform was announced to bring a much stricter benefit regime. This expectation seems to be confirmed by the observation that individuals do not appear to be selecting themselves into the month of December to avoid participation in the WIA scheme.

Are the control and treatment groups observationally equivalent?

Here we check if the sick individuals in the control and treatment groups are similar in their mean background characteristics and pre-reform mean labor and benefit outcomes. Table 1 presents sample means and balancing tests of background characteristics and outcomes in control and treatment before individuals fall sick and after they become eligible to apply for disability benefit. The top panel of the table presents the sample means of a number of background characteristics in control and treatment groups. A sample mean is calculated as the average of a given characteristic of all individuals in a given group at the time these individuals fall sick. In all groups, the average age is almost 40, the fraction of men is slightly higher than that of women, and the majority of the sample is native-born. About 60 percent hold a regular work contract, less than 20 percent hold a temporary contract or a contract through a temporary work agency, and about 20 percent is unemployed.

In columns 4 and 5 of the table, we present the mean differences between treatment and control groups based on linear regression, and indicate whether the difference is statistically

significant. In particular, a difference represents the estimated coefficient from the regression of the characteristic as the dependent variable, and the treatment indicator (whether member of the transitional WAO or the WIA group) as the explanatory variable. The differences are often small but statistically significant at the 1 percent level. At the bottom of the top panel of the table, the reported F-statistic tests whether the estimated coefficients of all characteristics are jointly zero in a regression where treatment indicator is the dependent and all the background characteristics are regressors. The statistic is significant at the 1 percent level. The significant mean differences as well as the F-statistic suggest that background characteristics affect which disability scheme the sick individual participates in. However, background characteristics are not likely to directly affect participation in a particular scheme, but affect it indirectly through the differences in the dates individuals fall sick. That is, as described above, sick individuals become eligible to participate in one of three disability schemes depending on the date they fall sick during the period from July 2003 to March 2004. The differences in the dates individuals fall sick are likely to drive the differences in observed background characteristics across the control and treatment groups. For example, the positive difference for age reflects the fact that the WIA and the transitional WAO groups fall sick at later dates and therefore become eligible to apply for the disability schemes they participate in. For the work status, the two treatment groups less often work on temporary contracts but are more often unemployed compared to the control group. This might be due to that fewer temporary jobs are available in the fourth quarter of 2003 and the first quarter of 2004, when individuals in the two treatment groups fall sick, possibly due to a seasonal change in demand for labor. With respect to the place of birth, foreign-born people are less likely to end up participating in the transitional WAO or the WIA. If there are fewer temporary jobs available in the fourth quarter of 2003 and first quarter of 2004, and foreign-born people have less favorable labor market opportunities and are therefore more likely to have temporary contracts, they will be less likely to end up participating in the transitional WAO or the WIA. The result on gender seems to confirm that background characteristics affect selection into one of three disability schemes through the differences in the dates individuals fall sick. The difference for gender is statistically insignificant, as we would expect, and since gender is time-invariant, it does not act to select individuals into different disability schemes. In our formal comparison of the group mean differences based on Equation (1), we control for time-invariant background characteristics, through individual fixed effects, to address the endogeneity of treatment status.

The bottom panel of Table 1 presents the sample means of labor participation and benefit receipt and income from labor and benefit programs for control and treatment groups, and the mean differences across these groups, during the pre- and post-treatment periods. During the pre-treatment period, a number of the mean differences are significant especially for the WIA group. These differences are not likely to be behavioral, however. By definition, members of the WIA group fall sick at least three months later than the members of the WAO group. Therefore, for a given outcome variable, the mean of the WIA group uses additional observations that in fact are associated with large or small values leading to a group mean that is higher or lower than that for the WAO group. For example, in Figure 1a, the plot for wage income shows that the WIA group earns a high wage for a number of additional months when the WAO group has already fallen sick, which leads to a higher mean wage for the WIA group. In our formal comparison of the group mean differences based on Equation (1), we control for monthly calendar time effects to capture the calendar time trends specific to the control and treatment groups.

During the post-treatment period, the mean differences are often statistically significant suggesting that the reforms induce behavioral changes. The differences are almost always larger for the WIA group than for the transitional WAO group reflecting the fact that the WIA regime

is stricter than the transitional WAO regime. These differences in group means are in line with the differences in the time profiles of group means across the months of the post-treatment period observed in Figure 2a and Figure 2b.

Are the pre-reform time trends common to control and treatment groups?

The main assumption of our identification strategy is that control and treatment groups share common time trends in pre-reform labor and benefit outcomes. Figures 1a and Figure 1b showed that control and treatment groups share very similar time trends until individuals fall sick providing visual evidence for this main identifying assumption. Here we use regression analysis to test this assumption. In particular, we use pre-reform data from January 1999 to June 2003 to estimate a regression where labor participation, benefit receipt, labor income, or benefit income is the outcome, and calendar month dummies, interactions of treatment and calendar month dummies, and time-invariant individual fixed effects are controls. January 1999 is chosen as the base month for comparison. Coefficient estimates of interaction terms that are not statistically different from zero provide evidence in favor of the common trends assumption.

Figures 4a and 4b plot the coefficient estimates of the treatment and calendar month dummy interactions from regressions where labor participation, benefit receipt, labor income, and benefit income are the outcomes. For both treatment groups, the coefficient estimates are not statistically different from zero throughout the pre-reform period in the plot of other benefits. The coefficient estimates deviate from zero to different extents through the end of the pre-reform period in the plots of other outcomes. Based on the F-tests presented in the notes below the figures, for these outcomes, we also often reject the hypothesis that all the interaction terms are jointly equal to zero. Besides this regression based test on the exact equality of the time trends across the control and treatment groups, the visual evidence based on the time trends in Figures 1a and 1b, however, suggest that control and treatment groups share very similar downward or upward time trends. In the section on sensitivity analysis, we provide additional analysis to show that the statistically significant differences in the time trends of control and treatment groups through the end of the pre-reform period are random, and have no notable impact on the baseline estimates of the reform effects.

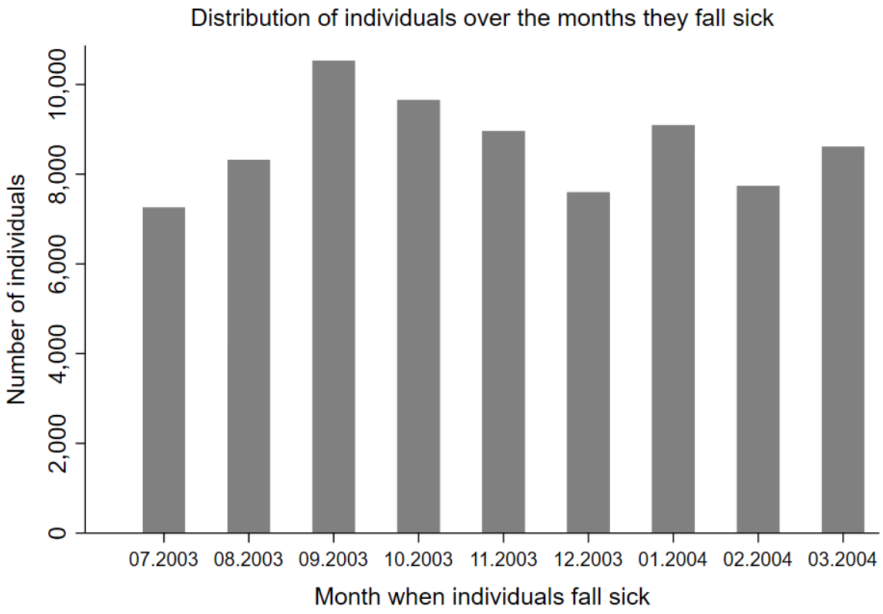

Figure 3: Distribution of individuals over the months they fall sick.

Table 1: Sample means and balancing tests of background characteristics and outcomes in control and treatment before and after the sickness period

	Before					After				
	WAO group (1)	Trans. WAO group (2)	WIA group (3)	Dif. trans. WAO and WAO (4)	Dif. WIA and WAO (5)	WAO group (6)	Trans. WAO group (7)	WIA group (8)	Dif. trans. WAO and WAO (9)	Dif. WIA and WAO (10)
Background characteristics										
Age	39.690	40.137	40.628	0.447^{***}	0.937^{***}					
Female	0.441	0.437	0.440	-0.004	-0.001					
Foreign-born	0.188	0.177	0.170	-0.011^{***}	-0.018^{***}					
Work characteristics										
Regular contract	0.581	0.574	0.593	-0.007	0.011^{***}					
Temporary contract	0.138	0.130	0.119	-0.008^{***}	-0.020^{***}					
Temporary contract via agency	0.048	0.049	0.034	0.001	-0.014^{***}					
Unemployed	0.188	0.204	0.213	0.015^{***}	0.025^{***}					
Other	0.044	0.043	0.041	-0.001	-0.003					
F-statistic				8.210^{***}	32.75^{***}					
Labor participation and benefit receipt										
Dis. and possibly une. ben.						0.198	0.191	0.141	-0.007^{**}	-0.057^{***}
Work	0.858	0.856	0.860	-0.002	0.003	0.534	0.536	0.539	0.002	0.004
Une. ben. (non-par. in dis. ben.)	0.033	0.036	0.037	0.002^{**}	0.003^{***}	0.064	0.065	0.069	0.002^{*}	0.005^{***}
General assistance	0.036	0.035	0.032	-0.001	-0.003^{***}	0.050	0.048	0.047	-0.002^{*}	-0.004^{**}
Other benefits	0.008	0.008	0.008	0.000	0.001	0.017	0.018	0.012	0.001^{*}	-0.005^{***}
Income from labor and benefit programs										
Dis. and possibly une. ben.						253.895	250.790	212.504	-3.105	-41.391^{***}
Wage	1768.763	1792.750	1820.622	23.987^{**}	51.859^{***}	1376.214	1407.520	1453.732	31.306^{**}	77.518^{***}
Une. ben. (exclusive of dis. ben.)	40.620	43.528	45.717	2.908^{**}	5.097^{***}	86.192	90.183	96.359	3.991^{***}	10.167^{***}
General assistance	27.657	26.730	25.061	-0.927	-2.596^{**}	47.287	44.936	44.009	-2.351	-3.277^{**}
Other benefits	5.650	6.486	5.922	0.836	0.273	20.346	23.393	20.335	3.048^{***}	-0.010
Individuals	26,111	26,217	25,441			26,111	26,217	25,441		
Observations	1,439,381	1,518,527	1,551,421			3,573,931	3,515,137	3,027,959		

Notes: 1. "Before" denotes the period that precedes sickness and starts on January 1999 and ends on June 2003 for the WAO group. It ends on September 2003 for the transitional WAO group, and on December 2003 for the WIA group. "After" denotes the period that succeeds sickness and starts on July 2004 and ends on December 2015 for the WAO group. It starts on October 2004 for the transitional WAO group, and on January 2006 for the WIA group. The transitional WAO and WIA groups are the treatment groups and consist of individuals who fell sick in the fourth quarter of 2003 and the first quarter of 2004, respectively, and are subjected to the reform. 2. Columns 1, 2, 3, 6, 7 and 8 present means of characteristics and outcomes of individuals in control and treatment before and after the sickness period. Columns 4, 5, 9 and 10 present the differences between treatment and control. A difference represents the estimated coefficient from the regression of the characteristic or outcome as the dependent variable, and the treatment indicator (whether member of the transitional WAO or the WIA group) as the explanatory variable. Standard error of the difference is clustered at the individual level. The reported F-statistic tests whether the estimated coefficients of all characteristics are jointly zero in a regression where treatment indicator is the dependent variable and all the individual's characteristics are included as explanatory variables. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively. 3. For background and work characteristics, a given mean represents the average of the individuals' given characteristic in the month the individual falls sick. Therefore, the number of observations used to calculate the mean is the same as the presented number of individuals. For the outcome variables, the number of observations used in the before period is smaller than that in the after period since the period that precedes sickness is shorter than the period that succeeds it.

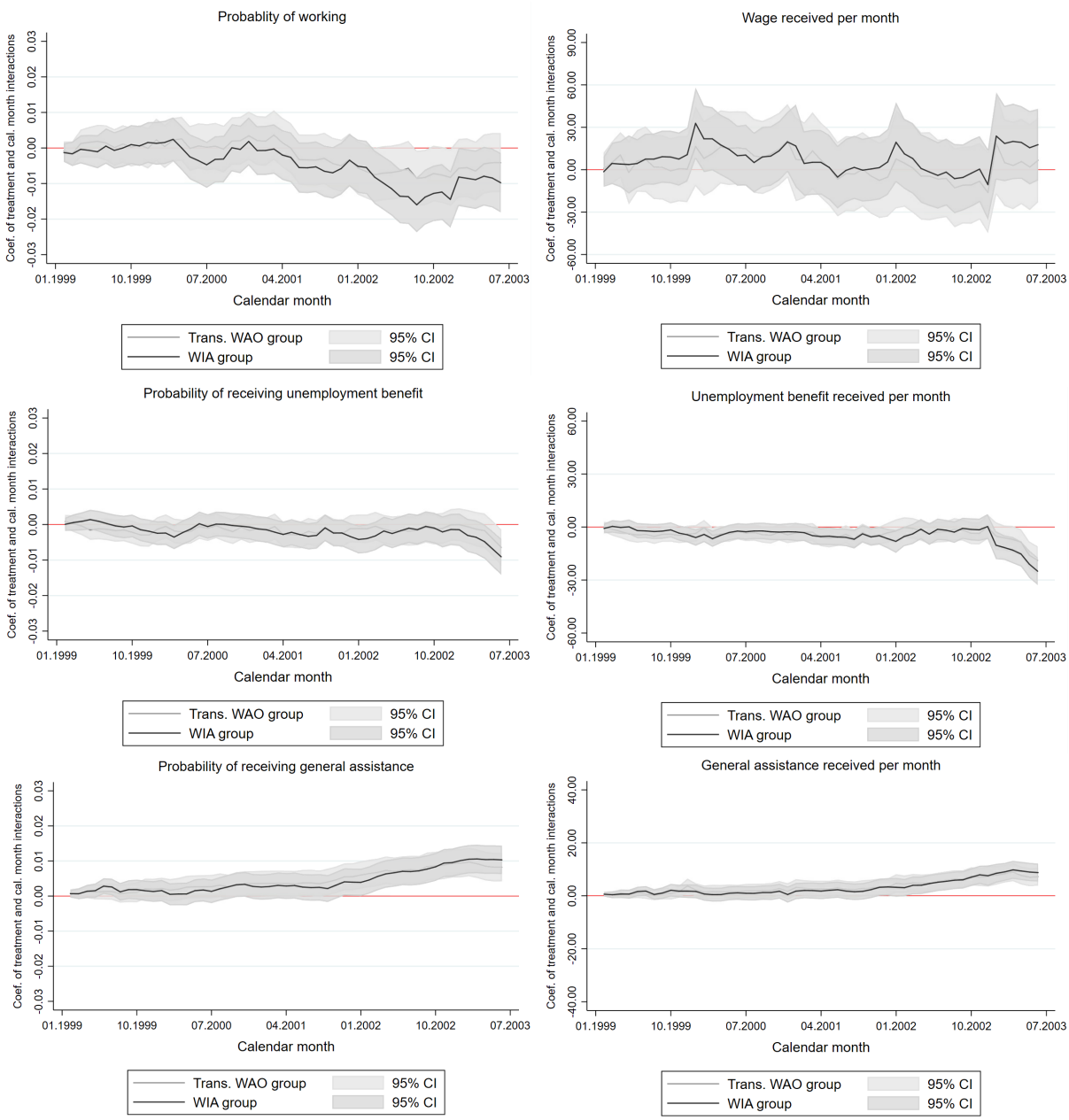

Figure 4a: Coefficient estimates of treatment and calendar month interactions for control and treatment in the pre-reform period from regressions for labor participation, benefit receipt (unemployment benefit and general assistance), labor income, and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on pre-reform data from January 1999 to June 2003. January 1999 is the base month for comparison (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and calendar month dummies, treatment and calendar month dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 4,199,526 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level. Test results based on an F-statistic of whether the coefficients of all interactions are jointly zero are as follows (p-values in parentheses). For labor participation, unemployment benefit receipt, and general assistance receipt outcomes, the results are respectively 1.100 (0.287), 1.190 (0.164), 0.940 (0.595) for the transitional WAO group, and 1.560 (0.006), 1.320 (0.061), 1.600 (0.004) for the WIA group. For labor, unemployment benefit, and general assistance income outcomes, the results are respectively 1.390 (0.030), 1.570 (0.005), 1.550 (0.006) for the transitional WAO group, and 1.330 (0.056), 1.990 (0.000), 2.070 (0.000) for the WIA group.

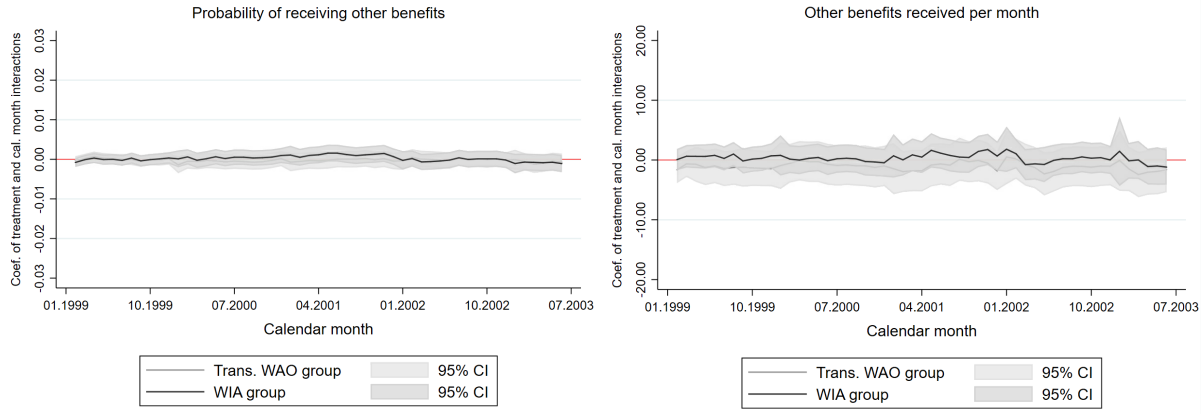

Figure 4b: Coefficient estimates of treatment and calendar month interactions for control and treatment in the pre-reform period from regressions for benefit receipt (other benefits) and income outcomes. Around each estimate is a 95 percent confidence interval. Regressions are based on pre-reform data from January 1999 to June 2003 (Section 5). January 1999 is the base month for comparison. Benefit receipt or income is the outcome, and calendar month dummies, treatment and calendar month dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 4,199,526 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level. Test results based on an F-statistic of whether the coefficients of all interactions are jointly zero are as follows (p-values in parentheses). For the other benefits receipt outcome, the result is 0.86 (0.747) for the transitional WAO group, and 1.220 (0.128) for the WIA group. For the other benefits income outcome, the result is 0.770 (0.885) for the transitional WAO group, and 0.96 (0.556) for the WIA group.

6 The effects of the disability reforms on labor and benefit outcomes

6.1 Baseline effects of the reforms

Here we present the baseline DID estimates of the effects of the transitional WAO and WIA reforms based on Equation (1). Table 2a presents the results for labor participation and benefit receipt, and Table 2b presents the results for income from labor and benefit programs. The estimated effects of the transitional WAO and the WIA reforms are always interpreted as the effects of the new rules of the transitional WAO and WIA regimes compared to the old rules of the WAO regime.

Table 2a shows that the transitional WAO reform reduced the probability of receiving disability benefit by 0.8 percentage points, and the WIA reform reduced it by 5.8 percentage points, on average, during the ten years after the reform has come into effect and where data is available on disability benefit. The larger effect of the WIA reform is expected because it brought a much stricter benefit regime than the transitional WAO reform did. The reduction of 5.8 percentage points corresponds to a reduction of 28.6 percent (0.058/0.203) in disability benefit rewards among the individuals who fell sick during the period from July 2003 to March 2004. This large effect of the WIA reform is in line with Van Sonsbeek and Gradus (2013) who showed that due to the WIA reform the number of disability benefit rewards in the working population decreased by 40 percent at the onset of the reform, although the impact of the reform slightly decreased over time.

As the reforms reduced disability benefit receipt, they increased labor participation. The probability of working is 0.7 percentage points higher among the individuals insured under the transitional WAO scheme compared to the individuals insured under the WAO scheme. This shows that the decrease in disability benefit receipt is matched by an almost same amount of increase in labor participation. The probability of working is 1.8 percentage points higher for the WIA group compared to the WAO group. This shows that the substantial decrease in disability benefit receipt is not matched by a comparable amount of increase in labor participation among the individuals in this group. It might be that disability benefit recipients are often labor participants at the same time. It might also be that labor supply responses to the WIA reform occur at the intensive margin instead of the extensive margin. These results show that, relative to the WAO regime, the transitional WAO regime is more effective than the WIA regime in terms of how individuals substitute disability benefit receipt with labor participation.

Individuals who could not access the disability benefit might have turned to benefits from other benefit programs. We find evidence that the reforms induced sick individuals to turn to the unemployment benefit. Compared to the comparison group, both treatment groups increase their unemployment benefit receipt, and the effect sizes are comparable to those of labor participation. Both treatment groups, however, become less likely to claim general assistance. It might be that the income earned above the subsistence minimum due to work resumption or benefit substitution limits the access to general assistance. Another possibility is that there is less need for claiming general assistance on top of the disability benefit because the disability benefit itself can more easily be topped up to the social minimum income (Section 2). Similar reasons can explain the decrease in receiving other benefits for the WIA group.

Earlier studies in other countries also find evidence that tightening the eligibility criteria leads to more take-up of other benefits (Karlström et al., 2008; Staubli, 2011). Studies on earlier disability insurance reforms in the Netherlands, however, show mixed results. Borghans et al. (2014) find that the disability reform in 1993 led to more benefit claims from other benefit programs, while Koning and van Vuuren (2010) and De Jong et al. (2011) find that the

experience rating reform in 1998 and the "Gatekeeper protocol" reform in 2002 had no spillover effects on unemployment benefit.[4]

Table 2b shows large effects for (monthly) disability and unemployment benefits received and wages earned. Compared to the WAO regime, under the WIA regime individuals received 50.8 euros less disability benefits, while they earned 55.7 euros more wages, received 22.8 euros more unemployment benefits, on average. They also received 4.9 euros less from general assistance. All the effects are statistically significant at the 1 percent level. This shows that, due to the WIA reform, the decrease in disability benefits received is compensated by higher wages and higher unemployment benefits so that, on balance, individuals earn a higher income (22.8 euros), on average. This result is remarkable because the increase in labor participation rate due to the reform is rather modest. All the effects are statistically significant at the 1 percent level. The effects of the transitional WAO reform are small and statistically insignificant except for the unemployment benefit. This shows that while for the WIA reform the effects are observed both at the extensive (labor participation and benefit receipt) and intensive (labor and benefit income) margins, for the transitional WAO reform the effects are limited to the extensive margin.

6.2 Long-term effects of the reforms

We analyze the long-term effects of the transitional WAO and WIA reforms in Figures 5a and 5b. In the figures, in a given plot we present the coefficient estimates of treatment and year dummy interactions over the post-treatment period of ten years based on the regression given by Equation (2). The WIA reform shows structural effects on disability benefit receipt and labor participation. During the whole post-treatment period, the effects of the reform on these outcomes are sizable and statistically significant. The effects of the reform on unemployment benefit, general assistance, and other benefits receipt are large at the onset of the reform but become smaller and eventually statistically indistinguishable from zero about seven years after it is introduced. This a consequence of the unemployment benefit being temporary. The impact of the transitional WAO reform on labor participation and benefit use is short-lived. For example, as discussed in Section Section 4, the reform has a significant effect on disability benefit receipt only during the first two years of the reform. After the first two years, all participants in the old WAO scheme who were younger than 45 years of age were re-examined based on the eligibility criteria of the transitional WAO scheme. Therefore, the differences between old and transitional WAO scheme diminished for the younger participants.

6.3 Heterogeneous effects of the reforms

Here we analyze whether the disability reforms affected sick individuals with different background and labor market characteristics differently. We consider differential effects for gender groups, three age groups, and across three types of work status. The age groups distinguish between individuals who are younger than 30, between 30 and 49, and older than 49. The work status groups distinguish between individuals who hold a regular contract, temporary contract or a contract through a temporary work agency, and who are unemployed. Age and work status are those at the time individuals fall sick. Tables 3a through 5b present the results for labor participation and benefit receipt and for labor and benefit income based on the estimation of the regression given by Equation (1) on these subgroups.

Tables 3a and 3b show that, due to the WIA reform, men are less likely to receive disability benefit and they receive a smaller disability benefit compared to women. Based on the Wald

[4]The Gatekeeper protocol reform mandates firms to report sickness on time and clarifies firms' responsibilities of re-integrating the sick employees back to work.

test, the difference between the coefficient estimates of the two gender groups is statistically significant at the 10 percent level for both disability benefit receipt and income. This result is expected. Men earn, on average, higher wages than women. In the Dutch disability scheme, a higher pre-disability wage means a higher probability of getting a partial disability benefit. This is because the pre-disability wage is compared to a (lower) fictitious new wage that can be earned given one's disability. A higher pre-disability wage implies more alternative work opportunities with lower fictitious new wages (Section 2).

Tables 4a and 4b show notable age effects. First, due to the WIA reform, compared to the younger and older age groups, the 50-54 age group relies on disability benefits the least in terms of benefit receipt and income. This seems to explain why the same group increases its labor participation the most and earns the highest wages. To a certain extent these findings might reflect a composition effect. Individuals of the 50-54 age group more often hold a regular work contract (68 percent versus 49 and 61 percent for the youngest and oldest age groups, respectively) which might make it more difficult to access the disability scheme. For example, it might be easier to find potential jobs these individuals can still perform leading to a smaller disability grade for them (Section 2). Another possibility is that with the WIA reform employers have become reluctant to allow access to disability insurance on the basis of the age of their employees. The cost of a sick worker for the employer has increased since the introduction of experience rating in 1998 which punishes firms with many employees on disability benefits by charging them with a higher premium. Since older individuals are more likely to fall sick, cost of a sick worker is higher if the share of older workers is higher in the firm. Since the WIA reform extended the experience rating period from 5 to 10 years for partially disabled individuals (Section 2), employers might have become selective in allowing their older workers to access disability benefits. However, this explanation is not supported by the smaller effects found for the oldest age groups.

Second, compared to the younger age groups, the oldest age group is substantially more likely to rely on disability and unemployment benefits in terms of both benefit receipt and income. For example, the amount of unemployment benefit this group receives is almost seven times larger than the amount the youngest age group receives. This group is also much less likely to increase labor participation and income. In 2015 the government has put into effect the law raising the legal retirement age. Our results show that sick individuals close to the legal retirement age struggle to respond to the work incentives of the reforms the most compared to younger sick individuals. This implies that a higher legal retirement age is likely to limit the remaining work capacity further and increase benefit claims or amounts among the future population of disabled workers close the legal retirement age. Labor participation of this group can therefore be expected to become important in the coming decades, and even more so as the size of this group is increasing since the law raising the legal retirement age has come into effect in 2015: the number of benefit claims in the WIA increased remarkably from 35,764 in 2015 to 43,419 in 2018 (Berendsen et al., 2019).

The effect of the WIA reform found for individuals younger than 45 years is similar to the effect of the disability reform in 1993 among the individuals of the same age. Borghans et al. (2014) analyzed the effect of re-examining existing disability benefit recipients below age 45 on labor participation and benefit substitution based on the stricter criteria introduced in 1993. They find that the 1993 reform increased the probability of working by 2.9 percentage points. During the years before the 1993 reform was introduced, exceptionally large numbers of people were receiving benefits due to the very generous disability scheme. There was therefore much potential for strong labor supply responses to the disability reform. We could expect that the potential for labor supply responses have become smaller in the aftermath of the 1993 reform due to the stricter eligibility criteria brought by this reform. Therefore, the labor supply responses

to the WIA reform could be smaller than the labor supply responses to the 1993 reform. The similar effect sizes of the 1993 and WIA reforms, therefore, suggests that the effect of the WIA reform on labor participation is not small.

The effects of the reforms show substantial heterogeneity with respect to work status. First, the WIA reform decreases use of disability benefit among the unemployed individuals more than it does among those who hold a temporary or regular contract. It might be that the unemployed are more often sick, without necessarily having severe health problems, and therefore are more likely to be affected by a stricter disability benefit regime. Furthermore, the reform decreases labor participation among the unemployed but increases it among those with regular or temporary work contracts. The results based on the income outcomes show that unemployed individuals receive much larger amounts of unemployment benefit due to loss of income from disability benefit and wages. It might be that the work resumption incentives brought by the WIA reform lead employers to reintegrate their employees back to their job while they prove ineffective if there is no employer (Koning and Lindeboom, 2015). On the employees' side, the incentives to resume working may be smaller for the unemployed individuals insured under the WIA, compared to those insured under the WAO, since they spend an additional year in the sickness scheme before they become eligible to apply for disability benefits. A longer unemployment spell may lead to more human capital loss or have a stronger "scarring effect" for the sick unemployed individuals, and decrease their prospects of finding a job (Arulampalam, 2001; Arulampalam et al., 2001). These results suggest that the reforms risk making the individuals who are unemployed particularly vulnerable to become disabled since they appear to have limited access to the disability scheme and lack the incentives other groups have to resume working.

A second notable result follows from comparing the results for workers with regular and temporary contracts. Table 5a shows that, due to the WIA reform, individuals with regular contracts show a slightly higher increase in labor participation compared to those with temporary contracts. This is expected because these individuals have an employer to return to after recovering from sickness within two years. However, Table 5b shows that these individuals are far better in coping with the loss of disability benefit as they increase their wages by a much larger amount and rely less on unemployment benefit. Possibly for the same reason, resuming work with their own employer gives them more flexibility in increasing their number of work hours to the former level compared to the individuals with temporary contracts who usually have to resume working with a new employer.

Table 2a: Linear model explaining the effects of disability reforms on the probabilities of labor participation and benefit receipt

	Disability and possibly unemployment benefit receipt	Labor part.	Unemployment benefit receipt (no disability benefit) receipt)	General assistance receipt	Other benefits receipt
$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.008^{**}	0.007^{**}	0.003^{**}	-0.003^{*}	0.001^{***}
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.058^{***}	0.018^{***}	0.014^{***}	-0.004^{***}	-0.005^{***}
	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)
Post	0.203^{***}	-0.368^{***}	0.010^{***}	0.009^{***}	0.003^{***}
	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Constant	-0.000^{***}	0.777^{***}	0.035^{***}	0.057^{***}	0.009^{***}
	(0.001)	(0.002)	(0.001)	(0.001)	(0.000)
R-squared (within)	0.124	0.182	0.012	0.006	0.002
Observations	14,626,356				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 2b: Linear model explaining the effects of disability reforms on the amounts of wage and benefits received per month

	Disability and possibly unemployment benefit received per month	Wages earned per month	Unemployment benefit received per month (exclusive of disability benefit)	General assistance received per month	Other benefits received per month
$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-5.528	19.284	5.319^{***}	-2.841^{*}	2.262^{*}
	(4.300)	(12.802)	(1.882)	(1.487)	(1.225)
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-50.813^{***}	55.659^{***}	22.824^{***}	-4.898^{***}	0.248
	(4.413)	(13.332)	(1.952)	(1.550)	(1.230)
Post	282.124^{***}	-389.573^{***}	17.368^{***}	19.042^{***}	11.393^{***}
	(3.411)	(11.262)	(1.838)	(1.371)	(2.721)
Constant	-0.000	$1,550.659^{***}$	37.578^{***}	42.494^{***}	6.612^{***}
	(1.222)	(5.814)	(0.911)	(0.654)	(0.640)
R-squared (within)	0.081	0.036	0.010	0.007	0.000
Observations	14,626,356				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

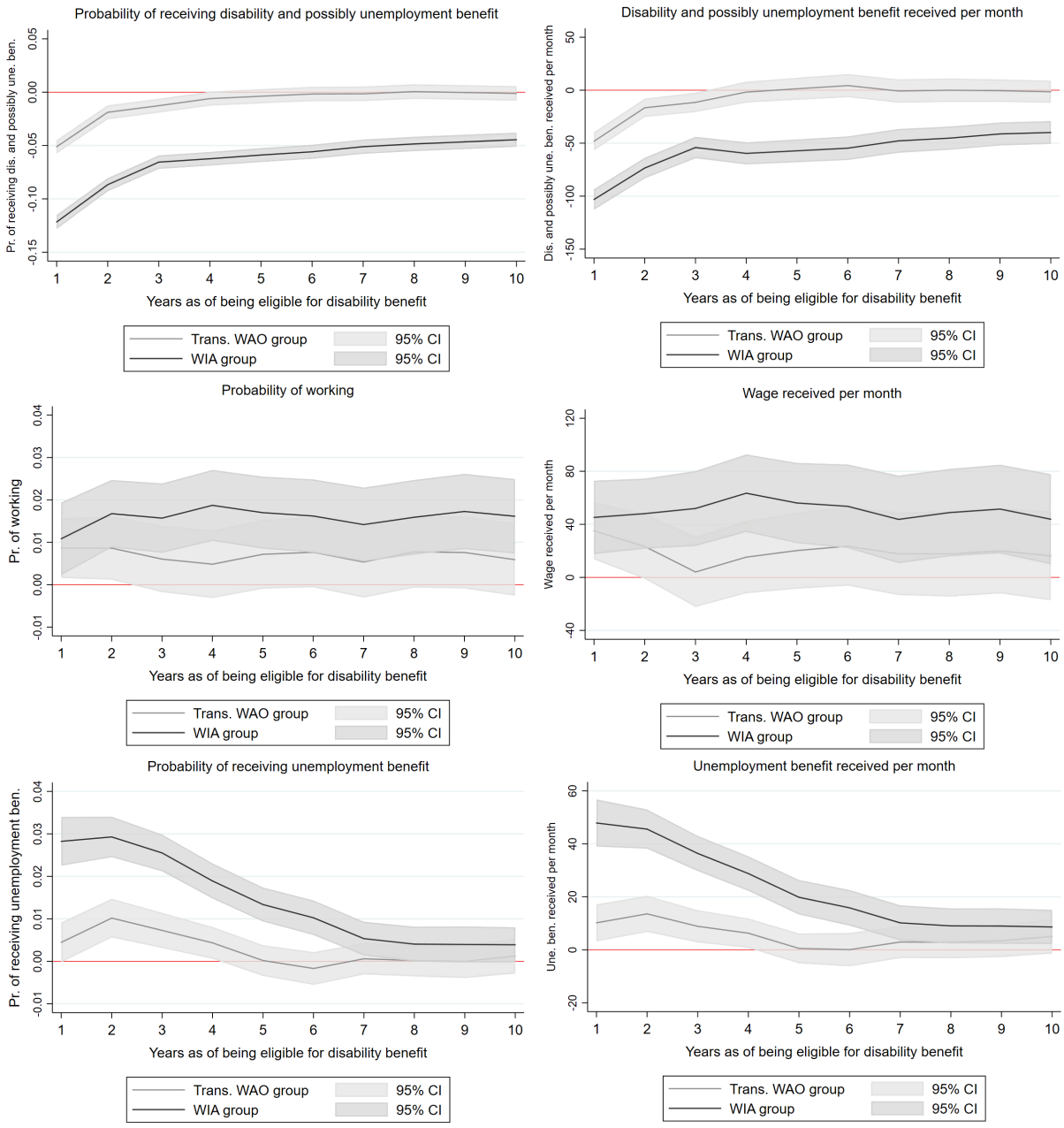

Figure 5a: Coefficient estimates of treatment and annual dummy interactions for control and treatment during the reform period from regressions for labor participation, benefit receipt (disability and unemployment benefit), labor income and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

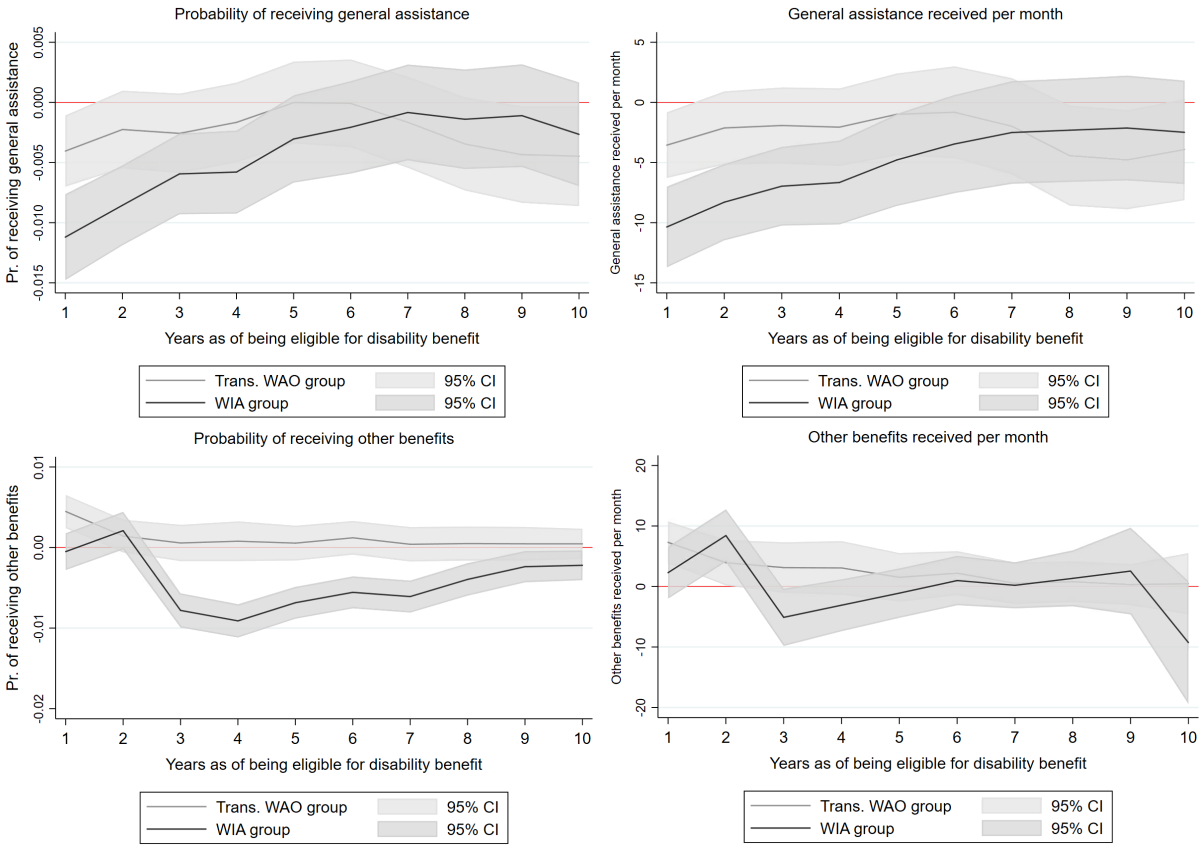

Figure 5b: Coefficient estimates of treatment and annual dummy interactions for control and treatment during the reform period from regressions for labor participation, benefit receipt (general assistance and other benefits), labor income and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

Table 3a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by gender groups

		Female	Male
Dis. and possibly une. ben. receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.006	-0.009^{**}
		(0.004)	(0.004)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.049^{***}	-0.066^{***}
		(0.004)	(0.004)
Labor part.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.009^{*}	0.006
		(0.005)	(0.005)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.019^{***}	0.017^{***}
		(0.005)	(0.005)
Une. ben. receipt (no dis. ben. receipt)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.004^{*}	0.002
		(0.002)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.013^{***}	0.015^{***}
		(0.002)	(0.002)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.002	-0.004^{**}
		(0.003)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.005^{**}	-0.003^{*}
		(0.003)	(0.002)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	0.001
		(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{***}	-0.007^{***}
		(0.001)	(0.001)
Observations		6,429,372	8,196,984
Individuals		34,186	43,583

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for men and women are as follows. For the transitional WAO reform, the p-values are 0.596, 0.671, 0.480, 0.579, and 1.000, respectively for the disability benefit receipt, labor participation, unemployment benefit receipt, general assistance receipt, and other benefit receipt outcomes. For the WIA reform, the p-values are 0.003, 0.777, 0.480, 0.579, 0.005 for the respective outcomes.

Table 3b: Linear model explaining the effects of disability reforms on labor and benefit income by gender groups

		Female	Male
Dis. and possibly une. ben. per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-9.903^{*}	-2.629
		(5.551)	(6.288)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-42.663^{***}	-57.375^{***}
		(5.720)	(6.454)
Wages earned per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	21.049	20.141
		(14.452)	(19.657)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	45.298^{***}	64.482^{***}
		(15.289)	(20.395)
Une. ben. per mon. (exc. of dis. ben.)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	5.544^{**}	5.125^{*}
		(2.277)	(2.844)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	15.630^{***}	28.468^{***}
		(2.358)	(2.953)
General assistance per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-1.970	-3.536^{**}
		(2.637)	(1.662)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-6.059^{**}	-3.996^{**}
		(2.738)	(1.739)
Other benefits per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	1.071	3.138
		(1.333)	(1.918)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-1.547	1.642
		(1.312)	(1.939)
Observations		6,429,372	8,196,984
Individuals		34,186	43,583

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for men and women are as follows. For the transitional WAO reform, the p-values are 0.386, 0.970, 0.908, 0.615, and 0.376, respectively for the disability benefit, labor, unemployment benefit, general assistance, and other benefits income outcomes. For the WIA reform, the p-values are 0.088, 0.452, 0.001, 0.525, 0.173 for the respective income outcomes.

Table 4a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by the age when individuals fell sick

		Fell sick before age 40	Fell sick between ages 40 and 44	Fell sick between ages 45 and 49	Fell sick between ages 50 and 54	Fell sick between ages 55 and 59	Fell sick after age 59
Dis. and possibly une. ben. receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−0.003	−0.015*	−0.002	−0.029***	−0.004	−0.011
		(0.004)	(0.008)	(0.009)	(0.010)	(0.008)	(0.011)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−0.046***	−0.067***	−0.081***	−0.091***	−0.051***	−0.031***
		(0.004)	(0.008)	(0.009)	(0.010)	(0.007)	(0.009)
Labor part.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.008	0.014*	0.010	0.030***	0.011	0.002
		(0.005)	(0.009)	(0.009)	(0.009)	(0.009)	(0.022)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.003***	0.021**	0.023**	0.043***	0.013	0.004
		(0.005)	(0.009)	(0.010)	(0.010)	(0.009)	(0.021)
Une. ben. receipt (no dis. ben. receipt)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	−0.000	0.004	0.004	0.010**	0.024*
		(0.001)	(0.003)	(0.004)	(0.004)	(0.005)	(0.014)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.007***	0.013***	0.016***	0.029***	0.030***	0.033**
		(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.014)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−0.004	−0.004	0.001	−0.002	0.000	−0.006
		(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.007)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−0.004	−0.006	−0.002	−0.004	0.002	−0.006
		(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.007)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	−0.001	0.002	0.000	0.003	−0.002
		(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−0.003***	−0.008***	−0.006***	−0.010***	−0.003	−0.002
		(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)
Observations		7,099,392	2,084,652	1,864,368	1,782,324	1,406,172	389,448
Individuals		37,724	11,087	9,915	9,481	7,486	2,076

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 4b: Linear model explaining the effects of disability reforms on labor and benefit income by the age when individuals fell sick

		Fell sick before age 40	Fell sick between ages 40 and 44	Fell sick between ages 45 and 49	Fell sick between ages 50 and 54	Fell sick between ages 55 and 59	Fell sick after age 59
Dis. and possibly une. ben. per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−0.533	−13.981	2.763	−46.739***	9.442	−17.428
		(5.080)	(11.689)	(14.824)	(16.094)	(13.364)	(12.893)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−37.139***	−48.792***	−81.611***	−103.513***	−44.784***	−22.614*
		(5.266)	(12.263)	(15.017)	(16.631)	(12.853)	(12.749)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	31.518**	50.331	32.396	108.105***	−3.782	−27.985
		(15.788)	(31.254)	(36.250)	(34.607)	(44.997)	(80.652)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	95.659***	83.539**	96.918***	157.964***	58.700	−31.836
		(16.581)	(32.711)	(36.458)	(37.188)	(43.227)	(78.772)
Une. ben. per mon. (exc. of dis. ben.)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	2.366	3.962	8.858	2.125	14.837*	47.416*
		(1.919)	(4.704)	(5.596)	(6.360)	(8.689)	(24.411)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	10.981***	22.582***	21.292***	41.984***	56.365***	69.712***
		(1.995)	(4.822)	(5.776)	(6.730)	(8.620)	(23.916)
General assistance per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−4.123*	−4.028	2.311	−1.908	1.315	−2.656
		(2.390)	(4.021)	(3.966)	(3.412)	(2.507)	(4.329)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−5.174**	−5.769	−2.447	−4.293	2.380	−5.393
		(2.520)	(4.227)	(4.087)	(3.604)	(2.434)	(4.470)
Other benefits per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	1.216	2.917	2.942	0.193	11.794*	−18.181
		(0.927)	(2.635)	(4.217)	(5.370)	(6.243)	(12.476)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−0.249	2.705	−2.649	−3.681	6.038	−2.834
		(0.987)	(2.606)	(3.641)	(6.050)	(5.817)	(10.497)
Observations		7,099,392	2,084,652	1,864,368	1,782,324	1,406,172	389,448
Individuals		37,724	11,087	9,915	9,481	7,486	2,076

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 5a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by work status

		Regular	Temp.	Unemp.
Dis. and possibly une. ben. receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.009^{**}	-0.016^{**}	-0.004
		(0.004)	(0.008)	(0.008)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.058^{***}	-0.055^{***}	-0.068^{***}
		(0.003)	(0.008)	(0.008)
Labor part.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.008^{*}	0.019^{**}	-0.000
		(0.004)	(0.008)	(0.008)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.032^{***}	0.026^{***}	-0.027^{***}
		(0.005)	(0.009)	(0.008)
Une. ben. receipt (no dis. ben. receipt)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	$0.007**$	0.007^{*}
		(0.001)	(0.003)	(0.004)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.005^{***}	0.023^{***}	0.043^{***}
		(0.001)	(0.003)	(0.004)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.001	$-0.009*$	-0.006
		(0.001)	(0.005)	(0.004)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{*}	-0.014^{***}	-0.003
		(0.001)	(0.005)	(0.005)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	-0.002	0.003
		(0.001)	(0.002)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.006^{***}	0.002	-0.009^{***}
		(0.001)	(0.002)	(0.002)
Observations		8,518,944	2,534,856	2,945,844
Individuals		45,312	13,445	15,682

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Work status indicates individuals who hold a regular contract, temporary contract or a contract through a temporary work agency, or are unemployed, and it is that at the time individuals fall sick. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals hold a temporary work contract and those who hold a regular work contract are as follows. For the transitional WAO reform, the p-values are 0.434, 0.219, 0.058, 0.117, and 0.180, respectively for the disability benefit receipt, labor participation, unemployment benefit receipt, general assistance receipt, and other benefits receipt outcomes. For the WIA reform, the p-values are 0.725, 0.560, 0.000, 0.031, 0.000 for the respective outcomes. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals who are unemployed and those who hold a regular work contract are as follows. For the transitional WAO reform, the p-values are 0.576, 0.371, 0.146, 0.225, and 0.371, respectively for the disability benefit receipt, labor participation, unemployment benefit receipt, general assistance receipt, and other benefits receipt outcomes. For the WIA reform, the p-values are 0.242, 0.000, 0.000, 1.000, 0.180 for the respective outcomes.

Table 5b: Linear model explaining the effects of disability reforms on labor and benefit income by work status

		Regular	Temp.	Unemp.
Dis. and possibly une. ben. per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−4.681	−12.497	−9.387
		(5.231)	(10.595)	(11.194)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−55.753***	−39.572***	−51.530***
		(5.292)	(11.544)	(11.454)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	30.114	45.490*	−9.080
		(18.607)	(25.687)	(23.914)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	125.315***	61.299**	−119.296***
		(19.240)	(28.120)	(23.816)
Une. ben. per mon. (exc. of dis. ben.)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	2.518	10.605***	18.325***
		(2.007)	(3.937)	(6.169)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	8.604***	32.785***	70.116***
		(2.140)	(4.301)	(6.107)
General assistance per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−1.532	−9.038*	−5.472
		(1.270)	(4.723)	(4.468)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−2.974**	−15.257***	−3.515
		(1.345)	(5.062)	(4.631)
Other benefits per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	3.493**	−0.336	2.592
		(1.753)	(2.254)	(2.677)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−3.611**	6.693**	6.562***
		(1.733)	(2.818)	(2.450)
Observations		8,518,944	2,534,856	2,945,844
Individuals		45,312	13,445	15,682

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Work status indicates individuals who hold a regular contract, temporary contract or a contract through a temporary work agency, or are unemployed, and it is that at the time individuals fall sick. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals hold a temporary work contract and those who hold a regular work contract are as follows. For the transitional WAO reform, the p-values are 0.508, 0.628, 0.067, 0.125, and 0.180, respectively for the disability benefit, labor, unemployment benefit, general assistance, and other benefits income outcomes. For the WIA reform, the p-values are 0.203, 0.060, 0.000, 0.019, 0.002 for the respective income outcomes. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals who are unemployed and those who hold a regular work contract are as follows. For the transitional WAO reform, the p-values are 0.703, 0.196, 0.015, 0.396, and 0.778, respectively for the disability benefit, labor, unemployment benefit, general assistance, and other benefits income outcomes. For the WIA reform, the p-values are 0.738, 0.000, 0.000, 0.911, 0.001 for the respective income outcomes.

6.4 Sensitivity analysis

Is the common trends assumption violated when alternative comparison groups are considered, and how much are the estimated reform effects affected when it is violated?

In Section 5, a statistically significant difference between the pre-reform time trends of the treatment and control groups provided evidence against the common trends assumption. We showed that the difference in the time trends of the treatment and control groups is due, in particular, to the statistically significant differences in mean outcomes in the months through the end of the pre-reform period, from January 2002 to June 2003. Here we first provide evidence that the common trends assumption is not systematically violated when alternative control groups are considered. Second, we show that the differences in pre-reform time trends have no notable impact on the baseline estimates of the reform effects. In these analyzes we focus on the WIA group since Section 5 showed that the pre-reform time trend of the WIA group is statistically different from that of the WAO group for almost all outcomes, while the pre-reform time trend of the transitional WAO group is statistically different from that of the WAO group only for wage and benefit income outcomes.

To investigate if the observed difference between the pre-reform time trends of the treatment and control groups is due to particular behavioral responses of the control group at hand (WAO group), we check if indeed a difference is absent when alternative control groups are considered. We consider several alternative control groups. First, we consider the transitional WAO group as a new control group. It is important to note that in this particular check our aim is not to show whether the baseline effects of the WIA reform are robust when the control group is changed from the WAO group to the transitional WAO group. As far as the post-reform period is concerned, behavioral differences between the WIA and WAO groups may differ from those between the WIA and transitional WAO groups since the benefit rules that apply to the post-reform period differ between the transitional WAO and WAO groups (Section 2). Instead, our aim is to show whether the behavior of the WIA group is statistically indifferent from that of "any" given alternative control group (in this case the transitional WAO group) in the absence of the reform as we should expect.

Exploratory graphical analysis in Section 4 showed that the differences between the pre-reform time trends of the WIA and the transitional WAO groups are often insignificant providing descriptive evidence for common trends in the pre-reform period. As a formal check, we consider the WIA group as the treatment group and the transitional WAO group as the new control group and, as in Section 5, use pre-reform data from January 1999 to June 2003 to estimate the regression given by Equation (1) for different outcomes. We then test for the common trends assumption using the F-statistic that tests whether the estimated coefficients of all interactions are jointly equal to zero. The results for the F-statistic are as follows (p-values in parentheses). For labor participation, unemployment benefit receipt, general assistance receipt, and other benefits receipt outcomes, the results are respectively 1.02 (0.443), 1.25 (0.107), 1.01 (0.460), and 0.89 (0.697). For labor, unemployment benefit, general assistance, and other benefits income outcomes, the results are respectively 0.99 (0.500), 1.11 (0.270), 1.07 (0.330), and 0.90 (0.677). The large p-values for all outcomes provide evidence in favor of the common trends assumption in that the observed difference in the pre-reform time trends of the WIA and the WAO groups is not likely to be related to the behavioral responses to the reforms, such as anticipatory responses, specific to a certain group.

In consideration of a second alternative control group, we distinguish across individuals in the WAO group who fell sick in July, August, and September 2003, and consider them as three alternative control groups. We check how the pre-reform time trend of the WIA group differs

from each of those of the three subgroups of the WAO group. We then investigate whether all the three differences are always at the same extent of the difference between the pre-reform time trend of the WIA group and that of the WAO group as a whole, suggesting that the observed difference in the pre-reform time trends of the WIA and WAO groups is systematic. To these purposes, we compute F-statistics (same as above), based on regressions estimated on the three subgroups of the WAO group, to test whether the pre-reform time trends are significantly different from each other across the treatment and different control groups. The results for the F-statistics are as follows (p-values in parentheses). For labor participation, unemployment benefit receipt, general assistance receipt, and other benefits receipt outcomes, the results for the July group, respectively, are 1.41 (0.026), 1.29 (0.075), 1.11 (0.267), 1.14 (0.221). The results for the August group, respectively, are 1.41 (0.027), 1.27 (0.091), 1.32 (0.057), 1.38 (0.035). The results for the September group, respectively, are 1.32 (0.060), 0.74 (0.924), 1.33 (0.053), 0.95 (0.572). For labor, unemployment benefit, general assistance, and other benefits income outcomes, the results for the July group, respectively, are 1.11 (0.274), 2.16 (0.000), 1.21 (0.137), 1.25 (0.108). The results for the August group, respectively, are 1.26 (0.097), 1.64 (0.002), 1.80 (0.000), 0.85 (0.768). The results for the September group, respectively, are 1.19 (0.165), 0.91 (0.653), 1.36 (0.041), 1.06 (0.366). These results show that for almost none of the outcomes the test statistics are significant at the 5 percent level for the September group. Besides, we do not observe any systematic rejection of the test statistics across outcome variables, or across individuals who fell sick in consecutive months in the control groups. These results suggest that the differences in the pre-reform time trends of the WIA group and the subgroups of the WAO are not systematic.

Tables 6a and 6b present results from regressions based on the July, August and September subsamples of the individuals insured under the WAO and considered as alternative control groups. The estimates of the effects of the WIA reform when the control group is the September group are close to the baseline estimates of the WIA reform in Tables 2a and 2b except for labor participation and income. In view of the finding that the common trends assumption is not violated based on the above presented F-statistics and a significance level of 5 percent for most of the outcome variables when the September group is considered as the control group, these regression results suggest that concerns for the violation of the common trends assumption when the WAO group is considered as the control group should be small.

Next, we investigate to which extent the observed difference between the pre-reform time trends of the treatment and control groups has an impact on the baseline estimates of the effects of the WIA reform. In particular, we check whether the baseline estimates of the reform effects remain largely unaffected when we drop in our study sample the months where the pre-reform time trends of the treatment and control groups differ, but keep those months where they do not. To determine the months where the pre-reform time trends do not differ between the treatment and control groups, we rely on the visual evidence in Figures 1a and 1b, and on an F-statistic (same as above) that tests whether the pre-reform time trends between the treatment and control groups are significantly different from each other. Figures 1a and 1b show that the time-trends of the treatment and control groups overlap throughout the period from January 1999 to December 2001 in all outcome variables. The results for the F-statistics are as follows (p-values in parentheses). For labor participation, unemployment benefit receipt, general assistance receipt, and other benefits receipt outcomes, the results, respectively, are 1.22 (0.180), 1.20 (0.193), 1.27 (0.135), 1.30 (0.114). For labor, unemployment benefit, general assistance, and other benefits income outcomes, the results, respectively, are 1.31 (0.105), 1.05 (0.394), 1.48 (0.035), 0.89 (0.652). These test results show that from January 1999 to December 2001 the pre-reform time trends of the treatment and control groups are statistically indistinguishable from each other for all outcomes except for general assistance income. Tables 7a and 7b present

estimation results based on the restricted pre-reform period from January 1999 to December 2001. The results are very close to the baseline results in Tables 2a and 2b except that the coefficient estimate for general assistance receipt becomes insignificant and the coefficient estimate for general assistance income becomes smaller and less significant. We conclude that the observed difference between the pre-reform time trends of the treatment and control groups has no notable impact on the baseline estimates of the effects of the WIA reform.

Are the estimated reform effects sensitive to the cut-off number of days spent in the sickness scheme?

As explained in Section 3, sickness cases that last shorter than 180 days are under-reported among the individuals who fell sick in the earlier months of the period from July 2003 until March 2004 where sickness data are available. The earlier the date individuals fell sick, the more severe under-reporting is in the data. We have restricted the initial sample of sick individuals to include only those who have spent at least 180 days in the sickness scheme to ensure that the three cohorts of sick individuals who fell sick in the third and fourth quarters of 2003 and the first quarter of 2004 are comparable in the number of days spent in sickness. The effects of the reforms could be under- or overestimated if we allowed our study sample to include the sickness cases that last shorter than 180 days and are under-reported. For example, among the individuals with short-term sickness, those who are insured under the WIA may be inclined to resume working earlier than if they were insured under the WAO if they perceive that they eventually will not become entitled to disability benefits due to a longer sickness period, and better chances of recovery, but also much stricter eligibility criteria. In this case, the imposed sample restriction will drop proportionally more short-term sickness cases of relatively healthier individuals in the WIA group, and lead to an underestimation of the effects of the WIA reform.

Here we examine to which extent the estimated effects of the reforms are sensitive to allowing in the study sample short-term sickness cases that are shorter than 180 days and under-reported to different extents. In Tables 8a and 8b we present results from the estimation of the regression given by Equation (1) on subsamples of individuals who have spent at least 90, 120, and 150 days in the sickness scheme. With respect to labor participation and benefit receipt, compared to the baseline results in Table 2a, the results in Table 8a show small differences for the WIA reform in all outcomes except when the outcome is labor participation and the number of days spent in the sickness scheme is at least 90 days. Larger differences are observed for the transitional WAO reform when the outcome is labor participation, or when it is disability benefit, and the number of days spent in the sickness scheme is at least 90 days. A potential reason is that, since under-reporting of short-term sickness cases occurs more often for the transitional WAO group compared to the WIA group, the estimates of the effects of the transitional WAO reform are more prone to be biased down- or upward due to reasons as exemplified above. With respect to wage and benefit income, similar differences are observed for wage and disability benefit income when baseline results in Table 2b are compared to the sensitivity results in Table 8b. These results show that the baseline estimates of the effects of the WIA reform are to a large extent robust to the sampling of sick individuals with respect to the number of days they spend in the sickness scheme. This is true to a lesser extent for the transitional WAO reform.

Are the estimated long-term effects of the reforms sensitive to the correlation between the event and calendar time dummies?

We analyzed the long-term effects of the disability reforms using the regression model given by Equation (2) in Section 5. In particular, the long-term effects are analyzed through the interactions of treatment and year dummies over a period of ten years. In this equation, the year

dummies of event time are correlated with the month dummies of calendar time by construction. This correlation structure could affect the point estimates or the standard errors of the main effects of interest given by the coefficients of the interaction terms. To check to whether the correlation between the year dummies of event time and month dummies of calendar time lead to biased estimates of long-erm effects, instead of year dummies, we consider dummies that indicate periods of three years of event time which have much weaker correlations with the month dummies of calendar time. Figures 6a and 6b present the results based on three-year period dummies. Using these dummies leads to estimates of the effects of the interaction terms very similar to those when year dummies are used in Figures 5a and 5b so that our qualitative results are not affected.

Table 6a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by subgroups of the WAO group as control groups

		Control group is the July WAO group	Control group is the August WAO group	Control group is the September WAO group
Dis. and possibly une. ben. receipt	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.064^{***}	-0.060^{***}	-0.057^{***}
		(0.005)	(0.004)	(0.004)
Labor part.	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.033^{***}	0.018^{***}	0.008^{*}
		(0.005)	(0.005)	(0.005)
Une. ben. receipt (no dis. ben. receipt)	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.017^{***}	0.014^{***}	0.012^{***}
		(0.002)	(0.002)	(0.002)
General assistance receipt	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.005^{***}	-0.003	-0.004^{**}
		(0.002)	(0.002)	(0.002)
Other benefits receipt	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.006^{***}	-0.004^{***}	-0.005^{***}
		(0.001)	(0.001)	(0.001)
Observations		11,006,388	11,210,676	11,635,380
Individuals		58,915	59,979	62,191

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 6b: Linear model explaining the effects of disability reforms on labor and benefit income by subgroups of the WAO group as control groups

		Control group is the July WAO group	Control group is the August WAO group	Control group is the September WAO group
Dis. and possibly une. ben. per mon.	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-59.919^{***}	-48.631^{***}	-49.470^{***}
		(6.693)	(6.267)	(5.761)
Wages earned per month	$\text{Treat}^{\text{WIA}} \times \text{Post}$	110.880^{***}	42.434^{***}	30.597^{*}
		(21.004)	(18.580)	(16.618)
Une. ben. per mon. (exc. of dis. ben.)	$\text{Treat}^{\text{WIA}} \times \text{Post}$	26.627^{***}	23.021^{***}	20.385^{***}
		(2.938)	(2.747)	(2.508)
General assistance per mon.	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-5.119^{**}	-5.004^{**}	-4.694^{**}
		(2.320)	(2.179)	(2.060)
Other benefits per mon.	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-2.678	2.716	0.564
		(1.857)	(1.738)	(1.623)
Observations		11,006,388	11,210,676	11,635,380
Individuals		58,915	59,979	62,191

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 7a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt when the pre-reform period is restricted

	Disability and possibly unemployment benefit receipt	Labor part.	Unemployment benefit receipt (no disability benefit) receipt)	General assistance receipt	Other benefits receipt
$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.008^{**}	0.005	0.003^{**}	-0.001	0.001
	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.059^{***}	0.015^{***}	0.013^{***}	-0.002	-0.005^{***}
	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)
Observations	12,916,711				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time. The pre-reform period corresponds to the months from January 1999 to December 2001. The post-reform period corresponds to the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 7b: Linear model explaining the effects of disability reforms on the amounts of wage and benefits received per month when the pre-reform period is restricted

	Disability and possibly unemployment benefit received per month	Wages earned per month	Unemployment benefit received per month (exclusive of disability benefit)	General assistance received per month	Other benefits received per month
$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-5.618	16.998	5.068^{***}	-1.236	2.194^{*}
	(4.300)	(13.296)	(1.827)	(1.608)	(1.257)
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-50.829^{***}	55.893^{***}	20.480^{***}	-3.024^{*}	-0.000
	(4.414)	(13.836)	(1.894)	(1.671)	(1.239)
Observations	14,626,356				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time. The pre-reform period corresponds to the months from January 1999 to December 2001. The post-reform period corresponds to the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 8a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by the number of days individuals spend in the sickness scheme

		Number of days spent in the sickness scheme is at least 90	Number of days spent in the sickness scheme is at least 120	Number of days spent in the sickness scheme is at least 150
Dis. and possibly une. ben. receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.021^{***}	-0.014^{***}	-0.010^{***}
		(0.002)	(0.002)	(0.003)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.052^{***}	-0.050^{***}	-0.053^{***}
		(0.002)	(0.002)	(0.003)
Labor part.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.017^{***}	0.011^{***}	0.008^{**}
		(0.003)	(0.003)	(0.003)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.028^{***}	0.021^{***}	0.018^{***}
		(0.003)	(0.003)	(0.003)
Une. ben. receipt (no dis. ben. receipt)	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.002^{*}	0.002^{***}	0.003^{***}
		(0.001)	(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.010^{***}	0.011^{***}	0.013^{***}
		(0.001)	(0.001)	(0.001)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.001	-0.002	-0.002^{*}
		(0.001)	(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{***}	-0.004^{***}	-0.004^{***}
		(0.001)	(0.001)	(0.001)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.000	0.000	0.001
		(0.001)	(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.004^{***}	-0.004^{***}	-0.004^{***}
		(0.001)	(0.001)	(0.001)
Observations		23,579,556	20,078,508	17,097,660
Individuals		125,535	106,807	90,918

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

Table 8b: Linear model explaining the effects of disability reforms on labor and benefit income by the number of days individuals spend in the sickness scheme

		Number of days spent in the sickness scheme is at least 90	Number of days spent in the sickness scheme is at least 120	Number of days spent in the sickness scheme is at least 150
Dis. and possibly une. ben. per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-24.601^{***}	-14.913^{***}	-8.820^{**}
		(3.009)	(3.362)	(3.804)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-52.127^{***}	-46.730^{***}	-46.764^{***}
		(3.084)	(3.462)	(3.914)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	41.900^{***}	23.891^{**}	16.379
		(10.100)	(10.895)	(11.799)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	77.722^{***}	59.118^{***}	53.001^{***}
		(10.528)	(11.407)	(12.339)
Une. ben. per mon. (exc. of dis. ben.)	$\text{Treat}^{\text{ Trans. WAO}} \times \text{Post}$	3.820^{***}	4.767^{***}	6.152^{***}
		(1.447)	(1.574)	(1.721)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	16.603^{***}	18.725^{***}	21.714^{***}
		(1.489)	(1.627)	(1.783)
General assistance per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	$-1.167)$	-2.180^{*}	-2.847^{**}
		(1.147)	(1.245)	(1.367)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-4.193^{***}	-4.766^{***}	-4.952^{***}
		(1.187)	(1.297)	(1.426)
Other benefits per mon.	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.977	1.612^{*}	2.009^{*}
		(0.883)	(0.971)	(1.081)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.567	0.095	0.145
		(0.888)	(0.983)	(1.088)
Observations		23,579,556	20,078,508	17,097,660
Individuals		125,535	106,807	90,918

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for disability benefit. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for disability benefit. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * indicate statistical significance at the 0.01, 0.05, 0.10 levels, respectively.

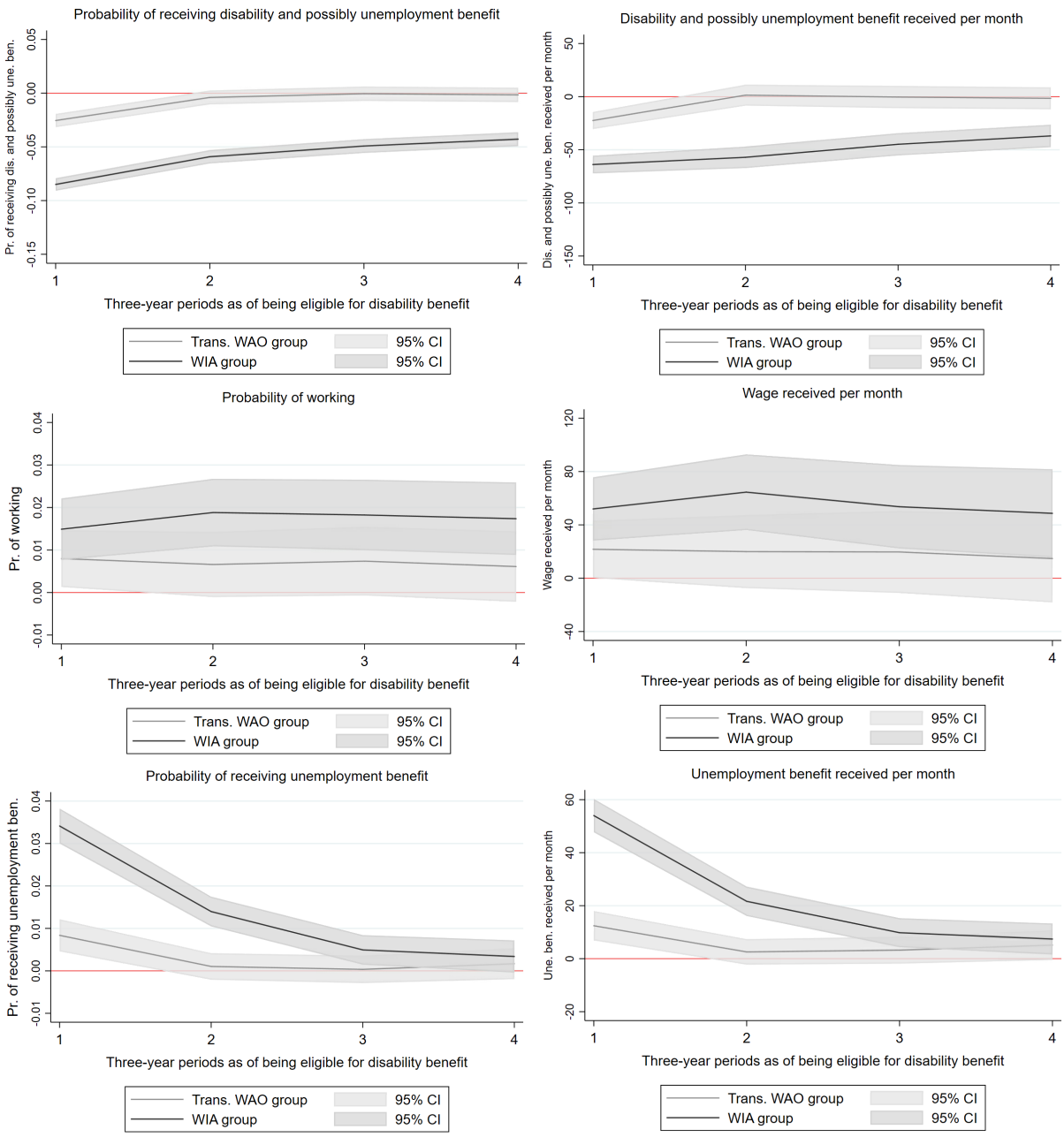

Figure 6a: Coefficient estimates of treatment and annual dummy interactions for control and treatment during the reform period from regressions for labor participation, benefit receipt (disability and unemployment benefit), labor income, and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

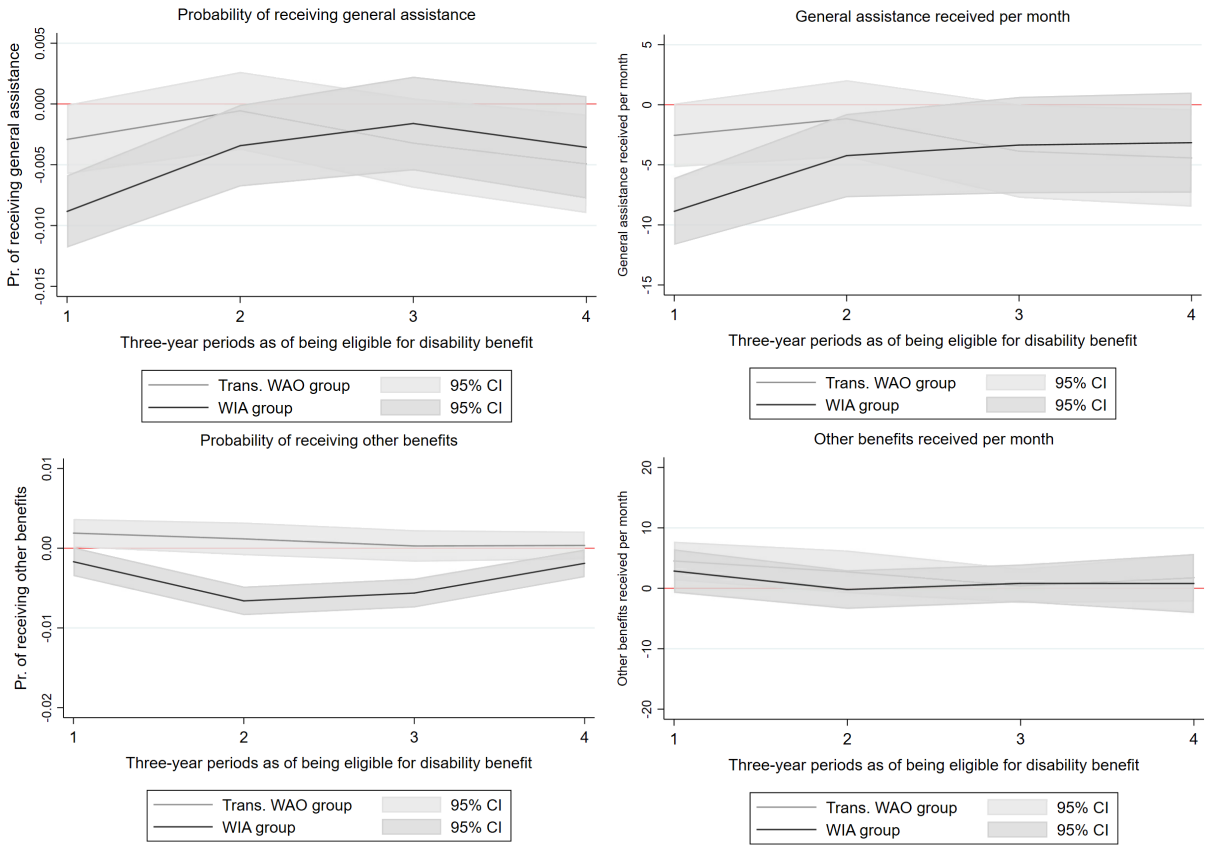

Figure 6b: Coefficient estimates of treatment and annual dummy interactions for control and treatment during the reform period from regressions for labor participation, benefit receipt (general assistance and other benefits), labor income and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

7 Conclusion

In the last decades disability insurance programs have grown in many western countries. Governments implement social security reforms to reduce enrollment in the disability insurance program and the cost of disability benefits. We evaluated a major disability insurance reform that introduced strong incentives for work resumption, not only for the disability scheme but also for the sickness scheme that precedes it. Using unique administrative data on agents who fall sick before and after the reform, we analyzed the effect of the reform on labor participation decisions and use of benefits from alternative social security programs.

Difference-in-difference analyzes provide notable findings. The new benefit regime substantially decreased disability benefit receipt, and its impact had been persistent during the ten years of the study period. It increased labor participation and unemployment benefit receipt although to a limited extent. On the other hand, it substantially increased wages earned and unemployment benefits received. Individuals fully compensate lost disability benefits by increasing their wages, but in addition, they also increase their income from unemployment benefit.

The impact of the new benefit regime is substantially heterogenous with respect to work status and age. Individuals who are unemployed are much more limited in their access to the disability insurance scheme and lack the incentives the group of individuals with regular or temporary contracts have to resume working. This suggests that unemployed individuals, who are already a vulnerable labor market group, face additional limitations in their access to the labor market when they fall sick. Besides the unemployed, older individuals appear as a second vulnerable group. They are less able to compensate the decrease in disability benefit receipt or income with higher labor participation or wages, whereas they more often rely on unemployment benefit.

A distinctive feature of the new disability scheme is the extension of the sickness period from one to two years. Participation in different sickness schemes could generate differences in health or economic behavior until the groups who participate in these schemes become eligible to apply for disability benefit. For example, due to the extra year spent in the sickness scheme, the health status of the treatment group could improve more until this group becomes eligible for disability benefit. We did not analyze to which extent the possible behavioral differences due to participation in different sickness schemes drive the behavioral differences we observe when individuals are eligible for disability benefit. To address this question, the comparison of the treatment and control groups can be conditioned on the health status of these groups at the end of the sickness period, and subsequent labor and benefit claiming behavior can be analyzed at the onset or during the disability period. This is left to future research.

Acknowledgments

This research is supported by the Network for studies on Pensions, Aging and Retirement (Netspar) under grant number LMVP 2014.03. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of Netspar. We thank the UWV, and in particular Lucien Rondagh and Roel Ydema, for providing the sickness data. We thank Jennifer Alonso-García, Pilar García-Gómez, Pierre Koning, Martin Salm, Arthur van Soest, seminar participants at Tilburg University, and conference participants at the Netspar Pension Day 2018 and at the Netspar International Pension Workshop 2019 for their helpful comments.

References

Arulampalam, W., 2001. Is unemployment really scarring? Effects of unemployment experiences on wages. The Economic Journal 111 (475), F585–606.

Arulampalam, W., Gregg, P., Gregory, M., 2001. Introduction: unemployment scarring. The Economic Journal 111 (475), F577–584.

Autor, D., Kostøl, A., Mogstad, M., Setzler, B., 2019. Disability benefits, consumption insurance, and household labor supply. American Economic Review 109 (7), 2613–54.

Autor, D. H., Duggan, M., Greenberg, K., Lyle, D. S., 2016. The impact of disability benefits on labor supply: Evidence from the VA's disability compensation program. American Economic Journal: Applied Economics 8 (3), 31–68.

Autor, D. H., Duggan, M. G., 2003. The rise in the disability rolls and the decline in unemployment. The Quarterly Journal of Economics 118 (1), 157–206.

Berendsen, E., van Deursen, C., van Sonsbeek, J.-M., 2019. Stijging WIA-instroom nog geen reden tot zorg. Economisch Statistische Berichten 104 (4773), 230–232.

Borghans, L., Gielen, A. C., Luttmer, E. F. P., 2014. Social support substitution and the earnings rebound: evidence from a regression discontinuity in disability insurance reform. American Economic Journal: Economic Policy 6 (4), 34–70.

Burkhauser, R., Marc, D., McVicar, D., Wilkins, R., 2014. Disability benefit growth and disability reform in the US: lessons from other OECD nations. IZA Journal of Labor Policy 3 (4).

Campolieti, M., 2004. Disability insurance benefits and labor supply: some additional evidence. Journal of Labor Economics 22 (4), 863–889.

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

Chen, S., van der Klaauw, W., 2008. The work disincentive effects of the disability insurance program in the 1990s. Journal of Econometrics 142 (2), 757–784.

Dahl, G. B., Gielen, A. C., 2018. Intergenerational spillovers in disability insurance. NBER Working Paper No. 24296.

De Jong, P., Lindeboom, M., 2004. Privatisation of sickness insurance: evidence from the Netherlands. Swedish Economic Policy Review 11, 121–144.

De Jong, P., Lindeboom, M., van der Klaauw, B., 2011. Screening disability insurance applications. Journal of the European Economic Association 9 (1), 106–129.

Deshpande, M., 2016. The effect of disability payments on household earnings and income: Evidence from the SSI children's program. Review of Economics and Statistics 98 (4), 638–654.

Fevang, E., Hardoy, I., Red, K., 2017. Temporary disability and economic incentives. The Economic Journal l127 (603), 1410–1432.

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

García-Gómez, P., Gielen, A., 2018. Mortality effects of containing moral hazard: evidence from disability insurance reform. Health Economics 27 (3), 606–621.

Gelber, A., Moore, T. J., Strand, A., 2017. The effect of disability insurance payments on beneficiaries' earnings. American Economic Journal: Economic Policy 9 (3), 229–61.

Groot, N. D., Koning, P. W. C., 2016. Assessing the effects of disability insurance experience rating. The case of the Netherlands. Labour Economics 41, 304–317.

Gruber, J., 2000. Disability insurance benefits and labor supply. Journal of Political Economy 108 (6), 1162–1183.

Karlström, A., Palme, M., Svensson, I., 2008. The employment effect of stricter rules for eligibility for di: Evidence from a natural experiment in sweden. Journal of Public Economics 92 (10-11), 2071–2082.

Koning, P. W. C., 2009. Experience rating and the inflow into disability insurance. De Economist 157 (3), 315–335.

Koning, P. W. C., Lindeboom, M., 2015. The rise and fall of disability insurance enrollment in the Netherlands. Journal of Economic Perspectives 29 (2), 151–172.

Koning, P. W. C., van Sonsbeek, J.-M., 2017. Making disability work? The effects of financial incentives on partially disabled workers. Labour Economics 47, 202–215.

Koning, P. W. C., van Vuuren, D. J., 2010. Disability insurance and unemployment insurance as substitute pathways. Applied Economics 42 (5), 575–588.

Kostøl, A. R., Mogstad, M., 2014. How financial incentives induce disability insurance recipients to return to work. American Economic Review 104 (2), 624–655.

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

Maestas, N., Song, J., 2011. The labor supply effects of disability insurance: Evidence from automatic conversion using administrative data. Michigan Retirement Research Center Research Paper No. 2010-247.

Mandicó, S. G., García-Gómez, P., O'Donnell, O., 2018. Earnings responses to disability benefit cuts. IZA Discussion Paper No. 11410.

Moore, T. J., 2015. The employment effects of terminating disability benefits. Journal of Public Economics 124, 30–43.

Mullen, K. J., Staubli, S., 2016. Disability benefit generosity and labor force withdrawal. Journal of Public Economics 143, 49–63.

OECD, 2010. Sickness, Disability and Work: Breaking the Barriers. OECD Publishing, Paris.

Ruh, P., Staubli, S., 2018. Financial incentives and earnings of disability insurance recipients: evidence from a notch design. CEPR Discussion Papers (12979).

Staubli, S., 2011. The impact of stricter criteria for disability insurance on labor force participation. Journal of Public Economics 95 (9-10), 1223–1235.

Van Sonsbeek, J.-M., Gradus, R. H. J. M., 2013. Estimating the effects of recent disability reforms in the Netherlands. Oxford Economic Papers 65 (4), 832–855.