Decomposing Employment Trends of Disabled Workers*

Pierre Koning † Heike Vethaak ‡

May 6, 2020

Abstract

Many OECD countries are facing decreases in the employment rates of disabled workers. To uncover the driving forces of these trends, this paper estimates Age-Period-Cohort (APC) models on administrative data of Disability Insurance (DI) application cohorts for the Netherlands between 1999 and 2013. We find that the substantial decrease in employment rates of applicant cohorts in this time period is almost fully explained by cohort effects – equalling about 30 percentage points – and that the impact of period effects is only small. In turn, cohort effects stem from changes in the observed composition of applicants, with increasing shares of workers without (permanent) contracts in the year before the application. These changes are largely confined to years following two major DI reforms that increased self-screening among potential applicants. We next expand the APC model with distinct APC-effects for awarded and rejected DI applicants. Assuming common compositional cohort effects for these two groups, difference-in-difference estimates of cohort effects indicate that the effect of changes in benefit conditions ('incentive effects') is limited. Disability reforms thus induced substantial self-screening in the sickness period before the DI decision, rather than changing individual employment rates.

JEL-codes: H75, J21, C23

Keywords: Employment, disability insurance, APC-models, self-screening.

*This research is sponsored by Instituut Gak. We are grateful to Patrick Hullegie for his indispensable work on the data. We also thank Nicole Maestas, Jan-Maarten van Sonsbeek, Eric French, Astrid Grasdal, Lawrence Katz, Marike Knoef, Carla van Deursen, Barend Barentsen and seminar participants at the Dutch Economists Day of 2019, the University of Bergen, the University of Nantes and the KVS New Paper Sessions of 2019 for useful comments to presentations and earlier drafts of the paper.

†Universiteit Leiden, Vrije Universiteit Amsterdam, TI and IZA. Email: p.w.c.koning@law.leidenuniv.nl.

‡Universiteit Leiden. Corresponding author. P.O. Box 9500, 2300 RA Leiden, The Netherlands. Email: h.t.vethaak@law.leidenuniv.nl.

1 Introduction

Over the last decades, many OECD countries have faced increased Disability Insurance (DI) inflow rates and declining employment rates of disabled individuals (OECD, 2010). Particularly for the US, there is strong evidence that the Social Security Disability Insurance (SSDI) program has become relatively more attractive for low-skilled workers (Maestas, 2019; Autor & Duggan, 2003; Bound et al., 2003, 2014; Von Wachter et al., 2011). Since the mid-eighties, the expansion of the SSDI program went together with higher fractions of applicants with mild impairments for whom the receipt of benefits has discouraged them from working.[1] Accordingly, the decline in employment among SSDI recipients can be attributed to both changes in the composition of (potential) applicants – with vulnerable labor market positions – and changes in incentive effects of the scheme.

This paper analyzes employment trends of Disability Insurance (DI) applicants in the Netherlands, a country that also experienced strong decreases in the labor force attachment of claimants. Unlike the SSDI program, however, drastic reforms have been implemented to curb the inflow into DI and increase work incentives for disabled workers. The potential impact of these reforms is twofold. On the one hand, increases in screening stringency and eligibility thresholds of the DI scheme may have changed the composition of DI applicants (Deshpande & Li, 2019). As a result, it is likely that the severity of DI claims has increased and the employment rate of new application cohorts has decreased (Godard et al., 2019; De Jong et al., 2011).[2] On the other hand, one of the reforms increased work incentives for DI recipients with residual earnings capacities (Koning & van Sonsbeek, 2017). This may have improved the employment rates of awarded applicants. The overall assessment of the employment effects of the Dutch reforms therefore should incorporate both the effects on the composition of DI applicants and on their individual behavior.

[1]To estimate the discouraging impact of SSDI benefits, Bound (1989), Chen & Van der Klaauw (2008), Von Wachter et al. (2011), Maestas et al. (2013) and French & Song (2014) compare accepted and denied SSDI applicants. Following the seminal article by Bound (1989), the resulting estimates form an upper bound of the employment rates of awarded applicants, since rejected applicants are considered to have more labor market attachment than accepted applicants.

[2]Contributions of Campolieti (2006) for Canada, **?** for the US, Markussen et al. (2018) for Norway and Liebert (2019) for Switzerland suggest that increased scrutiny and increased application costs have the potential to substantially lower DI inflow rates.

To provide such a broad assessment of employment trends of disabled workers, this paper is the first to estimate Age-Period-Cohort (APC) models on administrative data on DI applicant cohorts. We use data on DI applicant cohorts between 1999 and 2013 which are followed up to 2016. In the context of the APC model, 'age' corresponds to the elapsed duration since application, period effects capture business cycle and other calendar time effects and cohort effects resemble changes in employment rates that are specific to DI application cohorts. With reforms in the Netherlands affecting new applicant cohorts only, policy effects are thus captured by cohort effects. Using a conventional Deaton-Paxson (DP) specification, we first disentangle cohort effects from period and age effects. The cohort effects represent the joint effect of (i) compositional changes induced by gradual cohort-specific time trends in the demand for workers with disabilities; (ii) compositional changes induced by disability reforms that affected self-screening before application and (iii) individual changes in the employment rate of awarded applicants – or: 'incentive effects' – induced by changes in benefit conditions.

Our second aim is to provide a further decomposition of cohort effects into changes stemming from compositional changes and changes in the individual's employment probability stemming from DI reforms. In the spirit of Bound (1989), we follow a Difference-in-Difference (DiD) approach with (partially) awarded and rejected DI applicants as treatment and control groups, respectively. Assuming that compositional effects – both induced by reforms and gradual changes in the labor market – affected treatment and control cohort groups equally, the DiD estimates of the reforms indicate changes in the individual employment probability of awarded applicants. These changes can be characterized as 'incentive' effects of the reforms on awarded applicants.

We argue that the Netherlands provides an interesting setting to study the distinct effects of cohort and period effects, as drastic and seemingly effective changes both in the eligibility to the DI scheme and in worker and employer incentives were effectuated in 2003 and 2006. Both these reforms aimed to curb the high level of DI inflow and DI enrollment that prevailed at the turn of the century, amounting to about 12% of the working population. In 2002, the Gatekeeper Protocol (GKP) increased the reintegration responsibilities of employers and workers in the sickness period that precedes the DI application (De Jong et al., 2011; Koning & Lindeboom, 2015). In turn, this reform affected the size and composition of new DI application cohorts since 2003. With the evidence pointing at strong increases of screening and self-screening – see e.g. Koning & Lindeboom (2015) and Godard et al. (2019) – our expectation is that the GKP implied a decrease in the average employment rate of applicant cohorts since 2003. In 2006 a new disability law (WIA) was implemented for new cohorts of DI applicants. This new scheme implied (i) an extension of the waiting period before DI application from one to two years; (ii) a higher threshold for eligibility to partial DI benefits; and (iii) stronger work incentives for individuals with partial DI benefits. It is likely that these changes have altered both the composition of DI applicants and the work incentives of awarded applicants. This particularly holds for individuals with substantial residual earnings capacity that either are rejected benefits or awarded partial benefits.

Our main research findings can be summarized as follows. First, cohort effects of DI applicants are the main contributor to their observed decline in employment, amounting to about 30 percentage points in total. Contrasting to this, the effect of calendar time effects is negligible, suggesting that both business cycle effects or more gradual – and possibly non-stationary – time trends that affected all cohorts equally were not important. Second, a substantial part of the changes in cohort effects is explained by changes in demographic variables and the initial labor market position of applicants. As far as we can infer from the inclusion of observed controls, there is a general worsening in the labor market position of more recent cohorts. This finding resembles e.g. Autor & Duggan (2003), Von Wachter et al. (2011) and Maestas et al. (2013) who argue there is a declining demand for low-skilled workers with health conditions in the US. Third, changes in cohort effects are largely in tandem with the disability reforms of 2003 and 2006; it is only for the years after the 2006 reform that we observe a gradual and substantial further decline in cohort effects. Finally, our DiD-analysis provides limited evidence for employment rates to respond to changes in the work incentives of awarded applicants. This implies that the substantial changes in cohort effects are almost entirely driven by compositional changes of applicants. Again, this highlights the importance of self-screening among potential applicants as a driver of the observed changes in employment rates.

The remainder of the paper is organized as follows. The next section describes the DI system in the Netherlands with the legislative changes and the accompanying expected behavioral responses. Section 3 provides a description of the selected data and Section 4 contains the methodological framework for the analysis. Section 5 presents the results of the analysis before Section 6 concludes.

2 Institutional background

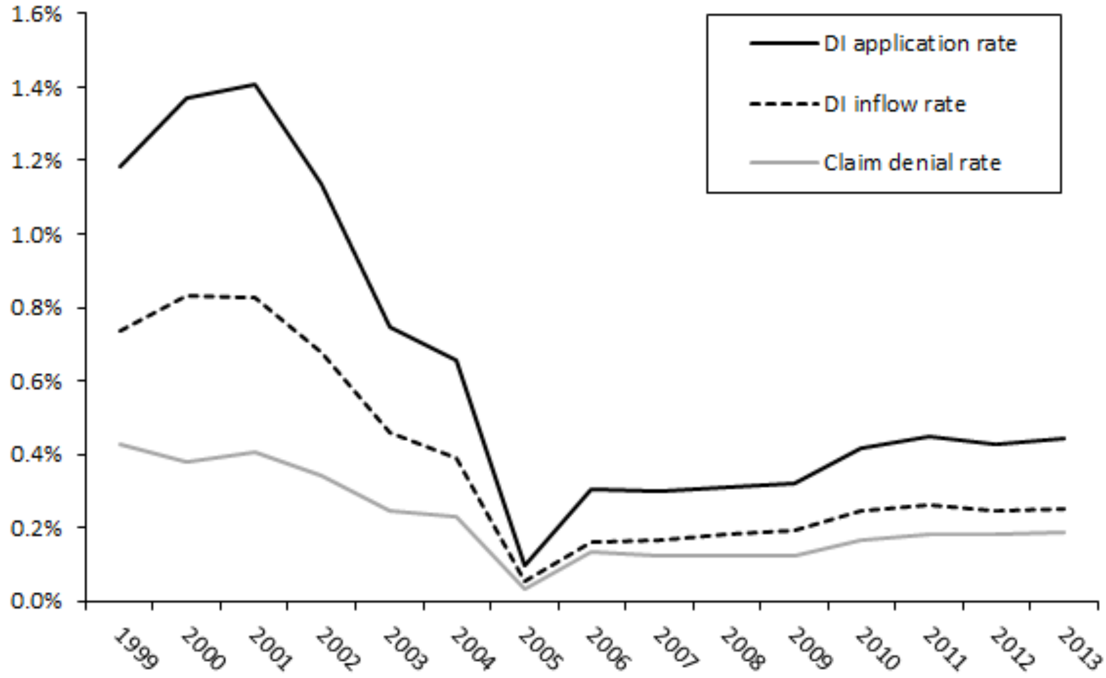
This section describes the main characteristics of the Dutch DI system and the two major disability reforms since 1999: the Gatekeeper Protocol (in Dutch: *Wet verbetering Poortwachter*) and the WIA (in Dutch: *Wet Werk en Inkomen naar Arbeidsvermogen*). From now on, we refer to these reforms as the GKP and WIA, respectively. When explaining these reforms, a particular interest lies in the distinction between expected compositional effects and incentive effects of these reforms. Specifically, we define compositional effects as changes in the average employment rates that result from changes in the composition of new cohorts of DI applicants. As to DI applicants, these changes stem from changes in self-screening and work resumption in the waiting period before the DI decision. In addition, incentive effects are defined as changes in individual employment rates as a response to changes in the work incentives for awarded DI applicants, measured after the DI award decision.

2.1 DI in the Netherlands

The Dutch DI program covers income losses resulting from both occupational and non-occupational injuries of all employed workers. Sick-listed workers apply for DI benefits at the end of the waiting period of absence. The employer is obliged to continue wage payments in this period, which was extended from one to two years in 2004 (thus affecting new applicant cohorts as from 2006).

After application, the National Social Insurance Institute (NSII) determines the disability degree of workers. To this end, medical examiners assess the limitations

Figure 1: Annual DI application rate, inflow rate and claim denial rate of total insured working population, 1999-2013



Source: Statistics Netherlands

of applicants and vocational experts subsequently select occupations with corresponding wages to determine the residual potential earning capacity. The degree of disability then equals the lost potential earning capacity as a fraction of the pre-disability earnings. Until 2006, the applicant was awarded DI benefits if the disability degree exceeded the minimum threshold of 15%. This threshold was increased to 35% as part of the WIA reform in 2006. Workers with disability degrees between 35 up to and including 80% are awarded partial DI benefits and those with losses of more than 80% receive full benefits. Partially disabled receive 70 percent of their loss of earnings capacity and fully disabled receive 70 percent of their pre-disability earnings.

With its broad coverage, its generous benefits and its limited role for self-screening, the Dutch DI system laid the ground for a continuous increase in DI enrollment. Around the turn of the century, DI enrollment peaked at about 12% of the insured working population (Koning & Lindeboom, 2015). As Figure 1 shows, in this period annual DI application rates ranged between 1.2 and 1.4% of the working population. Since then, the first substantial drop in both DI application and awards occurred in 2003, at the time the GKP affected DI claims. Using a discontinuity-in-time regression, Godard et al. (2019) find that the effect amounted to a 40 percent reduction in the DI applicant rate. The second major decrease in DI application and award rates is observed since 2005. While this drop initially demarcates the transitory effect of the extension of the sickness period to two years, the new disability law (WIA) has led to persistently lower DI inflow rates. In what follows, we discuss both the GKP reform and the WIA reform in more detail.

2.2 Stricter screening: the GKP reform (2003)

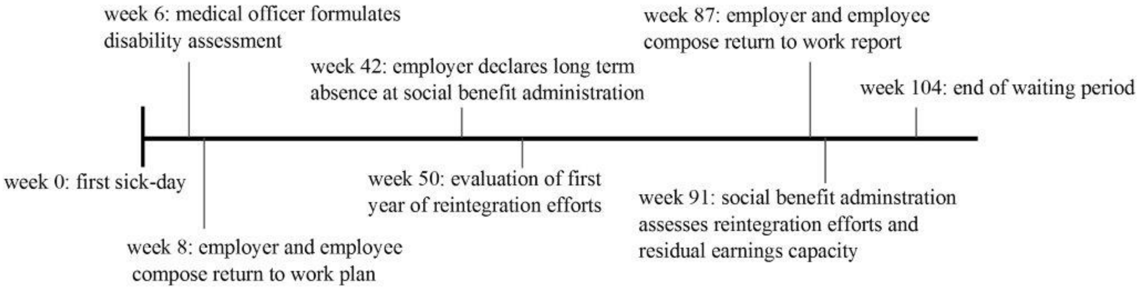
The GKP reform has affected the screening process for new DI application cohorts since 2003. The GKP stipulates the responsibilities of both the worker and the employer for sickness spells lasting at least six weeks. This means the responsibility of reintegrating sick workers during the waiting period was removed from the NSII, which since then acts as a gatekeeper at the moment of DI claim only. Figure 2 provides an overview of the steps of the application process towards entering DI under the GKP.[3] After six weeks of absence, the worker and the employer must draft a rehabilitation plan together which is based on an assessment of cause of disability, functional limitations and the likelihood of work resumption. The rehabilitation plan should be approved by a caseworker of the NSII in the eighth week of absence, after which it is binding for both parties. The worker can apply for DI benefits if work resumption is not established before the end of the waiting period and when all requirements of the GKP have been met. If not, the wage continuation period may be extended with one year at maximum.

There is strong evidence that the GKP changed the composition of DI applicants. The increased rehabilitation efforts did not only increase the likelihood of work resumption in the absence period that precedes DI claims for workers with better employment prospects, but also induced self-screening among those workers with less severe health conditions (De Jong et al., 2011; Godard et al., 2019).[4] Both

[3]Note that the figure is relevant under the (current) disability scheme with an absence period of two years. In the year the GKP came into force, the waiting period was one year.

[4]De Jong et al. (2011) evaluate a large-scale experiment in the Netherlands to study the effects

Figure 2: GKP conditions in the sickness waiting period



these mechanisms have resulted in a sample of DI applicants that are probably more deserving and with lower employment rates.[5]

2.3 The new disability law: WIA (2006)

The main goal of the WIA reform of 2006 was to stimulate workers with less-severe impairments to exploit their residual earnings capacity. The idea was that three policy changes would contribute to this: (i) increased self-screening through an extension of the waiting period, from one to two years; (ii) stricter eligibility, as the threshold for DI receipt was increased to 35%; and (iii) differentiated benefits for severely disabled and applicants with sufficient remaining earnings capacity.

First, the extension of the waiting period from one to two years implied another increase in the costs of wage continuation and all other costs inherent with the GKP. Following similar arguments as for the introduction of the GKP, one would thus expect this extension to increase work resumption and self-screening in the waiting period before DI application.

As a second part of the WIA, the threshold of the disability degree for eligibility was increased from 15 to 35 percent of pre-disability earnings. Van Sonsbeek & Gradus (2012) argue that this implied a drop in DI inflow rates of roughly 20

of increased screening. They find that this induces employers to increase reintegration activities, which in turn increases work resumption rates during sickness absenteeism. They argue that those higher rates are induced by self-screening among the potential applicants.

[5] Koning & Lindeboom (2015) argue that the increased application costs of the GKP may also have had adverse effects on the individual employment rates of disabled workers. The increased responsibilities and the risk of extension of wage sanctions – i.e., the increase of the wage continuation period – may have discouraged employers to hire workers with disabilities (see also Hullegie & Koning, 2018).

percentage points. With a substantially lower share of beneficiaries with partial benefits, it is expected that the average employment rate among the total group of beneficiaries has declined. This compositional effect may have been strengthened by increased self-screening among (potential) applicants with mild health conditions.

Third, the WIA differentiates between fully and permanently disabled workers (IVA) partially and/or temporary disabled workers (WGA) for which strong financial incentives were introduced. Workers in the WGA scheme receive 70 percent of their lost earnings during the first period of benefit receipt ('wage-related related benefits'). Depending on the work history, this period lasts 38 months at maximum. Next, WGA beneficiaries continue receiving the same benefit level if and only if they exploit at least 50 percent of their earnings capacity; if not, the benefit is based on the statutory minimum wage. Benefits for partially disabled workers thus function as a wage subsidy that incentivizes them to work.[6] Koning & van Sonsbeek (2017) find that the incentive change for partially disabled workers increases the employment incidence with 2.6 percentage points.[7] Still, the overall effect of the increase in incentives is probably smaller than this, as wage subsidies are targeted at partially disabled workers – constituting about one quarter of the total DI inflow – and are relevant in the second period of benefit receipt only (Koning & Lindeboom, 2015).[8]

Overall, the emerging picture is that the GKP and the WIA reform most likely affected the composition of the pool of new DI applicant cohorts. Increases in self-screening and increases in work resumption in the absence period probably have resulted in a smaller sample of DI applicants with more severe health conditions and lower employment rates. For applicants who were awarded benefits, the WIA reform also changed the incentive to work, albeit that effect is probably small.

[6]For a detailed explanation of the functioning and consequences of the wage subsidy, we refer to Koning & van Sonsbeek (2017).

[7]Kantarci et al. (2019) find somewhat smaller employment effects, comparing sick-listed worker cohorts that fell under the old and new disability scheme, respectively. In their study, the effect estimate of work incentives can be interpreted as an upper bound, as it also captures the effect of the waiting period extension from one to two years.

[8]The wage subsidy may have induced perverse work incentives for fully and temporary disabled workers in the WGA scheme, as switches to the partial scheme inhibit the risk of sizable declines in benefits (Koning & Lindeboom, 2015).

3 Data

3.1 Data sources

We use individual-level data on all DI applications between 1999 and 2013 from the administrative records of the NSII. Cohorts from these years are followed between 1999 and 2016. Records contain information on the award decision and date, the diagnosed impairment and the assessed degree of disability.[9] Medical diagnoses are grouped by impairment type (mental, musculoskeletal, respiratory, endocrine, cardiovascular, nervous system and other impairments).[10] The degree of disability is given by intervals (<15%, 15-34%, 35-44%, 45-54%, 55-64%, 65-79%, ≥80%). Note that from 2006 onward the group with the lowest disability degree is <35%.

We merge the application data with administrative data of Statistics Netherlands of the full Dutch population between 1999 and 2016. This yields individual-year data covering a sufficiently long period to assess the long-term effects of both the GKP and WIA reform. The Census Register contains information on the personal characteristics, such as gender, month of birth and death, and nationality. The tax records provide information (in 2015 Euros) on annual gross earnings and receipt of unemployment, disability, and social assistance benefits.[11] We define an individual as employed in a specific year when he or she received positive earnings. For employed individuals we also observe the contract type (permanent or temporary) and sector of employment (70 in total).

In total, we observe 1,183,186 individual applications between 1999 and 2013. For our empirical analysis, we exclude re-applications, workers that are younger than 18 or older than 65 at the time of application and workers for which the year of application or award decision was unknown. This reduces our sample to 962,356

[9] After 2007 we observe a shift in the data from rejections based on a too low degree of disability to rejections based on for 'unknown' reasons (see Figure A.10 in the Appendix). This probably reflects administrative changes, as the medical assessment was unchanged and rejection rates remained more or less constant. For this reason, our analysis does not differentiate between different reasons for rejection.

[10] The distribution of impairment groups by application cohorts is shown in Figure A.11.

[11] The records on disability insurance benefits also include information on the degree of disability in that year. The degree of disability of an individual can differ between years because of reassessment by the NSII.

observations. Attrition from this longitudinal sample stems from the occurrence of deaths and migration.

3.2 Descriptive statistics

Table 1 presents descriptive statistics of employment and earnings of rejected and awarded applicant cohorts, measured before and after the DI decision. We separate the total sample of applicants in three sub-samples or regimes: (i) cohorts unaffected by the reforms, 1999-2002; (ii) cohorts covered by the GKP but not by the WIA, 2003-2005; and (iii) cohorts subject both to the GKP and the WIA, 2006-2013.[12] The table shows that rejected and awarded DI applicants with full benefits have similar pre-disability employment rates two years before the DI assessment. Inherent with the eligibility conditions for DI, these rates are close to 100%. Applicants awarded partial benefits have higher pre-disability earnings and have more often a permanent contract than those rejected and those awarded full benefits. These higher earnings reflect the fact that percentage drops in earnings capacity are higher for applicants with higher pre-disability earnings. As expected, awarded applicants experience drops in income from earnings that are more sizable than for rejected applicants. Awarded applicants tend to be more often male, older and show higher mortality rates than rejected applicants. Over the years, we also observe substantial changes in the employment rates and the composition of DI applicants. Most notably, in the last time frame (2006-2013) applicant cohorts show markedly lower employment rates two years after application. This drop is most sizable for applicants awarded full DI benefits.

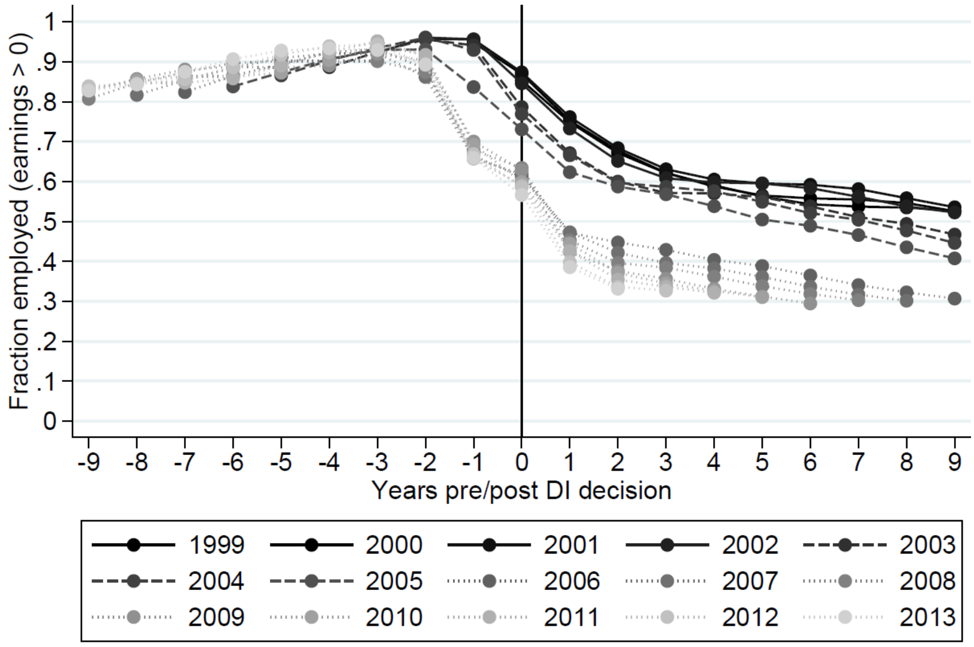
To shed more light on longitudinal patterns, Figure 3 depicts the evolution of employment rates of applicant cohorts before and after the award decision. Figure 4 shows a similar graph for separate samples of rejected, partially awarded and fully awarded applicants, with separate panels for the three regimes as in Table 1. From the figures, four general observations stand out. First, employment rates generally increase up to two years before the award decision and decline thereafter. While the

[12] The GKP affected sick-listed workers as from 2002. Hence, DI applicants of 2002 are not affected. Likewise, the extension of the waiting period from one to two years affected workers that became sick from 2004 onwards.

Table 1: Employment, earnings and demographic characteristics for rejected, and partially and fully awarded DI applicant cohorts

Application cohort	1999-2002			2003-2005			2006-2013		
Application status	Rejected applicants	Awarded partial benefits	Awarded full benefits	Rejected applicants	Awarded partial benefits	Awarded full benefits	Rejected applicants	Awarded partial benefits	Awarded full benefits
Labor supply and earnings 2 years before application									
Percent positive covered earnings	91.1	95.6	90.2	95.2	97.2	94.5	90.0	91.2	89.3
Average annual earnings (€1,000)	25,477	34,404	26,901	25,220	35,304	27,071	23,104	34,086	26,278
Median positive annual earnings (€1,000)	23,000	33,000	24,000	24,000	34,000	25,000	21,000	33,000	24,000
Percent permanent contract	—	—	—	72.9	86.4	79.3	71.8	78.4	76.8
Labor supply and earnings in year of application									
Percent positive covered earnings	88.2	91.4	76.8	81.6	86.6	68.7	64.9	70.5	57.8
Average annual earnings (€1,000)	24,149	32,353	22,957	23,441	32,105	22,479	20,648	26,953	18,709
Median positive annual earnings (€1,000)	22,000	32,000	20,000	22,000	31,000	19,000	18,000	25,000	15,000
Percent permanent contract	—	—	—	66.8	78.8	59.6	53.1	62.5	52.5
Labor supply and earnings 2 years after application									
Percent positive covered earnings	74.5	78.6	41.8	69.2	74.1	38.5	58.0	59.5	18.4
Average annual earnings (€1,000)	25,301	29,006	19,467	23,916	29,051	19,893	21,475	23,556	17,408
Median positive annual earnings (€1,000)	24,000	28,000	16,000	22,000	28,000	17,000	19,000	21,000	13,000
Percent permanent contract	67.3	74.9	38.1	60.3	69.4	33.3	47.5	54.1	16.0
Demographics									
Average age at application	40	44	43	41	45	42	43	46	46
Age at application < 40	49.0	32.3	39.9	46.7	30.6	40.4	39.1	27.5	29.3
Age at application 40 - 50	27.5	32.0	27.2	28.9	31.9	27.7	30.3	29.2	26.0
Age at application $50 \leq$	23.5	35.8	32.9	24.5	37.5	31.9	30.6	43.3	44.7
Percent male	37.3	51.6	40.8	41.4	53.4	46.6	46.1	54.3	48.8
Percent Dutch	77.1	83.0	76.7	73.9	81.2	74.4	70.3	77.3	72.6
Percent death after application									
Percent deceased 2 years after application	1.7	1.0	4.5	1.2	1.3	5.9	1.0	1.2	6.2
Percent deceased 4 years after application	2.3	1.9	6.1	2.2	2.4	7.8	1.8	2.4	8.9
Observations	126,323	115,639	158,558	55,909	41,496	52,019	145,677	44,709	142,835

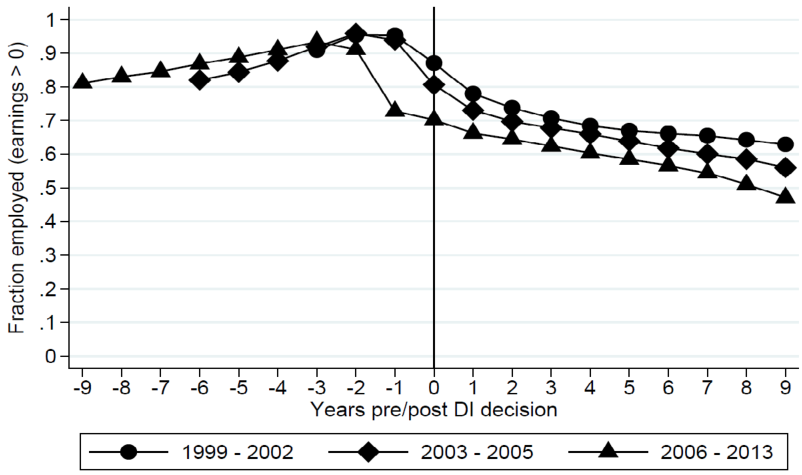
Figure 3: Annual fraction employed DI applicants before and after the DI decision, stratified by application year (1999-2013)



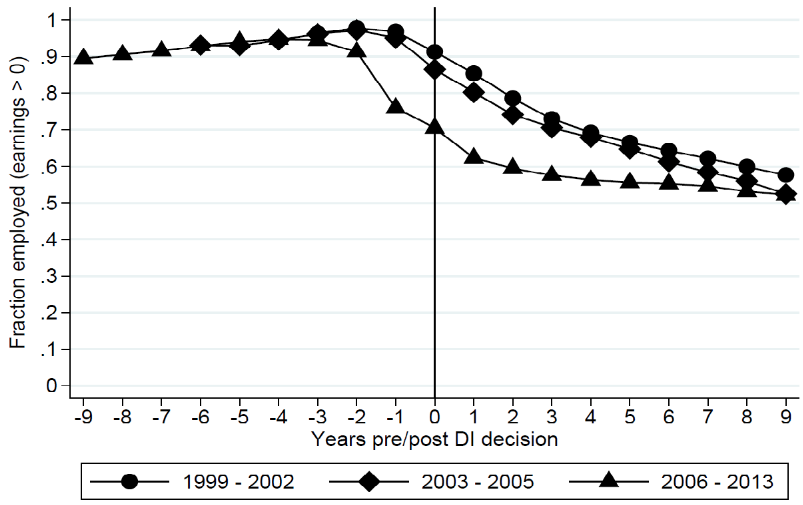
increase follows from the eligibility conditions inherent to the Dutch DI system, the subsequent decline follows from the start of the absence period that precedes the award decision. As expected, the declines are strongest for those awarded full DI benefits. Second, we observe large jumps in employment rates in the years the two reforms were implemented, but employment rates are roughly constant *within* the time periods of 1999-2002 and 2003-2005. This suggests that changes in employment rates until 2006 can largely be related to the GKP and WIA reform. Third, we observe important changes in the employment patterns of new cohorts after 2006, the year the new disability law came into force. Since then, a large share of the decline in employment is already observed in the absence period, two years before the disability decision. Finally, the employment of new applicant cohorts keeps on decreasing since 2006. Figure 3 suggests that the employment patterns of these subsequent cohorts have constant differences between successive cohorts. Following Voas & Chaves (2016), this indicates that either the impact of the period effects – or: business cycle – is limited, or the unlikely case that the 'age' effects (i.e., the elapsed duration) and period effects are almost perfectly balanced.

Figure 4: Annual average employment rates of rejected, and partially and fully awarded DI applicant cohorts for three time regimes, before and after application for DI benefits

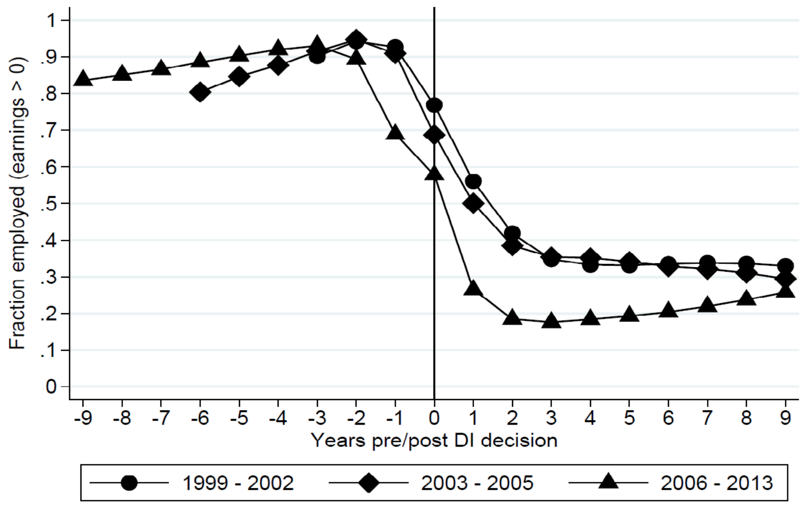
Panel A. Annual fraction employed of rejected DI applicants



Panel B. Annual fraction employed of applicants awarded partial DI benefits



Panel C. Annual fraction employed of applicants awarded full DI benefits



3.3 Comparing rejected and awarded applicants

As proposed by Bound (1989), the discouraging impact of DI benefits can be proxied by the difference in employment rate of rejected and awarded applicants. Considering the fact that the severity of health impairments is most likely stronger for accepted than rejected applicants, the Bound estimate provides an upper bound. As is pointed out by Maestas et al. (2013), however, a negative bias may stem from the fact that rejected applicants – with lower pre-application earnings – may also have less labor force attachment than awarded applicants.

Figure 5 presents yearly changes in the Bound estimate for the Netherlands. The Bound estimate is based on the employment rates of rejected and awarded applicants for each annual application cohort, measured three years after the award decision.[13] Panel A shows the evolution of the Bound estimate that follows from comparing rejected and *all* awarded applicants. Rejected applicants show a gradual decline in the employment rates three years after application, with a somewhat larger drop in 2006, when the WIA came into force. This contrasts to the change in employment rates for awarded applicants that shows a dramatic decline in the same year. After the WIA reform, the Bound estimate is about 30 percentage points. This estimate is in the ballpark of estimates obtained for SSDI benefits.[14]

To reduce the supposedly positive bias that follows from differences in the severity of impairments, we next limit the sample of awarded applicants with partial DI benefits, who are deemed to have substantial earning capacities. Panel B of Figure 5 shows that these two groups have very similar downward employment patterns. The corresponding Bound estimate becomes small and even negative, ranging between -2 and -5 percentage points. While the negative estimate may appear at odds with intuition, it can well be explained from the fact that applicants with higher pre-disability earnings are more likely to have a strong labor force attachment and experience a

[13] We argue that this gives a sufficiently long time delay to consider the long-term employment rates of these cohorts.

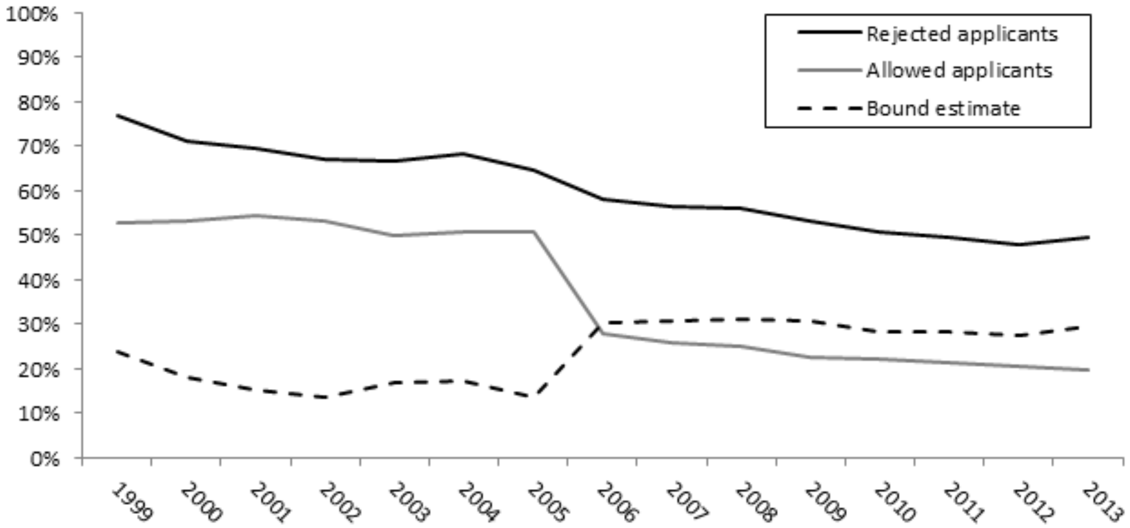
[14] Bound (1989) finds a difference in employment rates one year after application of between 26 and 30 percentage points for applicants aged 45-64. Von Wachter et al. (2011) shows that the Bound estimate amounts to more than 35 percentage points for applicants aged 30-44. Bound et al. (2003) estimates a difference three years after application of 20 percentage point. These results are similar to Chen & Van der Klaauw (2008) who show a reduction of the labor force participation of 15-18 percentage points.

higher percentage drop in earning capacity. Accordingly, individuals with higher pre-disability earnings have a higher probability to receive partial DI benefits than individuals with low pre-disability earnings. Note that similar arguments are put forward by Maestas et al. (2013), who show that rejected SSDI applicants typically have lower pre-employment rates.

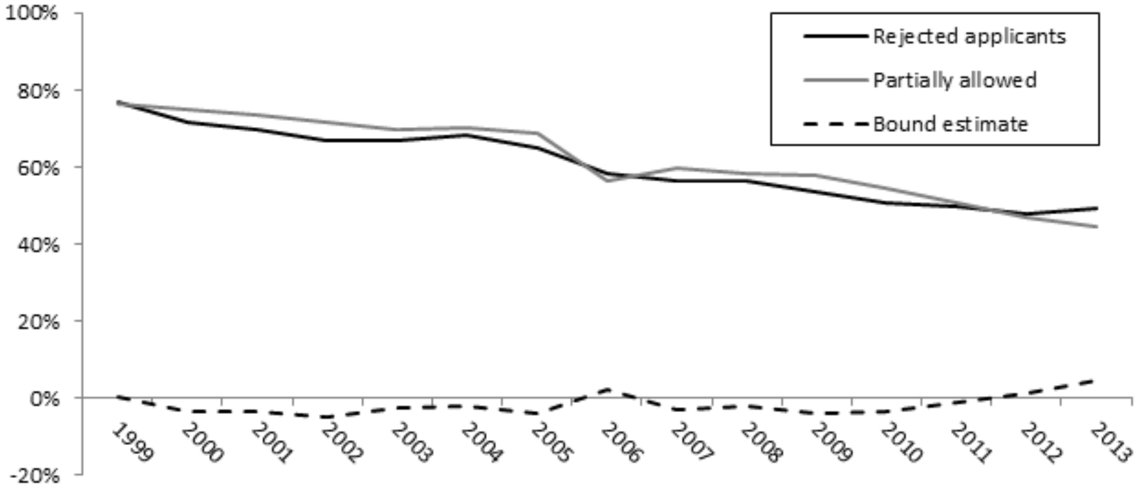
Panel C of Figure 5 finally compares the employment rates of DI applicants with disability degrees below 35% to those with disability degrees between 35 and 80%. As a result, the classification of the cohort samples is no longer affected by the increase of the disability degree threshold in 2006. Until 2006, we then find that the employment rates of both groups are virtually equal to each other. When the WIA came into force in 2006, we next see a somewhat larger decline in the employment rate for those with disability degrees below 35% than for those with disability degrees between 35 and 80%. This divergence may point at the increased work incentives for DI beneficiaries with partial benefits. In the longer term, however, the employment rates of these two cohort groups converge.

Figure 5: Annual employment rates and Bound estimates for different application cohort samples between 1999 and 2013, measured three years after the DI decision

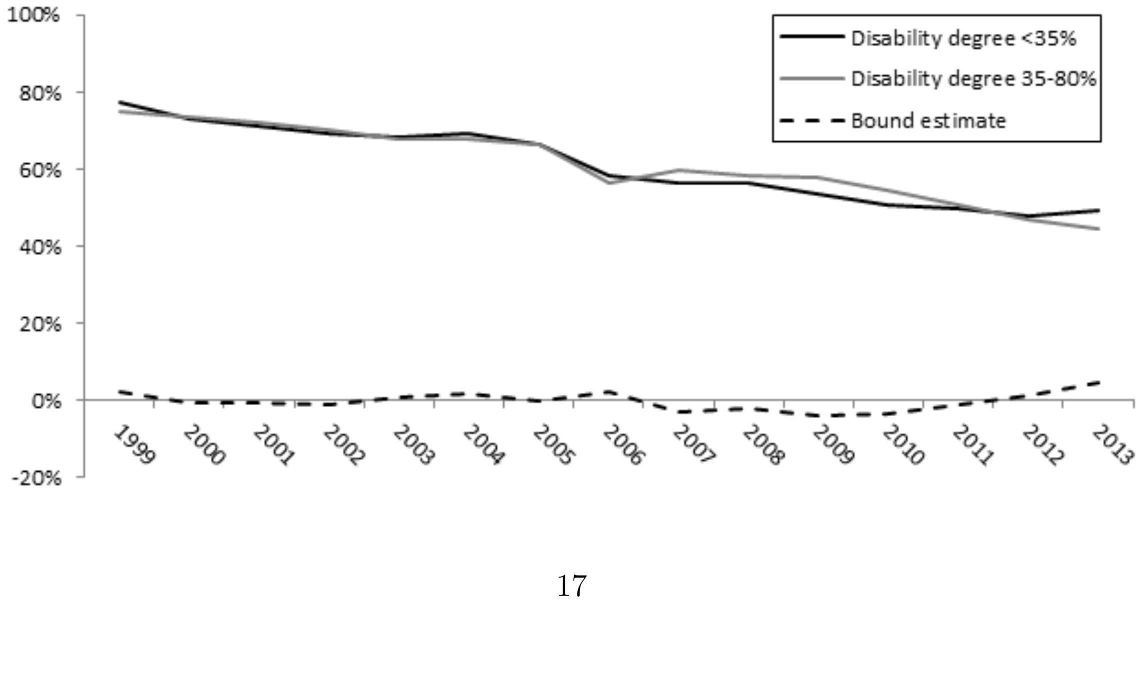
Panel A. Employment rates of rejected applicants and awarded applicants



Panel B. Employment rates of rejected and partially awarded applicants



Panel C. Employment rates of applicants with disability degrees of <35% and 35-80%



4 Empirical strategy

4.1 The Age-Period-Cohort (APC) model

The aim of this paper is to decompose the mechanisms underlying the substantial decline in the employment rates of DI applicants. To this end, we propose a two-step analysis with Age-Period-Cohort (APC) models. First, we decompose employment trends into changes in the effect of the elapsed duration since application (the 'age' effect), period effects and cohort effects. Second, we further decompose cohort effects into compositional and incentive effects, using a difference-in-difference approach that expands on the APC model. Contrasting to the compositional effects that affect all applicants, these incentive effects affect the individual employment rates of awarded applicants only.

We specify an APC model as a linear probability model that explains the prevalence of employment E for all DI applicants in our sample, measured for post-application years. E is equal to one while working, and zero otherwise.

$$E_{it,\tau} = \alpha_{t-\tau} + \pi_t + \gamma_\tau + \epsilon_{it}, \tag{1}$$

with $t \geq \tau$. In the above equation, the employment status E of individual i ($i = 1,..,N$) in year t ($t = 1,..,T$) with a DI decision in year τ ($\tau = 1,..,\mathcal{T}$) is determined by the number of years after application (i.e., the 'age' effect), a calendar year ('period') effect and a cohort effect. Note that we have $T = 18$ years (1999-2016) and $\mathcal{T} = 15$ cohort years (1999-2013) in our sample. Age, period and cohort effects are denoted by the vectors α, π and γ, respectively. Without controlling for the age of individuals, the 'age' effect captures both the effect of aging and the elapsed duration since application.[15] Finally, ϵ is an error term that is i.i.d. Equation 1 can be estimated with OLS, allowing for individual clustering effects.

As is well-known in the literature, the linear APC model does not allow for identification of all parameters of interest. Despite normalizing the starting values of age, period and cohort effects to zero, the linear relationship between age, period and

[15] We also will estimate model specifications that control for age. In doing so, the vector α can be interpreted as the genuine effect of elapsed duration.

cohort implies that another restriction is needed for identification. In our empirical analysis, we deal with this by following the well-known approach of Deaton & Paxson (1994). That is, we impose orthogonality constraints on period effects with respect to age and cohort effects. Specifically, in the Deaton-Paxson (DP) model the average effect of period effect is assumed to be equal to zero ($\Sigma_1^T\ \pi_t = 0$) and that there is no trend in period effects ($\Sigma_1^T\ t\pi_t = 0$). This captures the idea that time effects reflect business cycle effects that are transitory. In light of the long time period that is under investigation, we argue that these assumptions are not overly restrictive. To assess the overall importance of period effects in this setting, we compare the fit of the APC-DP model to a model with with age and cohort effects only (i.e., the Age-Cohort or AC model). Given the constraints imposed in the DP-model, the AC model will not result in major changes in cohort effect estimates, but differences in the fit of the AC and the APC-DP model may well provide insight into the overall importance of transitory period effects.

A major concern of the DP model is that it does not allow for non-transitory period effects. Structural trends in period effects, if existent, are absorbed by the age and cohort effect estimates. We therefore assess the stringency of the orthogonality assumptions of the DP model in two robustness tests. First, we will estimate an APC model with period effects that are specified as a quadratic function and age and cohort effects as (non-parametric) dummies. This specification enables us to estimate (part of) identifiable non-linear period effects that may be non-transitory. Specifically, the coefficient of the quadratic period effect, together with changes in age and cohort effect estimates, provide us with conservative tests on the existence of non-transitory period trends. Second, we consider model specifications where period effects depend on the ratio of vacancies to unemployment and employment rates of low-educated individuals. Arguing that low-educated individuals are over-represented among DI applicants, this auxiliary information can be used to proxy period effects that may also may show more structural trends e.g. arising from Skill-Biased Technological Change (SBTC).[16]

[16]For 2005 onwards, we observe employment rates of disabled individuals in the public scheme for disabled individuals that have no eligibility into the DI scheme (i.e., the Wajong). For this limited time period, we will also consider AC models with specified period effects.

Equation 1 yields patterns of cohort effects that can be related to specific reforms – particularly discontinuous changes – or to gradual changes in the composition of applicants that cannot be linked to reforms. To obtain insight in the sources of compositional changes, we also estimate model versions that include dummies for five-year age groups, gender, ethnicity, impairment types and the pre-disability employment status as controls.[17] If compositional effects are embodied by these variables, one would expect less sizable cohort effects. Changes in cohort effects then indicate self-screening among potential applicants.

4.2 Separating compositional from incentive effects

In the context of the APC model, changes in cohort effect estimates represent both compositional changes among applicants cohorts as a whole and incentive changes among the sample of awarded applicants. To disentangle changes in both effects, we need one closing assumption. Following Bound (1989), we do so by estimating APC models for pooled samples of awarded applicants that are affected by DI benefit reforms and rejected applicants that are not. Our key assumption is that compositional changes have equal effects on the employment rates of awarded and rejected applicants. This in turn requires two specific conditions to be met: (i) changes in self-screening that occur before the DI decision should affect the employment of awarded and rejected applicants equally and (ii) there should be no changes in the eligibility conditions that are linked to the employment opportunities of workers. Under these two assumptions, changes in the differenced cohort effects of these groups indicate changes in incentive effects.

To implement our approach, we define $A_{i,\tau}$ as a dummy that is equal to one if DI applicant i in the year cohort τ is awarded benefits, and zero otherwise. Expanding on Equation 1, we specify the following model:

$$E_{it,\tau} = (1 - A_{i,\tau}) \left\{ {\alpha^0}_{t-\tau} + \pi_t^0 \right\} + A_{i,\tau} \left\{ {\alpha^1}_{t-\tau} + \pi_t^1 \right\} + \gamma_\tau + (1 - A_{i,\tau}) \, \tilde{\gamma}_\tau + \epsilon_{it}, \quad (2)$$

[17] During our analysis we use the employment status in the year before application. However, we also estimated models using the employment status two years before application for cohorts after the WIA reform. By doing this, we take into account that these applicant cohorts face a longer waiting period of two years. Our results are robust to the different sets of control variables.

with α^0 and π^0 denoting age and period effects for the rejected applicants, respectively; α^1 and π^1 denoting age and period effects for the awarded applicants, respectively; and $\tilde{\gamma}_\tau$ as the cohort effect that is interacted with the award indicator.[18] Most notably, $\tilde{\gamma}_\tau$ can be interpreted as the Bound estimate for a specific cohort τ. This estimate controls for the fact that age and period effects may differ between awarded and rejected applicants.

We next impose restrictions on $\tilde{\gamma}$ that follow from the assumption of common compositional effects for awarded and rejected applicants. This assumption holds for all years without reforms, yielding the following DiD specification for $\tilde{\gamma}$:

$$\tilde{\gamma}_\tau = \tilde{\gamma}_0 + I(\tau \geq 2003)\ \tilde{\gamma}_{gkp} + I(2006 \leq \tau \leq 2009)\ \tilde{\gamma}_{wia,st} + I(\tau \geq 2010)\ \tilde{\gamma}_{wia,lt} \quad (3)$$

with $\tilde{\gamma}_{gkp}$, $\tilde{\gamma}_{wia,st}$ and $\tilde{\gamma}_{wia,lt}$ denoting the effect of the GKP reform and the short-term and long-term effect of the WIA reform on the Bound estimate.[19] In this context, it is important to stress that increases in the Bound estimate ($\tilde{\gamma}$) indicate equal decreases in incentive effects. This follows from the fact that the Bound estimate takes awarded applicants as a reference group.

Clearly, the assumption of common changes in compositional effects is more plausible if rejected applicants are compared to awarded applicants which are deemed to have substantial residual earnings capacity. It then becomes more likely that the first condition for identification is met, which states that changes in self-screening should affect the employment of awarded and rejected applicants equally.[20] At the same time, care should be taken of the (second) assumption that the screening process is unaffected by the reforms. In particular, the WIA reform implied a shift in the disability degree threshold for DI receipt from 15% to 35%. Arguing that higher disability degrees are associated with lower employment potential, this shift would both lower the average employment rate in the sample of rejected and awarded

[18] Similar to Equation 1, note that we impose orthogonality restrictions on α^0 and α^1 to estimate all parameters of Equation 2.

[19] In light of the long time period that is observed after the WIA reform, we allow for a more flexible specification that distinguishes short-term from long-term effects.

[20] Recall that this is confirmed from our graphical inference in the previous section, that shows common trends before the reform years.

applicants. The net biasing effect is thus ambiguous.

To address the concern that changes in the DI screening process would lead to inconsistent estimates, we will present the outcomes for different specifications and different samples. First, we estimate model specifications with and without the control variables that were discussed earlier. If changes in the screening process do not affect the composition of awarded and rejected applicants equally, these differences should be partially absorbed by the control variables and the DiD estimate should change. This provides a natural test on the compositional effects assumption. Second, we will re-define samples by disability degrees instead of the outcome of DI decisions, using a cutoff – for all years – of 35%. By construction, this ensures that there are no compositional changes that are inherent to changes in the eligibility condition. Still, this setup presumes that the incentive effect of DI benefit loss in the group applicants with disability degrees between 15% and 35% is limited. This implies we cannot rule out estimation bias here as well.

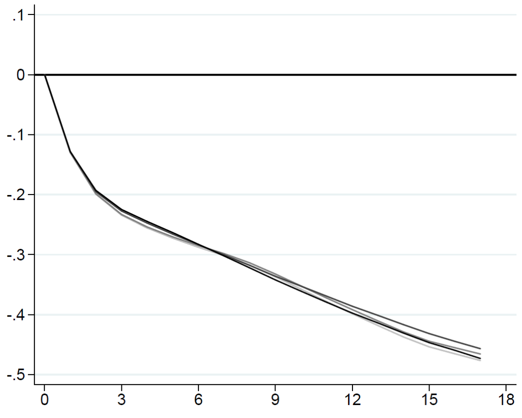
5 Estimation Results

5.1 The Age-Period-Cohort model

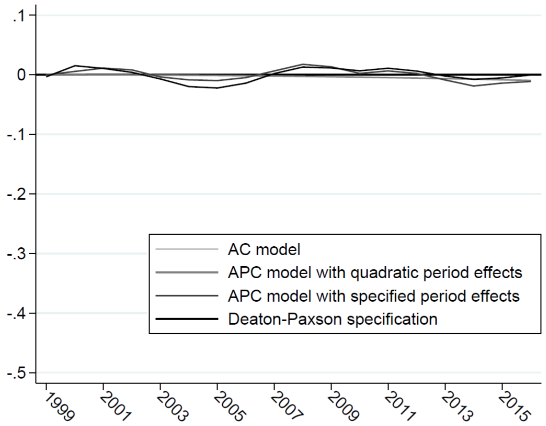
Figure 6 graphically presents the elapsed time (or: 'age'), period and cohort profiles of the employment for our full sample of DI applicants for the four model specifications: the Age-Cohort (AC) model, the APC model with quadratic period-effects, the APC model with time periods depending on the labor market tightness and employment of low-educated workers, and the APC-DP model. All four models do not include observed individual controls. We consider the APC-DP model as our preferred specification and show the results of the other models to assess the robustness of our findings. As individual controls are not included, the 'age' estimates do not only reflect the long-term effect of application over time but also the effect of aging of applicants. Likewise, the cohort estimates show the composite impact of *all* time-invariant variables that affect employment.

Figure 6 shows very similar age and cohort effects across model specifications. The elapsed time profile since the DI decision – i.e., the 'age'-effect – displays a

Figure 6: Elapsed time ('age'), Period and Cohort effects of AC model, APC model with quadratic time periods, APC model with specified period effects, and the Deaton-Paxson specification



(a) Age effects



(b) Period effects

(c) Cohort effects

kinked pattern for all four specifications. The drop in employment is largest in the first and second year after the DI decision, amounting to a decrease of nearly 20 percentage points. In this period, applicants awarded benefits may leave the labor market and a large fraction of rejected applicants is laid off by their employer.[21] Subsequently, the employment rate of applicants declines with approximately 2 percentage points per year, such that the total decrease after 17 years equals roughly 45 percentage points. Figure 6 also indicates sizable cohort effects, particularly when the GKP and WIA came into force. Changes in cohort effects add up to a 30 percentage points difference between 1999 and 2013. This difference largely stems from a 4 percentage points drop in 2003 and another drop of about 13 percentages

[21] Note that this contrasts to the SSDI system, where applicants typically have no (substantial) earnings from employment to begin with.

points in 2006. The cohort effects also show a continued decline in the years after the start of the WIA in 2006. In total, more than half of the change in cohort effects is confined to the reform years 2003 and 2006.

Alongside these findings, panel (b) shows relatively small period effects. The spread of period effects is below 5 percentage points, whereas the time and cohort effects add up to to about 45 and 30 percentage points, respectively. The comparison of the outcomes of the APC-DP model and the AC model (without period effects) also suggests that period effects do not explain a large part of the variation in the employment of disabled workers.[22] Still, the small period effects of the DP model mimic business cycle patterns seemingly well, with peaks in 2001 and 2008.

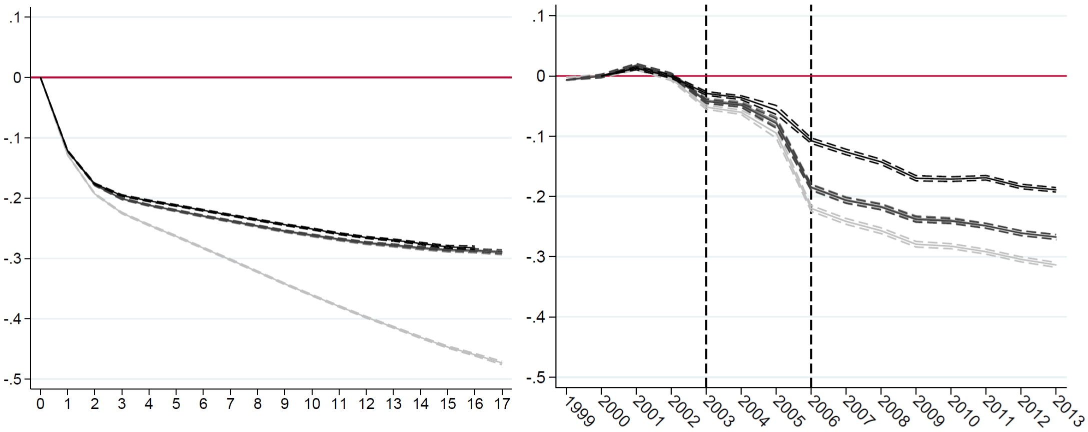
To test for non-transitory period effects, we also consider a specification with quadratic period effects. Albeit that the concerning coefficient is statistically significant, its magnitude is negligible and the accumulated cohort and age effects are very similar to those for the DP model. This conservative test thus suggests that period effects are transitory. Next, we specify period effects as a function of the annual vacancy-to-unemployment ratio and the annual average net employment rate of low-educated workers. While both variables have coefficients with expected signs and are statistically significant, the resulting range of period effects is comparable to those for the APC-DP model. Specifically, we find a one percentage point increase in the employment rate of low-educated workers to be associated with a 0.6 percentage point increase in the period effect.[23] Taken together, these additional findings essentially resemble the earlier eyeball tests that showed constant employment differentials of successive application cohorts.

To gain more insight in the sources underlying cohort and elapsed duration effects, Figure 7 shows the results for the DP model with various sets of control variables that are added sequentially: (i) individual characteristics that include dummies for five-year age groups, gender and ethnicity; (ii) impairment types; and (iii) the employment history in the year before application (employment status, UI

[22] The R-squared of the APC-DP model is 0.0683 and for the AC model 0.0680, respectively.

[23] We also re-estimated this model for 2005-2016, as we observe net employment rates of individuals in the second biggest disability scheme in the Netherlands, the Wajong. This variable did not yield a significant coefficient estimate.

Figure 7: Deaton-Paxson estimation results of elapsed time ('age') and cohort effects with step-wise inclusion of sets of control variables



(a) Elapsed time effects (b) Cohort effects

Note: The base specification (light grey line) is the model without control variables. We subsequently add: (i) dummies for age groups of five years, gender and ethnicity; (ii) impairment types; and (iii) employment status in the year prior application (employment status, UI benefit receipt and sector of employment). The dashed lines outline the 95-percent confidence intervals.

benefit receipt and the sector of previous employment). From the figure, it becomes apparent that the inclusion of control variables causes the elapsed time effect estimates to level out after the first two years.[24] Figure 7 also shows substantial reductions in cohort effects stemming from the inclusion of control variables, suggesting self-screening effects on the average employment that occur before the DI decision. Roughly speaking, about 40 percent of the 31 percentage points decline in employment rates of subsequent cohorts is explained by self-screening on observed variables. As we have a limited set of controls, this estimate should be interpreted as a lower bound for the total effect of self-screening. Interestingly, the inclusion of controls does not change cohort effects substantially until 2006. So while the GKP may have discouraged workers with less-severe impairments from applying, this does not imply these individuals had better labor market prospects. By contrast, the instantaneous drop in employment rates at the time of the WIA reform can largely be explained by the screening out of workers with better labor prospects, causing the remaining applicant pool to have less permanent contracts and a higher fraction

[24]The results are similar when we use 10-year age groups. The employment of disabled workers drops after the applicants reach their retirement age; this effect amounts to more than 20 percentage points.

being unemployed one year before application.[25] Finally, we find that the gradual further decline in employment after the onset of the WIA reform can partially be explained by gradual compositional changes in observed controls.

5.2 Incentive effects

Table 2 presents the estimation results for the incentive effect of the GKP and WIA reforms, $\tilde{\gamma}_\tau$, using the restricted (DiD) model of Equation 2. Recall that the incentive effect measures changes in the Bound estimate, with positive changes pointing at a worsening of the employment probability of awarded applicants (and reverse). The findings for the restricted model are complemented with the unrestricted Bound estimates for all annual cohorts – as shown in Figure 8. For both the restricted and the unrestricted model, we compare (differenced) cohort effects of the following groups: (i) rejected applicants versus awarded applicants with partial benefits in columns (1-2); (ii) applicants with disability degrees below 35% versus applicants with disability degrees between 35 and 80% in columns (3-4); (iii) applicants with disability degrees below 35% versus applicants with disability degrees between 35 and 55% in columns (5-6).

The DiD estimates in Table 2 suggest no incentive changes at the start of the GKP reform for all group comparisons. As the GKP aimed at changing the screening process before application, these results are in line with expectations and can be considered as Placebo-outcomes. The evidence for the incentive effects of the WIA reform, however, is less clear-cut. As to the effects in the first four years since the reform (i.e., 2006-2009), all model specifications without controls show negligible and only weakly statistically significant estimates of the incentive effects.[26] The estimates increase somewhat after the inclusion of controls, suggesting that the common compositional cohorts assumption may be violated. For the long-term incentive effects (i.e., 2010-2013), our results indicate decreases of work incentives

[25] The newer cohorts are also older (the last cohort is on average 5 years older than the first cohort), more often male (10%-points) and for a larger share non-native (8%-points). Perhaps strikingly, there is no changes in cohorts effects when including impairment types.

[26] Recall that both Koning & van Sonsbeek (2017) and Kantarci et al. (2019) also find only small causal employment effects of the WIA reform.

Table 2: DiD incentive effects of the Gatekeeper Protocol (GKP) and short-term and long-term incentive effect of the WIA reform

	Rejected vs partially allowed		Disability degree $< 35\%$ vs. $35 - 80\%$		Disability degree $< 35\%$ vs. $35 - 55\%$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{\gamma}_{gkp}$	-0.005*	-0.005*	-0.001	0.000	-0.001	-0.001
	(0.00)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
$\tilde{\gamma}_{wia,shortterm}$	0.009*	0.026***	-0.009*	0.013***	-0.005	0.013***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.006)	(0.005)
$\tilde{\gamma}_{wia,longterm}$	0.029***	0.032***	0.018***	0.025***	0.018***	0.022***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Separate age and period and common cohort effects	✓	✓	✓	✓	✓	✓
Controls	—	✓	—	✓	—	✓
Observations	6,730,460	5,561,737	6,736,052	5,567,329	6,193,528	5,095,192
R^2	0.0642	0.2026	0.0650	0.2030	0.0622	0.2001

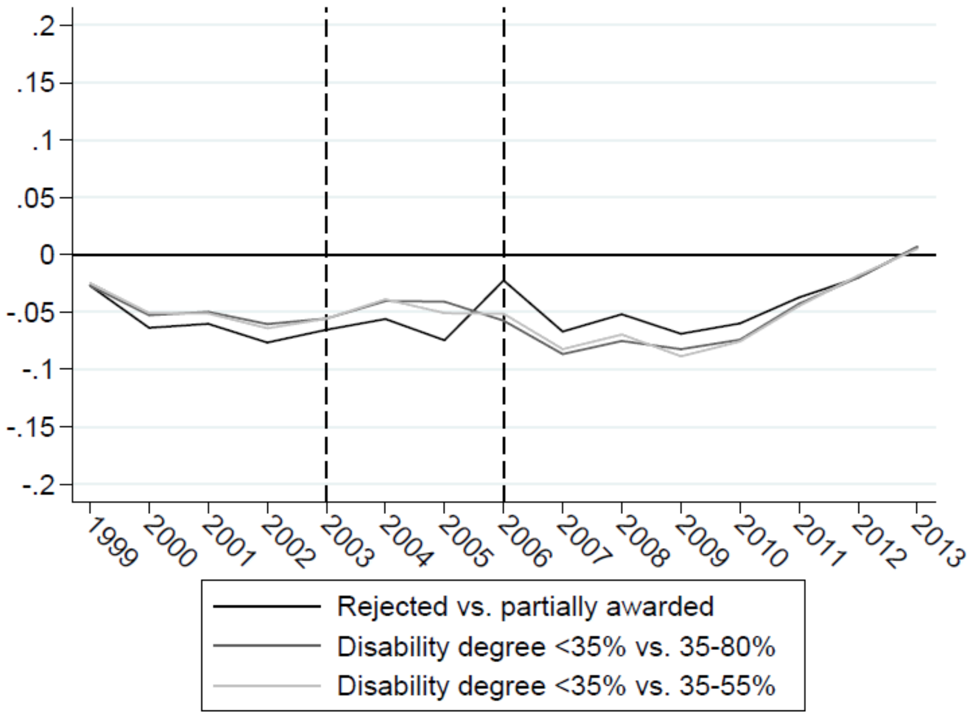
Note: Control variables include individual characteristics (age, gender, ethnicity), impairment types and employment history (employment status, UI benefit receipt and sector of employment). Individual clustered standard errors in the parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

ranging between 2 and 3 percentage points – i.e., an increase in the Bound estimate – for partially awarded applicants. While these findings may appear more robust than the short-term effects, the negative incentive effects are not in line with theoretical predictions.

We next move to the unrestricted Bound estimates as shown in Figure 8.[27] Similar to the graphical inference that was discussed earlier, the initial difference in cohort effects is negative and fairly constant up till 2005 for all group comparisons. This again underlines the notion that the GKP increased the reintegration responsibilities during the waiting period for all DI applicants. For the WIA reform, again, there is no clear pattern that emerges. Depending on the stratification of groups, the Bound estimate can either stay more or less constant or decrease in 2006 (which implies a positive incentive impact). If any, Figure 8 suggests that the incentive effects of the WIA reform are small. Moreover, from the figure it appears unlikely that the increases in the Bound estimate after 2010 can be interpreted as the effect of the WIA reform. Taking a broader perspective, we are safe to say that the accumulated changes in cohort effects by far cannot be explained by changes in DI

[27] All parameter estimates of $\tilde{\gamma}_{\tau}$, both without and with controls, can be found in Table A.3 in the Appendix, together with additional F-statistics which follow from multiple testing.

Figure 8: Annual Bound-estimates for the unrestricted pooled APC-DP models



Note: The vertical axis displays the parameter estimates of $\tilde{\gamma}_\tau$ from Equation 2.

benefit incentives.

5.3 Assessing cohort effects in more detail

As compositional cohort effects are the main driving force of the employment trends, we next investigate its origins and robustness in more detail. We re-estimate our preferred APC-DP model for samples that are stratified according to benefit decisions (rejected, partially awarded, fully awarded), gender, age groups (18-44 vs. 45-64) and impairment types (mental, musculoskeletal, cardiovascular and all other types). The estimated age and cohort effects of these groups are all shown in Figure A.13 in the appendix – both with and without controls.

In line with expectations, Figure A.13 shows larger and initially steeper age profiles for groups with higher disability degrees, older ages and those diagnosed with cardiovascular disorders. This contrasts with rejected applicants, partially awarded applicants and younger applicants that show more persistent employment profiles after the award decision. As to the cohort effects, the initial decline since the start of the WIA is more substantial among those awarded full benefits, but next the partially awarded applicants catch up and experience a similar aggregate decline.[28] Interestingly, changes in cohort effects are most substantial for workers with mental impairments and already materialize in the year the GKP reform took place.[29] This suggests that moral hazard was present among workers with mental impairments, as the GKP implied stronger screening before application.[30] At the same time, however, we do not find similar results for workers with musculoskeletal disorders.

We also re-estimate the APC model with specifications that relax the assumption of common age and period effects for all cohorts. Specifically, our interest lies in differences in age patterns and in cohort effects that are aligned with the three time periods: 1999-2002, 2003-2005, and 2006-2013. We therefore allow for different age effects and different effects of control variables that are linked to these periods. The implied changes in accumulated cohort effects are shown in Figure 9. In this figure, the first bar indicates the implied total change in cohort effects for the baseline DP model, the second bar the implied total change in cohort effects for the AC model (i.e., without period effects) and the third bar the implied total change for the AC model with distinct age and cohort effects for the three time periods.[31] The figure shows that the AC model yields cohort effects for the three time periods that are virtually equal to the DP model. More importantly, the implied absolute declines in cohort effects over the three time periods are robust to the flexible specification of age and control variable effects. In addition, the implied aggregate cohort effects of the time periods decreases with the inclusion of controls. This again underlines the importance of self-screening that went together with the reforms. The aggregate

[28] Koning & van Sonsbeek (2017) argue that the stronger work incentives induced by the WIA may have increased the relevance of a ‘cash-cliff’ to the fully and temporary disabled beneficiaries.

[29] In line with these results, Godard et al. (2019) find that increases in screening intensity in a field experiment that was conducted in the first year of the GKP reduced DI applications and this effect was largely confined to workers with mental disorders.

[30] Moral hazard may have been more important among workers with mental problems as it is a more heterogeneous group, with a high share of conflicts at work that prevent rehabilitation of sick-listed workers.

[31] Note that the estimation of APC models with distinct age effects would give rise to identification problems of period effects. Arguing that the period effects we find are generally small for the total period, setting these equal to zero is not a strong restriction to make.

Figure 9: Comparing implied absolute declines in cohort effects of three models, measured for 1999-2002, 2003-2005, and 2006-2013

Note: The three models: (i) Deaton-Paxson specification, (ii) AC model, and (iii) AC model with specific control effects per period. When controls are included, one cohort (1999) must be omitted. Control variables include individual characteristics, impairment types and employment history.

cohort effect for the period 2006-2013 does not change when we also control for the type of contract in the year before application. Finally, the negative cohort effects after 2006 represents either learning or adaptation effects of the WIA reform, or point at a more general trend in health and labor market conditions that are specific to new cohorts.

5.4 Other outcome measures

To analyze the welfare implications of the large cohort effects, we finally consider wage earnings, contract types, other social security schemes and the mortality of applicants after the DI decision as alternative outcome measures. The resulting age and cohort profiles for these outcomes are presented in Figure A.14 in the Appendix.

From panel A in Figure A.14 we infer that earnings show a similar pattern of cohort effects as the incidence of employment. Cohort effects accumulate to 10,000 Euro per year, corresponding to roughly 40 percent of the average income at the time of application in 1999. This corresponds to the extensive margin effect of employment, which also amounts to about 40%. Panel B shows that the decline in cohort effects of the probability on a permanent contract is roughly equal to the cohort effect for permanent and flexible contracts, suggesting the decline is fully confined to permanent contracts.

We also investigate the presence of substitution effects to other schemes than DI – see e.g. Koning & Van Vuuren (2010), Borghans et al. (2014) and Benitez-Silva et al. (2010). Between 1999 and 2013, we find that positive cohort effects for UI accumulate to about 10 percentage points. This rise is strongest for cohorts after the start of the WIA reform in 2006. This again suggests there were gradual changes in the composition of new applicant cohorts with increasing shares of more vulnerable groups with worse employment prospects. For the inflow into social assistance, this pattern is less apparent.

Finally, we estimate the APC-DP model on the mortality rate of the applicants (as show in panel E). In light of the drastic reductions in DI inflow and the large, negative cohort effects for employment, one may expect the remaining group of applicants to have more severe impairments and higher mortality rates. The cohort effect estimates for mortality without any controls seemingly confirm this hypothesis, with increases that are largely confined to the GKP and WIA reform.[32]

6 Conclusions

In this paper we expand on Age-Period-Cohort (APC) models to explain changes in the employment rates of Disability Insurance (DI) applicants. We use administrative data of DI application cohorts for the Netherlands, a country that has experienced major disability reforms that intensified the screening process before application, tightened eligibility and aimed to improve work incentives for benefit recipients. Using a conventional Deaton-Paxson specification, we first decompose cohort effects

[32]To calculate mortality rates, we follow the approach by Johansson et al. (2014) who use post-application mortality as proxy for ex-ante health.

from period and age effects. The resulting cohort effects represent the joint effect of (i) compositional changes induced by the disability reforms; (ii) compositional changes induced by general labor market and health trends; and (iii) individual changes in the employment rate of awarded applicants – or: 'incentive effects' – induced by DI reforms. To disentangle the incentive effects from compositional effects, we next propose a further decomposition that compares the employment rates of awarded applicants to those of rejected applicants. That is, we estimate APC models with distinct age, period and cohort effects for awarded and rejected applicants. Assuming that compositional cohort effects for employment – both induced by reforms and changes in the labor market – affected both groups equally, the Difference-in-Difference (DiD) estimate of the reforms indicates the change in the individual employment probability. These effects can then be characterized as incentive effects of the reforms on benefit recipients.

We find that cohort effects are the key driver of the observed decline in employment rates of DI applicants in the Netherlands. Both gradual changes in the labor market and large instantaneous self-screening effects induced by reforms affected new applicant cohorts, rather than period effects or changes in work incentives for awarded applicants. Even though the period effects mimic the business cycle quite well, its absolute importance in explaining employment changes is negligible. Likewise, our further decomposition of cohort effects into compositional and incentive effects suggests that changes in incentive effects are dwarfed by effects due to changes in the composition of applicants. This highlights the importance of self-screening effects that were inherent with the reforms, with sick-listed workers that have increasingly resumed work before DI application and/or have been discouraged to apply for DI benefits. This way of self-screening has dramatically changed the character of the DI scheme, with less room for workers with residual earnings capacities who complement their labor income with benefits.

Our results add to other international analyses that suggest a trend of more vulnerable, low-skilled labor market groups becoming applicants for disability benefits (Autor & Duggan, 2003; Von Wachter et al., 2011; Maestas, 2019). Specifically, we find changes in the initial labor market position and sector of employment of applicants as important drivers of the observed decline in employment. This change applies to new applicant cohorts, rather than affecting all individuals that have applied for benefits at some point in time. To some extent, the dominant role of cohort effects may stem from the relatively strict Employment Protection Legislation (EPL) that prevails in the Netherlands. In a similar vein, it is likely that gradual changes in the composition of applicant cohorts since the start of the new disability program in 2006 (WIA) cannot be reconciled from disability reforms alone. The higher share of vulnerable groups among applicants may point at a gradual sorting of low-skilled workers with health conditions into temporary and flexible jobs without employer obligations. We leave this topic for future research.

References

Autor, D. H. & Duggan, M. G. (2003). The Rise in the Disability Rolls and the Decline in Unemployment. *The Quarterly Journal of Economics*, 118(1), 157–206.

Benitez-Silva, H., Disney, R., & Jiménez-Martín, S. (2010). Disability, Capacity for Work and the Business Cycle: an International Perspective. *Economic Policy*, 25(63), 483–536.

Borghans, L., Gielen, A., & Luttmer, E. (2014). Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform. *American Economic Journal: Economic Policy*, 6(4), 34–70.

Bound, J. (1989). The Health and Earnings of rejected Disability Insurance Applicants. *The American Economic Review*, 79(3), 482.

Bound, J., Burkhauser, R. V., & Nichols, A. (2003). Tracking the Household Income of SSDI and SSI Applicants. In *Worker Well-being and Public Policy* (pp. 113–158). Emerald Group Publishing Limited.

Bound, J., Lindner, S., & Waidmann, T. (2014). Reconciling Findings on the Employment Effect of Disability Insurance. *IZA Journal of Labor Policy*, 3(1), 11.

Campolieti, M. (2006). Disability Insurance Adjudication Criteria and the Incidence of Hard-to-Diagnose Medical Conditions. *The BE Journal of Economic Analysis & Policy*, 5(1), 1–25.

Chen, S. & Van der Klaauw, W. (2008). The Work Disincentive Effects of the Disability Insurance Program in the 1990s. *Journal of Econometrics*, 142(2), 757–784.

De Jong, P., Lindeboom, M., & Van der Klaauw, B. (2011). Screening Disability Insurance Applications. *Journal of the European Economic Association*, 9(1), 106–129.

Deaton, A. & Paxson, C. (1994). Savings, Growth and Aging in Taiwan. In *Studies in the Economics of Aging* (pp. 331–362). National Bureau of Economic Research.

Deshpande, M. & Li, Y. (2019). Who Is Screened Out? Application Costs and the Targeting of Disability Programs. *American Economic Journal: Economic Policy*, 11(4), 213–248.

French, E. & Song, J. (2014). The Effect of Disability Insurance Receipt on Labor Supply. *American Economic Journal: Economic Policy*, 6(2), 291–337.

Godard, M., Koning, P., & Lindeboom, M. (2019). Targeting Disability Insurance Applications with Screening. IZA Discussion Paper 12343.

Hullegie, P. & Koning, P. (2018). Employee Health and Employer Incentives. *Journal of Health Economics*, 62, 134–146.

Johansson, P., Laun, L., & Laun, T. (2014). Screening Stringency in the Disability Insurance Program. *The BE Journal of Economic Analysis & Policy*, 14(3), 873–891.

Kantarci, T., van Sonsbeek, J.-M., & Zhang, Y. (2019). The Impact of the Disability Insurance Reform on Work Resumption and Benefit Substitution in the Netherlands. Netspar Discussion Paper 01/2019-013.

Koning, P. & Lindeboom, M. (2015). The Rise and Fall of Disability Insurance Enrollment in the Netherlands. *Journal of Economic Perspectives*, 29(2), 151–172.

Koning, P. & van Sonsbeek, J.-M. (2017). Making Disability Work? The Effects of Financial Incentives on Partially Disabled Workers. *Labour Economics*, 47, 202–215.

Koning, P. & Van Vuuren, D. J. (2010). Disability Insurance and Unemployment Insurance as Substitute Pathways. *Applied Economics*, 42(5), 575–588.

Liebert, H. (2019). Does External Medical Review Reduce Disability Insurance Inflow? *Journal of Health Economics*, 64, 108–128.

Maestas, N. (2019). Identifying Work Capacity and Promoting Work: A Strategy for Modernizing the SSDI Program. *The ANNALS of the American Academy of Political and Social Science*, 686(1), 93–120.

Maestas, N., Mullen, K. J., & Strand, A. (2013). Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *American Economic Review*, 103(5), 1797–1829.

Markussen, S., Røed, K., & Schreiner, R. C. (2018). Can Compulsory Dialogues Nudge Sick-listed Workers Back to Work? *The Economic Journal*, 128(610), 1276–1303.

OECD (2010). *Sickness, Disability and Work: Breaking the Barriers.*

Van Sonsbeek, J.-M. & Gradus, R. H. (2012). Estimating the Effects of Recent Disability Reforms in the Netherlands. *Oxford Economic Papers*, 65(4), 832–855.

Voas, D. & Chaves, M. (2016). Is the United States a Counterexample to the Secularization Thesis? *American Journal of Sociology*, 121(5), 1517–1556.

Von Wachter, T., Song, J., & Manchester, J. (2011). Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program. *American Economic Review*, 101(7), 3308–29.

A Appendix

Figure A.10: Fractions of awarded and rejected DI applicants by application cohort

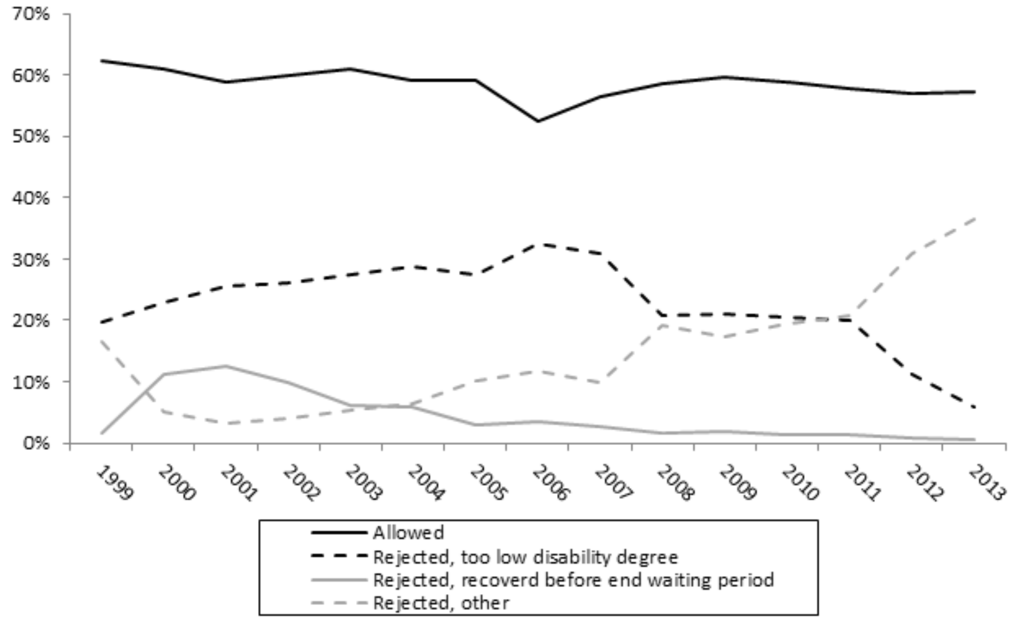


Figure A.11: Cumulative distribution of most important impairment groups of all applicants for disability insurance by application cohort

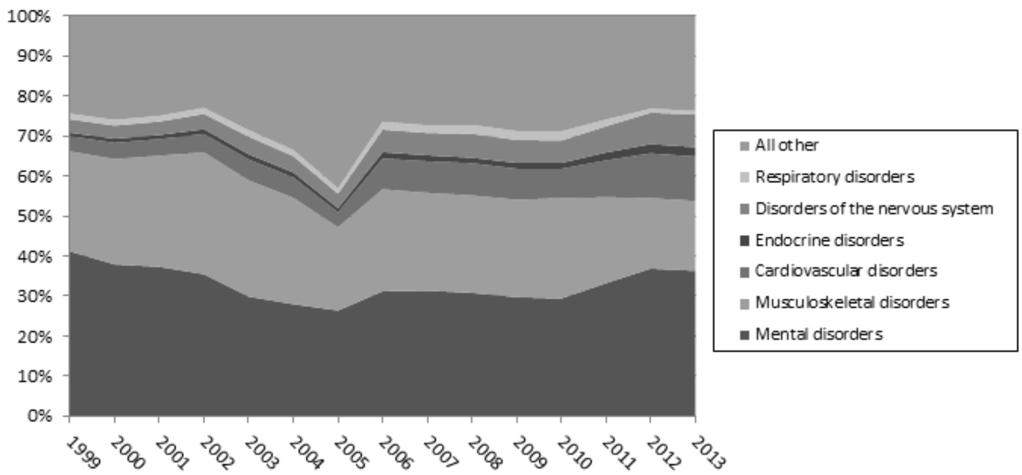
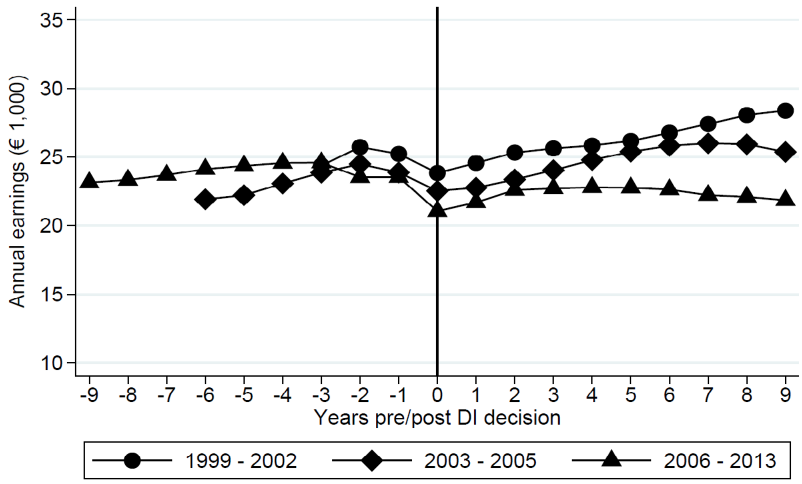
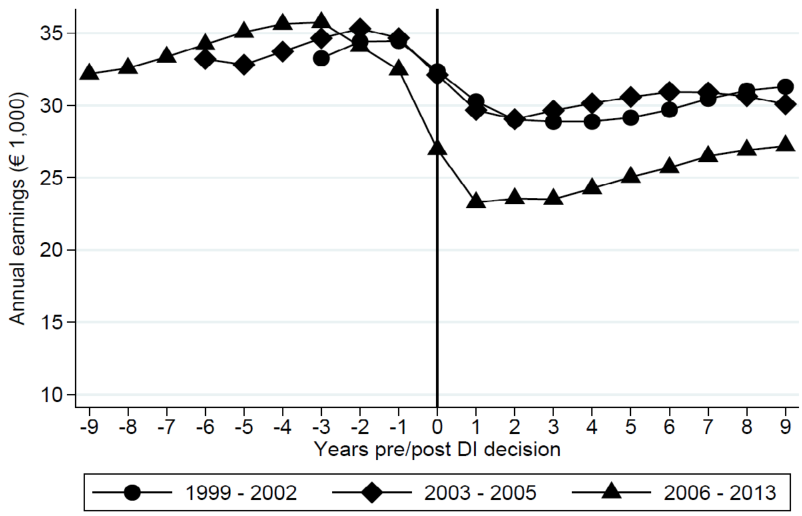


Figure A.12: Annual average earnings of rejected, and partially and fully awarded DI applicant cohorts for three time regimes, before and after application for DI benefits

Panel A. Positive annual earnings of rejected applicants



Panel B. Positive annual earnings of applicants awarded partial benefits



Panel C. Positive annual earnings of applicants awarded full benefits

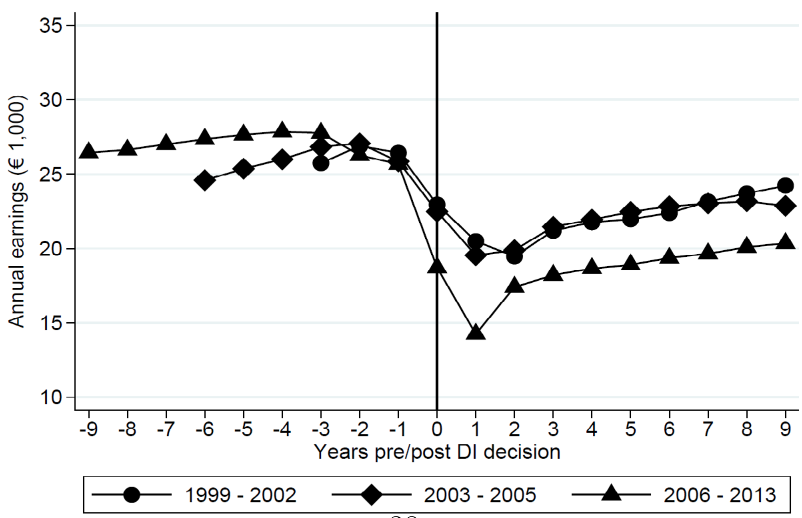


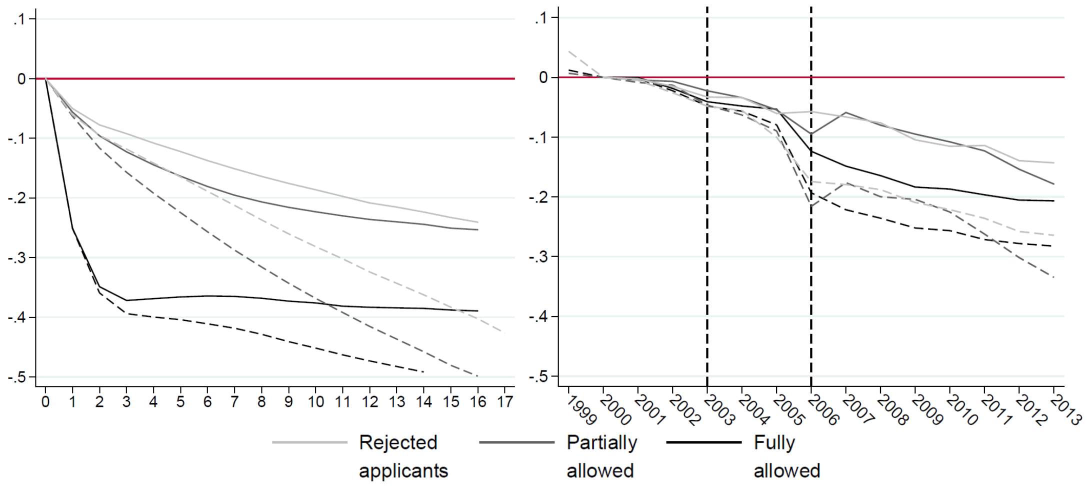
Table A.3: Estimated cohort differentials of rejected vs. awarded DI applicants

	Rejected vs partially awarded		Disability degree $< 35\%$ vs. $35 - 80\%$		Disability degree $< 35\%$ vs. $35 - 55\%$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{\gamma}_{1999}$	-0.027***	—	-0.027***	—	-0.025***	—
	(0.003)		(0.003)		(0.004)	
$\tilde{\gamma}_{2000}$	-0.064***	-0.035***	-0.053***	-0.028***	-0.051***	-0.023***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)
$\tilde{\gamma}_{2001}$	-0.060***	-0.032***	-0.050***	-0.028***	-0.051***	-0.024***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
$\tilde{\gamma}_{2002}$	-0.077***	-0.047***	-0.060***	-0.036***	-0.064***	-0.038***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
$\tilde{\gamma}_{2003}$ (GKP reform)	-0.065***	-0.046***	-0.056***	-0.036***	-0.056***	-0.038***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
$\tilde{\gamma}_{2004}$	-0.056***	-0.037***	-0.040***	-0.023***	-0.039***	-0.019***
	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
$\tilde{\gamma}_{2005}$	-0.075***	-0.042***	-0.040***	-0.020*	-0.051***	-0.025**
	(0.010)	(0.009)	(0.012)	(0.010)	(0.014)	(0.012)
$\tilde{\gamma}_{2006}$ (WIA reform)	-0.023***	0.011	-0.057***	-0.014**	-0.051***	-0.009
	(0.008)	(0.007)	(0.008)	(0.007)	(0.010)	(0.009)
$\tilde{\gamma}_{2007}$	-0.067***	-0.026***	-0.087***	-0.035***	-0.082***	-0.034***
	(0.008)	(0.007)	(0.008)	(0.007)	(0.009)	(0.008)
$\tilde{\gamma}_{2008}$	-0.052***	-0.018***	-0.075***	-0.031***	-0.069***	-0.033***
	(0.008)	(0.007)	(0.007)	(0.006)	(0.009)	(0.008)
$\tilde{\gamma}_{2009}$	-0.069***	-0.034***	-0.082***	-0.040***	-0.088***	-0.046***
	(0.008)	(0.007)	(0.007)	(0.006)	(0.009)	(0.008)
$\tilde{\gamma}_{2010}$	-0.060***	-0.029***	-0.074***	-0.033***	-0.075***	-0.037***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.008)	(0.007)
$\tilde{\gamma}_{2011}$	-0.037***	-0.013**	-0.043***	-0.013**	-0.044***	-0.016**
	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)
$\tilde{\gamma}_{2012}$	-0.020***	-0.009*	-0.020***	-0.004	-0.018***	-0.004
	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)
$\tilde{\gamma}_{2013}$	0.007	0.015***	0.007	0.019***	0.005	0.017***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)
F-statistic differenced cohort effects						
All cohorts	23.40	17.14	18.68	11.51	13.70	7.92
Regime 1	55.16	9.86	18.61	3.40	18.72	6.50
Regime 2	2.51	1.93	3.47	3.45	2.83	3.46
Regime 3	16.98	9.37	29.37	13.87	19.25	10.07
F-statistic differenced age effects	197.85	49.05	223.79	65.56	127.03	26.22
F-statistic differenced period effects	16.28	20.36	14.03	16.11	9.11	10.84
Age, period and common cohort effects	✓	✓	✓	✓	✓	✓
Controls	—	✓	—	✓	—	✓
Observations	6,730,460	5,561,737	6,736,052	5,567,329	6,193,528	5,095,192
R^2	0.0645	0.2027	0.0650	0.2030	0.0623	0.2002

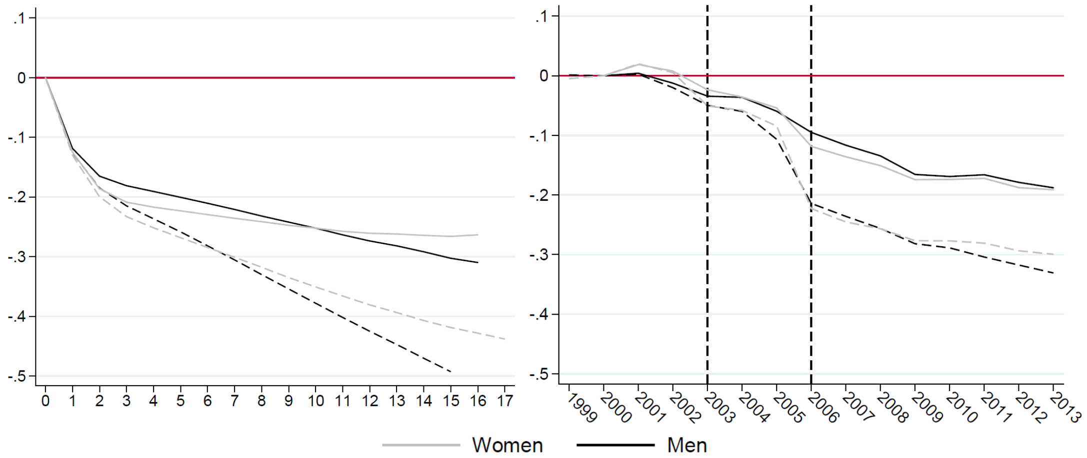
Note: Control variables include individual characteristics, impairment types and employment history. Reported F-statistics for multiple testing are Holm-adjusted. Individual clustered standard errors in the parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.13: Deaton-Paxson estimation results of age and cohort effects stratified by award decision, gender, age and impairment types, without (dashed line) and with (solid line) control variables

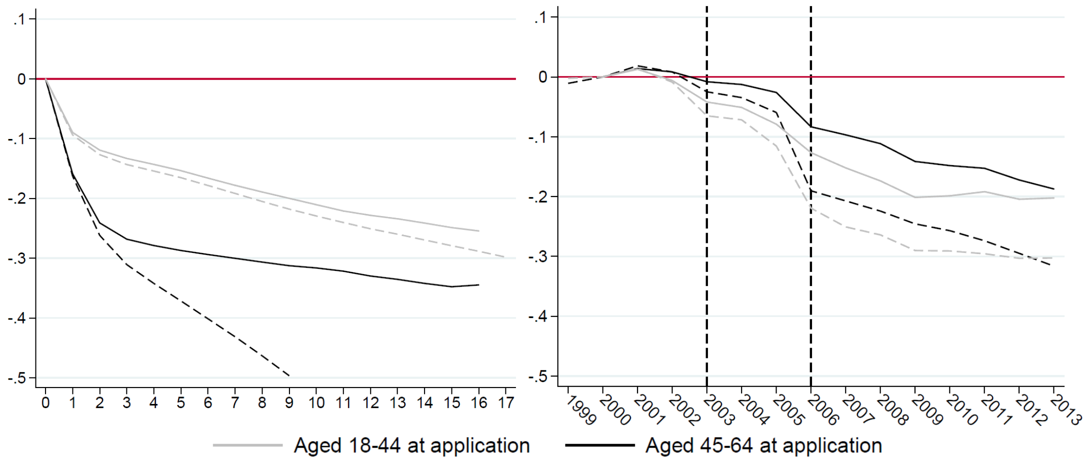
Panel A. Estimation results for rejected, and partially and fully awarded applicants



Panel B. Estimation results stratified by gender



Panel C. Estimation results stratified by age at application (18-44 vs. 45-64)



Panel D. Estimation results stratified by impairment types

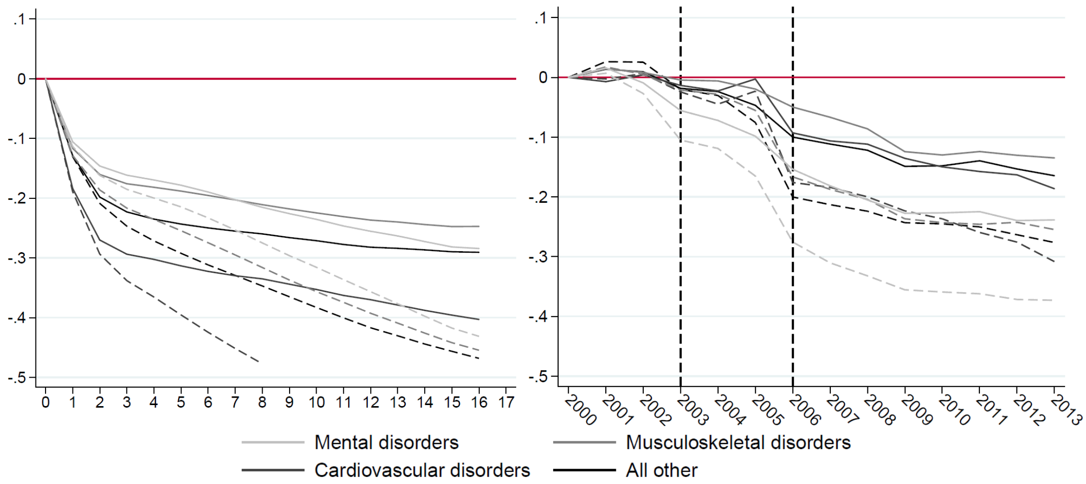
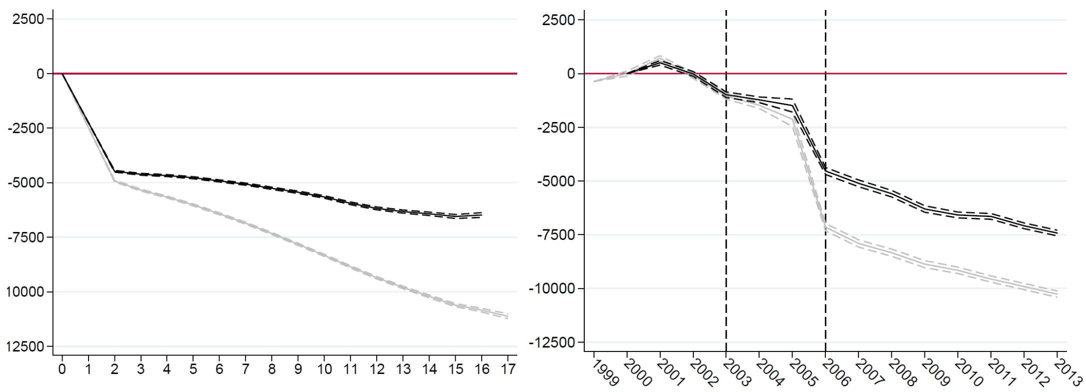
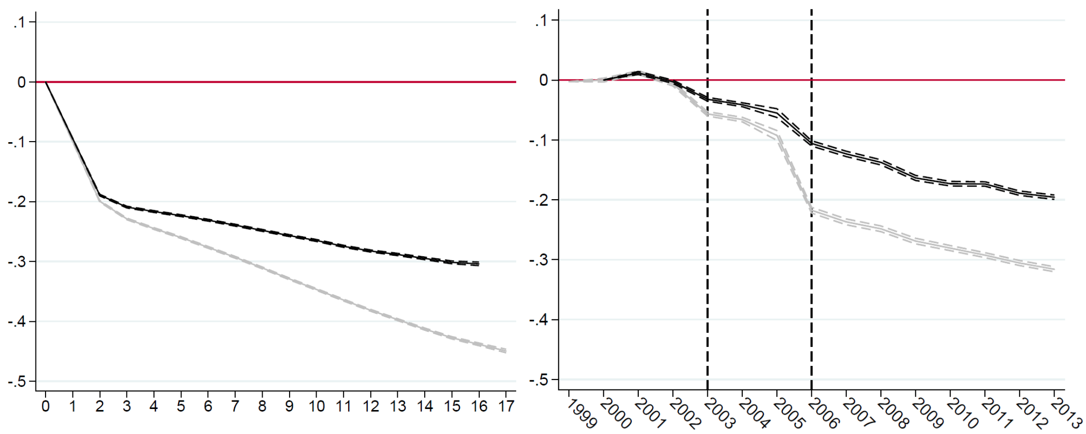


Figure A.14: Deaton-Paxson estimation results of age and cohort effect for other labor market and social security outcomes and mortality, without (grey) and with (black) control variables

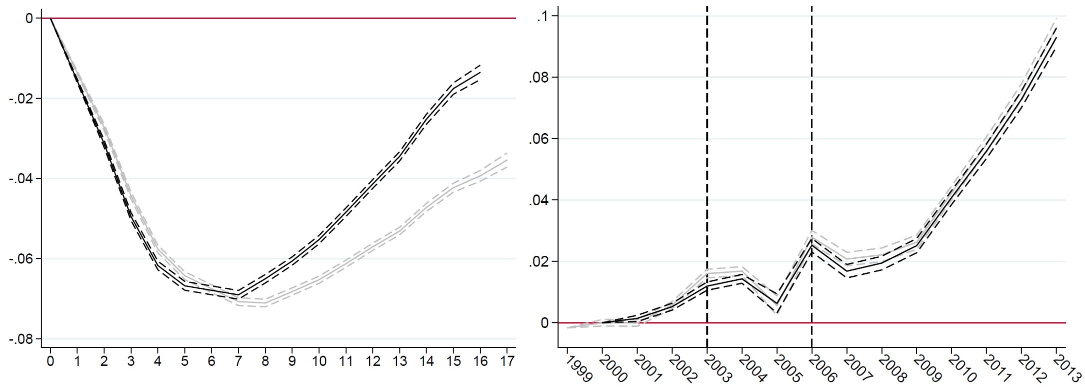
Panel A. Annual gross earnings (in 2015 Euros)



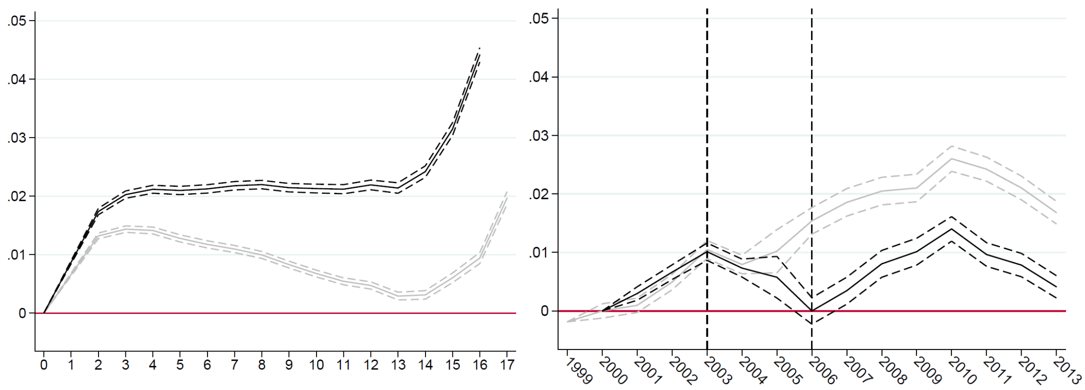
Panel B. Having a permanent contract



Panel C. Unemployment insurance benefit receipt



Panel D. Social assistance benefit receipt



Panel E. Deceased