RESEARCH REPORT
STUDY PAPER 204
MAY 2024

From Disability Insurance to Workfare: Evidence from a Danish Labour Market Reform

Jeppe Elholm Madsen og Anders Holm

THE ROCKWOOL FOUNDATION
RESEARCH

STUDY PAPER 204 MAY 2024

From Disability Insurance to Workfare: Evidence from a Danish Labour Market Reform

Published by:

Address:
The ROCKWOOL Foundation Research Unit
Ny Kongensgade 6
1472 Copenhagen, Denmark

Telephone +45 33 34 48 00
E-mail: kontakt@rff.dk
en.rockwoolfonden.dk/research/

May 2024

From Disability Insurance to Workfare: Evidence from a Danish Labour Market Reform*

Jeppe Elholm Madsen†

Anders Holm‡

March 11, 2024

Abstract

We examine a Danish labour market reform that replaced the Disability Insurance program (DI) with a "workfare program." Post-reform, 80 % of individuals who would have enrolled in DI before the reform were instead placed in the workfare program, while the remaining with severe health issues still enroll in DI. The reform had a positive impact on employment transitions, primarily in subsidized employment, but resulted in a 15 % reduction in long-term income. Our data enables us to focus on average rather than marginal enrollment. We observe only modest positive effects on employment and negative effects on income.

1 Introduction

In this paper, we demonstrate how a disability reform that replaced the Disability Insurance (DI) program with a Workfare Program (WP) affects employment and income of the enrollees. The WP provides lower benefits and assistance into employment compared to the DI program. Unlike other studies, Haller, Staubli and Zweimüller (2020), Maestas, Mullen and Strand (2013), we are able to address the effects of benefit reforms on the average rather than the marginal benefit recipient. We show that the effects of the reform vary considerably

*We would like to express our gratitude to Søren Leth-Petersen for his valuable comments on this paper.

†Department of Economics, University of Copenhagen, Email: jeppeelholm@hotmail.com

‡Department of Sociology, University of Western Ontario and The Rockwool Foundation, Denmark, Email: aholm@uwo.ca

between marginal and average recipients. While marginal recipients experience large positive employment effects, the average recipient experiences both a significant income loss and no improvement in employment. Our results shed new light on the employability of disability recipients.

Over the past thirty years, the number of individuals receiving disability insurance (DI) has significantly increased in Western countries, Burkhauser and Daly (2022). According to the OECD, approximately 15 percent of individuals are registered as disabled, and their employment rates remain considerably lower compared to those without disabilities, OECD (2022).

The cost of disability is rising rapidly. In the US alone, disability insurance rolls increased from 1.2 million in 1967 to 8.8 million in 2012. Since 2009, the DI program has paid out more in annual benefits than it has received in taxes and interest from its trust fund, Burkhauser et al. (2014). One significant issue is that the DI program is inherently long-term, as beneficiaries usually remain in the program indefinitely once they have been enrolled.

Consequently, the expansion of the program has led to cash benefits from the DI program being more than three times larger than those paid by unemployment insurance (UI). Specifically, DI paid out \$136.9 billion compared to \$42.7 billion for UI. Additionally, the expansion of the DI program is also prominent relative to the working-age population. From 1985 to 2012, the percentage of DI claimants in the United States more than doubled, rising from approximately 2.4 percent to 5.9 percent of the working-age population. Meanwhile, the DI program's share of total expenditures on old-age, survivors, and disability insurance (OASDI) increased from 10 percent to 17 percent.

The very nature of the DI program is to provide income replacement for people who cannot work due to health-related disabilities. However, it is puzzling that despite medical advancements that allow more people to remain employed, increased life expectancy, improved overall health, the predominance of less physically demanding jobs, and the implementation of laws prohibiting workplace discrimination against those with disabilities, the number of individuals receiving disability insurance continues to rise. This situation begs the question: Will any policy measure be able to motivate individuals, who might otherwise end up on DI, to return to work?"

To address this question, our paper examines the employment effects of a 2013 labour market reform in Denmark, which decreased the inflow to the DI program by 80 % for people under 40. Instead of being automatically admitted to DI, individuals were enrolled into a "workfare program" (WP) designed to assist them in finding employment, primarily by helping them search for subsidized jobs with regular employers.

The 2013 reform in Denmark serves as an ideal case study for examining DI policy for two reasons. Firstly, the policy shift in 2013 contrasts two diametrically opposed approaches. Prior to 2013, Denmark had one of the most lenient DI systems among OECD countries, evidenced by one of the highest ratios of individuals on DI relative to the workforce, OECD (2022). This approach is typically observed in Scandinavian countries, often referred to as the Nordic welfare model. After 2013, Denmark significantly restricted access to the DI program, nearly ceasing all new admissions and instead redirecting applicants to the WP. Thus, this policy experiment allows us to compare a policy approach to DI that aligns with the Nordic welfare model with one that is closer to the liberal welfare model, as observed in the US, UK, and other countries with a strong emphasis on work incentives and limited access to DI.

Secondly, the swift policy change in 2013 allows us to fix other factors that might significantly impact the DI program. For instance, labor market shifts and reduced demand for unskilled labor could drive more unskilled workers toward the DI program as their employment opportunities diminish. Similarly, an aging population and an increase in the retirement age might compel more individuals to continue working despite health impairments, thereby increasing DI program applications. The Danish policy shift in 2013 provides a unique opportunity to examine the effects of a DI policy transition from one extreme to another in a scenario where labor market and demographic factors remain fixed.

The WP is a long-term initiative designed to assist individuals who might otherwise enter the DI program in transitioning back to employment. To aid WP participants in their job search, they are guided, especially towards finding subsidized jobs. When in subsidized jobs, participants receive a basic income from the municipality. This setup allows employers to offer a lower hourly wage than in the regular job market while ensuring that participants earn an income comparable to DI. When not engaged in subsidized or regular employment, WP participants receive benefits lower than the standard DI program and are expected to seek either subsidized or non-subsidized employment.

We find that the effect of assigning individuals to the WP instead of DI is minimal, as participants struggle to find employment in the regular job market. Our paper's findings underscore the potential benefits of subsidized jobs as a means of facilitating employment for individuals with disabilities. We also find that even for subsidized employment only the healthiest and most qualified gain from the WP. Overall, the observed effects are modest. Our paper further concludes that a large fraction of those admitted to the WP will remain in the unemployment system, either staying on WP or eventually ending up on DI at a later stage.

Descriptive statistics of WP participants shed light on why they remain unemployed and have difficulty entering the unsubsidized job market, where subsidized jobs are unavailable. WP participants face multiple challenges, including health issues and limited education. Their health challenges and lack of education make them unappealing to employers, who are reluctant to offer wages above benefit levels for low productivity. Denmark does not have a formal minimum wage. Nonetheless, collective agreements and generous unemployment benefits effectively set a minimum wage rate for ordinary employment.

The findings of this study are significant in the context of two key areas of literature. Firstly, they contrast with the extensive body of research on active labour market policies for individuals receiving unemployment insurance benefits. Studies in this domain have documented that interventions like job search activities and on-the-job training programs significantly enhance short-term and long-term employment (see Card, Kluve and Weber (2010) for an excellent overview). In contrast, this paper concludes that for people with disabilities entering the WP, undergoing similar measures has only a minimal impact on employment.

Secondly, and relatedly, the existing literature on DI primarily focuses on individuals at the margin of qualifying for the DI program. For example, Bound (1989) uses rejected DI participants as a control group for DI participants, Maestas, Mullen and Strand (2013) uses exogenous variation among disability examiners, both of which involve people on the cusp of entering the DI program. These studies find substantial effects on employment for individuals not allocated to the DI program, with around 50 % of those marginally not admitted finding employment in the US case. The effects reported in these studies are "Local Average Treatment Effects" (LATE), see Imbens and Angrist (1994). LATE identifies the average effect of the treatment for those who comply with the instrument, in this context, individuals who are accepted into the DI program because they encounter a more lenient examiner. In Maestas, Mullen and Strand (2013), only 23 % of the DI applications comply with the instrument.

In contrast, the LATE effects in this study does not only include the marginal compliers and in turn does not find such sizeable effects. Since the inflow into the DI program was reduced by about 80 %, our study also encompasses people with significantly worse health than those at the margin. Consequently, the substantial effects observed in previous LATE studies are among a selective group with the best health, from which one can expect to find such sizable estimates. However, our study reveals that near-zero effects are found for groups with considerably worse health. Therefore, studies focusing solely on those at the margin of the DI program may overestimate the impact of restricting access to the DI program. To give context, we find employment effects of up to 30 % among people with the best health and the most educated group. Yet, this group only represents 4.5 % of the total sample. In contrast, we find no effect on the employment of those with poor health and no education, who comprise approximately 50 % of the sample.

We estimate the marginal treatment effect (MTE), see Heckman, Urzua and Vytlacil (2006). The MTE has the advantage of allowing for both heterogenous effects on observables, but also selection on gains on unobservables. The MTE identifies a broad range of effects including LATE, average treatment effects and average treatment effects on the treated all with different weights on the MTE's. With the MTE approach, we are able to show that the LATE estimates put considerable weight on those with high selection on gains, while those with low selection on gains receive less weight. The MTE approach suggests that studies which rely solely on identifying LATE effects might not be providing a complete picture. The LATE may put disproportional much weight on a specific subset of the population (those with high selection on gains), and may therefore not be representative of the overall effect of the treatment on the entire population.

Our findings also align with the literature which suggests that there are complex interplays between health and earnings among DI applicants, Bound (1989); Low and Pistaferri (2015), and that reforms in disability benefit programs should be carefully examined within the broader context of the policy landscape in various OECD countries, Burkhauser et al. (2014); Burkhauser and Daly (2022).

The remainder of this paper is organized as follows: Section 2 describes the Danish setting and the Disability Insurance (DI) system. Section 3 outlines the data used in this study. Section 4 discusses the methodology employed, while section 5 presents and discusses the results. Finally, section 6 draws conclusions from the findings and discusses their implications for policy and future research.

2 Institutional Setting

2.1 The Danish DI Program

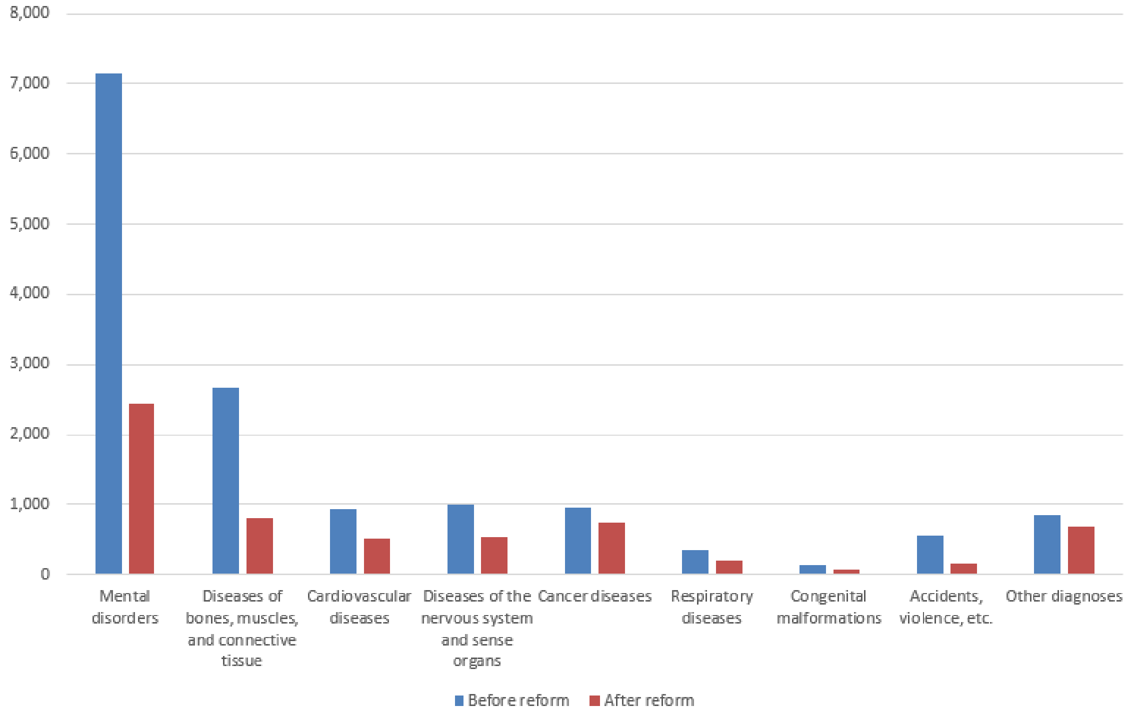
Disability insurance (DI) is a type of financial support provided by the municipality to individuals who are unable to work due to a long-term illness or disability. To qualify for DI, applicants must undergo a medical assessment to determine the severity and duration of their disability, which must render them unable to work, even with support or accommodations. Eligibility for DI is based on the individual's work history, age, and level of disability. Only individuals below the retirement age of 65 can apply for DI. If the applicant meets the requirements for DI, they can submit an application to their municipality, and the municipality will assess whether they should be enrolled in the program. Once an individual is accepted into the DI program, it is highly uncommon for them to exit the program. The instances where individuals do leave the program are typically limited to situations where the individual passes away or transitions into the old-age pension program.

2.2 The Introduction of The Workfare Program

In June 2012, the Danish parliament passed a reform that altered the Disability Insurance (DI) system. This reform, effective from January 2013, introduced stricter eligibility criteria with the goal of decreasing the number of new entrants into the DI program. Similar policy changes have also been seen or proposed in other OECD countries.

By principle, individuals under the age of 40 are generally deemed ineligible for DI regardless of their health. For those over 40, the municipality can grant eligibility for DI when there is conclusive evidence that attempts to enhance the individual's work capacity would be futile. To examine the impact of these new rules on the inflow into DI, Figure 1 demonstrates the effects of stricter eligibility criteria on the number of new DI recipients across different diagnostic groups for both under 40 and over 40. Inflow from all diagnoses was reduced after the reform; however, the reduction for people with mental disorders was the largest in absolute terms. Other diagnoses, such as inflow from people with cancer, were also reduced after the reform but not as much as for those with mental disorders in relative terms.

Figure 1: Inflow into Disability Insurance by Diagnosis Before and After the Reform for all age groups



Note: The figure displays the changes in the inflow into DI by diagnosis before and after the reform. The inflow from all diagnoses was reduced after the reform; however, the reduction from people with mental disorders was the largest in absolute terms.

In place of the DI program, a workfare program (WP) was established. Because the DI program was most severely affected for those under 40, our focus is on this group. To be eligible for the WP, the person must be at risk of ending up on DI. Consequently, individuals without health challenges, for example, those in the regular social assistance program, are not eligible.

These new rules led to a reduction of approximately 80 % in the inflow into the DI program for those under 40 in the years following the reform. The WP program, which lasts between 1 and 5 years, aims to help participants find employment. Compared to the DI program, the WP program offers a substantial benefit cut and includes a labour market activation component. About 80 % of WP participants engage in educational courses and job training programs, while the remaining 20 % take part in firm internships. The WP program also provides support for health and social issues, including counselling, substance abuse treatment, and rehabilitation. The reform also introduced subsidized work opportunities for those in the WP program. After completing the WP program, participants are expected to transition into regular or subsidized work. Subsidized work involves receiving a benefit transfer from the municipality along with a wage payment from the employer based on the number of hours worked. The income earned through subsidized work is higher than that provided by the WP program, making it an attractive option for participants.

2.3 The DI and WP schemes

Individuals enrolled in the DI program, both pre- and post-reform, are generally not expected to work due to their disability. However, there is some flexibility as they are permitted to engage in employment to a certain extent, often only as little as 10 hours per week, and this is usually within a subsidized program. When doing so, the municipality might reassessment their work capacity and they risk losing their DI. As such, individuals have a strong incentive not to work or reveal any work capacity. On the other hand, individuals on WP are encouraged to find employment. They are assisted in finding either subsidized or unsubsidized employment through job training or firm internships. The incentive to find employment (subsidized or not) is considerably larger on WP than on DI. Firstly, the benefit on WP is considerably lower than DI. Secondly, when leaving WP and transitioning into subsidized employment, individuals receive a benefit comparable to the DI benefit and additionally, an adjusted wage rate that reflects the amount of work provided, which is paid by the employer. Upon finding regular employment, benefits are entirely replaced with wages from that employment.

To illustrate the potential incentives (or disincentives) from DI and WP regarding employment, Table 1 displays the average and standard deviations of observed income for DI recipients before the reform, and for DI and WP recipients after the reform. This includes WP recipients in both subsidized and regular employment.

Table 1: Comparison of Mean Monthly Earnings (DKK) Before and After Reform for DI and WP Recipients.

		Before Reform		**After Reform**	
	Statistic	**Couple**	**Single**	**Couple**	**Single**
Disability Insurance	Mean	17,272	19,106	17,777	19,039
	SD	6,061	3,134	5,595	3,376
Subsidized Work	Mean	20,286	22,189	21,235	21,138
	SD	4,465	3,446	5,933	5,659
Unsubsidized Work	Mean	-	-	19,583	20,319
	SD	-	-	8,726	8,136
Workfare Program	Mean	-	-	14,477	13,362
	SD	-	-	3,678	3,160

From the table, it is evident that both single individuals and those with a spouse have a short-term incentive to choose DI, if possible. However, there is also an incentive to engage in subsidized employment and regular employment if the disutility of work is lower than the monetary gain from such employment. Nevertheless, it is clear that there is only a limited incentive, if any, to pursue regular employment over subsidized employment, as regular employment typically entails longer working hours.

Selection into different states in Table 1 is likely to occur, as we can anticipate that individuals who are employed are generally in better health and more motivated to work compared to individuals who are still on WP. The presence of this selection may induce selection bias and hence hamper group comparisons. If the earnings capacity of those working is higher than for those in WP, this may lead to a downward bias in the work incentive moving from WP to employment as apparent in the table. Hence, those in employment are presumably more productive than individuals on WP. Therefore, the differences in Table 1 are probably upper bound indicators of the incentive to be employed from WP. Finally, we find more heterogeneity in the subsidized work after the reform compared to before the reform. This is due to the less specialized nature of subsidized jobs after the reform.

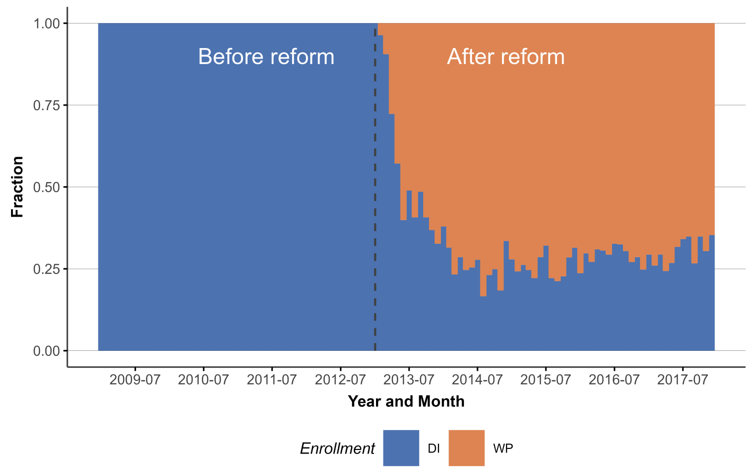
3 Data

This study makes use of panel data from administrative records provided by Statistics Denmark, which covers the entire Danish population between the years 2008 and 2019. The data is extensive and comprises information on various sources such as income, employment status, benefits, education, health, as well as demographic attributes, in addition to individual ID numbers. Income data are obtained from exact tax records, providing monthly information on individual income, including wage and income transfers. For information on individuals' entries into the Disability Insurance (DI), Workfare Program (WP), or other benefit programs, we use a database from the Danish Agency for Labour Market and Recruitment, providing weekly information for every person in Denmark regarding the benefits they receive. Health information is obtained from two sources. First, we use pharmaceutical data sourced from the Danish Medicines Agency's Prescription Database, which includes records of prescription drugs dispensed, the type of medicine (categorized by the ATC code - Anatomical Therapeutic Chemical code), the dosage, and the date of prescription. Secondly, we utilize hospital admission data from the Danish Health Data Authority's National Patient Register, which includes information on hospital admissions and diagnoses coded using the ICD-10 coding system, an internationally recognized classification system for diseases and health conditions. These two sources of health data are integrated into a health index, which is explained in subsequent sections. Educational data, which includes yearly updates on the highest educational attainment of individuals, is obtained from the Ministry of Higher Education and Science. Demographic data, including information such as age, gender, and ethnicity, is sourced from the Danish Population Register.

3.1 Data Description

In this section, we describe our data. We pay special attention to what happens in and around the reform. In figure 2, we show how the inflow into DI/WP changes from only DI before the reform and into DI/WP after the reform. The figure illustrates how DI inflow separates into DI and WP after the reform. After the reform, the fraction of individuals entering WP rapidly increases to a saturation level of around 75 to 80 percent. The path towards the saturation level for WP played out during the calendar year 2013. Hence, for comparisons before and after the reform, 2013 may be special and requires special attention.

Figure 2: Inflow into Disability Insurance and Workfare Program Pre and Post-Reform



Note: The figure displays the changes in the inflow into DI and WP. Before the reform, the inflow was only into DI. Post-reform, the inflow is split between DI and WP, with WP inflow quickly rising to a saturation point of 75-80 % during the calendar year 2013.

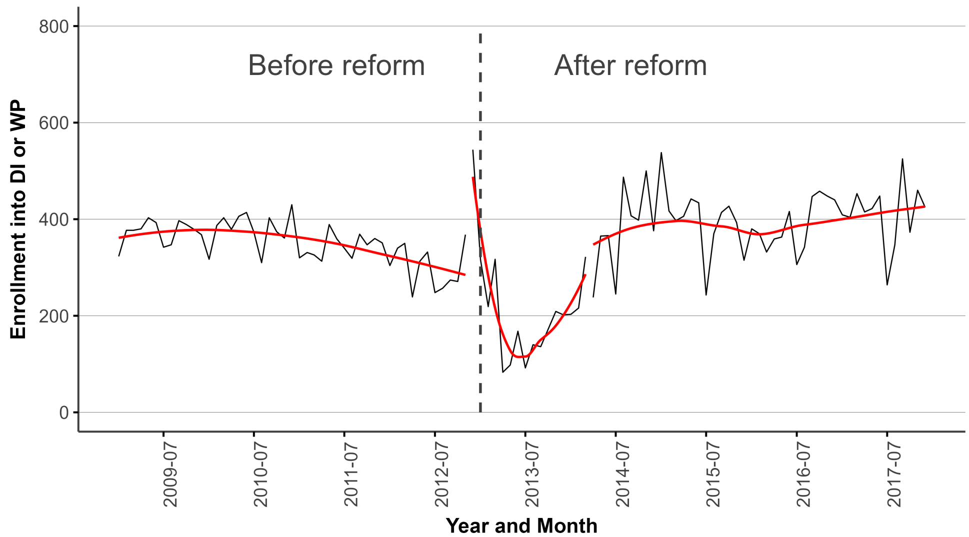
We employ two different analysis with two different identification strategies. First, we conduct a marginal treatment analysis (MTE), Bjorklund and Moffitt (1987). This analysis reacquires the availability of a continuous instrumental variable that will span the entire latent propensity to be in DI versus WP. We use the enrollment probability into WP as such an instrument.

Our second analysis of the effect of DI versus WP uses comparisons before and after the reform. Here DI's before the reform are used as a control group for the DI/WP's after the reform. As we want to use pre-reform behavior of DI recipients as a control group for post-reform DI and WP recipients, we need to show that the two groups are comparable on important observed characteristics and that selection into DI before the reform is comparable to selection into DI/WP after the reform.

First, we show that the inflow into DI/WP before and after the reform remained the same. This supports the hypothesis that the reform did not have lasting impacts on who entered DI/WP before and after the reform. Second, we show that, based on important observed outcomes, there is no difference in the composition of DI/WP recipients before and after the reform. In summary, our evidence suggests that the reform did not introduce lasting selection biases into the DI/WP program.

From figure 3 we find that the level of enrollment into DI/WP program stays steady throughout the implementation of the reform, with the exception of a noticeable spike immediately before and a dip after the reform takes place. This spike and subsequent dip are attributed to case workers' efforts to register as many DI clients as possible into DI before the reform is enacted, as a means to prevent them from having to enter WP post-reform. However, the level of enrollment rises back to its pre-reform levels within a year. In this respect enrollment in 2013 may be special and in our subsequent analysis we treat enrollees in 2013 separately.

Figure 3: Number of People Enrolling into Disability Insurance and Workfare Program Pre and Post-Reform



Note: The figure displays the number of enrollees into DI pre-reform and DI/WP post-reform. A steady inflow is observed over time, albeit with a noticeable spike just prior to the reform and a dip thereafter. However, these fluctuations normalize within a year post-reform.

Table 2, displayed below, presents a set of descriptive statistics. It provides data for three distinct time periods: immediately prior to the reform's implementation, the initial year of the reform (which was 2013 and is considered unique), and the years following the reform.

From the table, we observe that there are no noticeable differences among individuals enrolled before the reform, immediately after the reform, and in the subsequent years after the reform in terms of age, the proportion of natives, the proportion with no education beyond comprehensive school, pre-enrollment income, and health (we will discuss the construction of our health index later). Only the fraction of individuals employed before enrolling into DI/WP and the fraction of males are lower in the subsequent years after the reform (2014-2015) compared to right after the reform (2013) and before the reform (2011-2013). This is due to business cycle effects, in that the general employment situation is improving over the period in question. This is reflected in improved employment opportunities. It is well-known that men's employment tends to be more responsive to business cycle fluctuations than women's, see for instance Razzu and Singleton (2016). Therefore, when employment opportunities improve (as observed in 2014-2015), it is expected that the proportion of males among DI/WP participants would decrease. To factor this in, we also account for business cycle effects in our later regressions.

We now turn to the construction of our health measure. Our dataset is extensive and includes

Table 2: Descriptive Statistics of 2011-2012, 2013 and 2014-2015 samples

Variable	**2011-2012**		**2013**		**2014-2015**	
	Mean	**Std.**	**Mean**	**Std.**	**Mean**	**Std.**
Age	29.43	7.38	29.01	7.64	29.08	6.91
Male	0.49	0.50	0.50	0.50	0.44	0.50
Native Born	0.84	0.37	0.84	0.36	0.85	0.35
No Education	0.37	0.49	0.39	0.50	0.40	0.49
Pre-Income (DKK)	13,728	7,492	13,661	8,090	12,364	6,443
Pre-Employment	0.12	0.32	0.12	0.32	0.09	0.29
Health Percentile	57.17	32.99	56.20	33.03	58.76	32.42
N	5924		1792		7375	

Source: Statistics Denmark and own calculations.
Note: The table reports means and standard deviations.

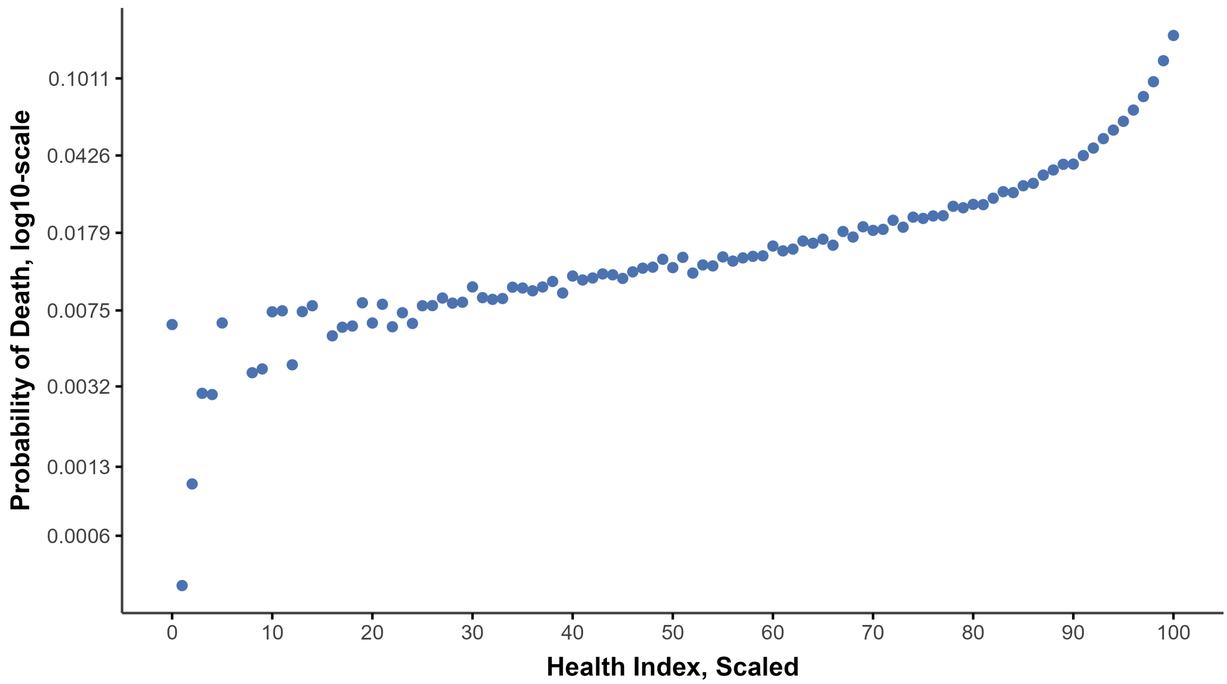
information on prescription drugs related to cardiovascular and psychiatric conditions, as well as opioids (identified by ATC codes A10, N02, N05, N06, N07, and C), and their dosage details. Additionally, we have data on the number of days each individual in the Danish population has been hospitalized and their corresponding diagnostic codes. Utilizing these variables, we conduct a Principal Component Analysis (PCA), see Pearson (1901). Following Bingley et al. (2014), we extract the first eigenvector from the PCA as an index for overall health. This method follows a similar vein to the one proposed by Charlson et al. (1987). However, they use diagnostic codes rather than medicine consumption and hospital admissions. We only have access to prescribed medical consumption and hospital admissions, not to diagnostic data. To illustrate that our index possesses the same merits as the measure proposed by Charlson et al. (1987), we demonstrate in figure 4 how our health index in year t predicts mortality in year $t+1$ for the general population.

From the figure, we find that our measure of health has good predictive characteristics in terms of predicting mortality. A higher index, indicating worse health, predicts a higher probability of mortality. It has better predictive characteristics compared to the measure developed by Charlson et al. (1987).

In figure 5, we illustrate the relationship between our health index, age, and gender. As expected, older individuals generally have poorer health than younger ones.

The figure reveals a clear and expected relationship between age and health, wherein older individuals exhibit significantly poorer health compared to younger individuals. For women we find a bump in the distribution in and around the age of 30 years. This is related to pregnancies, and being pregnant is, according to our measure, related to a temporary worsening of the health of women. This is corroborated by other studies, see for instance WHO (2019). Note that, according to our measure, younger men have better health than younger women, but this reverses around age 60. This may be due to a health selection effect in mortality, such that women with worse health have higher mortality than men with worse health.

Figure 4: Relationship Between Mortality and Health Index



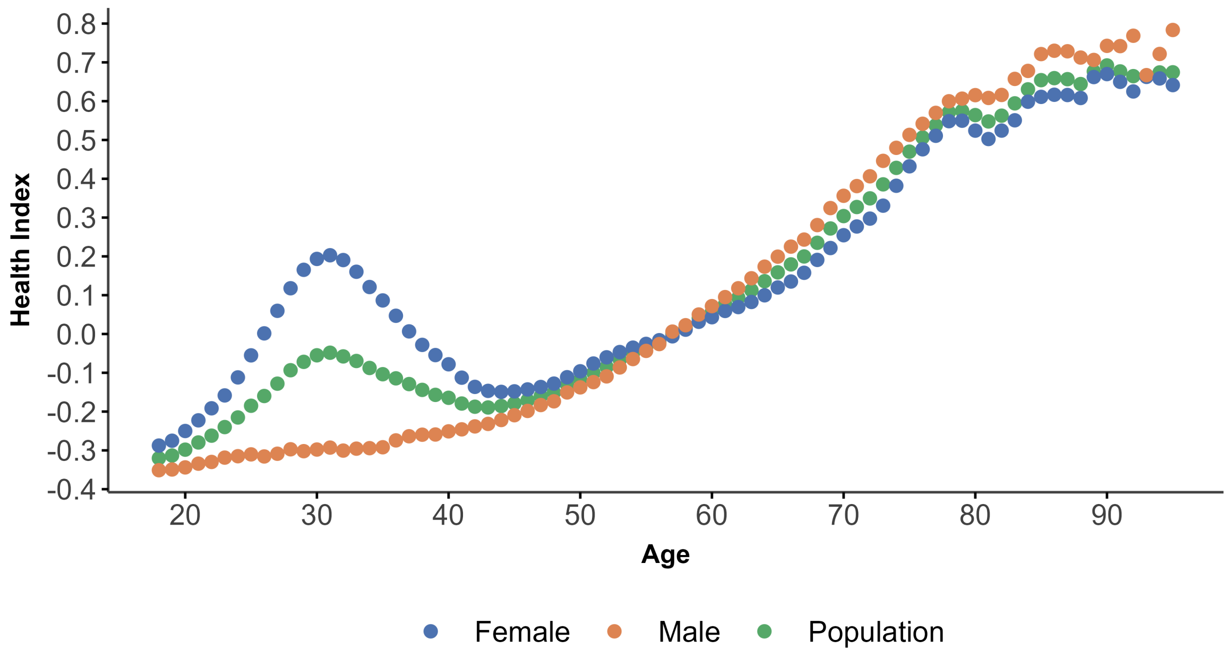
Note: The figure displays the relationship between the health index constructed using prescription drug information, hospitalization days, and the probability of mortality in the subsequent year. A higher health index value, indicating worse health, is associated with a higher probability of mortality.

To study the health condition of individuals on DI and WP, we now compare the health of the general population with those on DI and WP. In figure 6 below we show the cumulative health conditions of the general population, those older than 64 years (age at which one is eligible for public pensions) and those on DI and WP. The figure shows that individuals older than 64 have worse health compared to the population in general. This is expected based on general knowledge and figure 5. More interesting is the fact that DIs and DI/WPs have much worse health than both the general population as well as the population aged 65 and older, bearing in mind that we only look at DI and WPs younger than 40. Hence, the health conditions for DIs and WPs are significantly worse than individuals over the retirement age. Furthermore, we find that the distribution of health conditions of DIs before the reform is completely coinciding with the distribution of health for DIs and WPs after the reform. This is further evidence of no change in selection into DI and WP after the reform yielding perfect balancing for our control group (DIs before the reform) and the treatment group (DIs and WPs after the reform).

To further substantiate the balance between the treatment and control groups, we illustrate in figure 7 how the average key characteristics develop before and after the reform. In the figure, we display the trajectories of sample averages for our covariates, based on the enrollment date into either DI before the reform or DI/WP after the reform.

From the figure, we observe that the reform did not significantly alter the mean values of

Figure 5: Relationship Between Health Index, Age, and Gender



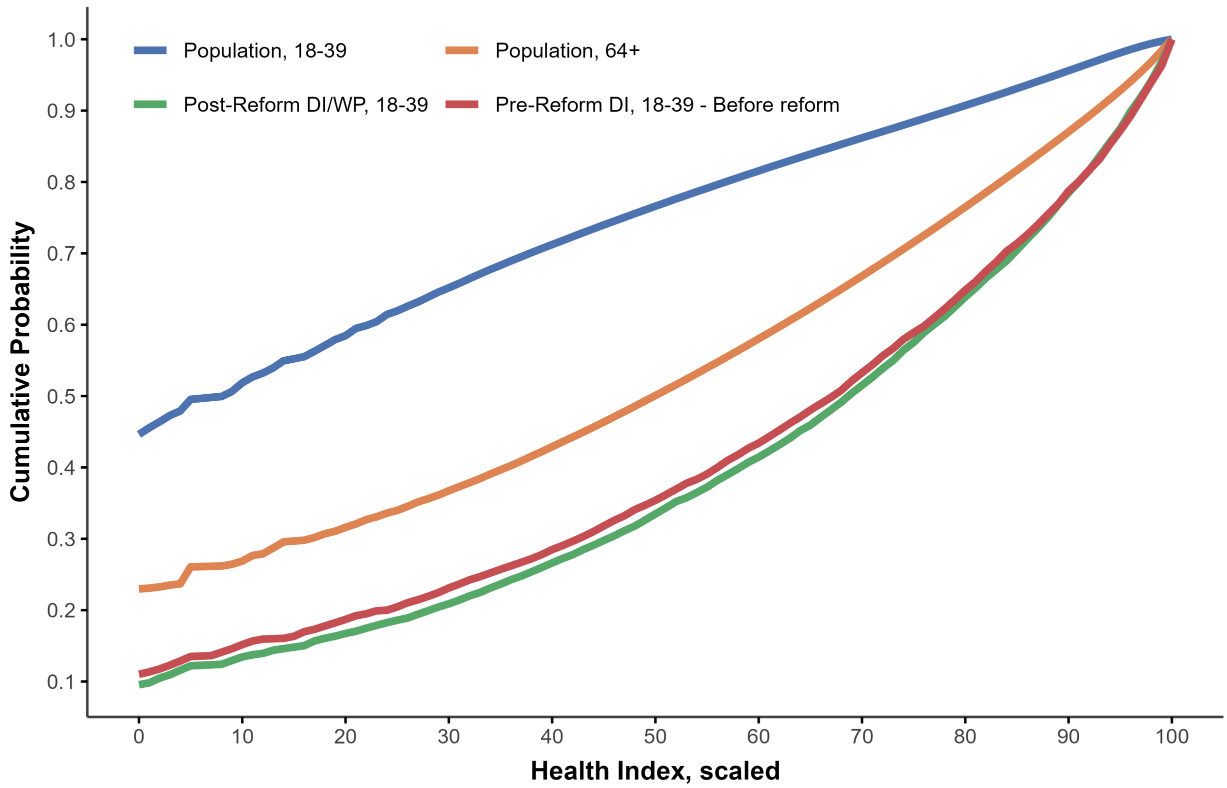
Note: The figure illustrates the relationship between the health index and age, differentiated by gender. It shows that older individuals generally have a poorer health. Additionally, a notable increase in the health index is observed among women around the age of 30 due to pregnancies.

the covariates. However, there are minor effects noticeable, especially during the first year after the reform.

Average age and pre-gross income have declined throughout the entire observation period. Shortly after the reform, there's a noticeable dip in both age and gender. This suggests that caseworkers might have directed older male applicants to alternative programs. A similar trend is observed for applicants without any further education.

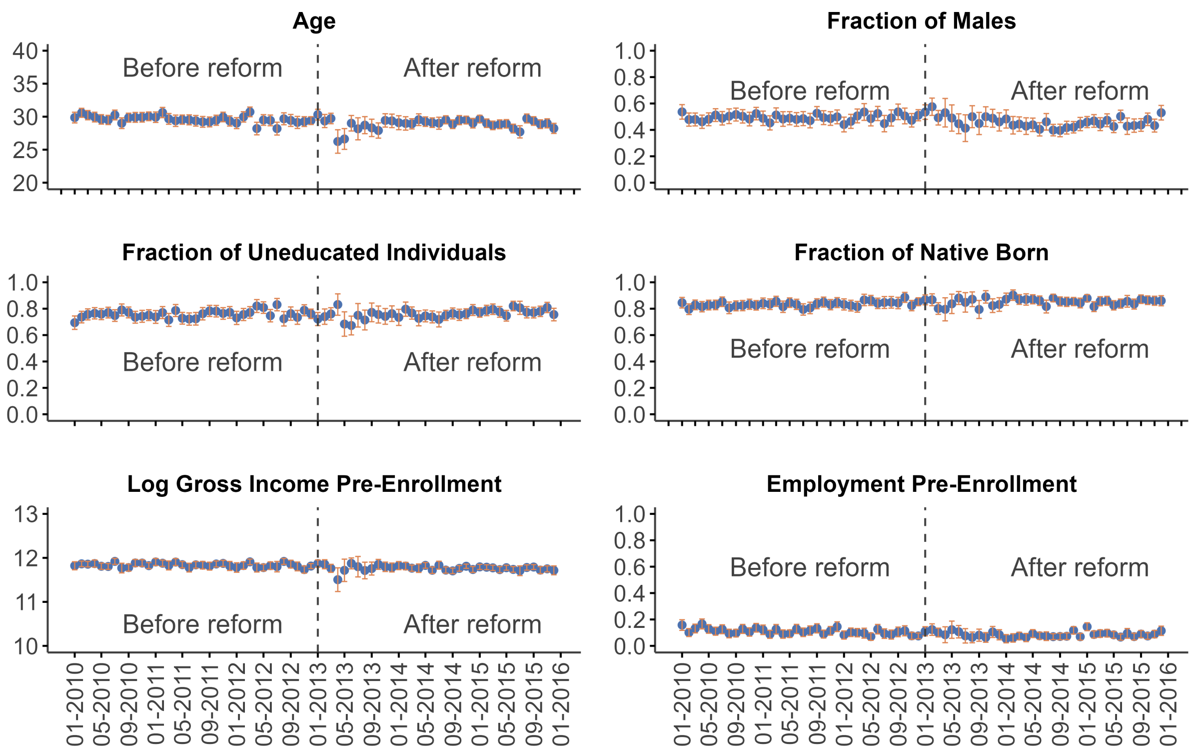
In summary, immediately following the reform, there seems to be a trend of 'inverse cream skimming' where more resourceful individuals are directed into the program, while less resourceful ones are channeled into other programs. However, this trend reverses as time progresses; later in the post-reform period, the average statistics return to their pre-reform levels.

Figure 6: Cumulative Health Distribution of the General Population, Individuals Older than 64 Years, and Individuals on DI and WP



Note: The figure illustrates the cumulative health conditions of the general population, individuals older than 64 years (age at which one is eligible for public pensions), and individuals on DI and WP. It reveals that individuals older than 64 generally have worse health compared to the general population. Additionally, it highlights that individuals on DI and WP exhibit significantly worse health compared to both the general population and the population aged 65 and older. The distribution of health conditions for DI recipients before the reform coincides with the distribution of health conditions for DI and WP recipients after the reform.

Figure 7: Balance before/after the reform on age, gender, education, ethnicity, pre-enrollment income and pre-enrollment employment.



Note: The figure presents the mean values of various covariates, based on the enrollment date into either DI or DI/WP. It shows minimal shifts in the mean of the covariates due to the reform.

4 Methods

In this section we present our two empirical approaches. First we conduct a marginal treatment analysis and next we do a before-after comparison using regression analysis.

4.1 Marginal Treatment Effects

To investigate how the outcomes of a treatment change across different groups, we construct a model that allows for both heterogeneous effects in unobservables and observables. Specifically, we introduce the concept known as the marginal treatment effect (MTE), as referred to in Bjorklund and Moffitt (1987). By examining the MTEs, we can derive a broad range of parameters including the Average Treatment Effects (ATE), Average Treatment Effects on the Treated (ATT), and Local Average Treatment Effects (LATE). Each of these parameters assigns different weights to the MTEs. In this way, we can examine if the LATE estimates are externally valid and understand if the estimated effect for the complier group is indicative of the effect for the entire population. To better understand the model, we introduce the following potential outcome equations:

$$\begin{aligned} Y_0 &= \mu_0(X) + U_0 \\ Y_1 &= \mu_1(X) + U_1 \end{aligned} \tag{1}$$

Here, Y_0 represents the outcome when the individual is on DI, while Y_1 denotes the counterfactual outcome when the individual enrolls in WP. The outcome is comprised of two components: $\mu_0(x)$ and $\mu_1(x)$, which are common effects of WP and DI for all individuals. Additionally, U_0 and U_1 are individual gains from WP and DI that are uncorrelated with X. In essence, the gains from the WP program, compared to the DI program, can vary both because of observed characteristics such as health, age, gender, and so on, but also because of individual gains from the WP. Next, we can connect the potential outcome to the observed outcome, Y, using a standard switching regression model, Quandt (1958):

$$\begin{aligned} Y &= (1 - D)Y_0 + DY_1 \\ &= \mu_0(X) + (\mu_1(X) - \mu_0(X))D + \eta D + U_0 \end{aligned} \tag{2}$$

Where D is a treatment dummy, set to 1 if the person is on WP and 0 otherwise, and $\eta = U_1 - U_0$. Further, consider D^* as the latent utility of WP.

$$\begin{aligned} D^* &= \mu_D(Z) - U_D \\ D &= \mathbf{1}[D^* \geq 0] \end{aligned} \tag{3}$$

$\mu_D(Z)$ represents the utility derived from the observed variable Z. This utility includes an instrument present in the latent utility equation but absent from the outcome equation, as well as observables, X, that appear in both the choice and outcome equations. U_D denotes an unobserved component symbolizing latent resistance to treatment. The second equation asserts that an individual opts for WP if its utility exceeds that of DI; this is the case when $\mu_D(Z)$ surpasses the latent resistance to treatment, represented by U_D. Without any loss of generality, U_D can be normalized to range between zero and one. $\mu_D(Z)$ can then be interpreted as a propensity score, denoted $P(Z)$. As a result, D is set to 1 if $P(Z) > U_D$.

To see how treatment varies over observables and unobservables, it is useful to define the so-called marginal treatment effects (MTE):

$$\begin{aligned} MTE(x,p) &= E(Y_1 - Y_0 | X = x, U_D = p) \\ &= \mu_1(x) - \mu_0(x) + \underbrace{E(U_1 - U_0 | X = x, U_D = p)}_{k(x,p)} \end{aligned} \tag{4}$$

The first component, $\mu_1(x) - \mu_0(x)$, captures how the observables influence the treatment effect. The second component, represented as $k(x,p)$, is the selection term associated with a specific value of latent resilience and x. The function tells us how the treatment effect varies over observables and resilience. Various causal parameters, including the Average Treatment Effect (ATE), the Average Treatment Effect on the Treated (ATT) and the LATE can be formulated as weighted averages of the $MTE(x,u)$, see Heckman, Urzua and Vytlacil (2006). For instance, for a given x:

$$\begin{aligned} ATE &= \int_0^1 MTE(u)\omega_{ATE}(u), du, \quad \text{where } \omega_{ATE}(u) = 1, \\ ATT &= \int_0^1 MTE(u)\omega_{ATT}(u), du, \quad \text{where } \omega_{ATT}(u) = \frac{P(p(Z) > u)}{P(D=1)}, \\ LATE &= \int_0^1 MTE(u)\omega_{LATE}(u), du, \quad \text{where } \omega_{LATE}(u) = \frac{1[p_0 < u < p_1]}{p_1 - p_0}. \end{aligned}$$

From the equations, it's evident that the LATE estimator is a weighted average of the MTE with weights switched on only in regions where individuals respond to the instrument. For instance, with an instrument that measures the varying strictness of disability examiners, the weights are only switched on for those who receive their DI from a lenient examiner, but wouldn't have if they had a stricter examiner.

To estimate the MTE, we use the method of local instrumental variables (LIV). To see how it works, let us define the expectation of Y given $P(Z) = p$ and $X = x$ using equation 2. The equation becomes the following:

$$
\begin{aligned}
E[Y|X=x,P(Z)=p] &= E[\mu_0(X)+(\mu_1(X)-\mu_0(X))D+\eta D+U_0|X=x,P(Z)=p] \\
&= \mu_0(x)+(\mu_1(x)-\mu_0(x))p+E[\eta|X=x,P(Z)=p,D=1]p \\
&= \mu_0(x)+(\mu_1(x)-\mu_0(x))p+\int_0^p E[\eta|X=x,U=u]du
\end{aligned} \tag{5}
$$

where we use the law of total expectation conditioning on $D = 0, 1$ and the assumption that U_0 is uncorrelated with X from the first to the second equation. Subsequently, we assume that both $\mu_0(x)$ and $\mu_1(x)$ are linear, and furthermore, that η is uncorrelated with X:

$$
E[Y|X=x,P(Z)=p] = \beta_0 x+(\beta_1-\beta_0)xp+\underbrace{\int_0^p E[\eta|U=u]du}_{K(p)} \tag{6}
$$

Following this, by taking the partial derivative of equation 6 wrt. p, we arrive at the MTE:

$$
\frac{\partial E[Y|X=x,P(Z)=p]}{\partial p} = (\beta_1-\beta_0)x+\underbrace{E[\eta|U=u]}_{K'(p)} \tag{7}
$$

In the next subsection, we describe how we estimate the MTE based on equation 6.

We conduct our analysis on four distinct sub-populations, dividing our data into two educational categories: those without any formal education beyond comprehensive school and those with vocational or some tertiary education. Additionally, we subdivide based on a health measure at the median to study heterogeneous effects. The distribution of the divisions by health and education is presented in table 3.

Table 3: Cross-tabulation of Education and Health

	Health		
Education	**Less-Sick-Half**	**Sickest-Half**	**Total**
Vocational/Higher Education	4.50	18.47	22.97
No Education	28.81	48.22	77.03
Total	33.31	66.69	100.00

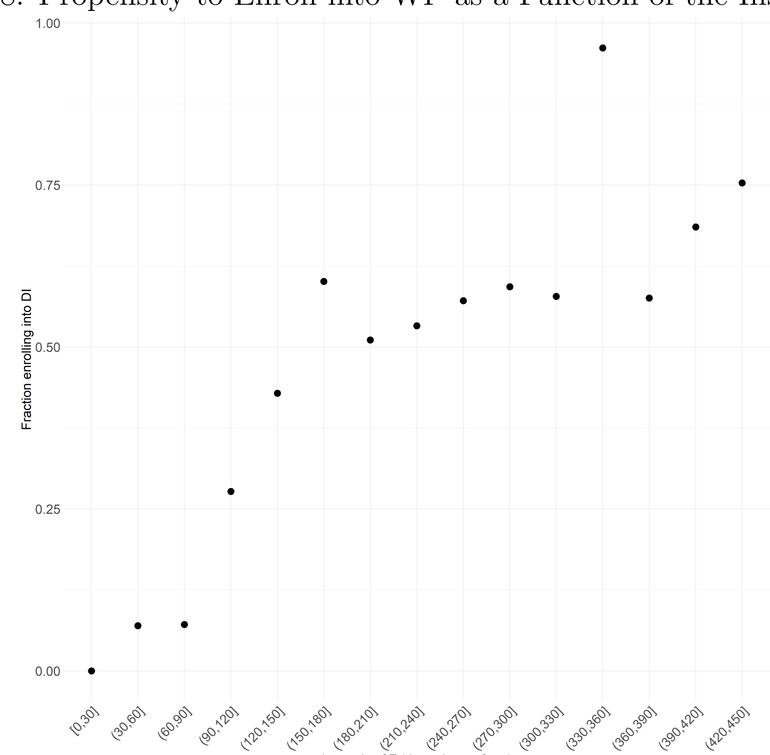
We select these specific sub-populations because we expect them to exhibit markedly different employment opportunities. The table shows that a majority of participants belong to the less healthy half of the health distribution and have no education. For context, about 2 out of 3 people are in the less healthy category, and 77 percent do not have any formal education. Both being in the healthier half and having a formal education are rare, with only 4.5 percent of participants fitting into this group.

We dis-aggregate into these four groups to align our findings with the IV-based results reported in Bound (1989) and Maestas, Mullen and Strand (2013). These studies highlighted significant employment effects due to DI reforms in the US. However, it's vital to understand that their results are local average treatment effects, which are potentially dependent on a specific group of compliers. This group might resemble our sub-population of individuals with better health and formal education.

4.1.1 Estimating the Marginal Treatment Effects

To estimate equation 6 and map out the entire MTE curve, we need an instrument, Z, that exogenously shifts the propensity of enrolling into WP. To do so, we construct a continuous instrument that draws from the gradual implementation of the reform, as shown in figure 2. Specifically, the instrument is the number of days since the reform was initiated. Figure 8 illustrates the increasing propensity to enroll in WP or, conversely, the decreasing propensity to enroll in DI for increasing values of the instrument. In the initial 0-30 days following the reform, enrollment into WP is 0 %. However, after over 400 days, the enrollment rate increases to 75 % and enrollments into DI are 25 %.

Figure 8: Propensity to Enroll into WP as a Function of the Instrument



Note: The figure depicts the declining propensity to enroll into DI as the instrument, representing the number of days since the reform's implementation, increases.

To get an estimate of the propensity score, we use a standard probit model that includes the instrument and control variables: health, health squared, gender, education, native-born,

Table 4: Probit Model with the Propensity to Enroll into WP

Variable	**Estimate**	**Std. Error**	**t value**	**P value**
z	0.007	0.0001	59.29	0.000 ***
N:	24267			

Note: Within the probit model, we control for health, health squared, gender, education, ethnicity, and a measure of aggregate GDP.

and an aggregate measure of GDP. The treatment is the propensity to enroll into WP. Table 4 shows the estimate of the instrument in the probit estimation. Similar to the figure, a higher value of the instrument, meaning the longer the time since the reform, the higher is the probability of enrolling into WP. The instrument is notably significant, evidenced by a t-value of 59.29.

Next, we use the propensities to estimate equation 6. We follow the procedure discussed in Zhou and Xie (2020), which was originally proposed in Heckman, Urzua and Vytlacil (2006). The procedure can be summarized as follows:

1. Regress Y, X, and $X\hat{P}$ on $\hat{P}$ to obtain the residuals e_Y, e_X, and $e_{X\hat{P}}$. We use a local linear regression for a non-linear fit. This method estimates a regression locally around a target point, denoted by x_0. It then applies a weighting function, $W_{i0} = W(x_i, x_0)$, to assign lower weights to observations, x_i, that are distant from x_0. The least squares method is then applied with the weights generated by this function.

2. Estimate a standard linear regression of e_Y on e_X and $e_{X\hat{P}}$ to obtain the residual e_Y^*. The coefficient corresponding to e_X provides an estimate of β_0, while the coefficient for $e_{X\hat{P}}$ provides an estimate of $\beta_1 - \beta_0$.

3. Fit a local quadratic regression of e_Y^* on $\hat{P}$ to estimate both $K(P)$ and $K'(p)$. The process for estimating a local quadratic regression mirrors that of step 1, but utilizes a quadratic function instead of a linear one.

4. Finally, the estimate of the MTE is obtained using equation 7, incorporating parameters from the preceding steps:

$$M\hat{T}E(x, u) = (\hat{\beta}_1 - \hat{\beta}_0)x + \hat{K}'(u) \tag{8}$$

The results of the MTE analysis are presented in section 5. Next, we discuss the other approach, namely a simple comparison before and after the reform.

4.2 Comparisons before and after the reform: A Regression Analysis

In this section, we present our before-and-after comparison analysis. We define $Y(0)$ as the potential outcome for an individual in the DI program, and $Y(1)$ represents the potential outcome for an individual admitted to the WP program. Because we do not observe the same person being admitted to both WP ($Y(1)$) and DI ($Y(0)$), we leverage the 2013 reform for our analysis. Let D be a binary dummy variable, with $D = 0$ for individuals before 2013 and $D = 1$ for those after 2013. This designation allows us to compare individuals assigned to the DI program before the reform with those admitted to the WP program after the reform. Next, X captures observed covariates relevant to our analysis. To identify the causal effect of admitting individuals to WP instead of DI, we adopt the Conditional Independence Assumption (CIA):

$$Y(0), Y(1) \perp D|X \tag{9}$$

The equation states that the potential outcomes $Y(0)$ and $Y(1)$ are conditionally independent of the assignment D, given the covariates X. In essence, we assume that, conditional on the observables X, the only distinction between the periods before and after 2013 is the admission of individuals to the WP program as opposed to the DI program.

As we have observed, our balancing on observables before and after the reform is somewhat influenced by business cycle effects. To be concrete, we run the following regression to account for these effects:

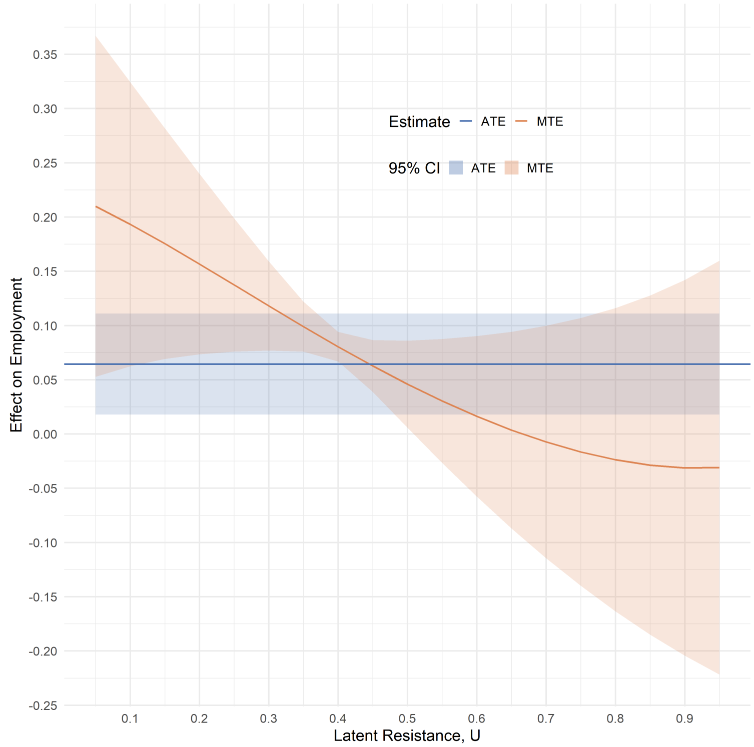
$$Y_{it} = \alpha + \beta X_{it} + \delta D_{it} + e_{it}, \quad t = -2, -1, 0, 1, 2, 3, 4. \tag{10}$$

In this equation, Y_{it} represents either employment or log-income, while α, β, and δ are regression coefficients. The vector X_{it} comprises control variables such as the aggregate unemployment rate, gender, age, ethnicity, and education. The error term is represented by e_{it}, and the subscripts i and t denote an individual and the number of years into the DI/WP spell, respectively. D_{it} acts as an indicator variable, taking the value 1 if the spell begins post-reform and 0 if it begins pre-reform.

5 Results

We now move on to presenting the results of the marginal treatment effects in the next subsection and then we report the results of the before-and-after comparison.

Figure 9: Marginal Treatment Effects and Average Effect of Treatment as a Function of Latent Resistance to Workfare Program



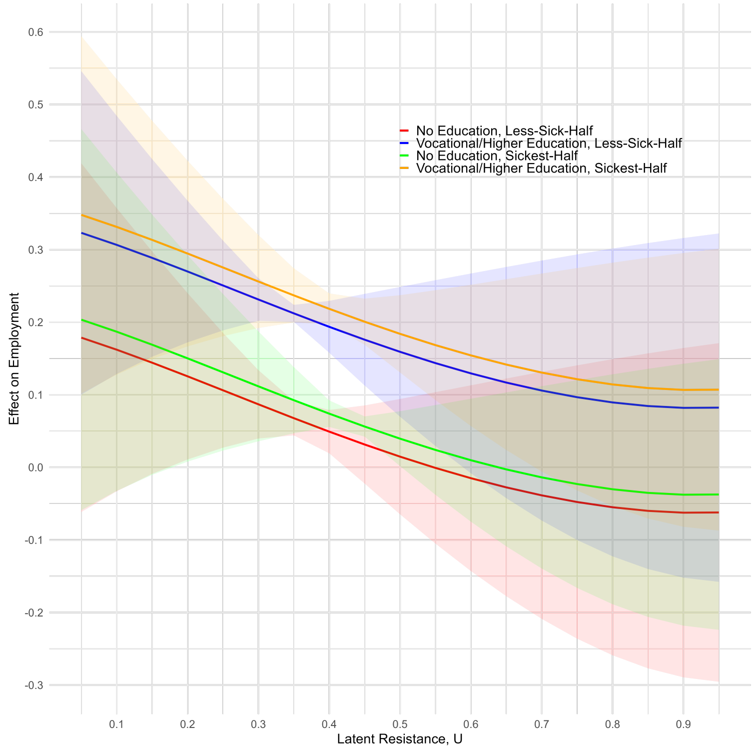
Note: The figure shows the Marginal Treatment Effects in relation to the latent resistance to WP, as well as the average effect of the treatment, both evaluated at the mean value of observable characteristics. The effect fluctuates considerably depending on the level of latent resistance – ranging from approximately a 20 % rise in employment for those with low resistance, to a slight negative effect for those who have a high resistance to WP.

5.1 Marginal Treatment Effects

In figure 9, we present the MTE as a function of latent resistance to WP, alongside the ATE. Our continuous instrument shifts the propensity across the entire unit interval; thus, our study population does not consist of individuals at the margin of entry into the program, as in Maestas, Mullen and Strand (2013), but rather the average enrollee. Consequently, we anticipate that our results will differ from those in Maestas, Mullen and Strand (2013), since our analysis focuses on individuals at the extensive margin, whereas their study examines those at the intensive margin.

From the figure, we observe strong evidence of heterogeneous employment effects from the program. Those with the lowest resistance to the WP experience a large positive effect from the reform, whereas those with high resistance exhibit zero or negative effects. Thus, the WP may be beneficial only for a relatively small and resourceful group of individuals, resulting in an overall effect, the ATE, that is close to zero. This contrasts sharply with the findings of Maestas, Mullen and Strand (2013), who report an ATE of about 30 percent.

Figure 10: Marginal Treatment Effects as Functions of Latent Resistance to a Workfare Program, Disaggregated by Health and Education



Note: The figure shows the Marginal Treatment Effects in relation to the latent resistance to WP, as well as the average effect of the treatment, both evaluated at the mean value of observable characteristics disaggregated by health and education. The shaded areas represent the 95 % confidence intervals. The effect fluctuates considerably depending on the level of latent resistance – ranging from approximately a 30 % increase in employment for those with low resistance and education to a slight negative effect for those with high resistance to WP and no education.

Thus, whether one analyzes a sample on the extensive or intensive margin plays a crucial role in determining the estimated effects.

To elaborate further on the point of whether one analyzes on the extensive or intensive margin, we display on figure 10 the MTEs for our four distinct groups of enrollees, cross-classified with respect to health and educational attainment.

Not surprisingly, we find large positive effects for individuals with the highest educational attainment, regardless of health and with low resistance to WP. On the other hand, for those with no educational credentials, only a small subset has positive effects.

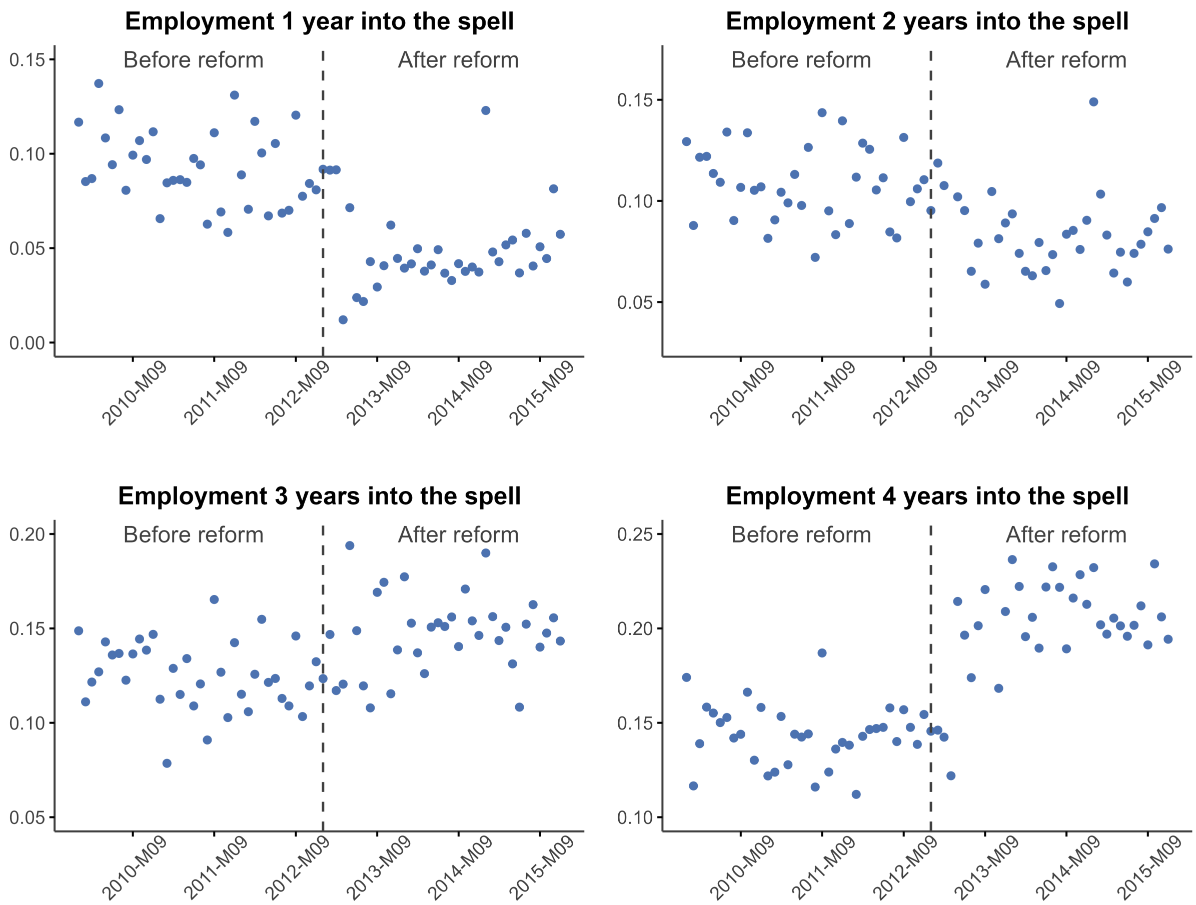
5.2 Comparisons before and after the reform: A Regression Analysis

In this section, we present the other methodological approach by simply comparing the group before and after the reform. Our control group comprises DIs prior to the reform, while the treatment groups include DIs and WPs post-reform. Using differences between the control and treatment groups, we derive Intention-To-Treat (ITT) estimates. This section offers two sets of results. First, we compare raw data between the control and treatment groups. Then, we present regression estimates that analyze the effects of the reform, accounting for factors such as business cycle variations, gender, age, ethnicity, and education. Figure 11 displays differences in employment outcomes both before and after the reform, merging subsidized and regular employment due to the negligible transitions into regular employment. We will delve deeper into this point when presenting the regression outcomes. The figure illustrates the employment fraction at one, two, three, and four years into the DI/WP spell based on the month of program enrollment. The figure reveals several notable observations. Initially, there's a clear lock-in effect post-reform. DIs and WPs entering the program after the reform (to the right of the vertical line) demonstrate lower employment one year into their DI/WP spell than those before the reform (to the left of the vertical line). However, this trend evolves. By the fourth year (as seen in the second panel), post-reform employment surpasses pre-reform levels. Thus, the reform appears to have a short-term negative effect on employment, but a positive long-term impact.

To assess how employment effects influence income, we present the log-incomes of DI and WP recipients both before and after the reform at one, two, three, and four years post-enrollment into the DI or DI/WP spell, as depicted in figure 12. The figure clearly indicates that the reform leads to an income reduction for both DI and WP recipients. Throughout the duration of the DI/WP spell, log-incomes after the reform remain consistently below those before the reform, despite the positive employment impacts of the reform. This initial drop in income can be ascribed to the reduced benefit levels under the WP in comparison to the DI. Nonetheless, this income gap narrows over time within the DI/WP spell, thanks to the employment effects of the reform.

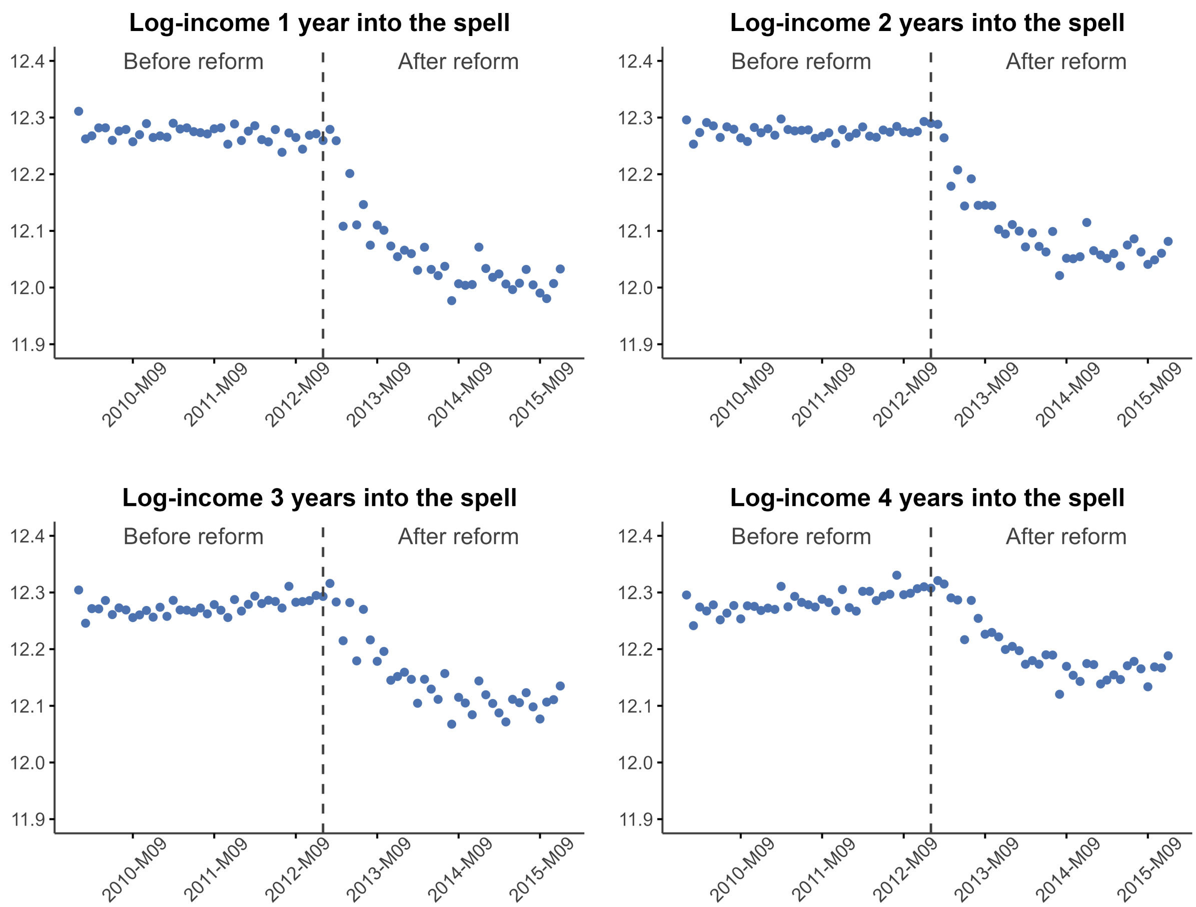
In the next subsection, we will demonstrate that these results remain robust when controlling for observables.

Figure 11: Employment for DIs and DIs/WPs, One, Two, Three, and Four Years into the DI/WP Spell Depending on Month of Enrollment



Note: The figure shows employment before and after the reform for individuals on DI and WP, at one, two, three, and four years into the DI/WP spell. The control group consists of DIs prior to the reform, while the treatment group includes DIs and WPs post-reform. Initially, individuals after the reform exhibit lower employment rates compared to those before the reform. However, over time, the post-reform group surpasses the pre-reform group in employment rates, indicating a positive long-term and negative short-term employment effect of the reform.

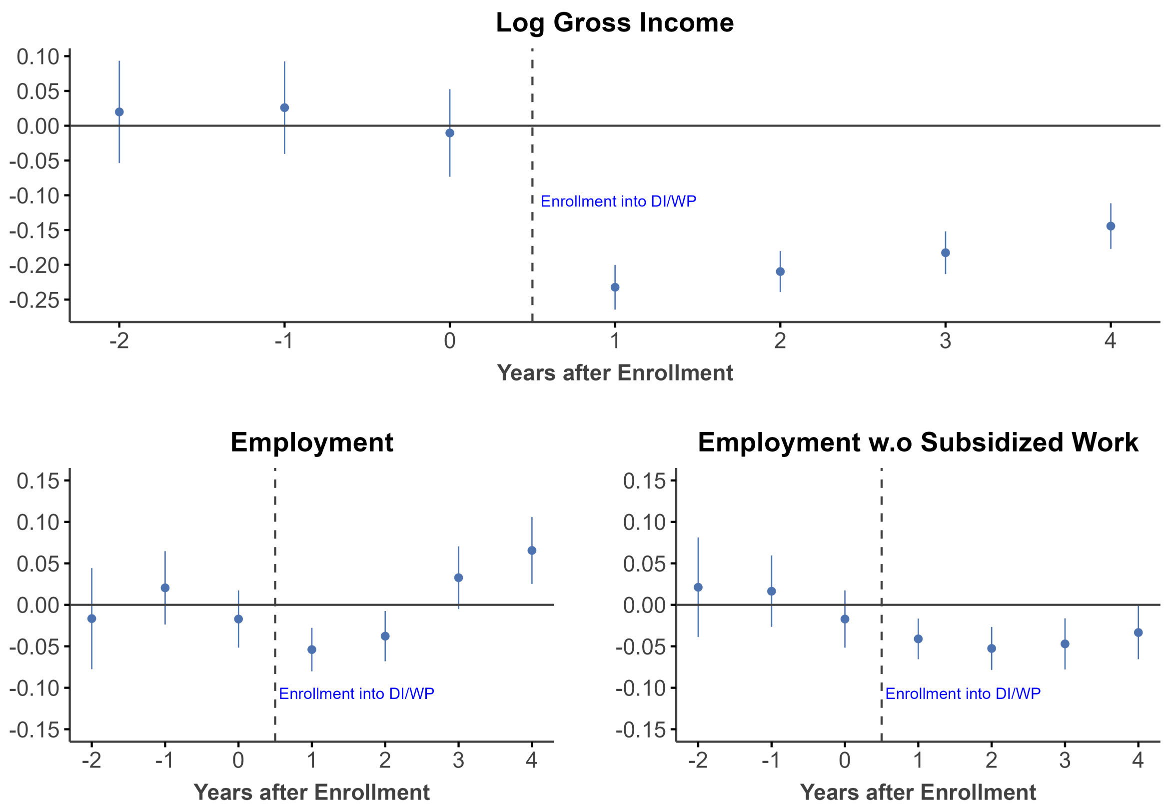
Figure 12: Log-income for DIs and DIs/WPs, One, Two, Three, and Four Years into the DI/WP Spell Depending on Month of Enrollment



Note: The figure displays the log-income for DI and DI/WP recipients both before and after the reform, from one to four years into the WP/DI spell. In all years, the log-income after the reform is lower because of the reduced benefit for WP recipients compared to DI recipients. However, over time, the decline in income starts to level off as WP recipients transition into employment.

First, we demonstrate differences in the effects of the DI/WP reform. We present the results overall and then break down the results into the four groups presented in table 3.

Figure 13: Difference in Log-income and Employment (with and without Subsidized Work) Between the Pre-Reform and Post-Reform Groups by Years after Enrollment into the DI or DI/WP Spell



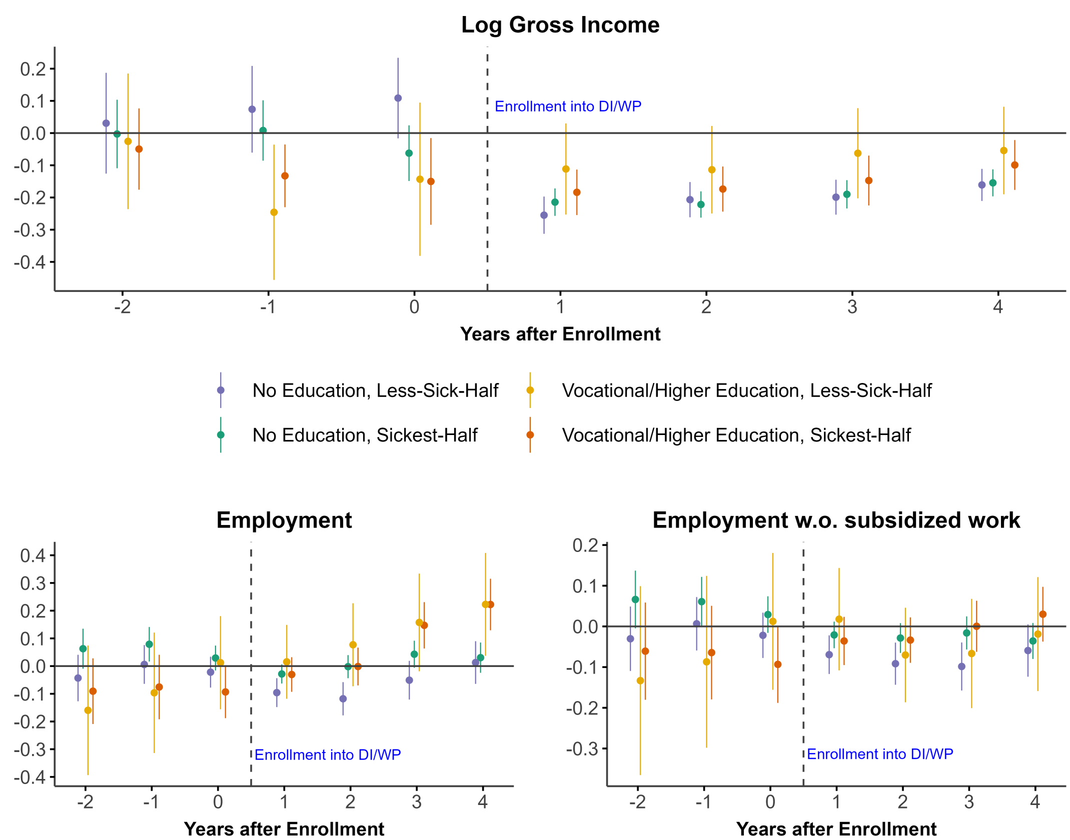
Note: The figure displays the difference in log-income and employment (with and without subsidized work) before and after the reform by years following enrollment into the DI or DI/WP spell. In every year, the log-income after the reform is lower because of the reduced benefits for WP recipients compared to DI recipients. However, as time progresses, the decline in income begins to wane as WP recipients transition into employment.

In figure 13, we depict the effect of the reform on log gross income, overall employment, and employment net of subsidized employment. The figure displays data for up to two years prior to entering DI/WP ($t = -2, -1, 0$) and up to four years after entering DI/WP ($t = 1, 2, 3, 4$). Prior to entering DI/WP, there are no statistically significant differences between enrollees before and after the reform, indicating no distinct selection effects into DI/WP due to the reform. However, differences emerge post-enrollment, with enrollees after the reform having lower income and lower employment net of subsidized employment. For income, there is a 25 percent short-term income loss and a 15 percent long-term income loss. Concerning overall employment, enrollees after the reform exhibit lower employment rates immediately upon entering DI/WP. This trend reverses after two years, where enrollees post-reform demonstrate higher overall employment. Since enrollees after the reform do not show increased employment net of subsidized employment, the higher overall employment rates in years 3 and 4 are attributable to subsidized employment.

In summary, the reform increases the likelihood of subsidized employment in the long run but also results in lower gross income.

Subsequently, we break down our findings from figure 13 across our four distinct groups of enrollees based on health and education, as represented in figure 14.

Figure 14: Difference in Log-income and Employment (with and without Subsidized Work) Between the Pre-Reform and Post-Reform Groups by Years after Enrollment into the DI or DI/WP Spell, Disaggregated by Health and Education



Note: This figure shows the differences in log-income and employment before and after reform, disaggregated by health and education. The most educated and healthiest group sees no income loss after reform, while others experience an income drop. Positive employment effects are evident in the healthiest, most educated group, while other groups show varied impacts.

From figure 14, we find clear evidence of the reform's heterogeneous effects. In terms of income, the most educated and healthiest group does not experience any income loss associated with entering DI/WP before and after the reform. However, for all other groups, the reform leads to an income loss after enrollment. The healthiest and most educated group also shows positive overall employment effects from the reform, whereas the other groups either experience a short-term employment loss or see no change due to the reform. Referring back to table 3, the groups with notable effects are also the smallest. Specifically, they make up only about 23 percent of the sample, while the remaining 77 percent of the sample shows no effect.

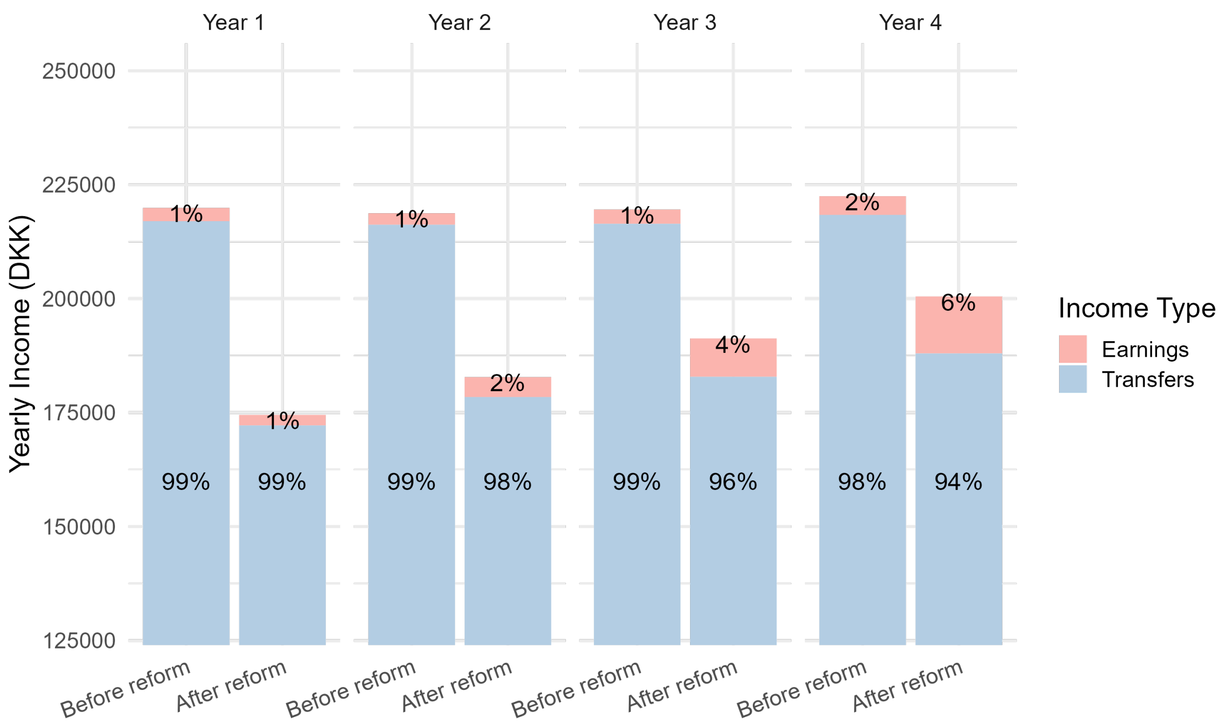
5.3 Decomposition of Income

In the previous section, we demonstrated no effects on regular employment and only small effects on subsidized employment for the after reform group. These effects are primarily driven by the healthiest and most educated group. Despite the small effects on employment, Figure 13 shows that the difference in gross income before and after the reform seems to diminish over time. In this section, we will further explain these results.

Figure 15 illustrates the changes in gross income, decomposed by income type, at 1, 2, 3, and 4 years into the spell, both before and after the reform. For the group before the reform, gross income remains roughly constant from 1 to 4 years into the DI spell, with 98-99 percent of the gross income consisting of transfer income, while 1-2 percent derives from earnings. In contrast, for the group after the reform, gross income increases over time. One year into the DI/WP spell, gross income comprises 99 percent transfers and 1 percent earnings. Subsequently, both earnings and transfers begin to increase. Despite the increase in earnings resulting from individuals entering subsidized employment, the majority of gross income still originates from transfer payments. By the fourth year, gross income is made up of 94 percent from earnings and 6 percent from transfers.

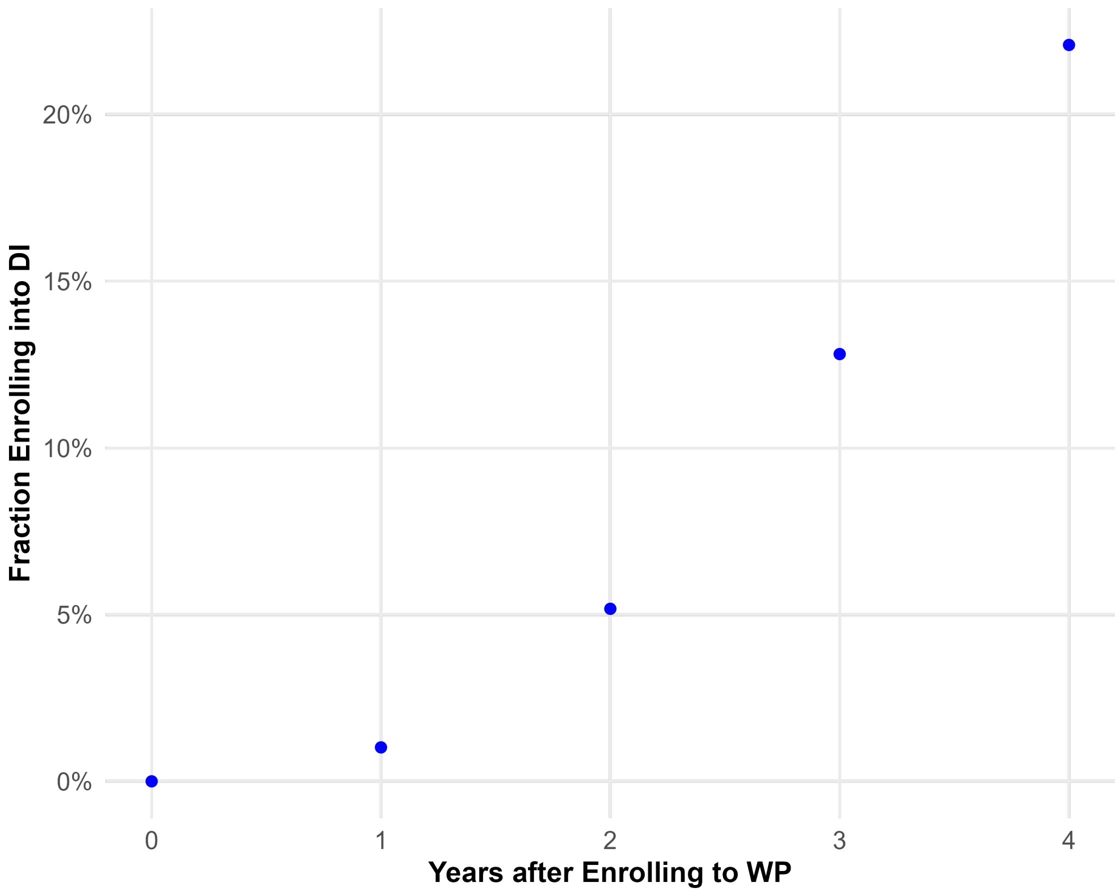
The main reason why transfer income increases over time and still dominates as the type of income in year 4 is due to a significant fallback into the DI program from WP. As people in the WP program fail to find work, they eventually get admitted to the DI program. This phenomenon is evidenced in Figure 16, which illustrates the fraction of those initially admitted to the WP after the reform but who ultimately end up in the DI program. One year into the spell, only 1 percent of those admitted to the WP program are admitted to the DI program. However, by year 4, this fraction grows to over 20 percent.

Figure 15: Decomposition of Income between the Pre-Reform and Post-Reform Groups by Years after Enrollment into the DI or DI/WP Spell



Note: This figure illustrates the evolution of gross income from 1 to 4 years into the spell, both before and after the reform. Before the reform, gross income remains constant, with the majority of it comprising transfer payments. Over time, the group after the reform exhibits an increase in both earnings and transfer income, although transfer payments continue to be a significant source of income throughout.

Figure 16: Transition of Individuals from the WP to the DI Program Over Time



Note: This figure shows the fraction of individuals initially admitted to the WP who are subsequently admitted to the DI program, increasing from 1 % in the first year to over 20 % by the fourth year, indicating a significant shift back to the DI program due to challenges in finding employment.

6 Conclusion

We analyzed the effect of a Danish reform that changed the structure of DI such that the majority of DI applicants were streamed into a WP program. The aim of the WP is to stream more DI applicants back into employment by offering job search assistance, by testing the employability of the applicants, and finally, by offering access to subsidized employment.

We find only modest positive effects of the program on employment. These effects are completely driven by entries into subsidized employment. As a result, the program has considerable negative effects on income. Entrants into WP have 15 percent lower income four years after enrollment into WP compared to enrollment into DI. Additionally, for those failing to find subsidized employment, we observe a significant fallback into DI. Thus, four years after enrolling in WP, more than 20 percent will eventually end up in the DI program anyway.

When we break applicants down by education and health status, we find that the employment effects are heterogeneous, especially by education and to a much lesser degree, health. Educated individuals gain from the program in terms of access to subsidized employment, whereas individuals with no education experience an initial negative employment effect. Disparities in health only matter for individuals with no education.

To further elucidate our findings of heterogeneous treatment effects, we estimate marginal treatment effects using the propensity of enrollment over time as an instrument. We find that less than 40 percent of recipients benefit from the reform in terms of employment. Upon further disaggregation based on observable characteristics, we find that for the least educated and least healthy part of the group of disability recipients (half of the population), less than 20 percent benefit.

Our findings support two conclusions. First, most disability recipients have very low labour market prospects on the ordinary labour market and are not capable of maintaining standard employment at all. Second, when using policy reforms that change the incentives or employment prospects for the marginal entrant into disability benefits, one may observe overly large and unrepresentative effects.

References

Bingley, Paul, Nabanita Datta Gupta, Michael Jorgensen, and Peder Pedersen. 2014. "Health, Disability Insurance and Retirement in Denmark." National Bureau of Economic Research Working Paper 20114.

Bjorklund, Anders, and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-selection." *The Review of Economics and Statistics*, 69(1): 42–49.

Bound, John. 1989. "The Health and Earnings of Rejected Disability Insurance Applicants." *The American Economic Review*, 79(3): 482–503.

Burkhauser, Richard V., and Mary C. Daly. 2022. "Lessons for US Disability Policy from other OECD countries." In *Work and the Social Safety Net: Labor Activation in Europe and the United States.* , ed. Dougla J. Besharov and Douglas M. Call. Oxford University Press.

Burkhauser, Richard V., Mary C. Daly, Duncan McVicar, and Roger Wilkins. 2014. "Disability benefit growth and disability reform in the US: lessons from other OECD nations." *IZA Journal of Labor Policy*, 3(4).

Card, David, Jochen Kluve, and Andrea Weber. 2010. "Active labour market policy evaluations: A meta-analysis." *The Economic Journal*, 120(548).

Charlson, Mary E, Peter Pompei, Kathy L Ales, and CR MacKenzie. 1987. "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation." *Journal of chronic diseases*, 40(5): 373–383.

Haller, Andreas, Stefan Staubli, and Josef Zweimüller. 2020. "Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits." National Bureau of Economic Research Working Paper 27602.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *The Review of Economics and Statistics*, 88(3): 389–432.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475.

Low, Hamish, and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive-Insurance Trade-Off." *The American Economic Review*, 105(10): 2986–3029.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review*, 103(5): 1797–1829.

OECD. 2022. *Disability, Work and Inclusion: Mainstreaming in All Policies and Practices.* Paris:OECD Publishing.

Pearson, Karl. 1901. "On Lines and Planes of Closest Fit to Systems of Points in Space." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11): 559–572.

Quandt, Richard E. 1958. "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes." *Journal of the American Statistical Association*, 53(284): 873–880.

Razzu, Giovanni, and Carl Singleton. 2016. "Gender and the business cycle: An analysis of labour markets in the US and UK." *Journal of Macroeconomics*, 47(PB): 131–146.

WHO. 2019. "Trends in maternal mortality: 2000 to 2017: estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division." *WHO Publication.*

Zhou, X., and Y. Xie. 2020. "Heterogeneous Treatment Effects in the Presence of Self-Selection: A Propensity Score Perspective." *Sociological Methodology*, 50(1): 350–385.