Empirical Economics (2022) 62:749–777
https://doi.org/10.1007/s00181-021-02018-2

Preventing NEETs during the Great Recession: the effects of mandatory activation programs for young welfare recipients

Emile Cammeraat[1] · Egbert Jongen[1,2,4] · Pierre Koning[1,3,4]

Received: 3 September 2020 / Accepted: 11 January 2021 / Published online: 16 February 2021

Abstract
We study the impact of mandatory activation programs for young welfare recipients in the Netherlands. What makes this reform unique is that it clashed head on with the Great Recession. We use differences-in-differences and data for the period 1999–2012 to estimate the effects of this reform. We find that the reform reduced the number of welfare recipients but had no effect on the number of NEETs (individuals not in employment, education or training). The absence of employment effects contrasts with previous studies on the impact of mandatory activation programs, which we argue is due to the reform taking place during a severe economic recession.

Keywords NEETs · Mandatory activation programs · Differences-in-differences

JEL Classification C21 · H31 · J21

We thank the editor Arthur van Soest, two anonymous referees, Leon Bettendorf, Richard Blundell, Matz Dahlberg, Sander Gerritsen, Bas Jacobs, Max van Lent, Daniël van Vuuren and numerous seminar and conference participants for very useful comments and suggestions. Remaining errors are our own.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00181-021-02018-2.

✉ Emile Cammeraat
e.cammeraat@law.leidenuniv.nl

Egbert Jongen
e.l.w.jongen@cpb.nl

Pierre Koning
p.w.c.koning@law.leidenuniv.nl

1 Leiden University, Leiden, The Netherlands

2 CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands

3 Tinbergen Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

4 IZA, Bonn, Germany

1 Introduction

Young individuals not in employment, education or training (NEETs) are a major policy concern (Carcillo et al. 2015). To reduce the number of NEETs, policymakers in many developed countries provide employment and training programs targeted at young individuals. There is strong evidence that making welfare conditional on job search effort and participation in employment programs may be an effective tool to decrease welfare claims and increase employment rates of young individuals (Blundell et al. 2004; Dahlberg et al. 2009; Persson and Vikman 2014; Kluve et al. 2016; Hernæs et al. 2017). Prominent examples of welfare conditionality targeted at young unemployed individuals include the New Deal for Young People in the UK and the Job Corps in the USA (Kluve et al. 2016).

A pertaining question is whether the intended effects of welfare conditionality for young individuals still stand in times of economic recession. While the requirements may still discourage young individuals from applying for welfare benefits, depressed labor demand may limit the potential employment effects. This would then imply only a decline in young NEETs on welfare, not a decline in the number of young NEETs. To shed light on this issue, this paper studies the effects of a mandatory activation program for young individuals during a severe economic recession. Specifically, we study the WIJ or 'Work Investment Act for Young Individuals' reform, introduced in the Netherlands at the end of 2009, which was just after the start of the Great Recession. The reform targeted individuals up to and including 26 years of age. For this group, the participation in public employment programs, apprenticeships or internships was mandatory. Compared to the old system, the reform implied both a strong shift to programs that were work-oriented and a shift from voluntary to mandatory programs. We study the effects of this reform on the following key outcome variables: NEETs claiming welfare benefits, NEETs not claiming welfare benefits, the overall NEETs rate, the employment rate and the enrollment rate in education.

To estimate the causal effect of the WIJ reform, we use differences-in-differences and a large administrative dataset, the Labor Market Panel (*Arbeidsmarktpanel*) of Statistics Netherlands (2015). The Labor Market Panel tracks 1.2 million individuals over the period 1999–2012 and contains a broad set of labor market outcomes and individual and household characteristics. We consider the treatment effect for three different age groups, 20–22, 23–24 and 25–26 years of age, while our base control group consists of individuals 27–28 years of age. Throughout our analysis, a key challenge is to control for potentially different time effects between the treatment and control groups, due to, e.g., differential trends or different business cycle responses (Bell and Blanchflower 2011). In our preferred specification, we therefore include demographic controls, a full set of unemployment–age dummies, age-specific trends and control-specific trends. We also present extensive placebo analyses, including placebo treatment dummies for the years just before the reform and placebo treatment dummies for the earlier economic downturn in 2002–2004.

Our main findings are as follows. First, the reform had a statistically significant and large negative effect on the number of young NEETs claiming welfare benefits of -24% in the age group 25–26 years of age, the only treatment group that passes all the placebo tests. Second, the reform had only a small and statistically insignificant effect on the total number of NEETs in this age group (and the other age groups). This means the reform pushed young individuals out of welfare, but did not increase the number of young individuals in employment or education. Third, we show that standard pre-reform placebo reform dummies may be insufficient to test for common time effects, as they may fail to reject differential business cycle responses. In light of these considerations, we argue that our estimation results of the WIJ effects on education enrollment do not pass all robustness tests.

Our paper is most closely related to studies that consider the effects of mandatory activation programs for young individuals. Blundell et al. (2004) use area-based piloting and age-related eligibility rules to identify the employment impact of a mandatory job search program targeted at individuals 18–24 years of age in the UK, the New Deal for Young People. They find that the program increased the probability to find employment by about five percentage points. Dahlberg et al. (2009) and Persson and Vikman (2014) analyze, respectively, the effect on the number of welfare recipients and entry and exit effects of a welfare reform in Sweden where city districts in Stockholm implemented mandatory activation programs at different rates.[1] The reform reduced welfare caseloads and increased employment rates of younger individuals, with the main effect being a reduction in the entry rate into welfare. Hernæs et al. (2017) exploit a geographically differentiated implementation of conditionality of welfare benefits for Norwegian youth and find that stricter conditionality reduces welfare claims and increases high school completion rates among 21-year-olds. These analyses suggest that the combination of welfare conditionality and welfare-to-work programs can reduce the number of NEETs and promote employment and enrollment in education among young individuals.

We make the following contributions to this literature. First, we show that the main effect of stricter conditionality combined with employment and training programs in the Netherlands was that it simply pushed young individuals out of welfare, not necessarily into employment. We argue that this is likely to be due to the state of the business cycle, as the reform clashed head on with the start of the Great Recession, during which it was hard for people, in particular young individuals, to find a job.[2] Second, we consider all potential outcome states, not only NEETs on welfare but also NEETs not on welfare, and the enrollment in education next to employment. Indeed, our analysis for young individuals in the treated group underlines the importance of studying the combined effects on the employment rate and the enrollment rate in education. Finally, we use an exceptionally long dataset that allows us to study and account for differential trends and test for differences in business cycle responses

[1] Both studies consider the effects for the total working-age population and for the younger age group 18–25 years old.

[2] Our analysis also contributes to a broader literature on the effect of active labor market policies (see, e.g., Kluve et al. 2016; Card et al. 2017). Card et al. (2017) conclude that active labor market programs are on average more likely to show positive impacts in a recession, which is not in line with our findings. In this respect, it is important to stress that two-thirds of their sample consists of training and job search assistance programs targeted at all age groups, which arguably can be expected to have a different effect from the work-oriented programs for young individuals that we consider. In the context of our analysis, however, both public and private sector employers may have less work opportunities available for young individuals during a recession. This renders it less likely for these type of programs to work during a recession compared to other active labor market policies.

across age groups in an earlier economic downturn. The latter turns out to be crucial, as standard pre-reform placebo treatment dummies may fail to reject the common time effects assumption.

The outline of the paper is as follows. Section 2 describes the institutional setting and the main features of the reform. Section 3 outlines the empirical methodology. Section 4 discusses the dataset and gives descriptive statistics. In Sect. 5, we then present graphical evidence, the estimation results and a large number of robustness checks. Finally, Sect. 6 discusses our findings and concludes.

2 Institutional setting and expected effects of the WIJ reform

Table 1 shows that there is considerable variation in the share of NEETs among the young across OECD countries, and the extent to which the share of NEETs has changed in the wake of the Great Recession. Panels A and B in Table 1 give indicators for individuals 20–24 years of age and individuals 25–29 years of age, respectively. The Netherlands has one of the lowest NEETs shares among OECD countries, in 2015 only 8.9% of 20–24-year-olds in the Netherlands were NEETs.[3] Over the period 2005–2015, there has been a moderate rise in the share of NEETs in the Netherlands. The low share of NEETs in the Netherlands is mirrored by the high share of 20–24-year-olds that are in education, as well as by the high share of 20–24-year-olds that are employed, whereas the share of unemployed 20–24-year-olds is relatively low.[4] Turning to individuals 25–29 years of age, the Netherlands also scores relatively favorable in terms of a low NEETs rate, a high enrollment rate in education, a high employment-to-population rate and a relatively low unemployment rate for this age group.

The reform we consider targets young individuals on welfare benefits. In the Netherlands, welfare benefits form a safety net that is provided by municipalities to support unemployed individuals who are not, or are no longer, entitled to other types of social insurance benefits like unemployment insurance. The vast majority of new welfare recipients consists of individuals with insufficient work history for entitlement to unemployment insurance.[5] Welfare benefits are means-tested and assets-tested.[6] The level of welfare benefits differs across household types and age groups. In 2008, just before the start of the WIJ reform, welfare benefits ranged from 220 euros per month for singles of 18–20 years of age to 1320 euros per month for couples with children (Ministry of Social Affairs and Employment 2008).

[3] In 2015, the only country in the OECD with a lower share of NEETs was Iceland (6.6%). Below, we will compare our results to studies for, e.g., Norway, Sweden and the UK. In this respect, it is relevant to note that Norway had a NEETs rate that was only slightly higher than in the Netherlands, the NEETs rate in Sweden was somewhat higher still, whereas the NEETs rate in the UK was considerably higher.

[4] The shares of individuals in education and individuals in employment add up to more than 100% because individuals in education can be employed, and employed individuals can also be in education.

[5] In 2014, only 22% of all new welfare recipients consisted of unemployed workers who exhausted their unemployment insurance benefits (UWV 2014).

[6] For single individuals, net worth should not exceed 5325 euros in 2008. For households with more persons, net worth should not exceed 10,650 euros in 2008.

Table 1 An international perspective on NEETs

	NEETs-to-population rate		Education-to-population rate	Employment-to-population rate	Unemployment-to-population rate
Year	2005	2015	2015	2015	2015
Panel A: Individuals 20–24 years of age					
Continental Europe					
Netherlands	8.1	8.8	57.7	69.4	6.7
Belgium	18.3	15.8	45.3	42.0	9.8
France	17.8	20.9	44.4	46.2	14.2
Germany	18.7	9.3	54.4	64.3	5.1
Scandinavia					
Denmark	8.3	12.4	59.1	63.4	7.6
Finland	13.0	18.3	47.8	52.5	14.7
Norway	9.6	10.2	42.1	66.6	5.8
Sweden	13.4	11.8	46.0	56.4	13.0
Anglo-Saxon countries					
Australia	11.6	13.1	44.5	71.5	7.3
Canada	14.4	14.4	41.6	64.7	8.3
UK	16.8	15.6	33.8	65.3	8.2
USA	15.5	15.8	38.5	64.1	6.5
OECD average	17.3	16.9	44.8	53.4	9.9

Table 1 continued

Year	NEETs-to-population rate 2005	NEETs-to-population rate 2015	Education-to-population rate 2015	Employment-to-population rate 2015	Unemployment-to-population rate 2015
Panel B: Individuals 25–29 years of age					
Continental Europe					
Netherlands	10.7	12.1	20.8	82.2	5.7
Belgium	17.7	20.2	8.5	74.4	11.0
France	19.8	23.4	8.5	72.1	12.5
Germany	21.2	12.8	20.8	77.9	5.0
Scandinavia					
Denmark	11.6	15.2	30.4	73.8	7.9
Finland	14.0	18.2	26.9	70.2	10.1
Norway	12.3	14.0	14.6	77.1	5.2
Sweden	10.0	10.8	25.1	75.6	8.7
Anglo-Saxon countries					
Australia	15.4	15.5	19.1	78.5	4.4
Canada	15.7	17.6	12.8	76.7	7.0
UK	16.6	16.2	12.7	79.4	5.0
USA	18.1	20.0	13.2	75.4	4.7
OECD average	19.0	19.3	16.3	73.5	9.4

Using data from OECD (2016a), OECD (2016b) and OECD (2016c). The education-to-population rate is the enrollment in education divided by the relevant age population. The unemployment-to-population rate is calculated as the unemployment rate multiplied by the labor force participation rate

The Work Investment Act for Young Individuals (In Dutch: *Wet Investeren in Jongeren*, WIJ) came into effect in October 2009. The reform followed from increased policy attention for NEETs and their welfare dependency. With about 70% of young welfare recipients being without a formal education degree, the idea was that work experience and training on the job would be key for the targeted group (Leenheer et al. 2011). The reform was designed before the start of the Great Recession, but implemented thereafter. Similar to, e.g., the New Deal for Young People in the UK, the aim of the WIJ was both to activate young welfare recipients and to foster their human capital formation.

The WIJ stipulated that for individuals below the age of 27, entitlement to welfare benefits became conditional on participation in a mandatory activation program. These mandatory programs typically consisted of public employment programs, apprenticeships and internships.[7] The introduction of these programs implied two important changes compared to the pre-WIJ period. First, participation in programs—which already was relatively high among young welfare recipients—became more intensive, more time demanding and became work-oriented. Second, non-participation into programs implied the full loss of welfare benefits, rather than temporary and partial benefit cuts in the pre-WIJ period—if case managers deemed this necessary.

Leenheer et al. (2011) show that almost half of the actual programs in the WIJ aimed at training skills to perform regular work, whereas 26% received support—e.g., wage subsidies—to be able to perform low-paid work with the continued receipt of welfare benefits.[8] Moreover, they find that 59% of the individuals perceive their program as employment while 56% perceive the program as education (individuals can give multiple answers). Leenheer et al. (2011) also show that the WIJ indeed intensified the effort and time spent on programs. In particular, on average programs in the WIJ took no less than half a year to be completed, whereas previous active labor market programs typically consisted of meetings with a job coach or job application training. In effect, the reform also reduced the discretionary room of case managers.

Regarding the timing of the effects, it is important to note that the WIJ implied an increase in the workload for municipalities.[9] The new law applied to all new entrants into welfare from October 2009 onward. However, municipalities were given an additional 9 months—until July 2010—to increase coverage of the WIJ to 100% of the stock of all welfare recipients. Hence, we may expect potential effects of the WIJ reform to show up in 2010 rather than 2009.

Since the WIJ reform implied the mandatory participation in programs, we hypothesize that it may have had both human capital effects and activation effects. The training

[7] The WIJ act prescribes that the programs could entail 'generally accepted labor, a provision aimed at employment, or a provision for education, training or social activation as well as support with work integration.'

[8] Any wage earnings were supplemented up to the level of welfare benefits when wage earnings fell short of the benefit level.

[9] Note that the additional workload may have varied across municipalities. For instance, apprenticeships, internships and public employment programs for young welfare recipients up to 23 were already provided in the (capital) city of Amsterdam (Board of Amsterdam 2009). The start of the WIJ thus implied an extension to 24–26-year-olds, together with the imposition of welfare conditionality for all young individuals below the age of 27. For other municipalities, however, the implementation of more intensive programs may have been more costly and may have taken more time to materialize.

and apprenticeship component of the reform aims to increase the human capital of the young welfare recipients, which is expected to increase future job market opportunities. The strict conditionality of welfare is expected to have an activation effect that may push welfare recipients out of welfare and potentially into employment or education. However, the WIJ reform clashed head on with the Great Recession, raising the question of whether stricter conditionality also increases employment during a recession. To investigate both human capital and activation effects, we will consider both employment and education for the age groups, 20–22, 23–24 and 25–26 years of age, as outcome measures. We expect that enrollment in education is a more likely alternative to welfare for the younger age groups, but probably less for the older age group.

Another relevant question to consider is whether the effect of the WIJ reform persists over time. For similar reforms, Blundell et al. (2004) and Dahlberg et al. (2009) find that most of the effect was in the beginning of the reform period, and then the effect diminishes in subsequent periods. As a potential explanation, Blundell et al. (2004) consider 'cleaning up the registers,' which has been observed in UK labor market reforms (Blundell et al. 2004, p. 594). We consider whether a similar mechanism may be at work in the Dutch case.

The WIJ law was abolished in January 2012. The government replaced the mandatory acceptance of activation programs with a 1-month mandatory job search period during which individuals did not receive welfare benefits. Faced with substantial budget cuts, the imposition of mandatory job search periods was generally considered a more efficient way to reduce welfare caseloads (Bolhaar et al. 2019). In that same year, another reform was implemented that changed eligibility for young welfare recipients. In particular, adult children living at home were no longer eligible to welfare benefits when they lived in a household in which first-degree relatives had sufficient income or assets (the 'household-income test'). To study to what extent these two additional reforms may affect outcomes since 2012, we also present treatment effects by individual treatment years, as well as the treatment effect on the probability of being an adult child living at home and the treatment effects for the subgroup of adult children living at home.

3 Empirical methodology

We use differences-in-differences (DD) to estimate the effects of the WIJ reform on a number of outcome variables. The reform was targeted at individuals up to 27 years of age and started in October 2009. Accordingly, we compare young individuals of 27 years and older to those below the age of 27, before and after 2009. To obtain consistent DD estimates of the reform, we need two key assumptions to hold: the Stable Unit Treatment Value Assumption (SUTVA) and the common trends assumption.

First, SUTVA implies that the WIJ reform does not yield spillover effects from the treatment group to the control group. If there are program effects that persist over time, one might be worried about 'treated' individuals reaching the age of 27 then becoming part of the control group. Likewise, we may be concerned that welfare recipients or their caseworkers might anticipate the 27th birthday of the welfare recipient, when participation in activation programs is no longer mandatory or that participation may continue after the 27th birthday of the welfare recipient. These spillover and anticipation effects may lead to underestimation of the WIJ effect. We therefore use DD with a sufficiently large sample of observations that are further away from the cutoff. At the same time, the SUTVA condition also explains why regression discontinuity (RD) analyses around the cutoff are only used as a robustness test and not as our preferred models.[10]

As argued earlier, the second key assumption is common time effects for the treatment and control group (in the absence of the reform). Our preferred treatment group consists of individuals 25–26 years of age and our preferred control group consists of individuals 27–28 years of age. For these groups, we will present eyeball tests on common trends in the next section, and placebo reform dummies in pre-reform years in the Results section. Our baseline model also considers the treatment effects for the treatment groups consisting of individuals 20–22 and 23–24 years of age, but we will show that changes in the enrollment in education over time and over the business cycle complicate the analysis for these groups.[11]

As potential outcome measures, we consider (i) the incidence of NEETs, defined as not being in employment or education[12], (ii) the incidence of NEETs on welfare, (iii) the incidence of NEETs not on welfare, (iv) the incidence of employment and (v) the incidence of education enrollment. The incidence rates of NEETs, employment and education enrollment sum to one. For all these outcome variables we estimate a linear probability model (Angrist and Pischke 2009), with y_{iat} as a dummy variable that is 1 if individual i in age group a is 'participating,' 'employed' or 'enrolled' in period t. In our preferred DD specification, we regress the outcome variable on a set of year fixed effects (α_t), age fixed effects (β_a), age-specific trends (with coefficients γ_a), an interaction term between age and the unemployment rate (u_t) with age-specific coefficients ϕ_a, a set of demographic controls X_i (gender and ethnicity) with coefficients μ_x, a set of demographic-control-specific trends with coefficients ψ_x, a treatment effect (DD_{gt}) for individuals in the treatment group g in a given year t in the post-reform period with coefficient δ_{gt}, and an error term ϵ_{iat}:

$$y_{iat} = \alpha_t + \beta_a + \gamma_a t + \phi_a u_t + X_i' \mu_x + X_{it}' \psi_x t + \delta_{gt} \mathrm{DD}_{gt} + \epsilon_{iat}. \qquad (1)$$

We are primarily interested in the treatment coefficients δ_{gt}. We include an interaction term between age and the unemployment rate to allow for different business cycle responses across age groups (Bell and Blanchflower 2011). Furthermore, we include age-specific and demographic-control-specific trends to allow for trend differences.[13]

In an extension to this model, we add placebo treatment dummies for the pre-reform years 2008 and 2009. As noted above, the coefficients on these placebo treatment dummies are informative about potential remaining differential time effects between

[10] To avoid underestimation of the WIJ effect, we also conduct 'donut' RD regressions where we leave out observations close to the threshold, see the Results section.

[11] Young individuals in the treatment groups have a choice of staying in education, while this is hardly a choice for individuals in the (older) control group.

[12] Similar to the OECD, we do not observe participation in training programs in our dataset.

[13] We have 10 years of pre-reform data to estimate the coefficients on these trends.

the treatment and control groups, for example, because of changes in group-specific trends or differences in business cycle responses not captured by the age-specific unemployment rate terms, and also about potential anticipation effects of the reform.

Finally, to allow for correlation in the error terms at a higher level than the individual and over time, we use cluster-robust standard errors (Bertrand et al. 2004; Donald and Lang 2007). The concern is that treatment is assigned according to age, whereas the different birth cohorts may also have been exposed to, e.g., different child care and education regimes. More generally, birth cohorts may have been exposed to different conditions before they enter the labor market and during the early part of their career, which in turn may result in time-varying birth-cohort specific differences in labor market outcomes. This essentially resembles an experimental design issue (Abadie et al. 2017). In our main specification, we therefore cluster the standard errors by month of birth. This results in 264 clusters in our base DD specification, which is deemed sufficiently large to use the large-sample properties of the estimator (Angrist and Pischke 2009). As a robustness check, we also consider clustering at year of birth. This does not affect the statistical significance of the results, as we will show in the Results section. As argued by Abadie et al. (2017), we will also assess the possibility that clustering at a too aggregate level leads to standard errors that are unnecessarily conservative. In line with this, we also consider clustering at the individual level. Again, this does not affect the statistical significance of the results, as we show in the Results section.[14]

4 Data

We use data from the Labor Market Panel (*Arbeidsmarktpanel*) of Statistics Netherlands (2015). The Labor Market Panel (LMP) is a large and rich household panel dataset, tracking 1.2 million individuals over the period 1999–2012.[15] The LMP is constructed by Statistics Netherlands for the CPB Netherlands Bureau for Economic Policy Analysis. The backbone of the LMP is the individuals observed in the Labor Force Survey in the years 1999–2012. To this data information is added on individuals and household characteristics like gender, ethnicity and household composition from the municipal registers (GBA) and income from wages, profits and several types of social benefits, as well as main source of income (SEC). We have data on the main source of income for the whole year and for the month of October (not for the other months). We use the years 1999–2009 as the pre-reform years, and 2010–2012 as the treatment years.

We consider three treatment groups: (i) individuals 25–26 years of age, (ii) individuals 23–24 years of age and (iii) individuals 20–22 years of age. Our main control group consists of individuals 27–28 years of age. As we will see below, the treatment

[14] The estimated standard errors are sometimes smaller and sometimes larger when using standard errors clustered at the individual level as opposed to using standard errors clustering at year of birth or month of birth level. The proverbial exception is the treatment effect on the enrollment rate in education for the youngest treatment group 20–22 years of age, which becomes statistically significantly different from zero when we use ’clustering’ at the individual level, see Table A.5 in the supplementary material.

[15] For a limited number of variables, not used in this study, the dataset also contains data for 2013.

group of individuals 25–26 years of age is the most similar to our main control group in terms of demographic characteristics, levels of the outcome variables and business cycle responses. The other two treatment groups with younger individuals are more likely to differ from the main control group, and hence, we have to be extra careful when interpreting the estimated treatment effects for these younger treatment groups.

The outcome variables are based on the social–economic classification (SEC) variable in the Labor Market Panel. The SEC variable classifies individuals according to their main source of income, where individuals in education are always classified as being in the state of education (even if their wage income is larger than their study grant) and individuals with profit income are always classified as being self-employed (even if their wage income exceeds their profit income). According to the SEC individuals can be in the following states: (1) employee, (2) owner of closely held company, (3) self-employed, (4) another type of employment, (5) on unemployment insurance, (6) on welfare benefits, (7) on disability or sickness benefits, (8) on retirement benefits, (9) on other social insurance, (10) in education with income, (11) in education without income, (12) without income. We count individuals in states (1)–(4) as employed, in states (10)–(11) as in education, and in states (5)–(9) and (12) as NEETs. Within the state of NEETs we count individuals in state (6) as NEETs on welfare and individuals in states (5), (7)–(9) and (12) as NEETs not on welfare. The outcome variables are averages for October each year, and the age variable is measured on the 1st of October of each year. As demographic control variables, we include gender and ethnicity (native/Western immigrant/non-Western immigrant).[16, 17]

Table 2 gives descriptive statistics for the respective treatment groups, along with the differences and normalized differences (for the demographic control variables) with the control group in the pre- and post-reform period. The differences in the demographic control variables gender and ethnicity are small for all treatment groups, in particular for the oldest treatment group with individuals 25–26 years of age. The same is true for the so-called normalized differences (mean differences divided by the square root of the sum of variances).[18] Furthermore, the differences in the demographic control variables hardly change from the pre- to the post-reform period. Hence, there is no indication of differential changes in the composition of the treatment and control group.[19]

Table 2 also gives descriptive statistics for the outcome variables. The NEETs rate on welfare in the oldest treatment group is very similar to the control group in the

[16] We only observe the highest obtained level of education in the year they are observed in the Labour Force Survey. Since the highest obtained level of education may change, in particular for young individuals, we do not use the highest obtained level of education as a control variable.

[17] We do not have information on the municipality in which individuals are living, so we cannot control for municipality fixed effects, or cluster standard errors by municipality. However, we do have information on the province in which individuals are living, and we use that information to study heterogeneous treatment effects, see the Results section.

[18] Imbens and Wooldridge (2009) argue that these normalized differences are an informative way to check if the treatment and control group have sufficient overlap in the covariates, and as a rule of thumb they suggest that when the normalized difference exceeds a value of .25, the regression results becomes sensitive to the functional form. The normalized differences for gender and ethnicity stay well below .25.

[19] Figure A.1 in the supplementary material plots the shares of the demographic control variables for the treatment and control group over time.

Table 2 Descriptive statistics treatment groups and control group

	Treatment Group (1999–2009)		Differences (treatment–control)		Normalized differences (treatment–control)	
	Mean	SD	1999–2009	2010–2012	1999–2009	2010–2012
Panel A: Treatment group 25–26						
Explanatory variables						
Female	0.506	0.500	– 0.006	0.000	– 0.009	0.000
Non-Western immigrant	0.102	0.302	0.001	– 0.004	0.003	– 0.008
Western immigrant	0.072	0.258	– 0.003	– 0.002	– 0.007	– 0.005
Dependent variables						
NEETS rate on welfare	0.025	0.155	– 0.001	– 0.004		
NEETs rate not on welfare	0.088	0.283	– 0.011	0.005		
Total NEETs rate	0.112	0.316	– 0.012	0.001		
Employment rate	0.818	0.386	– 0.036	– 0.065		
Enrollment rate education	0.069	0.254	0.048	0.063		
Panel B: Treatment group 23–24						
Explanatory variables						
Female	0.499	0.500	– 0.013	– 0.002	– 0.018	– 0.002
Non-Western immigrant	0.101	0.302	0.001	– 0.004	0.002	– 0.009
Western immigrant	0.069	0.253	– 0.005	– 0.004	– 0.015	– 0.010

Table 2 continued

	Treatment Group (1999–2009)		Differences (treatment–control)		Normalized differences (treatment–control)	
	Mean	SD	1999–2009	2010–2012	1999–2009	2010–2012
Dependent variables						
NEETS rate on welfare	0.022	0.146	– 0.004	– 0.007		
NEETs rate not on welfare	0.078	0.268	– 0.021	0.004		
Total NEETs rate	0.099	0.299	– 0.025	– 0.003		
Employment rate	0.714	0.452	– 0.140	– 0.212		
Enrollment rate education	0.187	0.390	0.165	0.215		
Panel C: Treatment group 20–22						
Explanatory variables						
Female	0.492	0.500	– 0.020	– 0.004	– 0.029	– 0.005
Non-Western immigrant	0.101	0.301	0.001	0.007	0.001	0.016
Western immigrant	0.067	0.249	– 0.008	– 0.006	– 0.021	– 0.017
Dependent variables						
NEETS rate on welfare	0.014	0.118	– 0.012	– 0.013		
NEETs rate not on welfare	0.063	0.243	– 0.035	– 0.012		
Total NEETs rate	0.077	0.267	– 0.047	– 0.025		
Employment rate	0.491	0.500	– 0.363	– 0.459		
Enrollment rate education	0.432	0.495	0.410	0.484		

Own calculations using the Labor Market Panel (Statistics Netherlands). Treatment groups: individuals 20–22, 23–24 and 25–26 years of age. Control group: individuals 27–28 years of age. Observations 1999-2012: treatment group 20–22: 582,364, treatment group 23–24: 375,182, treatment group 25–26: 376,083, control group 27-28: 391,627. Normalized differences are mean differences divided by the square root of the sum of the variances (see Imbens and Wooldridge 2009) pre-reform period, but drops relative to the control group in the post-reform period, suggesting a negative treatment effect on this outcome variable for this treatment group. The pre-reform differences in the NEETs rate on welfare are larger for the younger treatment groups, in particular for the youngest treatment group. Also for these groups, the difference becomes larger in the post-reform period, suggesting a negative treatment effect for the NEETs rate on welfare benefits for these groups. The NEETs rate not on welfare is also quite similar for the older treatment group and the control group before the reform, though somewhat lower for the treatment group than the control group, and lower still for the younger treatment groups. After the reform, the NEETs rate not on welfare move closer to the control group, suggesting a positive treatment effect on this outcome variable. The total NEETs rate again is quite similar for the oldest treatment group and the control group before the reform, though again somewhat lower for this treatment group, and lower still for the younger treatment groups. After the reform, the total NEETs rate of the treatment groups moves closer to the control group, which suggests a positive treatment effect for the total NEETs rate. The employment rate is lower for the treatment groups than the control group in the pre-reform period, and the difference becomes more negative in the post-reform period, suggesting a counterintuitive negative treatment effect on the employment rate. Finally, the enrollment rate in education shows the mirror image of the employment rate. The enrollment is higher in the treatment groups than in the control group in the pre-reform period, and this difference also becomes bigger in the post-reform period, suggesting a positive treatment effect on the enrollment in education. However, these simple treatment effects do not account for differential trends between the treatment and control groups. These differential trends will turn out to be important for some outcome variables, in particular for the younger treatment groups, in the empirical analysis below.

To gauge the validity of the common trends assumption and the presence of reform effects, Fig. 1 presents the developments in the outcome variables for the treatment and control group before and after the treatment. The solid black line denotes the control group of individuals 27–28 years of age, whereas the red, blue and green lines denote the treatment groups of 25–26, 23–24 and 20–22 years of age, respectively. The dotted lines denote the difference between the respective treatment groups and the control group. Figure 1a shows that the NEETs rate on welfare moves very much in tandem for the treatment groups 23–24 and 25–26 years of age and the control group in the pre-reform period, lending credence to the common trends assumption that is needed for identification of the reform effects. After the WIJ reform, we observe a clear negative treatment effect in 2010, which subsequently becomes smaller in 2011 and then remains roughly constant in 2012. For the youngest treatment group 20–22 years of age, the NEETs rate on welfare also shows a quite similar pattern to the control group prior to the reform, but there is no apparent treatment dip in 2010 (although the control group moves ‘up,’ presumably due to the Great Recession, and the treatment group 20–22 does not) nor is there an apparent recovery in the NEETs rate in 2011 or 2012 for this treatment group relative to the control group. Figure 1b–e makes clear that there are apparent trend differences between the treatment and control group for the other outcome variables, also for the treatment group 25–26 years of age. The main culprit here is the difference in trends in the enrollment in education by age groups,

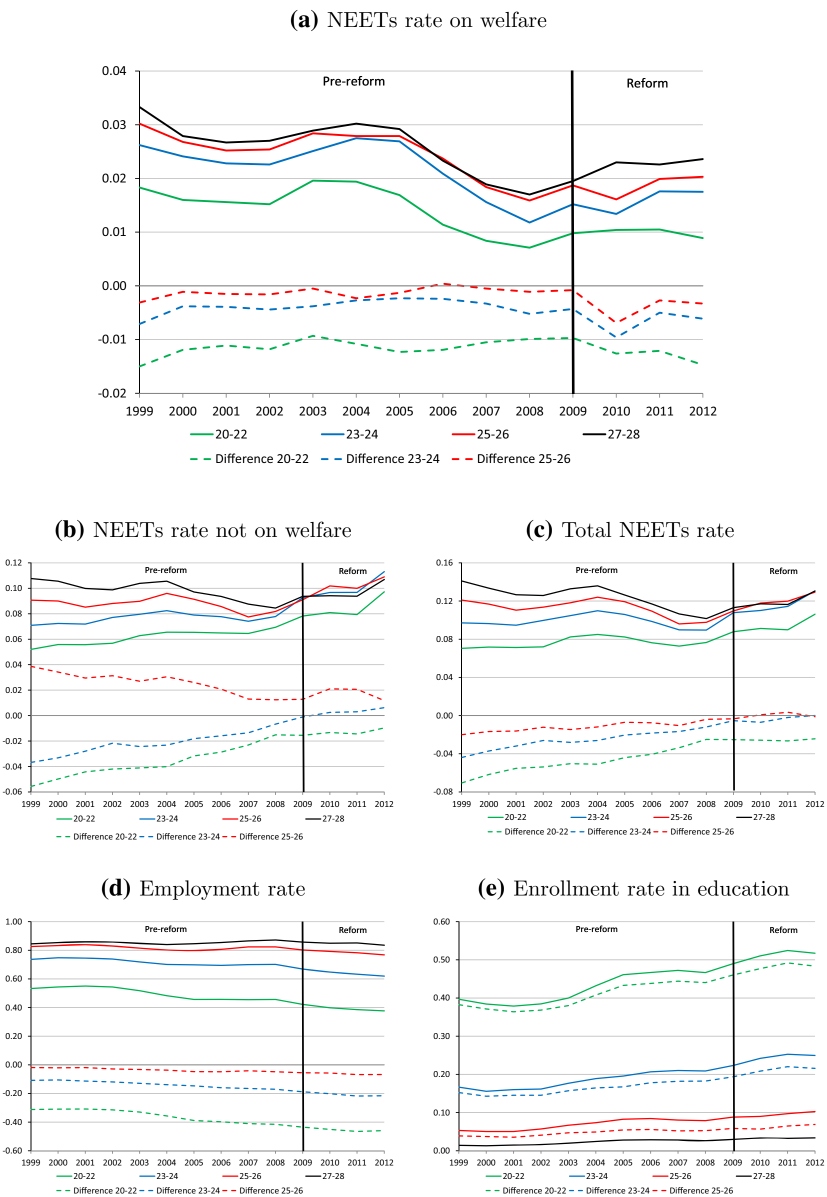

Fig. 1 Means outcome variables treatment and control groups: 1999–2012. *Notes*: Own calculations using the Labor Market Panel (Statistics Netherlands). The solid black line denotes the control group of individuals 27–28 years of age, the red lines denote the treatment group 25–26 years of age, the blue lines denote the treatment group 23–24 years of age and the green lines denote the treatment group 20–22 years of age. The dotted lines denote the difference between the treatment group and the control group. NEETs rates are individuals not in employment or education relative to the relevant age population, employment rates are individuals in employment relative to the relevant age population and enrollment rates in education are individuals in education relative to the relevant age population

see Fig. 1e. Hence, accounting for differential trends will be important to isolate the treatment effect of the reform for these outcome variables.

5 Results

5.1 Main results

Table 3 gives the main regression results. In all specifications, we use a single treatment dummy per treatment group for the post-reform years 2010–2012.[20] First consider the results for the treatment group 25–26 years of age in Panel A, the group that is the most similar to the control group in observable characteristics and means of the outcome variables. Column (1) shows the results of the basic DD setup, where we only include year dummies, a group dummy for each individual age group and a treatment dummy for the age group 25–26. This setup suggests a negative and statistically significant treatment effect of −0.30 percentage points on the NEETs rate on welfare. In column (2), we add demographic controls. Consistent with the observation that there were negligible compositional changes in these characteristics, this hardly affects the estimated treatment effect. In column (3), we add interaction terms for age and the national unemployment rate, to allow for a potential different business cycle response by age. Again, this does not substantially affect the estimated treatment effect for the NEETs rate on welfare. In column (4), we then also allow for age-specific trends, and this leads to a somewhat larger treatment effect in absolute terms (more negative) of −0.44 percentage points. Finally, column (5), our richest and preferred specification, shows that the inclusion of demographic-control specific trends gives a treatment effect that is very similar to the treatment effect in column (4). The treatment effect in column (5) of −0.46 percentage points also suggests a sizable negative treatment effect on the NEETs rate on welfare of −24% relative to a baseline of 1.9 percentage points in the last pre-reform year (2009).

As noted earlier, accounting for trend differences between the treatment and control group is important for the other outcome variables in Table 3. In particular, we find rather similar treatment effects for the specification in columns (1)–(3),[21] but allowing for differential trends in age in column (4) has an important impact on the treatment effects on the employment rate and the enrollment rate in education.[22] Our preferred specification is in column (5), with results suggesting a positive and statistically significant treatment effect on the NEETs rate not on welfare, but no effect on the total NEETs rate. Also, there appears to be no effect on the employment rate or the enrollment rate in education.

Hence, the reform seems to have pushed or kept the treated individuals in this age group out of welfare by stricter conditionality without higher employment and/or enrollment in education. Previous studies of related reforms show that these reforms

[20] Full regression results can be found in Table A.1 in the supplementary material.

[21] Although the ‘treatment effect’ for the employment rate and enrollment rate in education do vary in absolute size over the different specifications in columns (1)–(3).

[22] The inclusion of demographic-control-specific trends in column (5) again hardly affects the results when compared to column (4).

Table 3 Differences-in-differences: base regression results

	(1)	(2)	(3)	(4)	(5)
Panel A: Treatment group 25–26					
NEETs rate on welfare	−0.0030∗∗∗	−0.0028∗∗∗	−0.0032∗∗∗	−0.0044∗∗∗	−0.0046∗∗∗
	(0.0010)	(0.0010)	(0.0011)	(0.0014)	(0.0014)
NEETs rate not on welfare	0.0159∗∗∗	0.0161∗∗∗	0.0137∗∗∗	0.0061∗∗∗	0.0060∗∗∗
	(0.0017)	(0.0016)	(0.0019)	(0.0023)	(0.0023)
NEETs rate	0.0129∗∗∗	0.0133∗∗∗	0.0105∗∗∗	0.0017	0.0014
	(0.0019)	(0.0019)	(0.0020)	(0.0027)	(0.0028)
Employment rate	−0.0298∗∗∗	−0.0303∗∗∗	−0.0213∗∗∗	−0.0027	−0.0027
	(0.0032)	(0.0031)	(0.0035)	(0.0036)	(0.0036)
Enrollment rate in education	0.0169∗∗∗	0.0170∗∗∗	0.0108∗∗∗	0.0010	0.0013
	(0.0027)	(0.0027)	(0.0030)	(0.0032)	(0.0032)
Panel B: Treatment group 23–24					
NEETs rate on welfare	−0.0028∗∗∗	−0.0027∗∗∗	−0.0037∗∗∗	−0.0039∗∗∗	−0.0040∗∗∗
	(0.0011)	(0.0010)	(0.0011)	(0.0015)	(0.0015)
NEETs rate not on welfare	0.0248∗∗∗	0.0248∗∗∗	0.0209∗∗∗	0.0022	0.0022
	(0.0020)	(0.0019)	(0.0020)	(0.0026)	(0.0026)
NEETs rate	0.0220∗∗∗	0.0220∗∗∗	0.0172∗∗∗	−0.0016	−0.0017
	(0.0023)	(0.0022)	(0.0023)	(0.0030)	(0.0030)
Employment rate	−0.0728∗∗∗	−0.0730∗∗∗	−0.0598∗∗∗	−0.0145∗∗∗	−0.0145∗∗∗
	(0.0041)	(0.0040)	(0.0043)	(0.0047)	(0.0047)
Enrollment rate in education	0.0508∗∗∗	0.0509∗∗∗	0.0426∗∗∗	0.0161∗∗∗	0.0163∗∗∗
	(0.0036)	(0.0036)	(0.0040)	(0.0049)	(0.0048)

Table 3 continued

	(1)	(2)	(3)	(4)	(5)
Panel C: Treatment group 20–22					
NEETs rate on welfare	−0.0017*	−0.0024**	−0.0025**	−0.0043***	−0.0040***
	(0.0010)	(0.0010)	(0.0011)	(0.0014)	(0.0014)
NEETs rate not on welfare	0.0232***	0.0219***	0.0175***	−0.0061***	−0.0057***
	(0.0018)	(0.0017)	(0.0019)	(0.0021)	(0.0021)
NEETs rate	0.0215***	0.0195***	0.0151***	−0.0104***	−0.0097***
	(0.0022)	(0.0020)	(0.0023)	(0.0027)	(0.0027)
Employment rate	−0.0968***	−0.0946***	−0.0720***	0.0030	0.0027
	(0.0046)	(0.0045)	(0.0051)	(0.0054)	(0.0054)
Enrollment rate in education	0.0753***	0.0751***	0.0570***	0.0075	0.0070
	(0.0044)	(0.0044)	(0.0051)	(0.0055)	(0.0056)
Demographic controls	NO	YES	YES	YES	YES
Unemployment–age dummies	NO	NO	YES	YES	YES
Age-specific trends	NO	NO	NO	YES	YES
Control-specific trends	NO	NO	NO	NO	YES
Observations	1,725,256	1,725,256	1,725,256	1,725,256	1,725,256
Clusters	264	264	264	264	264

*Denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. Sample period 1999–2012. Treatment groups: individuals 20–22, 23–24 and 25–26 years of age. Control group: individuals 27–28 years of age. Cluster-robust standard errors in parentheses, clustered by month of birth (264 clusters), All specifications include age and year fixed effects. See Table A.1 in the supplementary material for the full regression results do indeed push young people out of welfare, but that this typically also translates into an increase in employment or enrollment in education (Blundell et al. 2004; Dahlberg et al. 2009; Persson and Vikman 2014; Hernæs et al. 2017). Since these other reforms were implemented at more favorable business cycle conditions, so it appears that the young people affected by the Dutch reform are struggling to find work in the context of the Great Recession. We will discuss the possible explanations of our results more extensively in the Discussion and conclusion section.

Panels B and C give the results for the younger age groups. We focus on our preferred specification controlling for differential trends in column (5). Similar to the age group of 25–26 years of age, negative and statistically significant treatments effects on the NEETs rate on welfare of about −0.4 percentage points are found for the age groups of 23–24 and 20–22 years of age. We find no effect on the overall NEETs rate for those aged 23–24, but a large statistically significant decrease for those aged 20–22. For the individuals aged 23–24, the suggested treatment effect on the enrollment rate in education is positive and the treatment effect on the employment rate is negative. An optimistic interpretation of this result is that this treatment group was stimulated to remain in (or return to) education following the WIJ reform. As we will show in our robustness tests, however, this finding should be interpreted with the appropriate care.

Turning to the placebo analyses, first consider the results in Table 4. In this table, we take specification (5) of Table 3 and add placebo treatment dummies for the years 2008 and 2009. For each treatment group, we also split the single treatment dummy (for 2010–2012) into single-year treatment dummies for 2010, 2011 and 2012. With this specification, we can both test for common time effects as well as for anticipation effects and the evolution of the treatment effect of the WIJ reform over time. From the table, the general picture that emerges is that the placebo dummies are small and statistically insignificant. It is only for the NEETs rate on welfare in the treatment group of 23–24 years of age that we find a significant placebo dummy for 2008. Another finding is that the treatment effects on the NEETs rate on welfare for 2011 and 2012 are often smaller than for 2010, which is consistent with the pattern in Fig. 1. Hence, most of the treatment effect seems to be confined to the first period of the reform, which is in line with the ‘cleaning up the registers’ mechanism observed in Blundell et al. (2004). Also for the NEETs rate not on welfare, most of the effect appears to be in 2010, after which the effect becomes smaller again. Finally, it should be noted that there is still no statistically significant treatment effect for the total NEETs rate, the employment rate nor the enrollment rate in education when we consider single-year treatment dummies.

We also exploit the richness of our data by conducting additional placebo analyses that capture the economic downturn in 2002–2004 in the Netherlands—see Table 5 for the estimation results. The general idea here is to detect possible differences in responses to the business cycle between the treatment groups and the control group not accounted for by the interaction terms between the unemployment rate and the individual age dummies. If such responses are different, this casts doubt on the common-time effects assumption underlying our DD approach. As the table shows, we do not find statistically significant placebo treatment effects for the treatment group 25–26 years of age. Hence, this lends support to the assumption of common time effects for this

Table 4 Differences-in-differences: pre-reform placebo's and annual treatment effects

	(1) NEETs rate on welfare	(2) NEETs rate not on welfare	(3) Total NEETs rate	(4) Employment rate	(5) Enrollment rate in education
Panel A: Treatment group 25–26					
Placebo 2008	−0.0023	0.0046	0.0022	−0.0037	0.0014
	(0.0022)	(0.0038)	(0.0044)	(0.0053)	(0.0044)
Placebo 2009	−0.0022	0.0027	0.0005	−0.0032	0.0027
	(0.0023)	(0.0039)	(0.0048)	(0.0056)	(0.0048)
Treatment 2010	−0.0086***	0.0114 ***	0.0028	−0.0019	−0.0009
	(0.0021)	(0.0036)	(0.0045)	(0.0057)	(0.0047)
Treatment 2011	−0.0045 *	0.0097 **	0.0051	−0.0102	0.0050
	(0.0024)	(0.0038)	(0.0047)	(0.0062)	(0.0051)
Treatment 2012	−0.0052 **	0.0039	−0.0013	−0.0038	0.0052
	(0.0022)	(0.0040)	(0.0047)	(0.0059)	(0.0054)
Panel B: Treatment group 23–24					
Placebo 2008	−0.0034 **	0.0006	−0.0029	−0.0030	0.0058
	(0.0017)	(0.0031)	(0.0035)	(0.0051)	(0.0052)
Placebo 2009	−0.0026	0.0044	0.0018	−0.0109	0.0090
	(0.0020)	(0.0037)	(0.0043)	(0.0070)	(0.0073)
Treatment 2010	−0.0082 ***	0.0056	−0.0026	−0.0171 **	0.0197 **
	(0.0024)	(0.0035)	(0.0046)	(0.0076)	(0.0078)
Treatment 2011	−0.0041	0.0030	−0.0011	−0.0273 ***	0.0284 ***
	(0.0027)	(0.0041)	(0.0050)	(0.0068)	(0.0077)
Treatment 2012	−0.0059 **	0.0038	−0.0021	−0.0154 **	0.0175 **
	(0.0026)	(0.0042)	(0.0049)	(0.0076)	(0.0076)

Table 4 continued

	(1) NEETs rate on welfare	(2) NEETs rate not on welfare	(3) Total NEETs rate	(4) Employment rate	(5) Enrollment rate in education
Panel C: Treatment group 20–22					
Placebo 2008	−0.0003	0.0015	0.0012	−0.0035	0.0023
	(0.0015)	(0.0026)	(0.0031)	(0.0050)	(0.0051)
Placebo 2009	−0.0004	−0.0019	−0.0023	−0.0077	0.0099
	(0.0017)	(0.0030)	(0.0037)	(0.0065)	(0.0063)
Treatment 2010	−0.0035 *	−0.0032	−0.0067 *	−0.0065	0.0132 *
	(0.0019)	(0.0030)	(0.0037)	(0.0077)	(0.0076)
Treatment 2011	−0.0034	−0.0087 ***	−0.0121 ***	−0.0080	0.0201 **
	(0.0022)	(0.0031)	(0.0040)	(0.0083)	(0.0082)
Treatment 2012	−0.0062 ***	−0.0067 **	−0.0129 ***	0.0128	0.0002
	(0.0022)	(0.0034)	(0.0042)	(0.0091)	(0.0090)
Demographic controls	YES	YES	YES	YES	YES
Unemployment–age dummies	YES	YES	YES	YES	YES
Age-specific trends	YES	YES	YES	YES	YES
Control-specific trends	YES	YES	YES	YES	YES
Observations	1,725,256	1,725,256	1,725,256	1,725,256	1,725,256
Clusters	264	264	264	264	264

*Denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. Sample period 1999–2012. Treatment groups: individuals 20–22, 23–24 and 25–26 years of age. Control group: individuals 27–28 years of age. Cluster-robust standard errors in parentheses, clustered by month of birth (264 clusters). All specifications include demographic controls, unemployment–age interaction terms, age-specific trends and control-specific trends

Table 5 Differences-in-differences: placebo treatment dummy economic downturn 2002-2004

	(1) NEETs rate on welfare	(2) NEETs rate not on welfare	(3) Total NEETs rate	(4) Employment rate	(5) Enrollment rate in education
Panel A: Treatment group 25–26					
Treatment 2010–2012	−0.0045***	0.0060***	0.0015	−0.0023	0.0008
	(0.0014)	(0.0023)	(0.0027)	(0.0035)	(0.0032)
Placebo 2002–2004	0.0001	0.0003	0.0004	0.0018	−0.0023
	(0.0012)	(0.0021)	(0.0026)	(0.0030)	(0.0020)
Panel B: Treatment group 23–24					
Treatment 2010–2012	−0.0039***	0.0025	−0.0015	−0.0139***	0.0153***
	(0.0015)	(0.0026)	(0.0030)	(0.0047)	(0.0049)
Placebo 2002–2004	0.0004	0.0027	0.0032	0.0047	−0.0078**
	(0.0015)	(0.0023)	(0.0028)	(0.0038)	(0.0032)
Panel C: Treatment group 20–22					
Treatment 2010–2012	−0.0038***	−0.0058***	−0.0096***	0.0046	0.0050
	(0.0014)	(0.0021)	(0.0027)	(0.0055)	(0.0056)
Placebo 2002–2004	0.0021**	−0.0006	0.0015	0.0239***	−0.0255***
	(0.0010)	(0.0019)	(0.0022)	(0.0039)	(0.0035)
Demographic controls	YES	YES	YES	YES	YES
Unemployment–age dummies	YES	YES	YES	YES	YES
Age-specific trends	YES	YES	YES	YES	YES
Control-specific trends	YES	YES	YES	YES	YES
Observations	1,725,256	1,725,256	1,725,256	1,725,256	1,725,256
Clusters	264	264	264	264	264

*Denotes significance at the 10% level, ** at the 5% level and *** at the 1% level. Sample period 1999–2012. Treatment groups: 20–22, 23–24 and 25–26 years of age. Control group: 27–28 years of age. Cluster-robust standard errors in parentheses, clustered by month of birth (264 clusters). All specifications include demographic controls, unemployment rate-age interactions, age-specific trends and control-specific trends treatment group and the control group. However, we do find statistically significant placebo effects for the two youngest treatment groups for the years 2002–2004. This casts doubt on the treatment effects on the employment rate and education enrollment rate we find for these groups during the WIJ reform period, which may also reflect different time effects for the control group and these younger treatment groups.

5.2 Robustness and heterogeneity analysis

The supplementary material presents some additional robustness checks. First, one may worry that the reform created spillovers for the control group via, e.g., the job-finding rate (Blundell et al. 2004; Gautier et al. 2018). In Table A.2, we address this concern by using individuals with 29–30 years of age as an alternative control group, and introduce 'treatment dummies' for our main control group of individuals 27–28 years of age. We then find rather similar treatment effects as in the base specification for the treatment groups 20–22, 23–24 and 25–26 years of age, and no statistically significant placebo treatment effects for our base control group (individuals 27–28 years of age).[23] Second, Table A.3 addresses the concern that treatment effects may persist as individuals age into the control group, another type of spillover effect that may bias our estimates. Here, we use individuals 30–31 years of age as the control group, as these were never in the treatment group during the WIJ reform period, and introduce 'treatment dummies' for individuals 27–29 years of age. Again, the results for the treatment groups 20–22, 23–24 and 25–26 years of age are (quite[24]) similar to the base specification, and the treatment effects for individuals 27–29 years of age are statistically insignificant. Third, Table A.4 shows that we obtain similar results when we narrow the treatment group down to individuals 26 years of age and the control group to individuals 27 years of age. Finally, Table A.5 shows that the different levels of clustering (at the individual level, by month of birth or by year of birth, respectively) (virtually[25]) does not affect the statistical significance of the results.

Table A.6 considers to what extent the changes in the stocks are driven by changes in the respective entry and exit rates.[26] When focussing on the older treatment group of individuals 25–26 years of age for which the baseline results turned out to be robust, we find that the effect on the NEETs rate on welfare runs entirely via an increased exit rate, with no effect on the entry rate, consistent with the mechanism of 'cleaning up the registers' (Blundell et al. 2004). Vice versa, we find that the effect on the NEETs rate

[23] The proverbial exception is the employment rate, which is 'borderline' significant at the 10% level.

[24] Of course, the control group becomes increasingly dissimilar to the main treatment groups, which results in some treatment effects (total NEETs rate and employment rate) for the age group 25–26 to become borderline significant at the 10% level, though with a counter-intuitive sign, also suggesting this is not a causal effect.

[25] As noted before, the proverbial exception is the treatment effect on the enrollment rate in education for the youngest treatment group 20–22 years of age, which becomes statistically significantly different from zero when we use 'clustering' at the individual level, see Table A.5 in the supplementary material.

[26] Specifically, for entry the dependent variable equals 1 when, for each state, the current state is 1 and the previous state was a different state, and zero otherwise. For exit, the dependent variable equals 1 when, for each state, the current state is a different state than the previous state, and the previous state is 1, and zero otherwise. We present results for our most elaborate specification, including demographic controls, unemployment–age interaction terms, age-specific trends and demographic-control-specific trends.

not on welfare is mainly due to an increase in the entry rate (although this coefficient is only statistically significant at the 10% level), with no effect on the exit rate. The exit and entry rates for the total NEETs rate, employment rate and the enrollment rate in education are statistically insignificant and typically small.

The supplementary material section also presents the outcomes for selected other outcome variables and by subgroups. In light of our earlier placebo results, we now focus on the treatment group of 25–26 years of age. Table A.7 shows that the effects of the WIJ on the enrollment rate in unemployment insurance (UI) and disability insurance (DI) are insignificant for this group. Next, Table A.8 gives the treatment effect on being in a particular household type. Distinguishing between adult children living at home, childless singles, single parents and couples, we do not find any statistically significant treatment effects. Given that being in a particular household type seems largely exogenous to the treatment, Table A.9 then studies the treatment effects by household type. Focusing again on the treatment group of 25–26-year-olds, the largest drop in the NEETs rate on welfare in absolute terms is for adult children living at home and single parents, −1.0 and −7.0 percentage points, respectively. In percentage terms however, the drop for single parents is −22% (relative to the 2009 level), which is comparable to the average treatment effect over all household types. But for adult children living at home, it is −45% (relative to the 2009 level), which can be explained by the additional reform in 2012, when adult children living at home were no longer eligible to welfare benefits when they lived in a household in which first-degree relatives had sufficient income or assets (see Sect. 2). The effect for childless singles is comparable to the average over all household types, whereas the effect for couples is close to zero. In line with the base results where we pool all household types, the NEETs rate not on welfare increases for all household types. The treatment effects for the other outcome variables are typically not statistically significant.

In addition to stratifying with respect to household types, Table A.10 gives the results by gender and ethnicity. The treatment effects for males and females are similar. The treatment effects for natives are somewhat smaller than the base results, whereas the results for immigrants are larger in absolute terms. But in percentage terms, the effects are much more comparable to the average, −29% for natives and −22% for immigrants for the NEETs rate on welfare (and a statistically insignificant effect on the total NEETs rate). Finally, Table A.11 considers the treatment effects for provinces that had a relatively low or a relatively high pre-reform unemployment rate. The treatment effect appears to be smaller (about half) in the provinces which had a lower pre-reform unemployment rate. However, in percentage terms the drop is almost the same in regions with low and high pre-reform unemployment rates, 25% (relative to the baseline in 2009) for low unemployment rate regions and 24% for high unemployment rate regions.

Finally, we also consider the more local treatment effect of the WIJ reform by considering outcomes around the cutoff age of 27, using regression discontinuity

(RD). The empirical specification is provided in the supplementary material.[27,28] Table A.12 in the supplementary material gives the RD regression results for the pre- and post-reform period, in Panels A and B, respectively. We find a small positive but statistically insignificant pre-reform treatment effect for the NEETs rate on welfare, the NEETs rate not on welfare and the total NEETs rate. In addition, the treatment effect on both the employment rate and the education enrollment rate is negative and statistically insignificant. For the post-reform period, we find a small but now negative treatment effect for the NEETs rate on welfare, though not statistically significant, a bigger positive and statistically significant treatment effect for the NEETs rate not on welfare (at the 5% level), and a small positive treatment effect for the total NEETs rate that is similar to the effect in the pre-reform period. Furthermore, the post-reform treatment effect is somewhat larger for the employment rate and somewhat smaller for the enrollment rate in education. Panel C of Table A.12 then gives the coefficient on a 'difference-in-discontinuity' (DRD) dummy, which is very close to the difference in the discontinuity between the pre- and post-reform period. This DRD dummy gives our preferred RD results because we are interested in the discontinuity that arises after the reform was implemented. The results of these preferred RD results are very similar to the DD analysis. There is a negative treatment effect on the NEETs rate on welfare, statistically significant at the 10% level, a positive treatment effect on the NEETs rate not on welfare and essentially no effect on the total NEETs rate (and the treatment effects for the employment rate and enrollment rate in education are insignificant).

Using the RD framework, we also investigated potential anticipation and persistence effects close to and shortly after reaching the age of 27. Specifically, Table A.13 gives results of a so-called donut RD (and DRD) analyses where we drop observations of individuals 3 months on either side of the cutoff.[29] These results are very similar to the base RD and DRD specifications (and even closer to the DD results than the base RD and DRD analysis).[30,31]

6 Discussion and conclusion

In this paper, we have studied the labor market effects of a Dutch mandatory activation program (WIJ) for welfare recipients up to 26 years of age in the Netherlands. We used

[27] Figure A.2 in the supplementary material shows the NEETs rate on welfare, the NEETs rate not on welfare and the total NEETs rate by month of birth of 25–28-year-olds, relative to the discontinuity—both for the pre-reform period (2007–2009, left panels) and post-reform period (2010–2012, right panels). Similar plots for the employment rate and the enrollment rate in education are given in Figure A.3.

[28] Figure A.4 shows that there is no manipulation in the running variable (age of the child in months).

[29] For an analysis of the implementation of donut RD designs, see, e.g., Barreca et al. (2011) or Barreca et al. (2016).

[30] Table A.13 gives the difference-in-discontinuity results for entry and exit probabilities. The DRD analysis also suggests a positive effect on the exit probability from welfare, in line with the DD analysis, significant at the 10 percent level. At the same time, however, it also suggests a negative effect on the entry probability into welfare, significant at the 10 percent level. Hence, the DRD analysis suggests there may have been some 'threat effect' of the WIJ reform.

[31] Cammeraat et al. (2017), the working paper version of this study, shows that our regression discontinuity results are robust with respect to a large number of additional robustness checks.

differences-in-differences and a long and rich administrative dataset to uncover the effect of the WIJ reform on the NEETs rate on welfare, the NEETs rate not on welfare, the total NEETs rate, the employment rate and the enrollment rate in education. We considered the separate treatment effects on individuals 20–22, 23–24 and 25–26 years of age, using individuals 27–28 years of age as the main control group. An extensive number of placebo tests suggests that we can interpret the effects on the group 25–26 years of age as causal, whereas the assumption of common-time effects seems questionable for the younger treatment groups. Focusing on the results for the group 25–26 years of age, we find that the reform reduced the number of NEETS on welfare with a substantial 24%, with most of the effect in the first year of the reform. At the same time, the reform did not reduce the overall NEETs rate, neither did it increase the employment rate nor did it increase the enrollment rate in education. The reform mainly pushed individuals out of welfare, where most of the effect appears to have come from an increase in the exit rate from welfare rather than a decrease in the entry rate into welfare.

Part of our findings are in line with previous studies on mandatory activation programs targeted at young individuals. Consistent with Blundell et al. (2004), Dahlberg et al. (2009), Persson and Vikman (2014) and Hernæs et al. (2017), we find a substantial negative effect on the number of young individuals on welfare. This effect is mostly geared by increased exit out of welfare—see also Blundell et al. (2004).[32] Similar to Blundell et al. (2004) and Dahlberg et al. (2009), we also find that most of the effect was in the beginning of the reform period, and then the effect diminishes in subsequent periods. As a potential explanation, Blundell et al. (2004) consider ‘cleaning up the registers,’ which have been noted of previous UK labor market reforms (Blundell et al. 2004, p. 594). A similar mechanism could be at work in the Dutch case. Also consistent with Blundell et al. (2004), we find no evidence of spillover effects to other groups. That is, we find no effects on the group of individuals that is 1 or 2 years older than the treatment group.

Part of our findings are also at odds with the literature. While mandatory programs for young individuals are usually associated with increased employment (Blundell et al. 2004; Dahlberg et al. 2009; Persson and Vikman 2014) or increased education enrollment (Hernæs et al. 2017), we find no evidence in this direction. One potential explanation for this is that we consider a country where the NEETs rate is relatively low, see Table 1. The findings of Hernæs et al. (2017) for Norway, a country with comparable level of NEETs rates, however, point at substantial program effects on employment and education enrollment. Therefore, a more likely explanation for the absence of program effects on employment is that the reform clashed head on with the Great Recession that started just prior to the start of the WIJ reform. The Great Recession made it inherently more difficult for individuals, especially young individuals, to find employment or to switch from apprenticeships and internships to regular employment contracts. Indeed, Fig. 2 shows the steep drop in the vacancy-to-unemployment ratio during the reform period. This was quite different for the reforms considered in previous studies. For

[32] For a reform in Sweden, Persson and Vikman (2014) find no significant effect on the exit rate from welfare, but a negative and statistically significant effect on the entry rate into welfare. We find that the effect on entry is insignificant in our DD setup, but is also negative and statistically significant in our DRD setup.

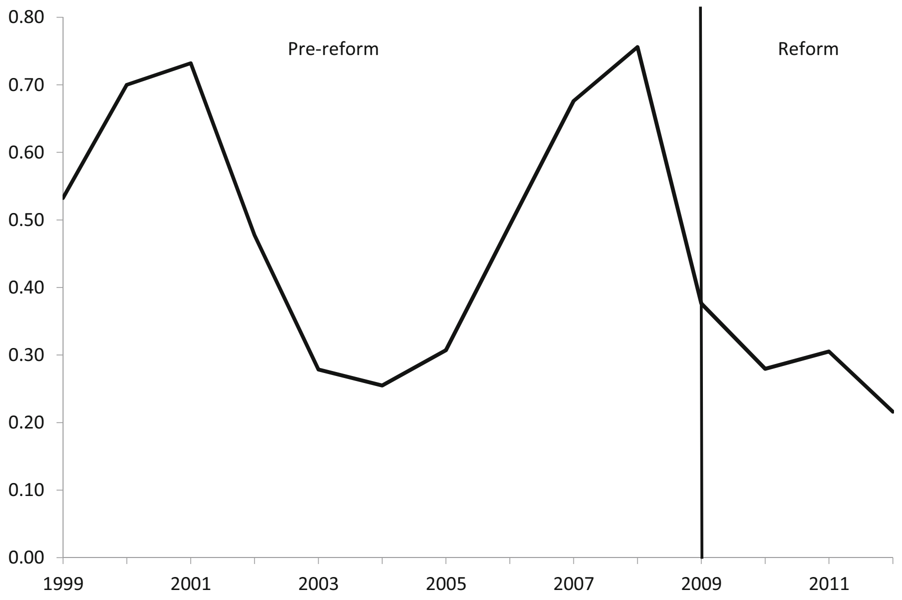

Fig. 2 Vacancy-to-unemployment ratio: 1999–2012. *Source*: Statistics Netherlands (Statline)

the UK, Blundell et al. (2004) note that the New Deal was introduced at a favorable point of the business cycle by historical standards, while stressing that '[C]learly, the program in this favorable climate may not apply to less favorable periods.' Likewise, the reforms in Sweden and Norway studied by Dahlberg et al. (2009), Persson and Vikman (2014) and Hernæs et al. (2017) were implemented in relatively favorable periods (the end of the 1990s). Our results thus suggest that mandatory activation programs may be a less effective policy tool for increasing employment during a recession.

Funding The authors declare that they have not received funding for the research described in this paper.

Availability of data and materials The paper uses anonymized data from Statistics Netherlands, accessible via remote access (contact microdata@cbs.nl for further details).

Compliance with ethical standards

Conflict of interest Emile Cammeraat declares that he has no conflict of interest. Egbert Jongen declares that he has no conflict of interest. Pierre Koning declares that he has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Code availability All the codes and programs used in the analysis are available on request, and researchers who obtain access to the datasets of Statistics Netherlands will be able to replicate all of the findings.

References

Abadie A, Athey S, Imbens G, Wooldridge J (2017) When should you adjust standard errors for clustering? NBER working paper no. 24003, Cambridge

Angrist J, Pischke J (2009) Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Princeton

Barreca AI, Guldi M, Lindo JM, Waddell GR (2011) Saving babies? Revisiting the effect of very low birth weight classification. Q J Econ 126(4):2117–2123

Barreca AI, Lindo JM, Waddell GR (2016) Heaping-induced bias in regression-discontinuity designs. Econ Inq 54(1):268–293

Bell D, Blanchflower D (2011) Young people and the Great Recession. Oxf Rev Econ Policy 27(2):241–267

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? Q J Econ 119(1):249–275

Blundell R, Dias MC, Meghir C, Reenen J (2004) Evaluating the employment impact of a mandatory job search program. J Eur Econ Assoc 2(4):569–606

Board of Amsterdam (2009) Vaststellen verordening tot wijziging van de toeslagenverordening, handhavingsverordening, afstemmingsverordening WWB, verordening clientenparticipatie DWI 2006 en re-integratieverordening i.v.m. de invoering van de Wet investeren in jongeren

Bolhaar J, Ketel N, Van der Klaauw B (2019) Job-search periods for welfare applicants: evidence from a randomized experiment. Am Econ J Econ Policy 11(1):92–125

Cammeraat E, Jongen E, Koning P (2017) Preventing NEETs during the great recession—the effects of mandatory activation programs for young welfare recipients. IZA discussion paper No. 11090, Bonn

Carcillo S, Fernández R, Königs S (2015) NEET youth in the aftermath of the crisis: challenges and policies. OECD social, employment and migration working papers no. 164, OECD, Paris

Card D, Kluve J, Weber A (2017) What works? A meta analysis of recent active labor market program evaluations. J Eur Econ Assoc 16(3):894–931

Dahlberg M, Johansson K, Mörk E (2009) On mandatory activation of welfare recipients. IZA discussion paper no. 3947, Bonn

Donald S, Lang K (2007) Inference with difference-in-differences and other panel data. Rev Econ Stat 89(2):221–233

Gautier P, Muller P, Van der Klaauw B, Roshlom M, Svarer M (2018) Estimating equilibrium effects of job search assistance. J Labor Econ 36(4):1073–1125

Hernæs Ø, Markussen S, Roed K (2017) Can welfare conditionality combat high school dropout? Labour Econ 48:144–156

Imbens G, Wooldridge J (2009) Recent developments in the econometrics of program evaluation. J Econ Lit 47:5–85

Kluve J, Puerto S, Robalino DA, Romero JM, Rother F, Stöterau J, Weidenkaff F, Witte M (2016) Do youth employment programs improve labor market outcomes? A systematic review. Ruhr economic papers no. 648, Essen

Leenheer J, Adriaens H, Mulder J (2011) Evaluatie wet investeren in Jongeren. Centerdata Tilburg, Tilburg

Ministry of Social Affairs and Employment (2008) Bevordering duurzame arbeidsinschakeling jongeren tot 27 jaar (Wet investeren in jongeren). Ministry of Social Affairs and Employment, The Hague

OECD (2016a) Education at a glance. OECD, Paris

OECD (2016b) Labour force statistics. OECD, Paris

OECD (2016c) Youth not in employment, education or training (indicator). OECD, Paris

Persson A, Vikman U (2014) The effects of mandatory activation on welfare entry and exit rates. In: Carcillo S (ed) Research in labor economics, vol 39. Emerald, Chennai, pp 189–217

Statistics Netherlands (2015) Documentatierapport Arbeidsmarktpanel 1999–2013. Statistics Netherlands, Den Haag

UWV (2014) Na de WW in de bijstand. UWV, Amsterdam

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.