Does Stricter Disability Screening Lead to Improved Targeting on Long-term Health and Employment Outcomes?

Work in Progress – Do not cite*

Chiara Campana† Pierre Koning‡ Maarten Lindeboom§

January 6, 2025

Abstract

While evidence points at potentially strong effects of tightening eligibility criteria for Disability Insurance (DI), little is known about their targeting effects: are workers screened out also those with better future employment, health and mortality outcomes? Knowing that DI receipt is mostly permanent for workers, it is key to understand the permanence of impairments and their implications for the ability to work. To shed light on this, we employ Regression-Discontinuity-in-Time regressions to compare the long-term health and mortality outcomes of DI application cohorts just before and after a reform in the Netherlands. This reform led to stricter screening in the sickness period before DI application and reduced the number of applicants by 33%. Up to 18 years after application, we find persistently lower survival rates, higher medical consumption and lower employment rates of post-reform cohorts. Using detailed information on the future chronic diseases of these cohorts, we next construct a Chronic Disease Index (CDI) that explains survival rates and a Work Ability Index (WAI) that explains long-term employment rates. Inferences on changes in the CDI due to the reform show that compliers to the reform have health conditions with high expected survival rates (i.e. high CDI-scores). For employment (the WAI), however, screening effects were less substantial. This reflects the effect of reduced applications from screened out workers with (mental) health conditions, having low mortality rates and relatively limited ability to work.

JEL code : H2, I3

Keywords : Disability Insurance, Mortality, (Mental) health, Screening, Targeting

*Disclosure statement: The authors herewith state to have no relevant or material financial interests that relate to research described in this paper. The authors use confidential administrative data owned by Statistics Netherlands. The data can be accessed only by remote access connection, after filing a request directly to Statistics Netherlands (https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/aanvraag-toegang-microdata). The authors are willing to assist in doing so. An IRB approval was not mandatory for this research project, as all privacy conditions and ethical standards are warranted by the owners of the data, Statistics Netherlands.

†Vrije Universiteit Amsterdam, Ca' Foscari University of Venice. Email : c.campana@vu.nl

‡Vrije Universiteit, Amsterdam, IZA and TI

§Vrije Universiteit, Amsterdam, TI, Centre for Health Economics, Monash University, IZA.

1 Introduction

Disability Insurance (DI) enrollment for workers has been on the rise in many OECD countries (OECD, 2022). This trend is difficult to explain from worsening health conditions of individuals alone and raises questions about the targeting properties of the DI program. Do they also provide income to workers with less severe health conditions, or also to workers with weaker labor market positions (Deshpande and Lockwood, 2022; OECD, 2022)? To address this, policy reforms in many countries have aimed at tightening the eligibility criteria of DI schemes for workers unable to work ("work disability"), or for workers whose severe health conditions increase the strain from work ("health disability") – see e.g. (Benítez-Silva et al., 2014; Godard et al., 2024; Haller et al., 2024; Liebert, 2019). There is little evidence on the targeting effects of these DI reforms, particularly on how mental and physical health conditions evolve in the years after DI application. This requires detailed individual information on changes in future health conditions, mortality, and labor market outcomes of DI applicants and awardees. Since DI benefit receipt is typically long-lasting but health conditions can be temporary, one would preferably have this information over a prolonged period of time. A pertaining question therefore is how reform induced changes in the composition of health conditions of workers are related to the long-term survival probabilities and work ability of workers. This is in essence what this paper is about.

In this paper we use a reform in the Netherlands that increased the screening intensity of DI applications and examine the long-term (up to almost 20 years after DI application) evolution of health, mortality, and employment. We analyze the characteristics of DI applicants who are screened out by the reform compared to those who continue to apply after its implementation. Additionally, we examine the evolution of their health, survival, and employment outcomes, assessing how the reform impacts "work disability" and "health disability" and their relative significance in explaining changes among DI applicants.

In January 2003, the so-called "Gatekeeper Protocol" (GKP) substantially increased the application costs for DI applicants and employers. During the one-year sickness/waiting period that precedes DI applications, workers and employers had to arrange more medical checks and more work resumption efforts to increase work resumption rates. This reform did not alter the DI eligibility criteria (Koning and Lindeboom, 2015). Therefore, changes in DI enrollment due the reform can be attributed to stricter screening in the waiting period. The reform discouraged a substantial group of workers from applying for DI benefits and led to a dramatic drop in DI

enrollment rates (Godard et al., 2024). We use data on DI applications and awards from the National Social Insurance Institute (NSII) to find that the introduction of the GKP led to a drastic reduction in the inflow into the DI scheme. This reduction is largely driven by higher work resumption rates in the one-year sickness/waiting period that precedes DI applications. We extend Godard et al. (2024) by linking NSII application data with detailed administrative data on health expenditures, medical prescriptions, mortality, and employment between 2000 and 2022. This allows us to assess in detail the long-term targeting properties of the DI application process of the GKP. We characterize "compliers" to the reform, i.e. those who no longer applied for benefits due to the stricter screening in the sickness/waiting period. Changes in applicant characteristics – both in terms of long-term health and labor market outcomes – can then be interpreted as changes in self-screening or "self-targeting" of workers facing higher DI application costs.

Next to studying the long-term effects of stricter screening, we construct a Chronic Disease Index (CDI) and a Work Ability Index (WAI) reflecting changes in the "health disability" and "work disability" of DI applicant cohorts. We utilize detailed data on medical prescriptions for DI applicants in the years following their award decisions and map these prescriptions to chronic diseases. These chronic disease indicators serve as explanatory variables to estimate 10-year survival rates. However, for younger cohorts, the mortality rates are low, prohibiting reliable estimates. To address this, we adopt the recently developed procedure proposed by Danesh et al. (2024) and estimate the regression model on a subsample of applicants who have reached age 65. The resulting (predicted) Chronic Disease Index (CDI) is then used to project the long-term consequences of current chronic diseases for younger cohorts. This approach allows us to identify mortality effects that would otherwise go unnoticed if we considered the observed survival rates for the full sample of applicants.

While the CDI serves as a proxy for the "health disability" of DI applicants, we also construct a "Work Ability Index" (WAI) based on employment outcomes up to 18 years after the DI award decision. To estimate the WAI, we use the same detailed chronic disease information derived from medical prescription data, alongside age and gender, to model employment outcomes. The resulting regression estimates are used to predict the WAI. Like the CDI, the WAI enables us to assess the impact of changes in the composition of chronic diseases following the GKP reform, providing insights into the evolving relationship between health conditions and work capacity among DI applicants.

With the CDI and the WAI scores that we infer for relevant DI application cohorts, we analyze how changes in the composition of chronic diseases of applicants induced by the GKP reform have affected the number of expected life years and the ability to work. We also analyze changes in the association between the CDI and WAI. In doing so, a special interest lies in the consequences of changes in the fraction of applicants with mental health impairments. This group typically shows low mortality but potentially large and diverse effects on the earnings capacity of workers. This is very different from e.g. cancers, which have large effects on mortality but to a lesser extent on employment.

Using Regression-Discontinuity-in-Time (RDiT) specifications that leverage the cutoff date of the GKP reform, we find a statistically significant decrease in survival rates up to 18 years after the DI award decision, amounting to a reduction of 1.1 percentage point (corresponding to 12.2% of the baseline mortality rate). Notably, this effect becomes negligible and statistically insignificant when controlling for changes in the fraction of impairment types diagnosed at the time of the award decision. This indicates that shifts in impairment types account for most of the observed decline in survival rates. For applicants who survived beyond the award decision, the fraction of individuals in poor health – defined as having medical expenditures in the top quartile of the entire Dutch population – was 1.6 percentage points higher 10 years after the DI decision. This effect is confined to medical expenditures on physical conditions and does not extend to mental health treatments. Additionally, the health and mortality impacts of the GKP reform are accompanied by a decline in employment, with employment rates among post-GKP applicants approximately four percentage points lower than those of pre-GKP applicants.

Based on the Chronic Disease Index (CDI) and Work Ability Index (WAI), we infer that the GKP reform led to stronger screening based on health than on work ability. This primarily reflects a reduction in the share of female applicants and workers with mental impairments – groups characterized by lower mortality rates but not higher employment rates. Applicants who complied with the reform and were screened out exhibited CDI outcomes that were 32% of a standard deviation higher than those who continued applying for DI (the "stayers"). Similarly, compliers had an average WAI score that is 23% of a standard deviation higher than that of stayers. When decomposing the decreases of the CDI and WAI attributed to the GKP reform, we find that only about a quarter of these changes can be explained by an increase in the prevalence of chronic diseases. Screening based on demographics played a more significant role: the decline in the CDI is largely attributable to the reduced share of female applicants, while

the decrease in the WAI is strongly driven by an increased proportion of older workers nearing retirement.

Our paper contributes to two strands of literature. First, it adds to the growing body of research examining the selection effects of stricter screening for DI applications or tightening eligibility criteria for DI benefits (Ahammer and Packham, 2023; De Jong et al., 2014; Haller et al., 2024; Johansson et al., 2014; Markussen et al., 2017), as well as changes in application costs (Deshpande and Li, 2019; Godard et al., 2024).[1] A key finding in this literature is that stricter eligibility criteria can significantly reduce the inflow into DI schemes. This reduction stems from two factors: the direct impact of stricter screening and the impact self-screening, where workers choose not to apply due to the perceived difficulty of meeting the tightened criteria. Moreover, a notable proportion of workers who are screened out subsequently return to the labor market, indicating improved targeting efficiency and minimal fiscal spillovers to other welfare programs (Haller et al., 2024). This highlights the potential for marginal workers, who might otherwise receive benefits, to re-enter the workforce. While these studies focus primarily on short-term employment effects, a comprehensive evaluation of targeting efficiency requires an analysis of long-term health and employment outcomes following DI award decisions. Our study aims to address this gap by examining both health and employment trajectories over an extended period.

Second, our paper contributes to the emerging literature on the causal effects of DI benefit levels and receipt on mortality (Black et al., 2024; Garcia-Gomez and Gielen, 2017; Gelber et al., 2023). DI programs compensate for the income loss associated with adverse health conditions, potentially preventing further health deterioration and reducing mortality rates. At the same time, receiving DI benefits may weaken work incentives, which could influence employment and, indirectly, subsequent health and mortality outcomes. Using exogenous variation in the assignment of medical examiners for SSDI applicants, Black et al. (2024) find that mortality increases for applicants at the margin of being allowed benefits but decreases for infra-marginal applicants who receive benefits. Similarly, Gelber et al. (2023) demonstrate that increases in SSDI benefits lead to lower mortality, with an estimated elasticity of mortality with respect to DI income ranging from 0.6 to 1.0. In contrast, Garcia-Gomez and Gielen (2017) show that the loss of DI income following medical reassessments raises mortality among women but decreases it among men. However, mortality is a coarse measure of health, especially for mental health conditions, which account for roughly one-third of DI inflows in most OECD countries (OECD,

[1] See Godard et al. (2024) for an extensive review.

2010). This underscores the need to examine health outcomes more comprehensively to better understand the full impact of DI programs.

In what follows, we will explain the GKP reform in more detail in Section 2 and discuss the administrative data sources we use in Section 3. Section 4 presents the empirical strategy. The estimation results are in Section 5. We conclude in Section 6.

2 The Gatekeeper Protocol

Workers in the Netherlands are subject to a waiting period before they can apply for DI benefits.[2] During the waiting period, permanent workers continue to receive full wage payments, while temporary workers are eligible for Sickness Insurance (SI) benefits, amounting to 70% of their wage earnings from the moment their contract ends. At the time of the GKP reform in 2003, the sickness waiting period amounted to 12 months.[3] At the 12-month claims assessment, a medical examiner from the National Social Insurance Institute (NSII) evaluates the worker's impairments, while a vocational expert identifies the highest-earning jobs to determine the worker's earnings capacity.[4] DI benefits were set at 70% of pre-disability earnings for individuals classified as fully disabled, having lost at least 80% of their earnings capacity. Partial DI benefits covered 70% of the earnings capacity lost between 15% and 80% of pre-disability earnings. Applicants with a disability degree below 15% were ineligible for DI benefits and were rejected.

The introduction of the GKP reform is widely regarded as the most effective measure in the Netherlands for reducing DI inflow and, consequently, DI enrollment (Koning and Lindeboom, 2015). While the reform did not alter the DI eligibility criteria, it significantly increased obligations and costs for both employers and workers during the waiting period.

With enrollment rates reaching 12% of the insured population, there was widespread agreement that insufficient screening of health conditions occurred during the waiting period before DI applications. To address this, the GKP reform introduced a series of mandatory steps for both employers and workers to qualify for DI application after 12 months of sickness. The GKP system has remained fundamentally unchanged since its implementation.

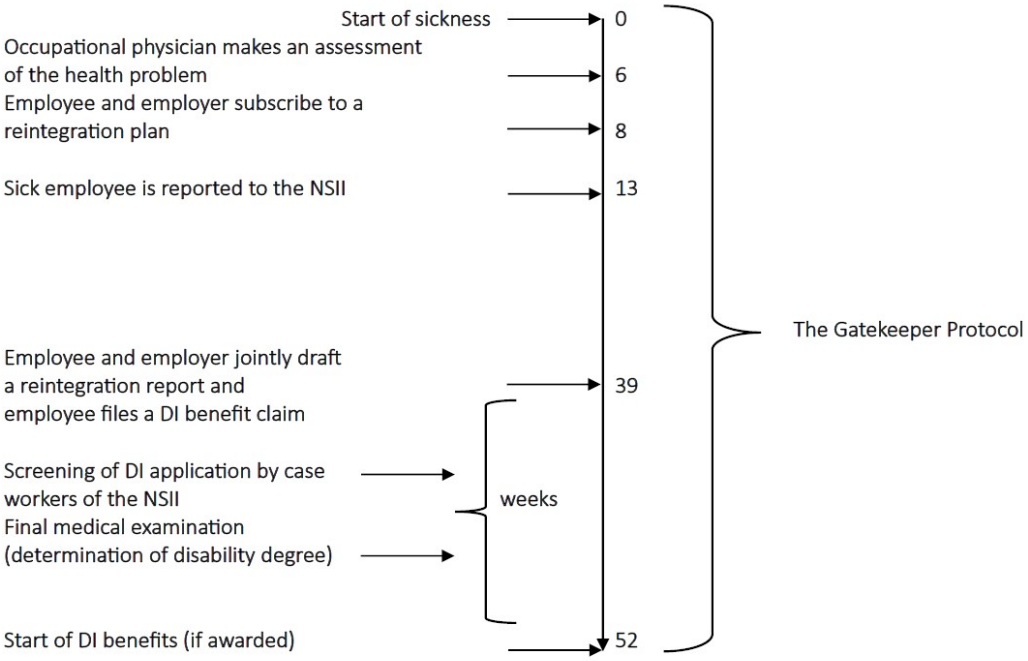
[2]This section describes the main elements of the Dutch institutional settings for sickness, DI applications and DI benefits at the time of the GKP reform. For more extensive information and related reforms, we refer to Koning and Lindeboom (2015), Godard et al. (2024) and Koning and Vethaak (2021).

[3]In 2004 the waiting period was extended to two years Koning and Lindeboom (2015). This effectively shifted the inflow of new DI awards of 2005 to 2006, and also mechanically lowered inflow rates since 2006.

[4]Before the WIA reform of 2006, the medical examiner did not assess the permanence of impairments at the time of the award decision.

Figure 1 illustrates the GKP process timeline, including key milestones: contacting an occupational doctor after six weeks of sickness, drafting a reintegration plan after eight weeks, timely reporting of sickness to NSII after 13 weeks, and submitting a completed report after 39 weeks. Under the GKP, both workers and employers are required to formulate and actively implement the reintegration plan. During the sickness period, employers may only dismiss workers who fail to collaborate. Additionally, employers risk being obligated to continue wage payments beyond one year if they do not comply with GKP requirements.

Figure 1: Timeline of one-year sickness period before DI applications



Given the substantial effort required to meet the obligations imposed during the one-year sickness period, compliers to the GKP reform can be characterized as workers who were discouraged from reporting sick. These workers may have either returned to work or deemed the obligations too burdensome relative to their health conditions. In the latter case, this could lead to what Parsons (1991) refers to as "perverse screening" based on health conditions. However, using data on hospitalizations prior to DI applications, Godard et al. (2024) find no evidence of such perverse screening effects. Their study reveals a dramatic reduction in DI applications – nearly 40% for permanent workers – with the most pronounced declines observed among women and applicants with hard-to-verify conditions, such as mental and musculoskeletal disorders.

For hard-to-verify conditions, such as burnout and depression, identifying the causes and assessing the severity is more challenging. While the GKP protocol may be effective for these cases, it also imposes significant additional effort on both workers and employers. In contrast, the GKP's impact may be more limited for applicants with easy-to-verify conditions. When work resumption is feasible, the protocol's obligations are relatively easy to implement. For severe cases, such as advanced-stage cancers or cardiovascular diseases, where the likelihood of returning to work is minimal, the added value of the GKP is similarly constrained.

Given the relatively high proportion of workers with hard-to-verify conditions who were discouraged from applying due to the GKP reform, an important question arises about the reform's targeting effects. From a "health disability" perspective, this can be explored by examining the long-term health outcomes of workers who complied with the reform. The onset or worsening of health conditions can lead to reduced well-being and a loss of expected residual life years in either good or poor health. When hard-to-verify diseases are temporary and associated with lower mortality than easy-to-verify diseases, the GKP reform may have improved targeting through self-screening in the sickness period. For instance, workers with temporary and milder mental conditions may have been discouraged from applying. However, if mental health conditions have a lasting impact on well-being, discouraging these workers from applying could be considered counterproductive. In such cases, the reform might undermine the objective of supporting workers with substantial and persistent health impairments.

An alternative perspective on self-screening and targeting is based on the concept of "work disability." Rather than focusing solely on health outcomes, this approach examines how health conditions affect the ability to work among workers who no longer apply for DI benefits. From this viewpoint, the implications of self-screening for targeting are less clear-cut. As such, this renders the impact of the GKP reform on targeting largely an empirical matter.

In summary, the GKP reform did not alter the medical screening process at the DI assessment but significantly influenced application costs and behaviors during the sickness period. The effects of the reform may thus have worked in different ways, and with distinct implications both for targeting on health and employment outcomes of workers who continued to apply for DI benefits after the one-year sickness period.

3 Data

3.1 Data sources

For our empirical analysis, we use data from different administrative sources. Disability Insurance (DI) application data are provided by the National Social Insurance Institute (NSII), and individuals are matched to these through a unique identifier with Statistics Netherlands databases.[5] We use administrative data on DI applicants aged 25-65 for the period 2001-2004.[6] The dataset contains the impairment types diagnosed by the NSII (musculo-skeletal, mental, cardiovascular, nervous, respiratory, endocrine or "other" non-categorized disorders), the assessed degree of disability (0–15%, 15–25%, 25–35%, 35–45%, 45–55%, 55–65%, 65–80% and 80–100%), and the date and outcome of the award decision (awarded or rejected). We observe the award decision, but not the exact date at which the worker started his/her sickness spell. As a consequence, the delay between the application date (in week 39) and the result of the award decision (which is taken at most after two months from the application) is not observed. In our empirical analyses we follow Godard et al. (2024) and apply a "donut hole" of one month on each side of January 2003, excluding applicants whose award date cannot be attributed with certainty neither to the period before nor after the reform.

The demographics of the applicants are obtained from the Municipal Personal Records Database. It contains information on gender, date of birth, nationality, marital status, household composition, place of residence of all individuals living in the Netherlands, and mortality dates, if relevant. Annual earnings from the policy administration register are tracked from 1999 to 2022, as well as the economic status of individuals (employee, self-employed, unemployment insurance (UI), social assistance (SA), pension benefit recipient, DI benefits, pension benefits, student, other without any income).

Annual medical spending data, observed from 2009 to 2021, include yearly healthcare expenses for all residents insured under the Netherlands' statutory basic medical insurance, which is mandatory. The costs reflect reimbursements by health insurers and are categorized by type of care: general practitioner care, hospital care, prescription costs, mental health care, physiotherapy, and specialist examinations. In addition, we observe yearly medical prescription data from 2006 to 2022, detailing the reimbursement of medications covered under the statutory basic

[5] A random identification number is assigned by Statistics Netherlands from the citizens' social service number (BSN).

[6] Godard et al. (2024) also consider a (maximum) time period of four years around the moment of the GKP reform.

medical insurance. The prescriptions are classified into ATC classes (Anatomical, Therapeutic, Chemical) based on the World Health Organization's (WHO) drug classification system. Following Huber et al. (2013) and Danesh et al. (2024), we map the prescription data to 22 chronic disease conditions, as outlined in Table B1 in the appendix of this paper.

3.2 Sample selection

We merge and select our data from different sources in a number of steps. First, the sample of DI applicant cohorts we follow is restricted to the period 2001-2004 to examine two years before and after the reform. This yields a sample of 449,017 DI applicants. We exclude workers who applied for DI benefits more than once or have conflicting information, as it is unclear which application corresponds to the final award decision. Similarly, we exclude applicants with multiple records for the same award year and award date, as well as those with unknown diagnoses or award decisions. This initial filtering reduces the sample to 355,479 individuals. Additionally, we select applicants aged 25–65 who were formally employed and not receiving DI benefits in the year prior to the award decision. This further reduces the sample to 290,813 workers. An individual is considered employed if they worked for at least six months (not necessarily consecutive) in a given year. Consequently, the sample includes part-time and temporary workers.

Table 1 shows descriptive statistics of DI applicants before and after the GKP reform. The comparison between columns (1) and (2) indicates that the reform was accompanied by changes in the characteristics of applicants. The difference in means is presented in column (3). After the GKP reform, applicants tend to be older, are more likely to be male and non-native, and they have lower annual earnings in the year of the award decision. In the year before the application, post-GKP applicants have slightly higher earnings than pre-GKP applicants. Regarding impairments, we observe a substantial decline in the share of applicants with mental disorders after the reform; this dropped from 37% to 28%. Finally, post-GKP applicants have higher physical and mental health care expenditure.

Table 1: Summary Statistics of pre- and post-reform DI applicant cohorts (2001-2004)

	2001-2002	2003-2004	Diff-in
	Pre-Gatekeeper Applicants	Post-Gatekeeper Applicants	-means
	(1)	(2)	(3)
Demographics			
Age	43.07	43.72	0.644***
	(9.65)	(9.71)	(0.006)
Male	0.40	0.47	0.071***
	(0.49)	(0.50)	(0.001)
Native	0.80	0.78	-0.021***
	(0.39)	(0.41)	(0.001)
Labour market outcomes[(a)]			
Employed (year award decision)	0.82	0.76	-0.058***
	(0.38)	(0.43)	(0.001)
Gross earnings (2015 euros)			
– in year of award decision	25.859	24.347	-1481.75***
	(23.303)	(20.036)	(87.01)
– in year before award decision	30.300	30.907	607.14***
	(21.369)	(18.182)	(79.56)
Health status (fractions)			
Impairments (grouped)			
– Musculo-skeletal	0.29	0.28	0.003***
	(0.45)	(0.45)	(0.001)
– Mental	0.37	0.28	-0.086***
	(0.48)	(0.45)	(0.001)
– Cardiovascular	0.04	0.06	0.012***
	(0.20)	(0.23)	(0.001)
– "Other" disorders	0.30	0.38	0.08***
	(0.45)	(0.48)	(0.001)
Health expenditure in 2015 (euros)			
– Total	4.507	5.208	-700***
	(12.382)	(14.146)	(54.02)
– Physical	3.759	4.201	441**
	(9.954)	(11.160)	(43.10)
– Mental	748	1007	259***
	(7.093)	(8.415)	(31.42)
Total number of observations	191,570	99,243	290,813

Notes:[(a)] Our sample considers yearly observations and excludes individuals not employed in the previous year. We consider workers aged 25-65. Labour-market characteristics are measured at the yearly level.

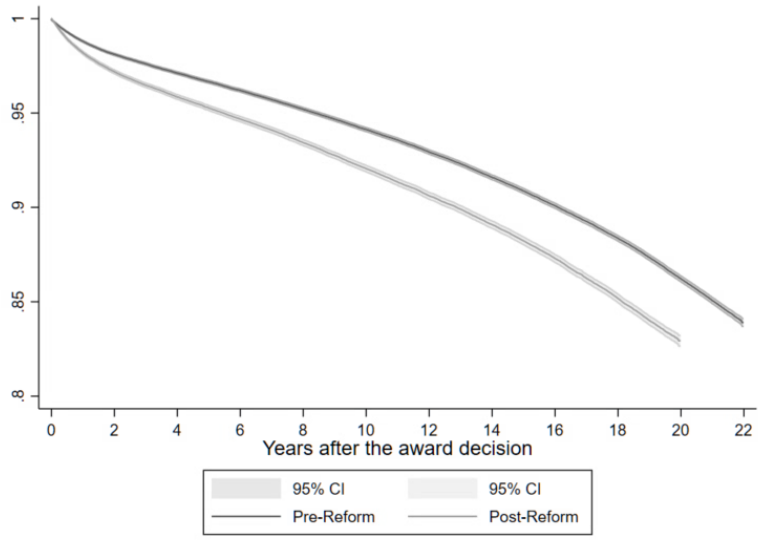
3.3 Long-term outcomes

Survival rates

To shed light on the survival rates of the applicants before and after the reform, we present the Kaplan-Meier curves of the corresponding cohorts (2001-2002 versus 2003-2004). Figure 2 indicates higher mortality rates among workers who applied for DI after the reform, with a difference of approximately three percentage points in most elapsed years since application. While this suggests that compliers to the reform had substantially lower mortality rates than

the "stayers", this difference could also partially reflect trends unrelated to the GKP reform. To explore potential drivers of this difference between cohorts, Figure A1 in the appendix presents survival estimates stratified by gender, age cohort, degree of disability, and impairment types. Notably, when conditioning on impairment types, the differences in survival rates between pre-GKP and post-GKP application cohorts become much smaller. This finding highlights the significant role of increased self-screening based on health conditions.

Figure 2: Kaplan-Meier survival estimates of pre- (2001-2002) and post-reform (2003-2004) DI applicant cohorts



Medical consumption

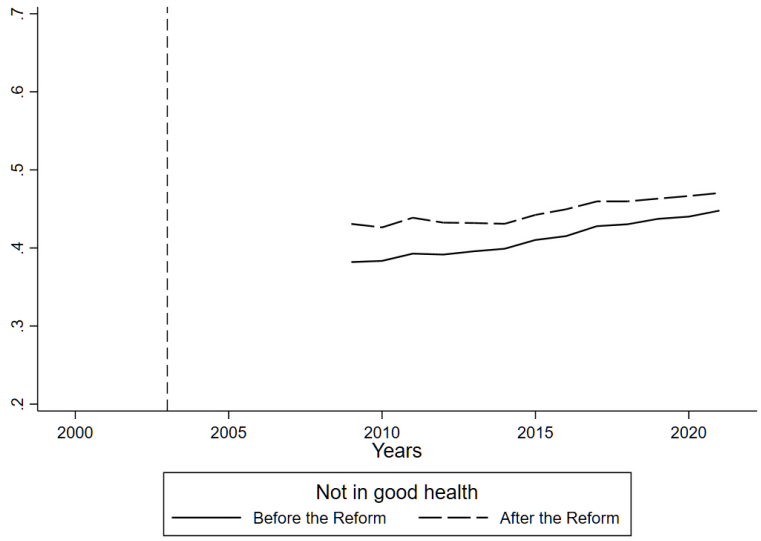
We proceed by examining changes in the composition of applicants based on their health outcomes. To begin, we use the Medical Consumption data from 2009 to 2021 as a proxy for the overall health status of applicants. Specifically, we calculate the distribution of annual medical consumption for all Dutch citizens and classify applicants in the top 25th percentile as "not in good health". For physical health, we apply the same approach using total physical health expenditures. For mental health, we use a binary indicator that equals one if an individual incurs any mental health costs in a given year and zero otherwise.[7]

Figure 3 compares the shares of the highest spenders in annual total health expenditures for both cohorts. Since we compare yearly data, the elapsed spell since the award decision is two years

[7]Note that any expenditures on mental health consumption correspond to the 75^{th} percentile (or higher percentiles), rendering the same approach infeasible.

shorter for the post-GKP cohort.[8] As expected, the share of DI applicants classified as "not in good health" is substantially higher than the 25% threshold applicable to all Dutch citizens. Furthermore, applicants after the GKP reform exhibit worse health conditions compared to those before the reform, with a maximum differential of three percentage points. Panels (a) and (b) of Figure A6 in the appendix also demonstrate that spending levels are higher for both physical and mental health expenditures among applicants. Notably, the fraction of applicants receiving mental health treatments ranges between 7% and 12%. This is significantly lower than the share of applicants diagnosed with mental impairments as the primary impairment at the time of the award decision by the NSII. The discrepancy arises because registered mental health treatments are typically associated with more severe cases of mental health conditions.

Figure 3: Annual fractions with not in good health status (physical and mental) among DI survivor applicant cohorts before (2001-2002) and after (2003-2004) the GKP reform



Medical prescriptions and chronic diseases

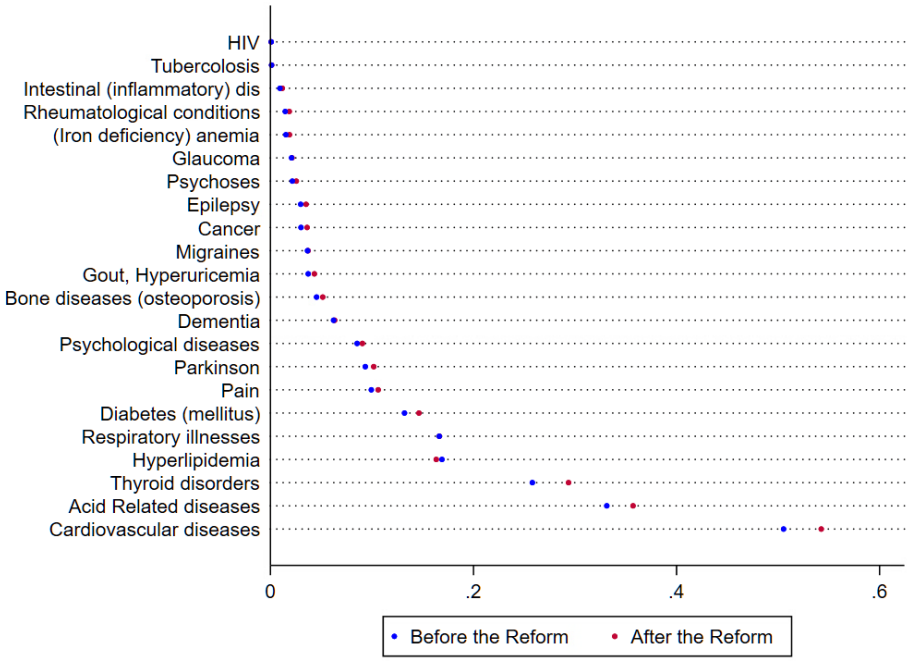
With the data at hand, long-term health outcomes of the DI applicant cohorts can also be tracked by the use of medical prescriptions. Table B1 in the appendix to this paper shows how these medical prescriptions can be mapped to chronic diseases (Huber et al., 2013).[9] Following from this, Figure 4 shows the prevalence of chronic diseases at the age of 65 in the pre- and post-

[8]Note that this difference in elapsed years is not problematic in the context of RDiT models where we zoom in to the month that the GKP started.

[9]Note that we map the chronic diseases with four-digit code, whereas Huber et al. (2013) uses five-digit codes.

GKP reform application samples. As such, the figure is based on older applicants that turn 65 at any year between 2006 and 2022.[10] For nearly all chronic disease types with prevalence rates exceeding 5%, we observe increases in prevalence among cohorts applying after the reform. The most notable differences between the pre- and post-reform cohorts occur in chronic diseases that are also the most common, such as thyroid disorders, acid-related conditions, and cardiovascular diseases. This pattern suggests that workers who continued to apply for DI after the reform tended to have more severe health conditions compared to those who stopped applying. However, these differences may also partially reflect the influence of other underlying time trends.

Figure 4: Prevalence of chronic diseases at the age of 65 in DI application sample: pre-GKP (2001-2002) and post-GKP (2003-2004) cohorts



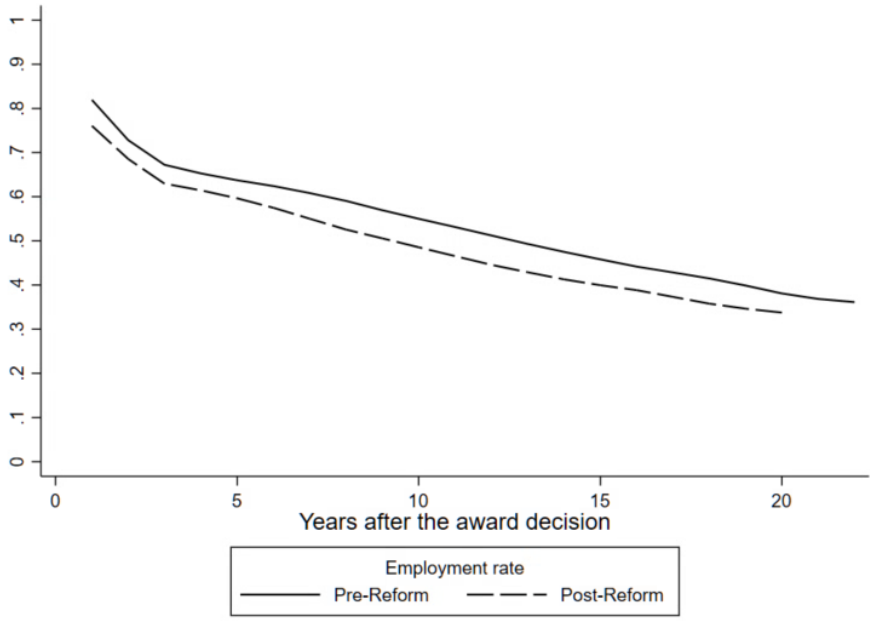
Employment rates

To gauge changes in the ability to work of cohorts before and after the reform, Figure 5 displays employment rates by the elapsed time since application. For both cohorts the largest drop in the employment rate is within the first four years after the award decision, amounting to around 15 percentage point. After 22 years, the (baseline) employment rate of the pre-GKP cohorts has decreased from 80% to around 35%. This mirrors both the (direct) effect of the loss of

[10] We restrict the subsample to those who turn 65 for reasons that will become clear later in Section 4, where we discuss the construction of the Chronic Disease Index (CDI)

employment in the first years after the award decision and the later – and more gradual – effect of retirement of applicants. We observe substantial differences between the employment rates, amounting to around five percentage point in most years after the award decision.

Figure 5: Employment rates of pre-GKP (2001-2002) and post-GKP (2003-2004) cohorts



4 Empirical Strategy

The GKP reform was implemented in April 2002, and started affecting the new DI applicant cohorts since January 2003. Our empirical strategy exploits this sharp discontinuity in January 2003 using a Regression Discontinuity-in-Time (RDiT) design (Hausman and Rapson, 2018). In what follows, we explain our strategy to estimate the effects of the stricter screening. We first compare DI applications before and after the GKP reform. Secondly, we estimate changes in the composition of DI applicants, analyzing survival rates, health conditions, and labor market outcomes of individuals. We use these results and the changes in the DI application rates, to infer and characterize "compliers" to the reform (those who no longer apply for DI benefits). Subsequently, we construct the Chronic Disease Index (CDI) and Work Ability Index (WAI) that are based on time-varying chronic conditions of applicants. With these measures, we analyze the relative importance of the drivers of changes in mortality and employment rates of pre- and post-GKP applicant cohorts.

4.1 The RDiT Model

Throughout our analysis, inferences on the compositional effects of the GKP reform are based on RDiT regressions. Our first interest lies in monthly (cohort) changes in DI application rates. We consider monthly DI applications for all employed workers in 2001-2004, providing us a 24-months-symmetric window around the reform cutoff date of January 2003. Since we do not observe the exact application date, but only the award decision date, award decisions observed around the cutoff cannot be ruled out as stemming from workers who applied in December 2002 (i.e., started their sickness spell before April 2002, under the old regime). We therefore drop observations in December 2002 and January 2003, constructing a "donut hole" around the cutoff. In doing so, we follow Godard et al. (2024).

Our preferred specification is a linear RDiT with a bandwidth of 24 months on each side of the cutoff and a donut hole of one month on each side of the threshold.[11] The linear probability model for DI applications is specified as follows:

$$Y_{it} = \alpha + \beta_Y \ T_t + \gamma_1 \ t + \gamma_2 t * T_t + \sum_{k=2}^{12} \delta_k M_{kt} + \epsilon_{it}, \tag{1}$$

where Y_{it} is a DI application dummy of worker i ($i = 1, ..N$) observed in application month t (with $t = -24, ..24$). T_t is a dummy variable that takes the value of one after December 31, 2002, and it is equal to zero for applicants before the reform. We allow the linear time trend to differ on each side of the cutoff by including t and the interaction term $t * T_t$, where t equals zero at the cutoff, and ranges from -24 to 24. The parameter γ_1 captures any pre-existing linear trend in DI applications, while γ_2 captures any differences with the linear trend after the policy change. M_{kt} with $k = 2, \ldots, 12$ represents a set of calendar-month indicators and δ_k month-of-year fixed effects. The month-of-year fixed effects are included to control for any seasonal and administrative effects that may be relevant at the end of the calendar year. Standard errors are clustered at the month-year level. The parameter of interest β_{DI} reflects the overall impact of the reform on DI application probability. As argued earlier, the reform may have induced workers not to enter the sick-pay period after 13 weeks of sickness or increased reintegration efforts may have increased work resumption.

Next, we analyze the effect of the GKP reform on the composition of the sample of applicants.

[11] As we will show later on, the estimates for the 12-24 months bandwidth are stable – see Figure A2 in Appendix A. Different specifications with varying bandwidths and higher order polynomials are shown in Figure A3 and in Figure A4 in Appendix A.

The RDiT regression is similar to Equation (1), but we now it is estimated on the sample of applicants, Y_{it} as the individual's survival status, health status or employment outcome at different times after the DI application. For survival and employment as outcomes, we consider elapsed times up to 18 years after the application date. For not-in-good (physical and mental) health as an outcome measure, we use medical consumption data starting from 8 years after the award dates (recall that we only observe the Medical Consumption data from 2009 - 2022).

4.2 Inferences on targeting

For insight into the targeting effects of the reform, we use a simple non-parametric method proposed by Almond and Doyle (2011). This method aims to characterize "compliers", i.e. workers who did apply before the reform but were deterred from applying after the reform. We cannot directly observe these two groups, but we know the average outcomes of the cohorts before the reform and of the "stayers" who continued applying. To derive the average outcomes of the compliers, two assumptions need to hold. First, it requires a monotonicity assumption, stating that the reform induced some workers to stop applying, but did not induce any workers to apply who would not have applied before the implementation of the reform. In light of the high DI inflow rates before the reform, this assumption is plausible (Koning and Lindeboom, 2015). Second, the reform should not have a behavioral effect on the future outcomes of the stayers. Specifically, workers not deterred from applying did not experience any changes in their employment outcomes due to the GKP obligations/reintegration measures during the sickness period. This is a reasonable assumption, as the increased work resumption obligations stipulated by the GKP did not lead to (full) work resumption in the sickness/waiting period for the "stayers".[12]

Under these two assumptions, we can infer the average outcome of compliers to the reform. For this we define $\overline{Y}_{NC}$ and $\overline{Y}_{C}$ as the averages of outcome Y for non-compliers ("stayers") and compliers to the reform, respectively. Also, we define π as the fraction of the number of compliers of the total number of compliers and stayers. β_Y denotes the RDiT coefficient for Y before and after the reform. Following Almond and Doyle (2011), the average of outcome Y_C equals:[13]

[12] We will more carefully investigate the validity of this assumption by considering the (smaller) sample of applicants that were awarded benefits. For these applicants, absence of treatment effects for stayers is more plausible than for rejected applicants.

[13] See also Godard et al. (2024) for a formal proof.

$$\overline{Y}_C \;=\; \overline{Y}_{NC} - \frac{\beta_Y}{\pi} \tag{2}$$

For inferences on the targeting properties of the reform, we compare the implied values of Y_C of compliers to those of the stayers, Y_{NC}. As Equation (2) shows, the difference between the two outcomes is derived from the RDiT estimate on Y and the fraction of compliers, π.

A broader perspective on the targeting effects of the reform can be obtained by considering the outcomes of workers who did not apply in the pre-GKP period – i.e., the so-called "non-applicants". Same as for the applicant cohorts, we observe the outcomes of non-applicant cohorts in the years after the reform. Improvements in targeting occur when health and employment outcomes are similar to those of non-applicants. In contrast, target improvements are limited when the outcomes of compliers remain close to those of the stayers. let's define $\overline{Y}_{NA}$ as the observed average value of Y of non-applicants. We then propose the following measure for the degree of targeting, DT, of the reform for a specific outcome variable Y:

$$DT_Y \;=\; \frac{\overline{Y}_C - \overline{Y}_{NC}}{\overline{Y}_{NA} - \overline{Y}_{NC}} \tag{3}$$

The degree of targeting DT_Y equals zero when the compliers have equal outcomes to non-compliers that kept on applying for DI, and it equals one when compliers have equal outcomes to non-applicants. In addition, negative values of DT may suggest "perverse" screening effects, e.g. when health conditions of the compliers are more severe than the non-compliers.

4.3 The Chronic Disease Index (CDI) and Work Ability Index (WAI)

Inferences on the reform effects on health typically analyze changes in mortality rates of workers (Black et al., 2024; Gelber et al., 2023). While survival rates provide us with an encompassing and often well-observable proxy for health, they however also come with important disadvantages. Most importantly, mortality is most prevalent at older ages and rarely observed among younger workers with existing health conditions. Knowing that workers can apply for DI benefits during their entire working career, this implies that very long time periods are needed to detect meaningful differences in mortality rates between pre- and post GKP applicant cohorts. More importantly, unobserved differences in the *expected* mortality between applicant cohorts may already materialize at younger ages. In our context, there may be important differences in health conditions that prevail at the moment of DI application and may persist over time, but

its consequences on mortality cannot be observed.[14]

In light of this, we follow Danesh et al. (2024) and construct a composite measure for mortality – the "Chronic Disease Index" (CDI) – that unpacks the later-life consequences of current chronic diseases. The intuition behind this approach is that we project the contributions of chronic diseases to the survival probability at later life. The projections are based on the sample of old-aged individuals for whom the consequences of chronic health conditions on mortality are more often observed. The CDI thus projects the survival rate that prevails if an individual reaches old-age.

The construction of a CDI of the DI applications cohorts consists of two steps. First, we select all workers in the DI applicant cohorts who become 65 in some year between 2006 and 2012 are observed.[15] Starting from that year, we estimate the 10-year survival probability for this sub-sample as a function of chronic conditions, gender, and year-fixed effects. Specifically, we estimate a linear probability model for the 10-year survival rate S of a worker i in year t aged 65 as a function of (22) chronic diseases CC, a dummy for gender F, and μ as year-fixed effects:

$$S_{i,t} = \beta_0 + \beta_F F_i + f(CC_{i,t})^* \beta_{CC} + \mu_t + \epsilon_{i,t} \qquad \text{if} \qquad AGE_{i,t} = 65 \tag{4}$$

with ϵ as i.i.d. with an expected value equal to zero and variance $\sigma^2_{\S}$. $f(CC_{i,t})^*$ denotes the set of chronic conditions and its possible interactions of individual i at time t that are found to contribute to the explanation of the 10-year-survival rate at the age of 65. We estimate Equation (4) using Lasso with CC as focal variables. With the coefficient estimates, we next predict the CDI *for all workers at any age* in the period between 2006 and 2022:

$$CDI_{i,t} = \hat{\beta}_0 + \hat{\beta}_F F_i + f(CC_{i,t})^* \hat{\beta}_{CC} + \hat{\mu}_t \tag{5}$$

With the CDI we open up differences in expected survival rates that may already matter in the first years after DI application. This also allows us to compare the persistency of differences in chronic diseases, as well as their (projected) impact on mortality. Related to this, differences in the CDI between pre- and post-reform cohorts can be decomposed in contributions stemming from differences in the prevalence of specific chronic diseases. For this, we need the RDiT

[14] One can see these unobservables as biomarkers that are unfortunately not observed in our administrative dataset.

[15] This period is chosen as we only observe medical prescriptions from 2006 until 2022

estimates of the GKP reform of the separate (22) chronic diseases.

To analyze how changes in chronic diseases impact the employment outcomes of workers, we complement the CDI with composite Work Ability Index (WAI) that is based on the same set of chronic diseases, CC, age and gender. Using current employment outcomes in Lasso specifications, the estimation is more straightforward than for the CDI. We now use the full sample of workers for which we observe CC between 2006 and 2022 and allow the effects of all explanatories to vary across calendar years. Specifically, we estimate the employment outcome E of worker i in year t with the following linear probability model using Lasso:

$$E_{it} = \gamma_0 + \gamma_F F_i + g(CC_{i,t})^* \gamma_{CC} + h(AGE_t) + \nu_i, \tag{6}$$

with ν that is i.i.d. with expected value zero and variance σ^2_ν. $g(CC_{i,t})^*$ denotes the union of chronic diseases and their relevant cross-products, while allowing for interactions for both with year effects. h is a step function in age of individual i at time t. Following from Equation (6), the predicted value of E equals the Work Ability Index, WAI:

$$WAI_{it} = \hat{\gamma}_0 + \hat{\gamma}_F F_i + g(CC_{i,t})^* \hat{\gamma}_{CC} + \hat{h}(AGE_{it}) + \hat{\nu}_t \tag{7}$$

With the predicted CDI and WAI scores, we can discern changes in health and the earnings capacity of workers as explained by changes in the health conditions of DI applicants before and after the GKP reform. This allows us to assess the effect of increased screening on the (relative) importance of "work ability" and "health disability" in explaining health and employment outcomes of DI applicants.

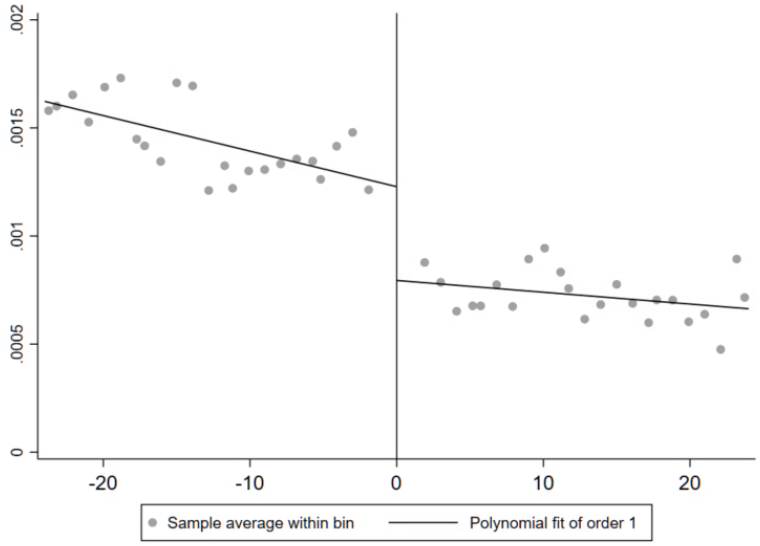
5 Results

5.1 DI applications

Figure 6 below displays the results from our preferred donut RDiT regression on DI applications. We consider the average monthly DI application rates for 2001-2004, providing us a 24-month-symmetric window around the reform cutoff date of January 2003. The monthly rate is calculated as the fraction of applicants divided by the total employed Dutch population in a given month. The vertical line at "0" indicates January 2003, the exact month in which the new regime started. The graph depicts a sharp and considerable decline of DI application rate at the cutoff that is similar in magnitude to the results obtained by Godard et al. (2024). The reduction in

the application rate of workers amounts to about one-third of the pre-GKP level. The effect is slightly smaller than obtained by Godard et al. (2024) who only consider permanent workers.

Figure 6: Regression Discontinuity plot for monthly DI application cohort rates, linear fit[a]



[a]: Monthly observations represent application cohorts centered around January 2003

Table 2 below shows our donut RDiT estimates using Equation (1) for our preferred model and model specifications with different donuts and smaller bandwidths. Standard errors are clustered at the month-of-the year level. The preferred model uses a 24-months bandwidth on each side of the threshold, whereas the second estimate uses a bandwidth of 15 months on each side of the cutoff. [16] To obtain conservative and efficient estimates of the reform, we therefore choose a maximum bandwidth of 24 months at each side of the cutoff. As mirrored by Figure 6, our preferred specification implies a 33% decrease in DI applications due to the GKP reform. The estimated relative effect size becomes larger (46.2%) if we choose the minimum bandwidth of 15 months or when we widen the donut support to two or four months around the cutoff (36.8% and 31.6%, respectively).

Table 3 complements our findings with RDiT estimates that are stratified by (four) broad impairment types. We then see that the drop in applications was most substantial for mental impairments (-41.3%) and "other" impairments (-53.2%), as against a decrease of 14.3% for cardiovascular impairments only. From this, we conclude that the decrease in application rates was most substantial among hard-to-verify conditions.

[16]Recall that our specification controls for calendar-month-fixed effects, which puts constraints on the minimum bandwidth size.

Table 2: RDiT Estimates of the Gatekeeper effect on DI applications

	Dependent Variable DI Application
	Coeff. [% change] (se)
RDiT estimate dropping those within one month of the cutoff 24 months bandwidth (=preferred model)	–0.0003*** [33.25%] (0.0001)
Mean of dependent variable	0.0009
Number of individual-month observations	260,783,292
RDiT estimate dropping those within one month of the cutoff 15 months bandwidth	–0.0004*** [46.18%] (0.0001)
Mean of dependent variable	0.0009
Number of individual-month observations	147,399,252
RDiT estimate dropping those within two month of the cutoff 24 months bandwidth	–0.0004*** [36.84%] (0.0001)
Mean of dependent variable	0.0010
Number of individual-month observations	249,444,888
RDiT estimate dropping those within four month of the cutoff 24 months bandwidth	–0.0003*** [31.62%] (0.0001)
Mean of dependent variable	0.0009
Number of individual-month observations	226,768,080

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) Estimates in brackets are presented in percentage change of the baseline DI application for the sample of interest. (3) ***/**/* Significant at the 1%/5%/10% level.

Table 3: RDiT Estimates of the Gatekeeper effect on DI applications by impairment types

	Dependent Variable DI Application
	Coeff. [% change] (se)
Panel A: All impairments	–0.0003*** [33.25%]
24 months bandwidth (=preferred model)	(0.0001)
Mean of dependent variable	0.0009
Number of individual-month observations	260,783,292
Panel B: Impairment types (preferred model)	
Musculo-skeletal	-0.0001*** [36.00%]
	(0.0000)
Mean of dependent variable	0.0003
Number of individual-month observations	251,237,832
Mental	-0.0001*** [41.30%]
	(0.0000)
Mean of dependent variable	0.0004
Number of individual-month observations	251,860,580
Cardiovascular	-0.0000*** [14.03%]
	(0.0000)
Mean of dependent variable	0.0001
Number of individual-month observations	249,779,126
Other	-0.0001*** [25.50%]
	(0.0000)
Mean of dependent variable	0.0002
Number of individual-month observations	250,123,436

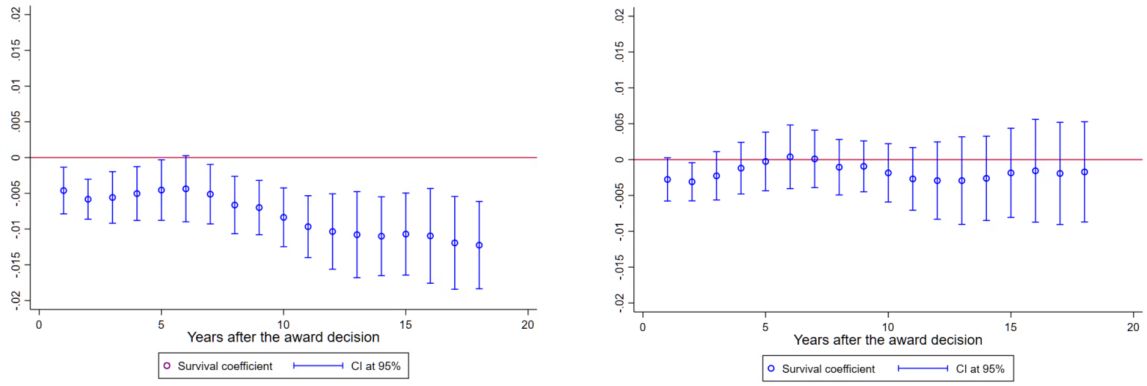
Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) Estimates in brackets are presented in percentage change of the baseline DI application for the sample of interest. (3) ***/**/* Significant at the 1%/5%/10% level.

5.2 Effects on survival rates, health status and employment

To see how the GKP reform affects the composition of applicants, we re-estimate Equation (1) on the survival rates, health status and employment outcomes of applicants. The models are estimated on the full sample and include the impairments diagnosed by the NSII at the award decision. We first present the GKP effects on survival rates in Figure 7, with estimates up to 18 years after the award decision. We see that there is a statistically significant mortality effect in the first year after DI application, which then stays more or less constant and accumulates after

five years to 1.1 percentage point after 15 years. This difference corresponds to about 12.2% of the average mortality rate of the pre-GKP cohort after 15 years. Interestingly, when controlling for gender, age, and impairments diagnosed by the NSII at the time of the award decision, all effects become negligible and statistically insignificant across all elapsed years. This indicates that the observed higher mortality rate can largely be attributed to changes in these factors. Tables B2 and B3 in the appendix provide detailed coefficient values for the survival rate models after 5, 10, 15, and 18 years, both with and without controls.

Figure 7: GKP effect on survival estimates rate of DI applicants over elapsed time



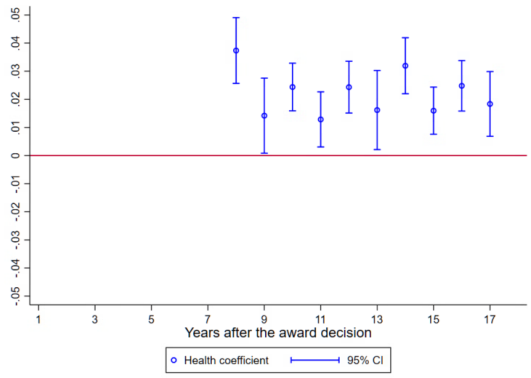
(a) Survival estimates without controls

(b) Survival estimates with controls

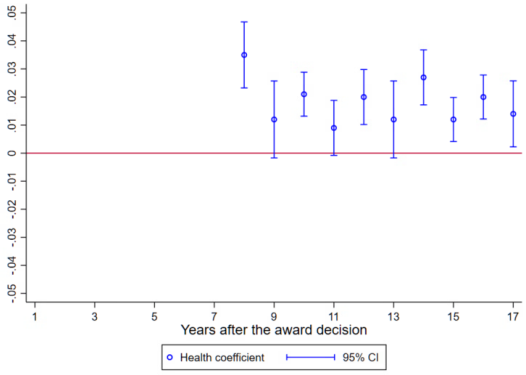
Panels (a) and (b) Figure 8 present the RDiT results for health status of the applicants as an outcome measure with and without controls.[17] Recall that this outcome measure equals the share of top medical spenders. Given that Medical Spending data are observed as from 2009 onwards, we only can estimate elapsed time effects of 8 year and later. Our results show that DI applicants after the reform are 2.4 percentage point more likely to be in bad health ten years after the award decision, and 1.6 percentage point after 15 years. The estimate only slightly declines when controlling for gender, age cohorts and impairments diagnosed by the NSII at the award decision, and remains positive and statistically significant. These findings are roughly similar when we consider our separate proxy for workers who are not in good physical health – as shown in Panel (c) and (d) of Figure 8.

[17]Table B4 and Table B5 in Appendix B show the GKP effect coefficients on not in good (physical, mental) health status applicants after 10 and 15 years, with and without controls.

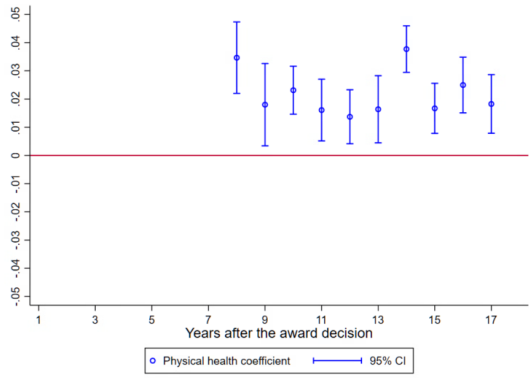
Figure 8: GKP effect on fraction with "Not in good health" status of DI applicants



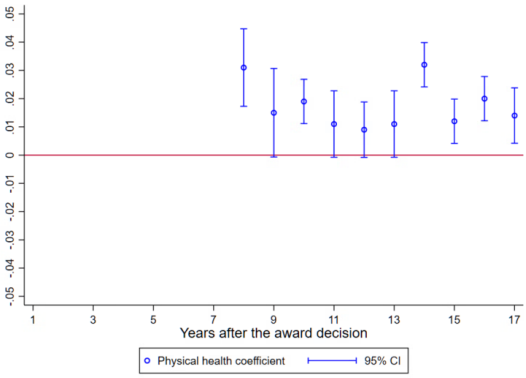
(a) Total health consumption without controls



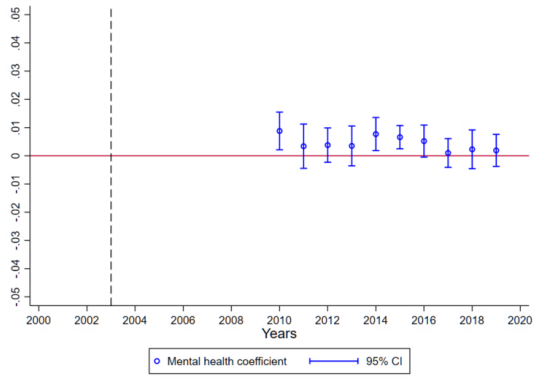
(b) Total health consumption with controls



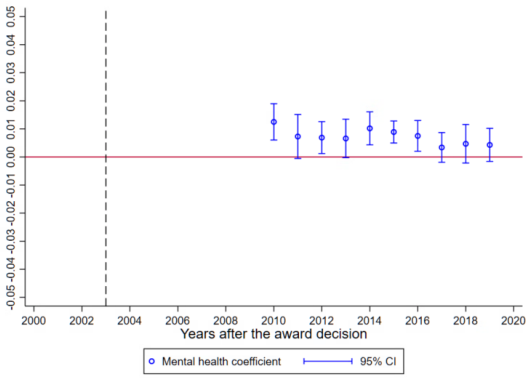
(c) Physical health without controls



(d) Physical health with controls



(e) Mental health without controls



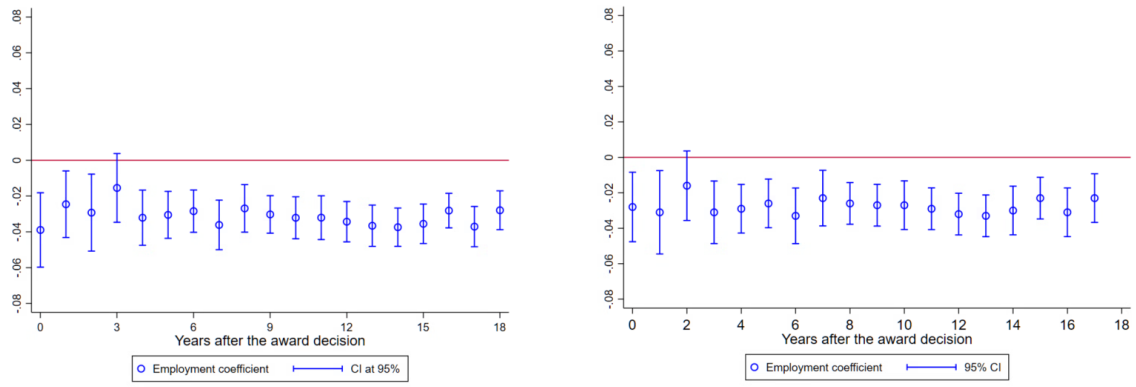
(f) Mental health with controls

Panels (e) and (f) of Figure 8 primarily show small and statistically insignificant effects on mental health outcomes between 10 and 18 years after the DI award decision. Although the share of applicants with mental conditions drops significantly at the time of the GKP reform, it is important to note that only a small fraction of the remaining applicants received mental health treatments recorded in our data. These treatments, which are relatively rare and typically provided to individuals with more severe mental conditions, suggest that the reduction in

applications was as pronounced for cases with severe long-term mental health conditions as for those with less severe or absent conditions. Consequently, within the group of potential DI applicants with mental conditions, there is no evidence to support improved (long-term) targeting properties due to the reform.

Finally, Figure 9 displays the employment effects up to 18 years following the award decision.[18] Employment is defined as having positive labor earnings in a given year, while the non-employed category may include individuals who retired during the observed period. The RDiT estimate indicates a four percentage point reduction in employment in the year of the award decision. This effect remains relatively stable over time. These findings suggest that compliers to the reform who were deterred from applying for DI benefits were, on average, more able to work than those who continued to apply. Interestingly, when controlling for impairment types diagnosed at the time of the DI award decision, the estimates are nearly identical to those without controls. Unlike the survival rate results, this indicates that increased self-screening on impairment types reported by the NSII is not strongly associated with the work ability of applicants.

Figure 9: GKP effect on the employment of DI applicants



(a) Employment estimates without controls

(b) Employment estimates with controls

5.3 Effects on targeting

To put our results in broader perspective, Table 4 shows the implied targeting effects of the GKP reform for the survival rate, the health status and the employment outcomes. The implied values of the stayers and compliers can be inferred from Equation (2), together with the value averages of the 2002 DI application cohort. For the cohort of employed non-applicants, we take

[18] Tables B6 and B7 in Appendix B provide the corresponding coefficients for the GKP effect on employment, with and without controls, for several specific years.

2002 as the reference year. With this information, we infer the degree of targeting that followed from the reform for different elapsed time windows.

Table 4: Implied targeting effects of the GKP reform

	within 1 year	**within 5 years**	**within 10 years**	**within 15 years**	**within 18 years**
Survival rate					
Stayers (%)	98.20	95.28	94.49	88.25	85.21
Compliers (%)	99.72	96.80	96.91	91.58	88.85
Non-applicants (%)	99.81	99.25	98.04	95.98	94.08
Degree of Targeting	0.94	0.38	0.68	0.43	0.40
Employment rate					
Stayers (%)	68.34	57.35	46.32	38.45	34.23
Compliers (%)	75.92	65.83	56.02	46.93	42.41
Non-applicants (%)	91.87	81.14	71.73	62.23	56.26
Degree of Targeting	0.32	0.35	0.38	0.35	0.37

Interestingly, the upper panel of Table 4 reveals that targeting effectiveness is nearly perfect in the first year after the award decision. The survival rate of compliers to the reform closely matches that of non-applicants during this period, suggesting that almost all workers who died within the first year after applying were "stayers" – those who continued to apply for DI benefits after the reform. However, these effects diminish over time. After 18 years, the long-term mortality outcomes of compliers become somewhat closer to those of stayers than to non-applicants, with a degree of targeting measured at 0.40. While the targeting properties of the reform improved in terms of survival rates, compliers still resembled stayers more than non-applicants in the long term.

The lower panel of Table 4 highlights a substantial decline in employment rates for the groups of stayers, compliers, and non-applicants over the 18-year period. These declines largely reflect the retirement of workers who reached pension age during the observed timeframe. Concurrently, the degree of targeting remains relatively stable at approximately 0.35. This indicates that the GKP reform screened out workers with higher employment rates than those of the stayers, but these workers still share more similarities with stayers than with non-applicants. However, compared to the findings on survival rates, it is important to note that employment rates among compliers and non-applicants likely reflect more than just health differences. The relatively low degree of targeting may also be attributed to the weaker labor market positions of compliers.

5.4 The Chronic Disease Index (CDI) and Work Ability Index (WAI)

The picture that emerges from our results so far is that the GKP reform induced substantial self-screening of (potential) applicants during the waiting time before the DI application. These effects were stronger for workers with hard-to-verify conditions, such as mental and musculo-skeletal diseases. Compliers to the reform had higher survival rates, lower (physical) medical consumption and higher employment rates. The effects persist up to almost 20 years after the award decision, suggesting that the targeting properties of the reform improved to some extent and stayed constant in the years after the award decision.

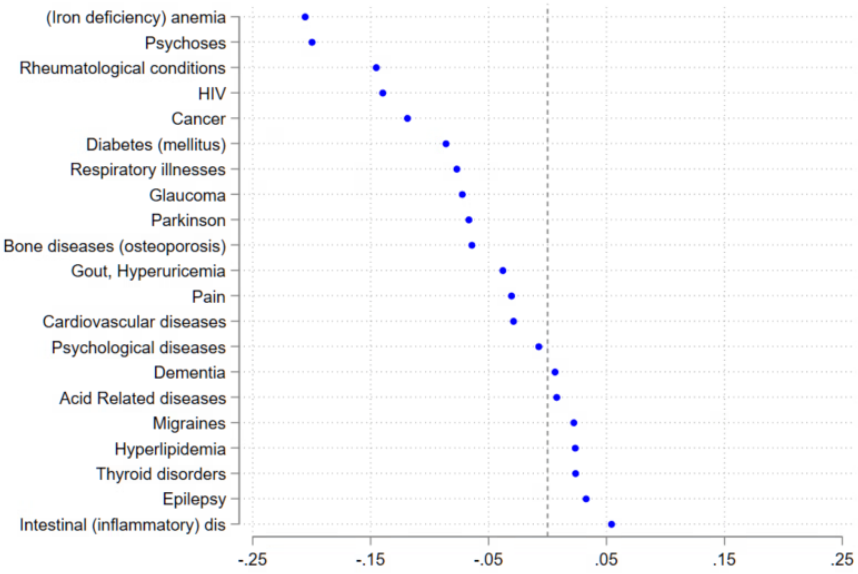
Having said this, two other findings raise questions on the types of health conditions that were screened out. First, we find no significant impact on the fraction of applicants with mental health treatments. One may argue that this finding can be explained by the fact that applicants with mental diseases were disproportionally screened out. But this argument overlooks the fact that mental health treatments concern only a small fraction of the applicants with more severe mental diseases. Second, we find that the impairments that are diagnosed by the NSII at the award decision cannot explain the lower employment fractions of cohorts that continued to apply for DI. This may imply that there was self-screening on health conditions that were more linked to the "health disability" than to the "work disability" of workers. We need more detailed information to discern such effects.

The impact of chronic diseases on the CDI and WAI

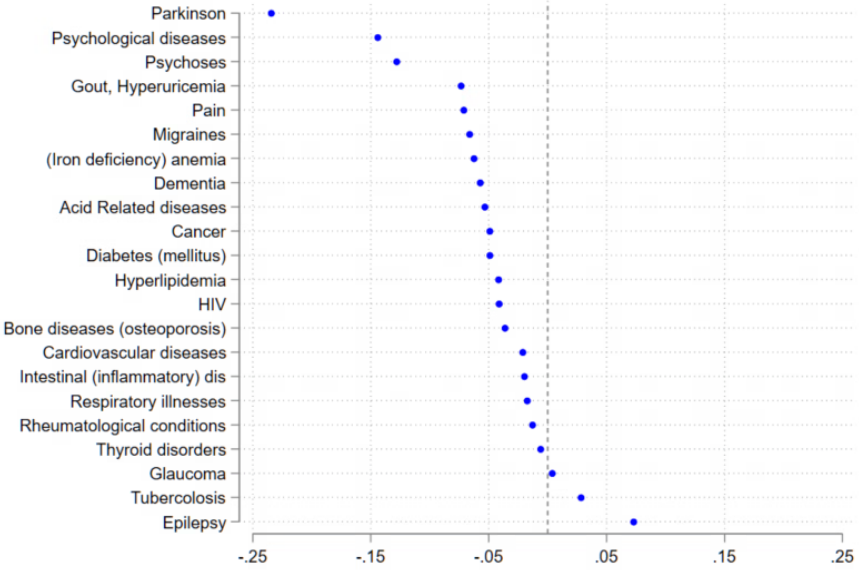
To explore the implications of changes in applicant cohort health conditions after the reform, we utilize the detailed Medical Prescription data records. Following the methodology outlined in Section 4.3, we construct two indices: the Chronic Disease Index (CDI) for survival and the Work Ability Index (WAI), both derived from the 22 chronic diseases identified in the data. These models are estimated using Lasso regression. The upper panel of Figure 10 presents the estimated coefficients of chronic diseases contributing to the CDI for survival. This index is based on data for workers aged 65 at some point between 2006 and 2012, using the 10-year survival rate as the outcome measure. The lower panel displays the impact of chronic diseases on the WAI for 2013, which corresponds to an average elapsed time of 10 years after the DI award decision. In both panels, chronic diseases are ranked by the magnitude of coefficient estimates.

Regarding the CDI, Figure 10 makes apparent that there substantial differences in survival

Figure 10: Lasso coefficients of chronic diseases as focal variables for (a) the Chronic Disease Index (CDI) for survival and (b) the Work Ability Index (WAI)



(a) Coefficients for CDI of years 2006-2012 (65-year old)



(b) Coefficients for WAI of year 2013

rates of old-aged workers across types of chronic diseases. In the sample of DI applicants who reach the age of 65, individuals with iron deficiencies, psychoses, HIV, rheumatic conditions, and cancers have CDI (10-year) survival rates that are between 15 and 25 percentage point lower than for individuals without any observed chronic diseases. The more prevalent cardiovascular and psychological diseases have a substantially lower impact on survival, amounting to a reduction of around 2.5 percentage point. We also observe a few chronic diseases, like migraines or intestinal

diseases, for which the survival rates exceed those of applicants without chronic diseases. At this point, it should be stressed that our applicant sample most likely also consists of some individuals with conditions that cannot be linked to medical prescriptions in our data. It would therefore be misleading to characterize the baseline as for individuals without any health conditions.

The lower panel of Figure 10 shows that chronic diseases also have a potentially strong impact on WAI, but that the ranking of the size contributions of the conditions to the index is different. WAI scores are lowest for individuals with Parkinson disease, amounting to an employment rate that is almost 25 percentage point lower than applicants without chronic diseases. More importantly, we now see strong negative effects of – the more prevalent – psychological diseases and psychoses on the WAI. This contrasts to the effects of cancers and cardiovascular diseases of which the effects on employment are limited to -5 and -2 percentage points, respectively.

RDiT estimates and decomposition

We next estimate the differences in the CDI and WAI induced by the GKP reform with the RDiT model; the results for different elapsed time windows since application are shown in Figure 13. Given that the Medical Prescription data are available since 2006, the estimates start (on average) three years after the award decision. For the CDI, the RDiT estimate is fairly constant and around -1.0 percentage points for all years. This suggests that that the implications of pre- and post-GKP differences of cohorts for later-life survival rates are already there in the first years after application and that these persist over time. Recall that we measure the 10-year survival rate at the age of 65. As such, the effect is not (directly) comparable to the earlier RDiT results for the observed survival rates.

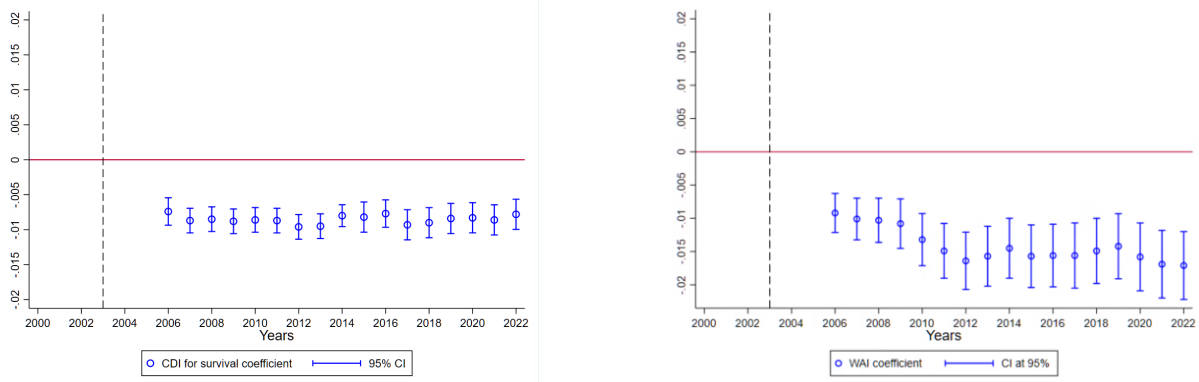
For the WAI index that explains changes in the ability to work, the interpretation of the effects is more straightforward. Specifically, estimates between -1.0 and -1.7 percentage points between 2006 (three years after the cutoff date) and 2022 (19 years after the cutoff date) reflect a decrease in employment rates ten years after the award decision that can be explained by changes in chronic diseases due to reform. Roughly speaking, the effect on the WAI explains 25% to 35% of the effects on employment. Contrasting to the CDI, we observe a gradual widening of the difference over time. This probably reflects the fact that the post-GKP cohort has a higher share of older workers that retire earlier than the pre-GKP cohort.

Table 5 provides insight into the drivers of the changes in CDI and WAI induced by the GKP reform. For each variable used as inputs for the CDI and WAI, we use the product of its

Table 5: Decomposition of GKP-reform effect on Chronic Disease Index (CDI) and Work Ability Index (WAI) 10 years after DI application

Category	**CDI contribution (x 100)**	**(%)**	**WAI contribution (x100)**	**(%)**
Chronic Diseases				
No chronic disease (reference)	–	–		
Acid related disorders	-0.0026	-0.26	-0.0473	3.01
Bone disorders (osteoporosis)	-0.0030	0.32	-0.0021	0.14
Cancer	-0.0222	2.33	-0.0094	0.60
Cardiovascular diseases	-0.0367	3.86	-0.0335	2.14
Dementia	0.0002	-0.02	-0.0011	0.72
Diabetes (mellitus)	-0.0480	5.05	-0.0286	1.82
Epilepsy	0.0003	-0.03	0.0011	-0.07
Glaucoma	-0.0059	0.62	0.0005	-0.03
Gout. Hyperuricemia	-0.0041	0.43	-0.0094	0.60
HIV	-0.0024	0.25	-0.0009	0.06
Hyperlipidemia	0.0198	-2.08	-0.0326	2.07
Intestinal (Inflammatory diseases)	0.0066	-0.69	-0.0023	0.15
(Iron deficiency) anemia	-0.0113	1.18	-0.0036	0.23
Migraines	-0.0008	0.09	0.0045	-0.29
Pain	-0.0053	0.55	-0.0085	0.54
Parkinson's disease	-0.0323	3.40	-0.1096	6.98
Psychological disorders	-0.0028	0.30	-0.0191	1.22
Psychoses	-0.0305	0.30	-0.0214	1.36
Respiratory illnesses	-0.0201	2.11	-0.0043	0.28
Rheumatological conditions	-0.0183	1.92	-0.0020	0.13
Thyroid disorders	0.0002	-0.22	-0.0069	0.44
Tuberculosis	-	-	0.0010	-0.07
	-0.2194	23.09	-0.0041	26.27
Demographics				
Female	-0.5830	61.39	-0.0730	4.67
25-34 years (reference)	-	-	-	-
35-44 years	-		-0.0751	6.04
45-54 years	-	-	-0.0992	6.32
55-64 years	-	-	-0.0619	3.94
65-74 years	-	-	-0.7450	47.45
75-84 years	-	-	-0.1375	8.76
Residual	-0.1476	15.54	0.0341	-2.17
Total	-0.95	100	-1.57	100

Figure 11: RDiT estimates for Chronic Disease Index (CDI) and Work Ability Index (WAI)



(a) Chronic Disease Index (CDI)

(b) Work Ability Index (WAI)

.

change in prevalence (with RDiT estimation) and the coefficient in the index to obtain separate contributions. Since both models use flexible Lasso specifications, the "residual" contribution stems from variation induced by cross-products of chronic diseases.

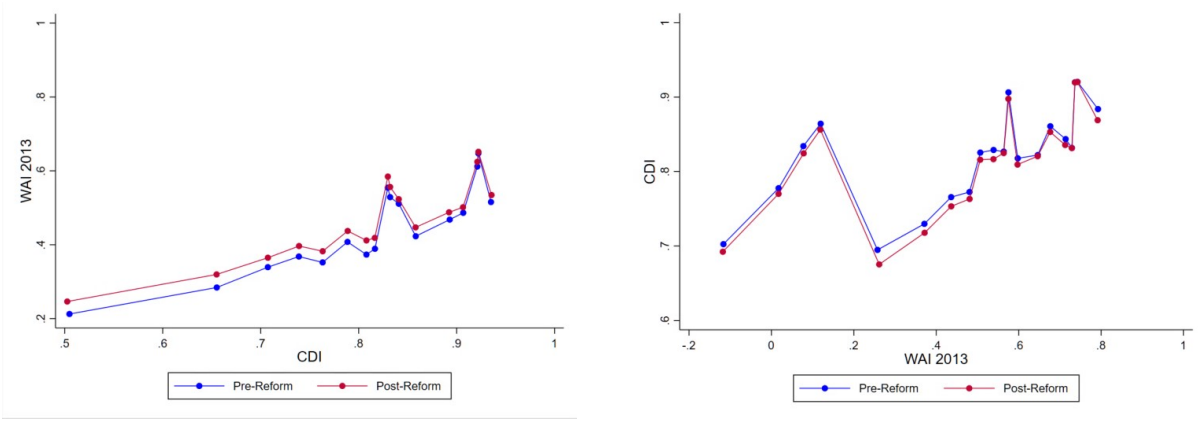
Although Figure 10 has shown that chronic diseases can make a strong difference in the survival and employment rates of applicants, Table 5 shows limited changes in the composition of chronic diseases of applicant cohorts before and after the reform. Specifically, they only explain 21.9% of the decrease of the CDI and 26.3% of the decrease in the WAI. Contrasting to this, screening on gender (for the CDI) and age (for the WAI) was more important. For the CDI, the decrease in the survival rate is largely (61%) explained by the decrease in the fraction of females induced by the reform. Presumably, for them the detrimental impact of chronic diseases on the ten-year-survival rate at the age of 65 is substantially smaller than for males. This finding confirms our earlier results on observed survival rates, for which the change induced by the GKP reform could be explained by a lower fraction of females and applicants diagnosed with mental impairments by the NSII. For the WAI, however, the lower ability to work largely stems from the higher share of older applicants after the reform. Since these older workers are close to retirement, the WAI of post-GKP applicant cohorts decreases more quickly than the pre-GKP applicant cohort.

CDI and WAI in perspective

The RDiT estimates we have shown so far provide information on the effect of changes in chronic diseases affecting the health and employment outcomes of applicant cohorts, but not on their relative importance. To see whether the GKP reform predominantly led to additional screening on "health disability" or "work disability", Figure 12 shows bin-scatters of the WAI and the

CDI for pre- and post-GKP cohorts. In both graph panels, averages are calculated for 20 bins for the joint pre- and post-GKP cohorts.

Figure 12: Bin-scatters of CDI and WAI pre- (2001-2002) and post-GKP (2003-2004) cohorts, ten years after the award decision (20 bins)



The panels in Figure 12 display a substantial variation in CDI and WAI scores. The 5^{th} percentile and the 100^{th} percentile of the CDI score for the (10-year) survival rates are 50% and 95%, respectively. For the WAI, we observe values between -10% for the 5^{th} percentile and around 85% for the 100^{th} percentile, respectively. This mirrors the fact that the variation in chronic diseases – and their cross-products – and demographics absorbs a substantial amount of variation in survival and employment rates. Specifically, the standard deviation of the CDI scores amounts to 8.9 percentage point, and 20.7 percentage point for the WAI scores. The (normalized) decrease in the CDI score due to the reform is equal to 10.6% of a standard deviation on the distribution of CDI scores, and the implied average CDI of compliers to the reform is 32% of a standard deviation higher than for the "stayers" who continued to apply. For the WAI score, the normalized decrease corresponds to 7.6% of a standard deviation. Consequently, compliers to the reform have WAI scores that are 22.8% of a standard deviation higher than stayers. Expressed as relative changes over the distribution of index outcomes, screening on the "work ability" was thus less strong than on "health disability".

The second panel of Figure 12 also makes clear that there is a non-monotonous relationship between WAI scores and CDI scores. While the overall correlation between the WAI and CDI is positive, the connected lines also suggest the presence of a substantial group of applicants with a low ability to work (low WAI scores) and high survival rates. This reflects the effect of applicants who pass the statutory retirement age within 10 years after application (i.e., they

were 55 years or older at application). After retirement, this group has very low employment rates but may have health conditions that are associated with low mortality.

The distances between the pre-GKP (blue) and post-GKP (red) lines in both panels of Figure 12 highlight that screening was more pronounced on the CDI than on the WAI. This reflects the more substantial relative decrease in the CDI compared to the WAI. As previously discussed, this outcome primarily stems from significant screening effects observed among females and workers with mental impairments, as diagnosed by the NSII at the award decision. While these groups typically have low mortality rates, their work ability is relatively limited. These findings are corroborated in the panels of Figure A7 in the Appendix, which display CDI and WAI scores stratified by four grouped main impairment diagnoses: musculoskeletal, mental, cardiovascular, and "other." Conditional on work ability, as measured by the WAI, survival rates (reflected in the CDI) remained largely unchanged after the reform for applicants with cardiovascular impairments. However, for those with mental impairments and the heterogeneous "other" category, the distances between the pre- and post-GKP lines suggest that screening focused more on health status than on work ability.

6 Conclusion

In this paper we investigate the long-term health and employment outcomes of workers who were discouraged from applying for DI benefits by stricter screening. Specifically, in January 2003 the so-called Gatekeeper Protocol (GKP) in the Netherlands drastically increased the efforts that needed to be made by sick-listed workers and their employers in order to be eligible for a DI assessment. The reform drastically reduced the number of DI applicants, particularly among workers with hard-to-verify conditions, such as mental and musculo-skeletal impairments. This raises questions on the targeting properties of the reform: given that the composition of medical conditions of applicant cohorts changed, what were the effects on the long-term health and employment outcomes? And related to this: what were the outcomes of the compliers of the reform that no longer applied for DI benefits?

To answer these questions, we exploit the discrete cutoff date of the GKP reform. Using RDiT regressions, we estimate an overall decrease in DI applications of 33% due to the reform. This substantial reduction is accompanied by a decrease in the survival rate, a persistently higher fraction of applicants with high medical consumption, and lower employment rates up to 18

years after DI application. As a result, compliers to the reform who were discouraged from applying show higher survival and employment rates. These findings suggest that the (overall) targeting properties of the DI screening system have improved, both in terms of screening on "health disability" and "work disability". However, from these findings alone it is difficult to discern the changes in health conditions of applicant cohorts that drive these changes, as well as how they have impacted health and work outcomes.

To delve deeper into the chronic diseases that explain the changes in the long-term health and employment outcomes, we construct a "Chronic Disease Index" (CDI) and "Work Ability Index" (WAI). Both the CDI and WAI are based on the same set of (22) chronic diseases that are used to explain the (10-year) survival rates and the employment rates of applicant cohorts in the years after application, together with the gender and age as controls. We then find that the impact of chronic diseases on the CDI and WAI is very different; for instance, psychological conditions have a substantial detrimental effect on the WAI, but not on the CDI. Perhaps more surprisingly, we infer that changes in both scores induced by the GKP reform are largely driven by changes in demographics. Specifically, higher mortality (i.e. a lower CDI score) largely stems from a lower remaining share of female applicants and lower work ability largely stems from a higher remaining share of older applicants who retire in later years after application. Finally, we conclude that the effect of screening, measured as relative changes in the CDI and WAI, was stronger on "health disability" than on "work disability".

While these results suggest that the broad targeting properties of the Dutch DI screening system have improved due to the GKP reform, our findings are more nuanced for workers with mental conditions for whom the drop in DI applications was most substantial. This raises important questions about the principles underlying the design of DI benefit systems. Given the trend of increases in DI applications from workers with mental conditions in many countries, governments may be inclined to tighten eligibility conditions to conditions that are less difficult to verify. These reforms can be effective in reducing inflow, but may also screen out workers with low ability to work (Deshpande and Lockwood, 2022).

Bibliography

Ahammer, A. and A. Packham (2023): "Disability Insurance Screening and Worker Outcomes," Institute for the Study of Labor (IZA) Discussion Paper no. 16235.

Almond, D. and J. J. Doyle (2011): "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays," *American Economic Journal: Economic Policy*, 3, 1–34.

Benítez-Silva, H., R. Disney, and S. Jiménez-Martín (2014): "Disability, capacity for work and the business cycle: an international perspective," *Economic Policy*, 25, 483–536.

Black, B., E. French, J. McCauley, and J. Song (2024): "The effect of disability insurance receipt on mortality," *Journal of Public Economics*, 229, 1–19.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017): "Rdrobust: Software for Regression-discontinuity Designs," *The Stata Journal: Promoting Communications on Statistics and Stata*, 17, 372–404.

Danesh, K., J. Kolstad, W. Parker, and J. Spinnewijn (2024): "The Chronic Disease Index: Analyzing Health Inequalities Over the Lifecycle," Mimeo.

De Jong, P., M. Gielen, and V. Haanstra-Veldhuis (2014): "Verzekeringsgraad kleine werkgevers," APE report number 1209.

Deshpande, M. and Y. Li (2019): "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 11, 213–248.

Deshpande, M. and L. M. Lockwood (2022): "Beyond Health: Nonhealth Risk and the Value of Disability Insurance," *Econometrica*, 90, 1781–1810.

Garcia-Gomez, P. and A. C. Gielen (2017): "Mortality effects of containing moral hazard: Evidence from a disability insurance reform," *Health Economics*, 27, 606–621.

Gelber, A., T. Moore, Z. Pei, and A. Strand (2023): "Disability Insurance Income Saves Lives," *Journal of Political Economy*, 131, 3156–3185.

Godard, M., P. Koning, and M. Lindeboom (2024): "Application and Award Responses to Stricter Screening in Disability Insurance," *Journal of Human Resources*, 59, X–Y.

Haller, A., S. Staubli, and J. Zweimuller (2024): "Tightening Disability Screening Or Reducing Disability Benefits? Evidence and Welfare Implications," *Econometrica*, 92, 79–110.

Hausman, C. and D. S. Rapson (2018): "Regression discontinuity in time: Considerations for empirical applications," *Annual Review of Resource Economics*, 10, 533–552.

Huber, C. A., T. D. Szucs, R. Rapold, and O. Reich (2013): "Identifying patients with chronic conditions using pharmacy data in Switzerland: an updated mapping approach to the classification of medications," *BMC Public Health*, 13, 1030.

Johansson, P., L. Laun, and T. Laun (2014): "Screening Stringency in the Disability Insurance Program," *B.E. Journal of Economic Policy and Analysis*, 14, 1–19.

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

Koning, P. and H. Vethaak (2021): "Decomposing employment trends of disabled workers," *BE Journal of Economic Analysis and Policy*, 21, 1217–1255.

Liebert, H. (2019): "Does external medical review reduce disability insurance inflow?" *Journal of Health Economics*, 64, 108–128.

Markussen, S., K. Roed, and R. Schreiner (2017): "Can compulsory dialogues nudge sick-listed workers back to work?" *The Economic Journal*.

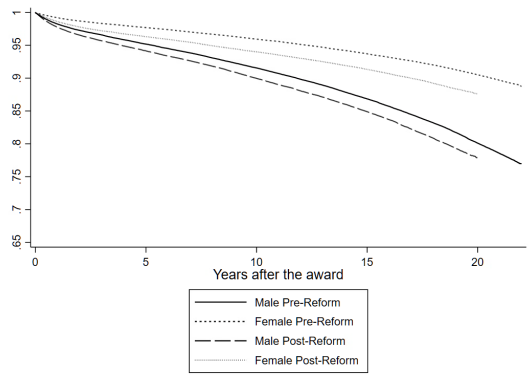
OECD (2010): "Sickness, Disability and Work: Breaking the Barriers. A Synthesis of Findings Across OECD Countries," OECD.

——— (2022): "Disability, Work and Inclusion. Mainstreaming all Policies and Practices," OECD.

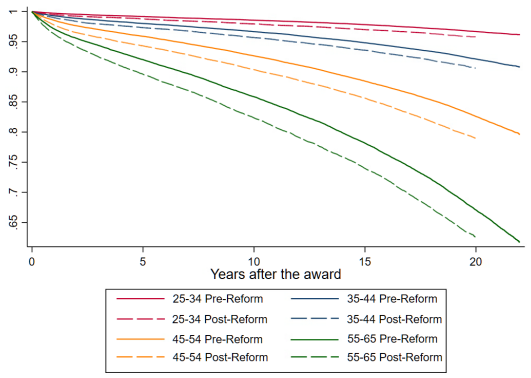
Parsons, D. O. (1991): "Self-screening in targeted public transfer programs," *Journal of Political Economy*, 99, 859–876.

A Online Appendix : Figures

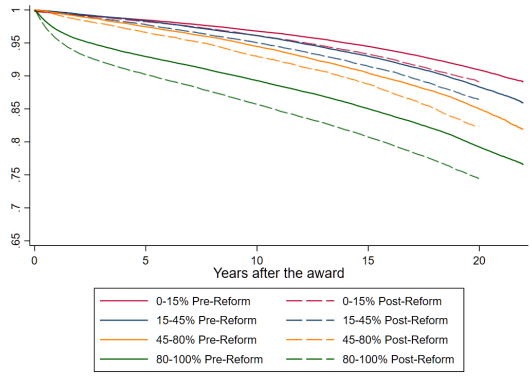
Figure A1: Pre- (2001-2002) and post-reform (2003-2004) Kaplan-Meier survival estimates by DI applicant cohorts stratified by gender, age cohort, degree of disability, and type of impairments



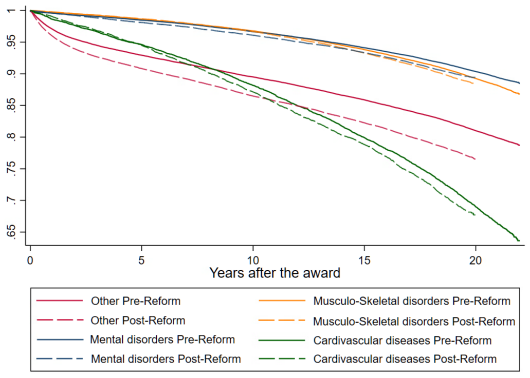
(a) Gender



(b) Age cohort



(c) Degree of disability



(d) Impairments

Figure A2: Donut Regression-Discontinuity estimates on DI application rates with varying bandwidths

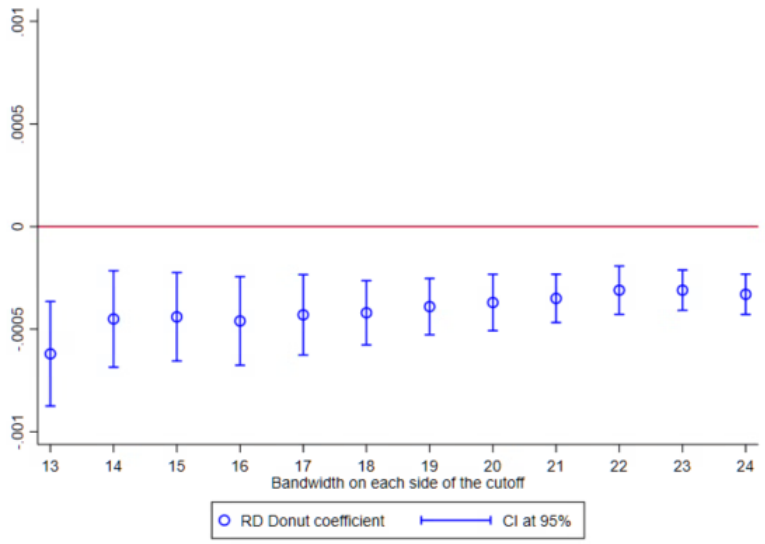
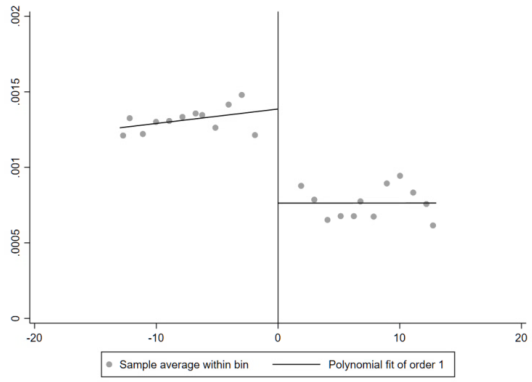
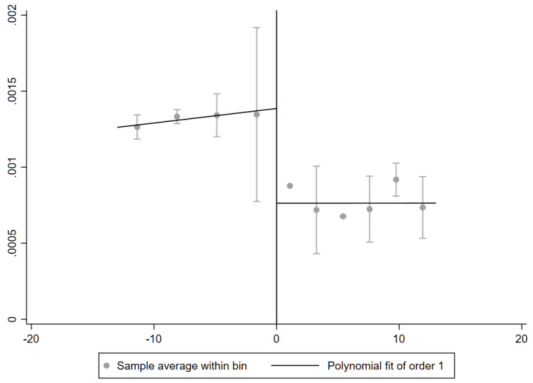


Figure A3: Donut Regression Discontinuity plots with 13 months bandwidth on each side of the cutoff.



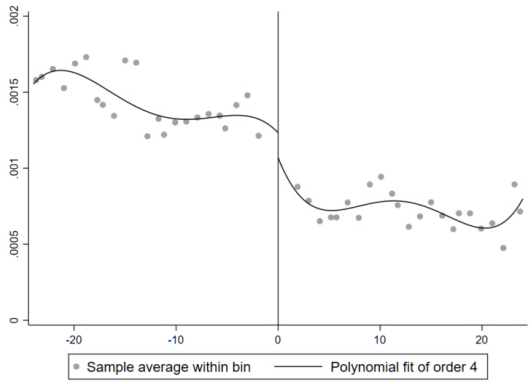
(a) Polynomial of order 1



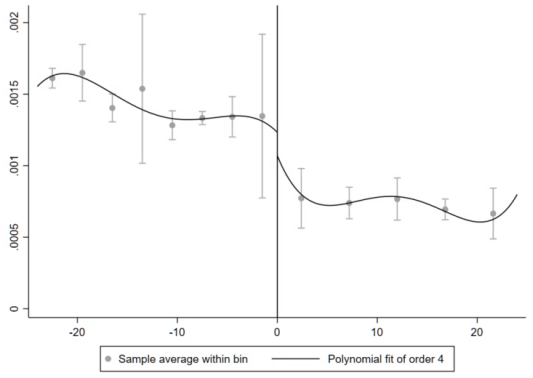
(b) Polynomial of order 1

Notes: (i) Donut RD dropping individuals within one month of the cutoff. (ii) Figure A2 (b) has 95% confidence intervals around the binned means calculated following Calonico et al. (2017).

Figure A4: Donut Regression Discontinuity plots with 24 months bandwidth on each side of the cutoff.



(a) Polynomial of order 4



(b) Polynomial of order 4

Notes: (i) Donut RD dropping individuals within one month of the cutoff. (ii) Figure A3 (b) has 95% confidence intervals around the binned means calculated following Calonico et al. (2017).

Figure A5: Fraction of the entire Dutch population with "Not in good mental health"-status

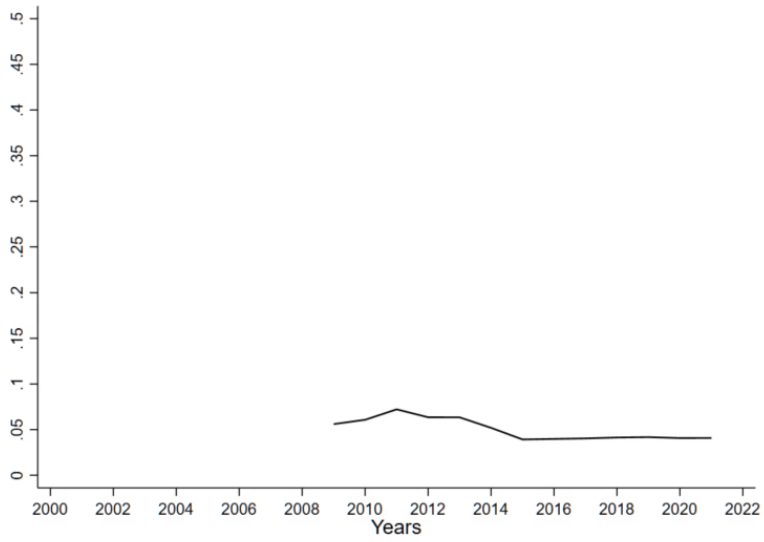
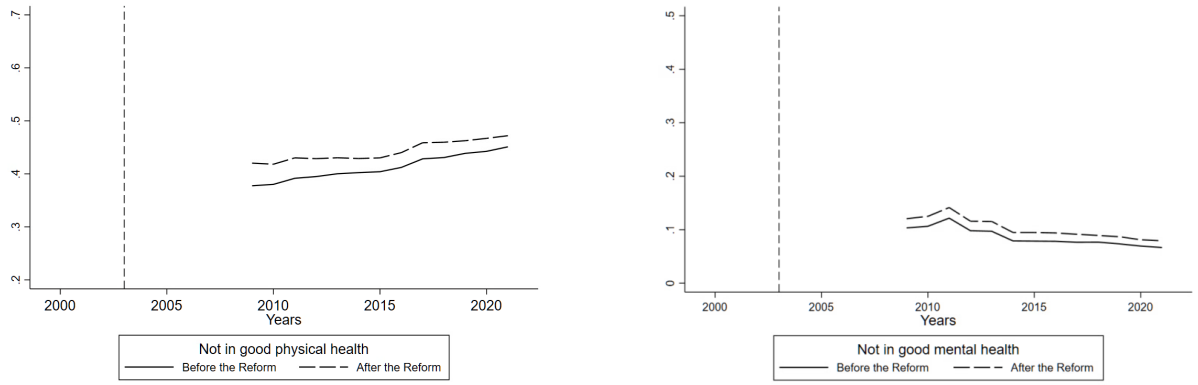


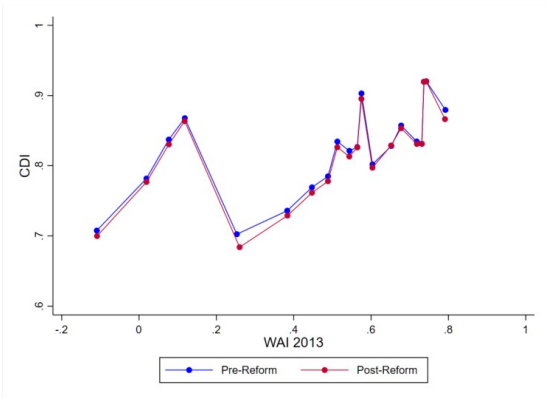
Figure A6: Fraction with "Not in good health status" among DI survivor applicant cohorts before (2001-2002) and after (2003-2004) the GKP reform (by year)



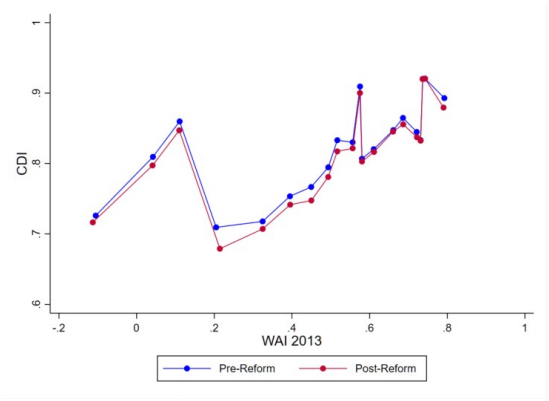
(a) Physical health

(b) Mental health

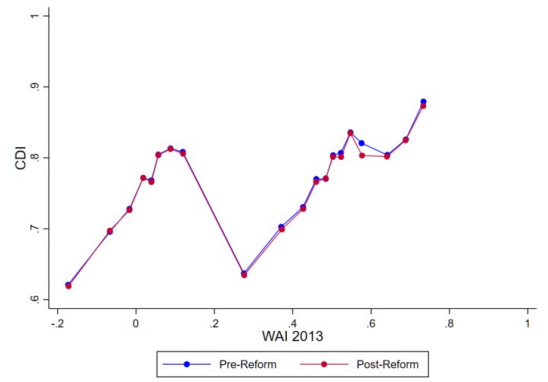
Figure A7: Bin-scatters of CDI and WAI pre- (2001-2002) and post-GKP (2003-2004) cohorts, 10 years after application (20 bins) and by impairment type at moment of award decision



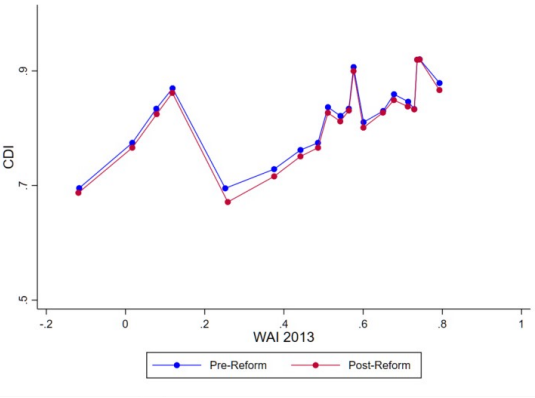
(a) Musculo-skeletal



(b) Mental



(c) Cardiovascular



(d) Other impairments

B Online Appendix : Tables

Table B1: Mapping of pharmaceutical data to chronic diseases

Chronic Disease	**ATC Code(s)**	**Medicine Description**
Acid related disorders	A02	Drugs for acid related disorders
Bone diseases (osteoporosis)	M05	Drugs for treatment of bone diseases
Cancer	L01	Antineoplastic agents
Cardiovascular diseases (inc, hypertension)	B01A, C01, C04A, C02, C07, C08, C09	Antithrombotic agents, cardiac therapy peripheral vasodilators, antihypertensives, beta blocking agents, calcium channel blockers, agents acting on the renin-angiotensin system
Dementia	N06D	Anti-dementia drugs
Diabetes (mellitus)	A10A, A10B, A10X	Insulins and analogues, blood glucose lowering drugs, other drugs used in diabetes
Epilepsy	N03	Antiepileptics
Glaucoma	S01E	Antiglaucoma preparations and miotics
Gout, Hyperuricemia	M04	Antigout preparations
HIV	J05A	Direct acting antivirals
Hyperlipidemia	C10	Lipid modifying agents
Inflammatory diseases	A07E	Intestinal antiinflammatory agents
(Iron deficiency) anemia	B03A	Iron preparations
Migraines	N02C	Antimigraine preparations
Pain	N02A, N02B	Opioids, other analgesics and antipyretics
Parkinson's disease	N04, N05B, N05C	Anxiolytics, hypnotics and sedatives
Psychological disorders	N06A	Antidepressants
Psychoses	N05A	Antipsychotics
Respiratory illnesses	R03	Drugs for obstructive airway diseases
Rheumatological conditions	L04A	Immunosuppressants
Thyroid disorders	H03	Thyroid therapy
Tuberculosis	J04A	Drugs for treatment of tuberculosis

Note: Classification is based on Huber et al. (2013).

Table B2: RDiT estimates of the Gatekeeper Protocol on the Survival rate of DI applicants

	within 5 years	**within 10 years**	**within 15 years**	**within 18 years**
GKP effect	-0.005*	-0.008***	-0.011***	-0.012**
	(0.002)	(0.002)	(0.003)	(0.003)
Constant	0.962***	0.932***	0.900**	0.874***
	(0.002)	(0.001)	(0.001)	(0.002)
Number of observations	282,421	282,421	282,421	282,421

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) ***/**/* indicate significance at the 1%/5%/10% level.

Table B3: RDiT estimates of the Gatekeeper Protocol on the Survival rate of DI applicants with controls

	within 5 years	**within 10 years**	**within 15 years**	**within 18 years**
GKP effect	-0.000	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.003)	(0.004)
Male	-0.015***	-0.027***	-0.042***	-0.055***
	(0.000)	(0.001)	(0.001)	(0.001)
Age at application (ref. = 25-34)				
35-44	-0.007***	-0.013***	-0.022***	-0.029***
	(0.000)	(0.000)	(0.001)	(0.001)
45-54	-0.025***	-0.047***	-0.078***	-0.102***
	(0.000)	(0.000)	(0.001)	(0.002)
55-65	-0.059***	-0.108***	-0.171***	-0.220***
	(0.001)	(0.002)	(0.002)	(0.003)
Diagnosis (ref. = diseases of the blood)				
cardiovascular diseases	0.203***	0.231***	0.213***	0.204***
	(0.009)	(0.009)	(0.009)	(0.009)
diseases of the skin	0.128***	0.176***	0.186***	0.194***
	(0.006)	(0.006)	(0.008)	(0.011)
endocrine disorders	0.190***	0.225***	0.221	0.213***
	(0.008)	(0.010)	(0.011)	(0.013)
diseases of pregnancy	0.209***	0.265***	0.278	0.288***
	(0.009)	(0.009)	(0.008)	(0.008)
diseases of the ear	0.234***	0.295***	0.318	0.339***
	(0.010)	(0.011)	(0.011)	(0.013)
nervous system	0.171***	0.207***	0.203	0.204***
	(0.008)	(0.009)	(0.009)	(0.010)
respiratory system	0.007	0.020*	0.004	-0.000
	(0.010)	(0.011)	(0.009)	(0.009)
digestive system	0.034***	0.074***	0.089***	0.106***
	(0.010)	(0.009)	(0.009)	(0.010)
genitourinary system	0.040***	0.048***	0.042	0.046***
	(0.008)	(0.009)	(0.008)	(0.011)
diseases of the eye	0.214***	0.266***	0.280	0.297***
	(0.007)	(0.010)	(0.009)	(0.008)
mental disorders	0.217***	0.275***	0.290	0.301***
	(0.009)	(0.009)	(0.008)	(0.008)
musculo-skeletal	0.225***	0.286***	0.301	0.312***
	(0.009)	(0.009)	(0.008)	(0.009)
other impairments	0.206***	0.260***	0.276***	0.287***
	(0.009)	(0.010)	(0.009)	(0.009)
Constant	0.790***	0.731***	0.718**	0.707***
	(0.009)	(0.009)	(0.009)	(0.009)
Number of observations	282,421	282,421	282,421	282,421

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) ***/**/* indicate significance at the 1%/5%/10% level. (3) The reference category for the age cohort is 25-34, the baseline category for impairments is diseases of the blood.

Table B4: The effect of the GKP reform on the composition of DI applicants: health estimates

	Not in good health		**Not in good physical health**		**Not in good mental health**	
	(1)	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**
RDiT Estimates	0.024***	0.016***	0.023***	0.017***	0.004	0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)
Constant	0.401***	0.442***	0.407***	0.444***	0.101***	0.082***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.002)
Number of observations	257,327	246,508	257,327	246,508	257,941	246,784

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) ***/**/* indicate significance at the 1%/5%/10% level. (3) Columns (1) and (2) represent results after 10 and 15 years from the award year for not in good health and not in good physical health estimates, and they represent year 2012 and 2017 for not in good mental health estimates.

Table B5: The effect of the GKP reform – with controls: health estimates

	Not in good health		Not in good physical health		Not in good mental health	
	(1)	(2)	(1)	(2)	(1)	(2)
RDiT Estimates	0.021***	0.012**	0.019***	0.012**	0.007**	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Male	-0.016***	0.002	-0.027***	-0.005**	0.002	0.004***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)
Age at application (ref.= 25-34)						
35-44	0.021***	0.049***	0.044***	0.061***	-0.033***	-0.029***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)
45-54	0.069***	0.117***	0.121***	0.143***	-0.077***	-0.072***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
55-65	0.016***	0.002	0.194***	0.233***	-0.114***	-0.095***
	(0.003)	(0.002)	(0.005)	(0.004)	(0.002)	(0.001)
Diagnosis (ref. = diseases of the blood)						
cardiovascular diseases	0.010	0.025**	0.007	0.021*	0.002	0.003
	(0.018)	(0.011)	(0.017)	(0.011)	(0.006)	(0.003)
diseases of the skin	-0.074**	-0.042	-0.098***	-0.053**	0.008	0.013
	(0.025)	(0.026)	(0.025)	(0.023)	(0.010)	(0.008)
endocrine disorders	0.109***	0.089***	0.098***	0.088***	0.012	0.006
	(0.024)	(0.011)	(0.022)	(0.012)	(0.008)	(0.008)
diseases of pregnancy	-0.0186***	-0.169***	-0.213***	-0.177***	0.003	0.003
	(0.019)	(0.015)	(0.020)	(0.014)	(0.010)	(0.008)
diseases of the ear	-0.102***	-0.115***	-0.118***	-0.120***	0.027	0.013
	(0.022)	(0.020)	(0.020)	(0.019)	(0.008)	(0.011)
nervous system	-0.023	-0.029	-0.40*	-0.025*	0.011	0.010
	(0.020)	(0.011)	(0.019)	(0.012)	(0.008)	(0-007)
respiratory system	0.078***	0.064***	0.062***	0.058**	0.022***	0.015*
	(0.023)	(0.019)	(0.019)	(0.019)	(0.006)	(0.007)
digestive system	0.033*	0.041***	0.015	0.038***	0.014	0.012
	(0.018)	(0.011)	(0.015)	(0.011)	(0.011)	(0.009)
genitourinary system	-0.039*	-0.046***	-0.060***	-0.054***	0.002	0.008
	(0.019)	(0.009)	(0.017)	(0.008)	(0.008)	(0.006)
diseases of the eye	-0.0114***	-0.090***	-0.131***	-0.098***	0.011	0.006
	(0.018)	(0.022)	(0.019)	(0.022)	(0.012)	(0.008)
mental disorders	-0.137***	-0.130***	-0.196***	-0.153***	0.088***	0.076***
	(0.018)	(0.010)	(0.017)	(0.011)	(0.007)	(0.006)
musculo-skeletal	-0.127***	-0.113***	-0.143***	-0.118***	0.000	0.006
	(0.019)	(0.009)	(0.018)	(0.010)	(0.007)	(0.006)
other impairments	-0.135***	-0.128***	-0.169***	-0.140***	0.035***	0.035***
	(0.018)	(0.011)	(0.017)	(0.011)	(0.007)	(0.006)
Constant	0.451***	0.458***	0.447***	0.452***	0.120***	0.097***
	(0.018)	(0.012)	(0.017)	(0.012)	(0.008)	(0.006)
Number of observations	257,327	246,508	257,327	246,508	257,941	246,784

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) ***/**/* indicate significance at the 1%/5%/10% level. (3) The reference category for age cohort is 25-34, the baseline category for impairments is diseases of the blood. (4) Columns (1) and (2) represent results after 10 and 15 years from the award year for not in good health and not in good physical health estimates, and they represent year 2012 and 2017 for not in good mental health estimates.

Table B6: RDiT estimates of the Gatekeeper reform on the employment rate of DI applicants

	within 1 year	within 5 years	within 10 years	within 15 years
Employment estimate	-0.025***	-0.028***	-0.032***	-0.028***
	(0.009)	(0.006)	(0.006)	(0.005)
Constant	0.691***	0.597***	0.491**	0.411***
	(0.006)	(0.006)	(0.006)	(0.006)
Number of observations	279,819	268,862	259,913	249,631

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

Table B7: RDiT estimates of the Gatekeeper reform on the employment rate of DI applicants

	within 1 years	**within 5 years**	**within 10 years**	**within 15 years**
Employment estimate	-0.028***	-0.029***	-0.027***	-0.030***
	(0.010)	(0.007)	(0.006)	(0.007)
Male	0.014**	0.010***	-0.004	-0.008***
	(0.004)	(0.002)	(0.002)	(0.002)
Age at application (ref. = 25-34)				
35-44	0.037***	0.010***	-0.011***	-0.040***
	(0.004)	(0.003)	(0.003)	(0.003)
45-54	0.041***	-0.084***	-0.211***	-0.376***
	(0.006)	(0.002)	(0.002)	(0.006)
55-65	-0-094***	-0.466***	-0.641***	-0.626***
	(0.016)	(0.009)	(0.007)	(0.008)
Diagnosis (ref. = diseases of the blood)				
cardiovascular diseases	0.004	-0.019*	-0.057***	-0.069***
	(0.011)	(0.010)	(0.011)	(0.010)
diseases of the skin	0.026*	0.059***	-0.011	-0.034*
	(0.013)	(0.016)	(0.015)	(0.015)
endocrine disorders	0.032**	0.013	-0.043***	-0.063***
	(0.012)	(0.017)	(0.011)	(0.012)
diseases of pregnancy	0.038**	0.068***	0.065***	0.072***
	(0.015)	(0.015)	(0.013)	(0.014)
diseases of the ear	0.030	0.002	-0.063***	-0.077***
	(0.018)	(0.019)	(0.019)	(0.016)
nervous system	-0.034***	-0.096***	-0.0128***	-0.135***
	(0.010)	(0.013)	(0.013)	(0.010)
respiratory system	-0.034**	-0.052***	-0.075***	-0.080***
	(0.013)	(0.014)	(0.013)	(0.012)
digestive system	0.012	0.021	-0.026	-0.050***
	(0.017)	(0.012)	(0.015)	(0.013)
genitourinary system	0.030***	0.050***	0.026**	-0.009
	(0.015)	(0.010)	(0.010)	(0.010)
diseases of the eye	-0.004	-0.052**	-0.090***	-0.088***
	(0.014)	(0.021)	(0.012)	(0.015)
mental disorders	0.024	0.036**	-0.002	-0.026**
	(0.015)	(0.012)	(0.012)	(0.011)
musculo-skeletal	0.038***	0.050***	0.004	-0.017
	(0.013)	(0.012)	(0.011)	(0.011)
other impairments	0.163***	0.136***	0.089***	0.052***
	(0.014)	(0.012)	(0.011)	(0.011)
Constant	0.634***	0.680***	0.692**	0.686***
	(0.011)	(0.013)	(0.011)	(0.011)
Number of observations	279,819	268,862	259,913	249,631

Notes:[(a)] (1) Standard errors are clustered at the month of the year level. (2) *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level. (3) The baseline category for age cohort is 25-34, the baseline category for impairments is diseases of the blood.