Return-to-Work Policies for Disability Insurance Recipients: The Role of Financial Incentives

Guida Ayza Estopà[1] & Ilan Tojerow[2]

October 2024
(Reviewed August 2025)

Abstract: What is the impact of reducing the financial incentives of Belgium's Adapted Work Program (AWP), which allows individuals on Disability Insurance (DI) to combine salary and benefits? Does it encourage employment resumption or, conversely, does it push individuals back into full benefit dependency? This paper analyzes the impact of an increase in implicit marginal taxation on labor earnings of DI recipients (which translates into a reduction of financial incentives to work) on the probability to either resume full employment or complete benefits dependency. Using a rich set of administrative data, our study leverages a kink in the AWP's design and applies a regression kink design to infer the causal impact of a 30% increase in marginal taxation. Our findings reveal that the probability of DI recipients exiting the AWP to return to full benefit dependency at the kink increases by 3.5 p.p., while no significant impact is found on the probability of resuming full-time employment. We rely on a labor-leisure decision model to explain these findings. Further analysis of heterogeneous effects shows that men, blue-collar workers, and individuals with mental health pathologies are more likely to resume DI. These findings hold significant relevance for the design of return-to-work policies for DI recipients, shaping a path toward more effective and inclusive strategies.

JEL Classification: I13, I38, J14, J22.
Key words: Disability Insurance, Return-to-Work Policies, Financial Incentives, Regression Kink Design.

[1] Université libre de Bruxelles (Dulbea, Cebrig), Belgium
[2] IZA, Bonn, Germany

* This work was supported by the Belgian National Institute for Health and Disability Insurance (NIHDI).

1. Introduction

Disability Insurance (DI) plays a central role in the social protection systems of developed economies and is the most important welfare program in many of these countries. Over the past years, the number of DI beneficiaries has risen significantly, representing, on average, in OECD countries, 2% of GDP in public expenditure, 3.4% in Belgium (OECD, 2021). While DI is crucial for supporting individuals with disabilities, it also imposes significant economic pressure on welfare states and can increase the risk of poverty and social exclusion for beneficiaries. Given its impact on both societal and individual levels, previous research has long focused on various dimensions of DI (Autor & Dugan, 2006; Marie & Vall Castelló, 2012; Maestas et al., 2013; Kostol & Mogstad, 2014). However, a key question remains: how can policy be designed to encourage DI recipients to fully re-enter the labor market (thus completely abandoning benefit dependency) without compromising their well-being?

One strategy policymakers use to try to limit DI dependency is to allow individuals to work part-time while retaining a portion of their benefits, thereby creating financial incentives to work. However, as individuals increase their work earnings, their benefits are gradually reduced, effectively imposing an implicit tax on their earnings. A large body of literature suggests that higher taxation tends to disincentivize work by reducing financial rewards (Meyer & Rosenbaum, 2001; Saez, 2010; Chetty et al., 2013). Yet, within the DI framework, evidence on this relationship is more limited. While the existing literature provides valuable insights into how financial incentives influence the labor supply decisions of DI recipients already working part time, showing that reducing taxation positively affects work at the intensive margin (Campolieti & Riddell, 2012; Weathers & Hemmeter, 2011; Krekó et al., 2024), and some evidence of the work decisions at the extensive margin of DI recipients who are not yet working (Kostol & Mogstad, 2014; Vall-Castelló, 2017; Ruh & Staubli, 2019; Kostol & Myhre, 2021; Zaresani & Olivo-Villabrille, 2022), the extensive margin responses of individuals on DI who were already working part time and transition to either full-time employment or to complete DI dependency remain unknown. Knowing that the percentage of DI recipients engaged in part-time work is quite low—only 3.8% in Belgium in 2013 (NIHDI, 2013)—it is crucial to assess whether return-to-work policies truly achieve their main objective of full employment reintegration. These policies are often designed to serve as a stepping stone towards full work resumption, but it is also important to consider whether, conversely, they might act as a trap, limiting recipients' ability to fully re-enter the workforce.

In this paper, we explore the effects of a combination scheme that allows DI recipients to work part time while retaining a portion of their benefits: the Adapted Work Program (AWP). Two unique aspects of the pre-2018 Belgian setting allow us to investigate this previously unexplored scenario. Although the policy context has since evolved, the insights derived from this analysis remain highly relevant today. First, DI benefits were progressively reduced depending on the salary level, with a kink where the marginal taxation rate changes abruptly.

Specifically, this shift occurs when crossing a threshold of 24.9€, calculated after dividing the monthly salary by the 26 official working days in Belgium, resulting in an increase in the marginal taxation rate from 20% to 50%. Second, and most importantly, unlike in previous studies, individuals in this setting could not easily adjust their earnings by changing their working hours. In the Belgian system, those who began part-time work under the DI scheme could not freely determine their own working hours. Instead, they were limited by an advisory doctor, who considers the individual's health status and workplace characteristics to set a maximum of hours worked. With this limitation in place, the individual should also agree the number of working hours with the employer. The main implication of this feature is that for individuals was very difficult to bunch at the optimal side of the threshold. Consequently, the only feasible response to a higher marginal taxation rate, if any, was to abandon the combination scheme and transition either to full-time employment or to full DI dependency.

To understand the behavioral responses to changes in financial incentives, we draw upon the neoclassical model of labor-leisure choice. This model posits that individuals allocate their time between labor and leisure based on their preferences and the utility derived from income and free time. In our context, assuming that income is derived solely from salary, DI benefits, or a combination of both, the model helps illustrate how individuals may adjust their labor supply in response to changes in their budget constraint, such as those induced by shifts in taxation rates. Assuming Cobb-Douglas preferences, the model shows how changes in the budget constraint influence individuals' choices, particularly when only three discrete options are available in the Belgian setting: continuing to work part time, returning to full DI dependency, or fully re-entering the labor market.

We use a Regression Kink Design (RKD) to evaluate the effect of a higher marginal taxation rate—which reduces financial incentives to combine work and benefits—on individual labor supply decisions. We take advantage of the fact that individuals on DI who are already participating in the labor force cannot easily adjust their working time, making their response to changes in financial incentives likely to be a binary decision. This empirical strategy leverages the kink in the combination scheme, where the marginal taxation rate increases from 20% to 50% at a daily salary of 24.9 euros. As shown in Figure 1, this creates a discontinuous change in the marginal taxation rate at the threshold set by the Belgian Social Security Administration. The random assignment, stemming from the inability of recipients to strategically position themselves around this kink (an assumption that we properly test), effectively addresses selection bias, enabling us to employ quasi-experimental methods.

We use this discontinuity to identify the causal effect of benefit adjustments by testing for changes in the slope of the relationship between our outcomes of interest (i.e., returning to full DI dependency or resuming full-time employment) and the assignment variable (i.e., wage) at the kink. This empirical strategy allows us to explore the effects of implicit taxation on labor supply at the extensive margin. Importantly, the unique characteristics of the Belgian setting enables a more accurate isolation of extensive margin responses. A key contribution of this paper is that the Belgian setting allows us to more accurately isolate extensive margin responses, in contrast to more common settings where individuals can adjust their working hours to remain below the threshold and avoid the taxation increase.

Our dataset comprises comprehensive administrative records from the National Institute for Health and Disability Insurance (NIHDI). These records include data on disability recognitions, benefits, wages, and employment histories, as well as various socioeconomic and demographic characteristics of all DI recipients working part time in 2013 through the AWP. Using this rich dataset, we estimate the effect of a reduction in financial incentives on the labor supply of individuals who are combining part-time salary with benefits. Additionally, since our data also include information on gender, health conditions, social status, age, and the duration of DI spells, we can conduct a thorough heterogeneous analysis to explore how different subgroups respond to changes in financial incentives.

We find that a 1€ per working day increase in the implicit marginal taxation rate leads to a significant 4.4 percentage points decrease in the probability of remaining in the AWP by the end of the year. Additionally, this reduction in financial incentives results in a significant 3.5 percentage points increase in the probability of returning to full DI status, whereas there is no effect on the likelihood of resuming full-time employment. This translates into an elasticity of 0.8[1], which means that a 10% increase in the implicit tax per working day increases the probability of reverting full DI benefits by 8%. These findings align with the labor-leisure decision model, suggesting that increased taxation, which reduces financial incentives to work, pushes individuals to decrease their participation in the labor force. This implies a negative response at the extensive margin for DI recipients who were already working part time. These findings highlight the crucial role of financial incentives in such programs, as they demonstrate how their effectiveness in encouraging DI recipients to return to work becomes entirely insignificant when there is an increase in implicit taxation. Instead of fostering full work resumption, the result is the opposite: individuals who were already working part-time are pushed back into full DI dependency.

Our heterogeneous analysis reveals that the response to the increase in marginal taxation is primarily driven by men, blue-collar workers, and individuals with mental health disorders. While both genders react to the change, men exhibit a notably stronger response, particularly in terms of leaving the program and returning to full DI status. Age and duration of incapacity do not appear to be major differentiating factors, as similar patterns are observed across all age groups and between short- and long-term DI recipients—both tend to return to full DI rather than resume full-time employment at the kink. However, pathology type presents an

[1] The elasticity is obtained by multiplying 0.035 (the coefficient, see Table 3) by the average implicit tax paid per day at the threshold (= T(24.97)= [(24.97-15.61)*0.2]=1.872 €) and dividing the result by the average probability of returning to full disability benefits at the threshold (=Y(24.97)=0.0816). The elasticity is therefore equal to 0.035*1.872/0.0816 = 0.80.

interesting exception: while individuals with mental health conditions are more likely to revert to full DI, those with musculoskeletal disorders exhibit the opposite pattern, showing no significant return to full DI and even a slight increase in the probability of resuming full-time employment. This result points out that the diagnosis is a relevant factor to consider when designing return-to-work policies for DI recipients.

This paper primarily contributes to the literature examining how financial incentives affect the labor supply decisions of DI recipients, adding new evidence on extensive margin responses among DI beneficiaries who are already working part time. Evidence on labor supply at the extensive margin within the DI context is scarce, especially when focusing on transitions out of DI for recipients who are already participating in the labor force. Most of the existing literature primarily addresses the intensive margin, examining how beneficiaries adjust their working hours or earnings in response to changes in financial incentives. For example, Campolieti and Riddell (2012) report a significant increase in employment among Canadian Pension Plan Disability (CPPD) beneficiaries following the introduction of an earnings disregard policy, but find no significant impact on extensive margin decisions. Specifically, they do not observe any increase in the uptake of CPPD benefits nor in individuals exiting the DI rolls after the policy change.

Similarly, Weathers and Hemmeter (2011) evaluate the impact of a benefit offset program in the US. The authors find that while the program significantly increased the likelihood of beneficiaries earning above the Substantial Gainful Activity threshold, it had no statistically significant effect on overall employment rates or decisions to exit DI. Zaresani and Olivo-Villabrille (2022) also show that within a combination scheme, marginal changes in taxation had a substantial effect on the intensive margin—leading those already working to work more—but only modest effects on the extensive margin. Specifically, these changes prompted a 0.79 percentage point increase in the labor force participation rate among DI recipients who had not previously been working. More recently, Krekó et al. (2024) find that a policy reform in Hungary that reduced the earnings limit also significantly reduced labor supply on the intensive margin, yet had no sizable effect on program entry and persistence in the DI program.

However, there are a few notable exceptions where significant responses at the extensive margin have been found. Vall-Castelló (2017) show that a reform in Spain that eliminated disincentives for DI recipients to work increased the probability of disabled men working by 6.5 percentage points, but they were not previously working and decided to enter the labor force. Kostol and Mogstad (2014) also find that increasing financial incentives in Norway's DI program led to a substantial increase in labor market participation among beneficiaries, showing that financial incentives can indeed promote transitions into work among DI recipients who were not working before. Myhre (2021) finds significant responses at the extensive margin using a regression discontinuity design (RDD), but again, this involves DI recipients who moved from non-participation to having positive earnings. Similarly, Ruh and Staubli (2019) find that relaxing earnings restrictions in Austria led to a 2.3% to 6.7% increase in labor force participation among DI recipients. However, their analysis also focuses on individuals who initially were not working and decided to begin employment.

The existing literature therefore suggests that while there are some cases of significant extensive margin effects, they are solely observed among DI beneficiaries who were not initially in the labor force and decided to start working. The key contribution of our study is the examination of extensive margin decisions among DI recipients who are already working part time, specifically looking at their decisions to either transition back to full DI dependency or move toward full labor force participation in response to an increase in marginal taxation. This is possible due to the unique features of the Belgian setting, where we do not observe bunching around the threshold as commonly found in other contexts. Bunching, often used to study intensive margin responses, is helpful for estimating how individuals adjust their labor supply within the existing framework (i.e., working more or less to stay below certain taxation thresholds). However, it falls short when it comes to understanding how individuals make extensive margin decisions, such as whether to stop working entirely or to fully exit DI. The absence of bunching in our setting allows us to more clearly isolate and identify the extensive margin responses of individuals already working part time, providing new insights into how higher marginal taxation rates influence the overall participation of DI recipients in the labor force.

Our study also relates to a larger literature on the effects of marginal taxation on labor supply. Saez (2010) demonstrates that individuals tend to bunch at kink points in the tax schedule where marginal tax rates increase, highlighting the impact of taxation on intensive labor supply decisions. Chetty, Friedman, and Saez (2013) further explore these responses and find that behavioral reactions to marginal tax changes are influenced by both frictions (such as adjustment costs) and attention to taxation. Meyer and Rosenbaum (2001) provide evidence that decreases in marginal tax rates and increases in work incentives, particularly through programs like the Earned Income Tax Credit, have significantly boosted labor force participation among low-income individuals, particularly single mothers. These findings collectively suggest that, as this paper also shows, taxation plays a critical role in shaping labor supply decisions, both at the intensive and extensive margins, for different subsets of the population.

Lastly, this paper contributes to the debate on which measures most effectively boost labor supply. The literature reviews several possibilities. For example, tightening eligibility rules for disability benefits, as highlighted by Haller et al. (2024), can increase labor force participation but may negatively impact the welfare of vulnerable individuals. Monitoring strategies, such as compulsory dialogues (Markussen et al., 2017), effectively nudge sick-listed workers back to work but may not support those with complex barriers. Additionally, activation measures, such as vocational rehabilitation (Bastiaans et al., 2024), show promise in improving employment outcomes but require significant resources and may have mixed effects on mental health. Last, job coaching (Fontenay & Tojerow, 2022), which offers individualized support, helps with reintegration but is resource intensive and not easily scalable. We add to this debate by showing that relaxing implicit taxation on earnings for individuals already in the labor force may prevent them from returning to full benefit dependency, allowing them to continue their path toward complete work resumption.

The rest of the paper is structured as follows. Section 2 describes the Belgian disability scheme and the AWP. Section 3 presents the data, and Section 4 discusses the theoretical model. Section 5 explains the empirical strategy. Section 6 presents the results and the heterogeneity analysis, Section 7 presents robustness checks, and Section 8 concludes.

2. Institutional Setting

2.1. The Belgian DI System

In Belgium, every worker, regardless of whether they are currently employed or unemployed,[2] has the right to be insured through the payment of disability benefits if they cannot work due to health reasons. The public insurance system for work incapacity operates at the federal level,[3] with the scheme depending on the duration of the sickness absence and the characteristics of the employer.[4] The start of sick leave can be decided by any treating practitioner on the first day of sickness, and it entails a period of one month (14 days for blue-collar workers) of guaranteed wage.[5]

Individuals must meet three criteria to be considered as unable to work. First, they must have stopped all productive activity. Second, they must have stopped this productive activity as a direct consequence of a health deterioration that is not directly linked to their professional activity.[6] Third, their ability to work must be reduced by at least 66% with respect to their previous occupation. After six months, this third criterion changes, and the ability to work is

[2] To qualify, full-time and unemployed workers must have fulfilled a minimum of 180 working days (or active days of job searching for the unemployed) during the last 12 months. Part-time workers must have worked at least 800 hours over the last 12 months.

[3] The DI that we describe in this paper differs from work accident and occupational disease insurances, which are covered by other institutions.

[4] Specifically, whether they are employed or unemployed and whether they are white- or blue-collar. Self-employed workers have a distinct sickness insurance program that we do not cover in this paper.

[5] The periods of guaranteed wages are as follows. For white-collar workers, it is 30 days of guaranteed wages, not bounded and paid by the employer. For blue-collar workers, guaranteed wages, not bounded, and paid by the employer from the 1st to the 7th day; 85.88% of the wage paid by the employer from the 8th to the 14th day; and 25.88% of the part of the wage below the limit fixed by the mutuality and 85,88% of the part of the wage over this limit paid by the employer + 60% paid by the mutuality from 15th to the 30th day.

[6] This distinction is necessary to differentiate the short-term disability program from other programs, such as the occupational injuries fund and the occupational diseases fund.

then evaluated with respect to any occupation that the worker could perform given their age, education, and experience.

After one month, an advisory doctor[7] decides whether the individual qualifies for short-term work incapacity. During this period, the individual will receive DI benefits from their health insurance fund ("mutuality").[8] For a regular employee, the benefits correspond to 60% of their gross salary, irrespective of the work regime.[9] This amount is bounded by maximums and minimums that slightly differ depending on the family situation. Unemployed workers receive the same amount from their mutuality that they used to receive from unemployment benefits.

During the first year of illness, individuals are categorized as having short-term work incapacity and are covered under a program called "primary incapacity." After one year, the applicant's doctor at her mutuality (who oversaw the applicant during the short-term period) submits an application to the NIHDI to transition the individual into the long-term work incapacity program, referred to as "invalidity." An advisory doctor then assesses whether the individual qualifies for long-term DI (invalidity), based either solely on the documentation provided by the general practitioner or through the advisory doctor's own evaluation.

Both programs are financed by the NIHDI. The main differences between them include the methods for evaluating residual work capacity and calculating the replacement rate. For short-term incapacity, the replacement rate is 60% of the last salary, while in the long-term program, it increases to 65%. This rate is subject to variations based on family status.[10] For the unemployed, the last wage received before unemployment is used to determine the applicable rate.

2.2. The AWP

In Belgium, salaried workers on DI can return to work by adjusting their working hours to suit their health condition, with approval from an advisory doctor from their mutuality, through the AWP.[11] This program is designed to facilitate the full reintegration of individuals with temporary or permanent illness into the labor market. It allows them to work part time while receiving a combination of salary and disability benefits until they are fully recovered and ready to resume full-time work. Participants can return to their previous job or transition to a

[7] The advisory doctors are physicians from the Belgian mutuality funds, and aim assess if a person is eligible to receive benefits (from the NIHDI) due to a health-related work incapacity. They can also recommend the worker to gradually return to work and must provide official authorization.
[8] In Belgium, the health care system is financed by both social security contributions and compulsory health insurance funds ("mutualities"). The latter are funded by the NIHDI (INAMI in French) and act as intermediaries between this institution and individuals with disabilities. They are responsible for reimbursing medical expenses and providing short-term disability benefits.
[9] The gross salary is calculated daily (a week consisting of six working days).
[10] This percentage is 65% for individuals with dependents, 60% for single households, and 40% for cohabitants, with defined floor and ceiling amounts.
[11] "Reprise de travail adapté" in French. Note that self-employed individuals participate in another program.

different job, as long as the new role is compatible with their health condition. Despite the AWP's significant advantages for DI beneficiaries, only 3.84% participated in the program in 2013 (NIHDI, 2023).

For an individual receiving short- or long-term disability benefits to re-enter the labor market, they must have a disability rating of at least 50% and approval from an advisory doctor. If these conditions are met and the selected job is compatible with the individual's health status, they may ask the advisory doctor for approval to work part time (i.e., enroll in the AWP), or the advisory doctor may suggest it. In either scenario, the advisory doctor limits the number of hours that can be worked, based on the individual's health condition and the job characteristics. Thereafter, and within this limit, the individual must negotiate the exact number of working hours with the employer. Once enrolled in the AWP, the individual receives a combination of income from employment and (reduced) disability benefits. The disability benefits are reduced only if the labor income exceeds certain thresholds.

The follow-up of AWP participants is managed by the advisory doctor, who must renew consent every two years. If an individual wishes to modify their approved working hours, they must request an appointment with the advisory doctor, who will reassess their health condition and decide whether the change can be authorized. The institutional feature that workers cannot freely choose their working time supports our no-bunching assumption and represents a key distinction between the Belgian DI scheme and those commonly studied in the international literature. Any choice of working hours requires agreement from both the advisory doctor—who may deny an increase in hours—and the employer.

2.2.1 Financial Design of the AWP

In Belgium, the law established by the Royal Decree of July 1965 and coordinated in July 1994 governs the reduction of disability benefits for those working while on a claim. This law has undergone several reforms, influencing decisions about program participation[12]. This section focuses specifically on the system as it was in 2013. However, from 2018, the system changed: disability insurance reductions are no longer calculated on the basis of income but rather on the number of hours worked. More specifically, no reduction applies as long as the beneficiary does not exceed a threshold equivalent to 20% of full-time employment. Beyond this threshold, benefits are gradually reduced according to the number of additional hours worked. While this reform provides a clearly identifiable institutional change from an empirical strategy perspective, it complicates the analysis of behavior with respect to the implicit tax rate on work. In this regard, the pre-2018 regime offers a more suitable framework for ensuring the validity of the empirical analysis.

[12] Table A1 in the appendix reports all the actualizations on the brackets since 2012.

Following adjustments in 2012, the system stabilized in 2013 with a structure that adjusts the marginal tax rate on benefits through three kinks based on the daily salary from adapted work. The first kink occurs at a daily salary of 15.61€ with a marginal tax rate of 20%, the second at 24.97€ with an increase from 20% to 50%, and the third at 34.33€ with an increase to 75%. Our analysis focuses on the second kink, which represents the most significant change in slope—an increase of 30 percentage points—and maintains all the necessary assumptions outlined in Section 5.

The relationship between the benefits received and the salary from adapted work can be defined by equation (1):

$$B_{AW} = \begin{cases} B_0 & if \quad \omega \leq 15.61 \\ B_0 - [(\omega - 15.61) \times 0.2] & if \quad 15.61 < \omega \leq 24.97 \\ B_0 - \begin{bmatrix} \left((24.97 - 15.61) \times 0.2\right) \\ +\left((\omega - 24.97) \times 0.5\right) \end{bmatrix} & if \quad 24.97 < \omega \leq 34.33 \\ B_0 - \begin{bmatrix} \left((24.97 - 15.61) \times 0.2\right) \\ +\left((34.33 - 24.97) \times 0.5\right) \\ +\left((\omega - 34.33) \times 0.75\right) \end{bmatrix} & if \quad \omega > 34.33 \end{cases} \qquad (1)$$

where B_0 refers to the benefits received before starting the adapted work, B_{AW} denotes the net benefits after applying the reduction (in brackets), and ω is the salary from the adapted work. The calculations are made using the daily salary.[13]

As indicated in equation (1), the reduction in benefits is determined by a function of the daily salary, which exhibits kinks at three specific points. The daily salary is computed by dividing the monthly salary by 26, the officially recognized average number of working days in a month in Belgium. Consequently, the salary used in our analysis is not affected by differential daily working hours or weekly workdays. The drawback of this approach is that we cannot differentiate between individuals working fewer hours with a higher salary and those working more hours with a lower salary. Nevertheless, this limitation does not affect the applicability of our methodology.

To illustrate the reduction at the second kink, consider a worker earning a daily salary above 24.97 euros. Up to 15.61 euros, there is no reduction in benefits, while for earnings between 15.61 euros and 24.97 euros, the marginal tax rate is 20%. Once earnings surpass 24.97 euros, they are taxed at a 50% marginal rate, resulting in a greater reduction in benefits. Figure 1 illustrates this reduction, showing the change in the slope of the policy rule at the second kink, and Appendix Figure A1 details the marginal tax rate for each salary level.

[13] Daily salary is computed by taking the monthly salary and dividing it by the official number of working days in the month (25 in months with 28 or 30 days, and by 26 in months with 31 days).

3. Data

3.1. Sample Description

We use three administrative datasets from the NIHDI that comprise data on short- and long-term DI payments, disability recognitions, and the AWP. The datasets are linked at the individual level using anonymized social security identifiers. We draw a large dataset covering all the salaried individuals who enter and exit the AWP in Belgium in 2013. The data include work incapacity, such entrance and exit dates, payment amounts, pathology, and reason of exit; AWP data, such as entrance and exit dates, salaries, and exit paths; and a rich set of demographic characteristics, including gender, age, region, and social status.

As we are interested in the effects of the financial incentives that stems from the combination of salary from part-time work and disability benefits, we narrow the sample to salaried individuals. Therefore, we exclude self-employed individuals (as they are subject to a different return-to-work program) and individuals working in non-remunerated activities (i.e., voluntary work). This results in a sample of 47,775 individuals. The baseline results are computed using the observations for which we have complete information on salaries (i.e., no missing value) from January to December of 2013, constituting a sample of 36,778 individuals.

Due to the richness of the dataset, we can observe behavioral reactions to taxation of people participating in the AWP in terms of labor supply, and perform different heterogeneity analysis considering characteristics such as gender, age, region, social security status, and pathology.

3.2. Descriptive Statistics

Table 1 presents sample characteristics, showing means and, in parentheses, standard deviations. It also provides descriptive statistics of the subsample around the kink, which refers to all individuals in the AWP with a salary within 8 euros gross of the kink point (10,008 observations). In the next section we explain why this is our preferred bandwidth for the estimations. Figure 2 displays the distribution of daily salary relative to the kink. The bandwidth of 8 euros includes mostly individuals in the second quartile of the salary distribution.

The table shows that in the sample, 63% are women, the average age is 46 years, and 55% of individuals are blue-collar workers. The mean daily salary is around 39 euros, and the mean daily benefits received are around 37 euros gross. An important observation is that the variability of salaries is significantly greater than that of benefits. This reflects how benefits are limited by minimums and maximums, thus reducing the deviation from the mean, while salaries are freely set. The table also reports the same statistics for the subsample around the kink: there are slightly more women than men around the kink, while age is very stable, reporting similar values. Not surprisingly, the ratio of blue-collar workers increases somewhat around the kink, as most white-collar workers should be earning higher salaries. In fact, the mean daily salary is notably lower in the sample around the kink, but mean daily benefits are equal.

Table 2 illustrates the most common pathologies among AWP participants, their reasons for exiting the AWP, and the regions where they live. It also reports data for the full sample and the subsample around the kink. For both, musculoskeletal disorders and mental health conditions are the two main reasons for being disabled, with stable prevalence percentages for all conditions. Additionally, there are no noticeable differences in terms of region. When examining reasons for exiting the AWP, the most common are returning to full-time work (40%) or reverting to full DI status (24%). The reasons for exiting appear to be highly influenced by proximity to the kink. The probabilities of returning to full DI status and exiting to the labor market have reversed—now it is more likely to return to DI than to reincorporate to the labor market, though both probabilities are around 30%. This shift represents a significant decrease in the labor market exit path, accompanied by an increase in the odds of returning to full DI status.

4. Theoretical Framework

This section develops the theoretical framework that supports our analysis. We build on the neoclassical model of labor-leisure choice, which suggests that individuals may reallocate their hours of labor and leisure in response to changes in their budget constraints, depending on their preferences. In our specific context, however, we extend the canonical model by incorporating constraints that reflect institutional rules, resulting in discrete labor supply choices. Below, we outline the fundamental components of the model and explain how it adapts to our particular setting.

4.1. Budget Constraints

We define two budget constraints: one that represents the scenario under the DI scheme, which includes the possibility of working via the AWP, and another that represents the scenario outside of the DI scheme in the formal labor market. We assume that income can be derived from a salary, DI benefits, or a combination of both.

Figure 3 illustrates these two budget constraints. BC1 represents the budget constraint for an individual outside the DI scheme, working in the formal labor market. This budget line follows the canonical neoclassical labor-leisure model (Boeri & van Ours, 2013). It is a straight line starting from a consumption level of zero, which occurs when the individual is dedicating all their available time to leisure, meaning that no income is generated if no work is performed. The slope of this line is negative and equal to the wage (-w).

BC2 is the budget constraint for individuals within the DI scheme. It builds on the canonical model but includes our specific institutional assumptions. Starting on the right of the graph, BC2 indicates that individuals can receive DI benefits while dedicating all their time to leisure, where consumption equals the amount of DI benefits received. As an individual begins working through the AWP, their total income increases, and the slope of BC2 becomes negative, reflecting the wage rate at that point. An increase in marginal taxation reduces the effective wage, which flattens the slope of the budget line. Additionally, BC2 features a kink where marginal taxation increases by 30%, leading to a flattening of the line beyond this point.

Thus, an individual at point A is fully reliant on DI without working, while an individual at point B is working part time through the AWP, combining benefits and salary (taxed marginally at 20%). An individual at point C is also working through the AWP but is now facing a higher marginal tax rate of 50% (30% higher than before). Any point on BC1 represents a condition where the individual is working outside the DI scheme in the formal labor market.

4.2. Agent Preferences

To simplify, we distinguish between two types of agents: those with a stronger preference for leisure and those with a stronger preference for consumption (and thus labor). According to the standard consumption theory, agents prefer a combination of leisure and consumption, rather only one good and zero of the other, implying that their indifference curves never touch the axes. Therefore, in both cases, we use Cobb-Douglas preferences but where each agent weighs leisure and consumption differently.

Due to the institutional constraints in our setting, individuals in the AWP cannot freely adjust their working hours. Instead, they can only choose between three options when faced with a change in marginal taxation: i) remain in the AWP (i.e., do not move), ii) return to full DI (move to point A), and iii) exit the DI scheme entirely and enter the formal labor market (move to any point on BC1). The shape of each individual's indifference curve determines which of these options maximizes their utility.

Example of agent type I: Stronger preference for leisure. Panel A of Figure 4 shows this individual’s indifference curves, which are more vertical compared to an individual with a stronger preference for consumption. Suppose that this individual initially starts at point B (AWP before the kink). Since point B lies on the highest attainable indifference curve, it is their equilibrium point and they have no incentive to move.

If marginal taxation increases, moving this individual to point C (AWP after crossing the kink), they must choose between staying in the AWP, returning to full DI (point A), or moving to the formal labor market (BC1). Given their strong preference for leisure, points A and C yield the same utility level, making either one a viable equilibrium. For any point to the left of point C, the individual would maximize utility by choosing point A, thus reverting to full DI.

Example of agent type II: Stronger preference for consumption (and labour). Panel B presents this individual's indifference curves, which are more horizontal. In this case, assuming that point B is the individual's endowment point, they will prefer to stay there rather than moving to either point A or any point in BC1. After an increase in marginal taxation, their new endowment point will become C. They will gain even more utility than from point B and will decide to remain there too, constituting the new equilibrium point. Their optimal decision is to remain in the AWP.

The consumption theory indicates that indifference curves never have positive slopes, and even this agent, who displays a high preference for labor, prefers to remain in the AWP rather than transitioning to the formal labor market (BC1). The insights from this model suggest that, despite the marginal taxation increase, moving back to full DI remains a rational choice for many agents, especially those with high leisure preferences. Remaining in the AWP is also plausible, either because the agent has a stronger preference toward work or due to information frictions, as the literature shows that individuals are often not fully informed about tax changes, which can limit their ability to make optimal decisions (Kostol & Myhre, 2021). Ultimately, these findings suggest that an increase in marginal taxation does not significantly encourage DI recipients to transition back to the formal labor market.

To test these theoretical insights, we proceed with an empirical strategy to assess the causal impact of higher marginal taxation on labor supply decisions using our dataset. We further extend this analysis by examining heterogeneous effects to determine whether specific individual or socioeconomic characteristics influence these decisions.

5. Empirical Strategy

5.1. RKD

In our empirical strategy, we use the RKD methodology, proposed by Card et al. (2012), which exploits exogeneous changes in the slope of a policy rule. This approach requires an explanatory variable—in our case, the reduction in benefits—to be a deterministic and known function of an assignment (or running) variable, which in our case is the daily salary through the AWP. This methodology is similar to an RDD, but instead of exploiting a discontinuity in the assignment rule, the policy rule is expected to be kinked (with a change in the slope when relating it to the assignment variable) at one or more points. One of the most typical scenarios to apply RKD is when a marginal tax rate shows discrete jumps, as in our case. One of the most typical scenarios for applying RKD is when a marginal tax rate shows discrete jumps, as is the case in our study.

The main goal of this methodology is to examine potential changes in the slope of the relationship between the outcome of interest and the assignment variable upon crossing the kink. Based on the same principles as an RDD, the RKD assumes that individuals are randomly located around the kink, and any behavioral differences are solely attributable to their locations on different sides of the kink, thereby being impacted differently by the policy. Thus, the methodology estimates the elasticity of the outcome in response to changes in the policy. Like an RDD, the RKD provides local estimates and relies on some assumptions.

5.2. Identifying Assumptions

One basic condition to implement RKD is that the function must have at least one kink point. This means that the function has segments where it is continuous and differentiable, but in at least one point, it is continuous yet not differentiable, having unequal left and right derivatives. This condition is easy to identify from the policy rule and is illustrated in Figure 1. Additionally, the RKD relies on two important assumptions: no sorting and smoothness. In this subsection we discuss and test these assumptions.

Assumption 1: No sorting. The density of the assignment variable, the daily salary, is smooth around the kink. This smoothness means that individuals cannot manipulate their salary to position themselves in the optimal part of the function, which in this case is just before a kink. If there were observable bunching near the kink points, it would invalidate this assumption. However, the local random assignment condition holds in our context; manipulation of earnings seems unlikely since individuals cannot decide their salary. While they might be able to adjust it by changing their working hours, such changes would require approval from both the advisory doctor and the employer, which may take some time. Moreover, as detailed in Section 2, the advisory doctor approves the number of hours the individual can work based on a health assessment. Thus, it is very unlikely that individuals can sort themselves in the salary distribution.

We start by providing graphical evidence of the absence of bunching before the kink in Figure 2; however, we even go one step further and, following the literature, present formal tests for this assumption below (Kostøl & Mogstad, 2014; Card et al., 2015; Landais, 2015; Fontenay, 2022). By plotting the kernel density function of the salary in the AWP, one can already see that the distribution looks smooth, and we cannot observe any discontinuity or bunching around the kinks. However, we also conduct formal tests to assess the no-sorting assumption. We start by providing the results of the McCrary test. McCrary (2008) presents a test of manipulation related to continuity of the running variable density function in RDD, a widely used methodology in the RDD and RKD literature. He explains that histograms and kernel density estimations are badly biased and do not allow for point estimation or inference. The test that he proposed complements these methods by estimating potential discontinuities in the density functions of the assignment variable at the kink. Panel A of Figure 5 reports the results of the McCrary manipulation test, which is insignificant and confirms the lack of discontinuity at the kink.

Additionally, we extend our analysis by performing the test proposed by Card et al. (2015), which examines the continuity of the first derivative of the probability density function (PDF) at the kink. We present the coefficient estimate and standard error of this test in Panel A of Figure 5 as well. The estimated coefficient of the first derivative indicates a significant increase in the slope of the PDF at the kink. This result provides further evidence supporting the validity of the RKD in our analysis.

More recently, Cattaneo et al. (2020) developed a novel discontinuity in density testing procedure that we also implement in this study. This procedure is based on their proposed local polynomial density estimator, and it is fully data-driven: bandwidth selection methods are formally implemented along with inference methods based on robust bias correction. Panel B of Figure 5 presents the density of the assignment variable on both sides of the kink as well as the formal results for the Cattaneo et al. (2020) manipulation test, which confirms that we cannot detect manipulation around the kink. Specifically, the test statistic is 1.3, with a corresponding p-value of 0.19, indicating that we fail to reject the null hypothesis of no manipulation.

Taken together, both the graphical evidence and the formal tests results suggest that the no-sorting assumption holds on this setting, indicating that individuals cannot manipulate their earnings to locate at one or the other side of the kink.

Assumption 2: Smoothness. In the absence of kinks in the benefit function, there would be a smooth relationship between outcomes and the assignment variable. If this assumption holds, evidence of a change in the slope would imply a causal effect of the benefit amount on the outcome. However, it is not empirically testable, as we cannot infer what the counterfactual would be absent the policy rule. As a result, we instead test the smoothness of different control variables at the kinks. If the covariates are not smooth around the kinks, this assumption fails. This procedure is commonly used in the RDD literature (Imbens & Lemieux, 2008) but is applicable to the RKD context (Fontenay, 2022). Figure 6 shows how various control variables are smooth around the kink, and thus the assumption holds.

5.3. Formal Model

The causal effect in the RKD is calculated as the ratio between the change in the slope of the outcome variable (remaining in the AWP by the end of the year) and the exogeneous change in the slope of the policy rule (the implicit marginal tax rate per working day). It is expressed as a fraction, where the numerator represents an estimation and the denominator is a deterministic value.

Formally, let Y_i be a dichotomous variable equal to 1 if individual $i \in \{1, 2, \ldots, n\}$ is in the AWP by the end of the year, w the daily salary in the AWP, B_0 benefits before the reduction, B_{Aw} benefits after the reduction (R), and T the tax:

$$\left.\begin{array}{l} B_{AW}=B_0-R \\ T=f(w)=B_0-B_{AW}(w) \end{array}\right\} B_{AW}=B_0-f(w), \tag{2}$$

where $f(w)$ is a defined function of the salary at the kink we are investigating; w_k=24.97.

Our parameter of interest is the change in the slope of the conditional expectation function:

$$m(w)=E[Y_i \mid w_i=w], \text{ at } w=w_k. \tag{3}$$

Thus, the probability of staying in the AWP by the end of the year for individual i, depending on their daily salary (which determines their implicit marginal tax rate). This term is then divided by the change in the slope of the deterministic (known) function:

$$T=f(w)=B_0-B_{AW}(w), \text{ at } w=w_k. \tag{4}$$

The general model of interest is of the form

$$Y_i=a(w_i,T,\varepsilon_i), \tag{5}$$

which explains the probability of staying in the AWP by the end of the year depending on the salary, the implicit marginal tax rate per working day (which also partly depend on salary), and the error term.

Card et al. (2012) show that τ, the change in marginal tax, is identified at $w=w_k$ if the following assumptions hold:

(i) $Y_i=a(w_i, T, \varepsilon_i)$ and its derivatives with respect to w_i and T, are continuous,
(ii) $f(w)$ has a kink, and
(iii) the density of w is smooth at w_k.

Under these assumptions, $E[\varepsilon_i \mid w_i=w_k]$ is a smooth function, and

$$\beta=\frac{D_+ m(w_k)-D_- m(w_k)}{D_+ f(w_k)-D_- f(w_k)}, \tag{6}$$

where

$$D_j\, m(w_k) = \lim_{w \to w_k^j} \frac{d\, m(w)}{dw},$$

$$D_j\, f(w_k) = \lim_{w \to w_k^j} \frac{df(w)}{dw}, \qquad j \in \{+,-\},$$

where j∈{+,-} refers to the left and right sides of the threshold.

The full equation will be

$$\beta\ = E\left[\frac{dY}{dB}|w = w_k\right] = \frac{\lim_{w \to w_k^+} \frac{d\,E[Y|w=w_k]}{dw} - \lim_{w \to w_k^-} \frac{d\,E[Y|w=w_k]}{dw}}{\lim_{w \to w_k^+} \frac{dT(w)}{dw} - \lim_{w \to w_k^-} \frac{dT(w)}{dw}}, \qquad (7)$$

where β is the weighted average of marginal treatment effects across the population. The weight is the relative likelihood that an individual has w_i=w_k, given ε_i. Individuals with a higher likelihood of being at the threshold receive higher weights. Weights can be applied differently, depending on the chosen kernel function.

From equation (7), one can see that the RKD estimator is a ratio of two terms. The numerator is the change at the kink point in the slope of the relationship between outcome Y (probability of staying in the AWP) and salary w. The denominator is the change at the kink point in the slope of the tax function. The resulting estimate can be interpreted as a local treatment-on-the treated parameter, as there may be some selection into the program.

The numerator of equation (7) is estimated semi-parametrically using the following local power series estimation:[14]

$$\mathrm{E}[Y_i|w_i = w] \approx \alpha_0 + \sum_{p=1}^{p}\left[\alpha_p (w - w_k)^p + \beta_p D_i (w - w_k)^p\right], \ (8)$$

where p is the chosen polynomial and D is the treatment status: $D_i(z)=1$, if z>0, $D_i(z)=0$ otherwise (corresponding with the two sides of the link, + and -). The power series is a local approximation of *m(w).*

Note that $|w-w_k| \leq h$, where h is the bandwidth chosen for the estimation. The denominator of the equation represents the change in the slope of the deterministic policy rule *f(w)* at the kink point. Since the NIHDI enforces the rules for benefit reduction and there are no cases in which an individual in the AWP is exempt from this reduction, we can follow Landais (2015) and proceed with a sharp RKD.

---

[14] Local power series estimation allows us to approximate a function by a polynomial, at least locally. Linearizing a function offers a good approximation near the point of tangency (by using with p(2) and p(3), we can further refine our approximations). To fully represent an analytical function, however, we must employ an infinite number of terms in a power series (i.e., a Taylor series).

As previously explained, our analysis focuses on the second kink, as the other two are not suitable for this analysis. The first kink presents challenges due to the smaller magnitude of the taxation change, the limited sample size around it, and its targeting of individuals with particularly low salaries. Regarding the third kink, as shown in Figure 2, the salary density function decreases immediately after the kink. This drop is due to the natural normal distribution of salaries, as this kink is located near the center of the salary distribution. Such a scenario may compromise the validity of Assumption 1 and potentially lead to confounded effect estimates.

5.4. Implementation

We measure the kink in the outcome variable by estimating three key parameters: a local polynomial regression of order p to the left and right of the kink, a kernel density function K, and bandwidth h.

Choice of polynomial order. In the previous literature, it was assumed that in the RKD context, a local quadratic approach is preferred to a local linear approach as it is expected to provide an asymptotically smaller bias. However, Card et al. (2017) warn against making the quadratic model a universal choice and argue that one should also account for the characteristics of the dataset of interest, including the sample size. In fact, some studies have shown that in certain scenarios the local linear polynomial performs better. Given our context and the linear nature of the policy rule (i.e., the reduction in benefits is linearly affected by the daily earnings), we select the local linear polynomial as our preferred specification (see, e.g., Fontenay, 2022).

Choice of kernel. Cheng et al. (1997) show that the triangular kernel is boundary optimal. However, Card et al. (2017) later demonstrate that the losses from using a uniform kernel (i.e., no weighting) are small, easing the RKD implementation. Following the literature, we choose the uniform distribution as our preferred specification. As a robustness check, we show the results of applying a triangular density function in Section 7.

Choice of bandwidth. The choice of bandwidth involves balancing variance and bias. Increasing the bandwidth size includes more observations, which reduces variance but can increase bias. Conversely, a smaller bandwidth reduces bias but can increase variance. If the bandwidth is "too large," it may introduce a significant bias in the estimator of the conditional expectation function. As Landais (2015) points out, the RKD tends to perform poorly with small samples and is therefore quite demanding in terms of bandwidth size compared to RDD. In our context, avoiding overlap between kinks is essential to prevent confounding effects, and so we limit bandwidths to a maximum of 9.36.

To explore the optimal bandwidth for each outcome, we use the data-driven procedure of Calonico et al. (2014), which is specifically designed for fuzzy RKD. The bandwidths selected by this procedure range from approximately 8 euros to 10 euros around the kink, depending on both the outcome and the selection method used. For comparison purposes, having a consistent bandwidth across different outcomes is important. Therefore, and considering our limit of 9.36, we use a common bandwidth of 8 euros as our baseline. For robustness checks, we also present results using each outcome's optimal bandwidth and additional bandwidth variations in Section 7.

6. Results

In this section, we present our findings on the effect of an increase in implicit marginal taxation. We begin by providing graphical evidence and then proceed to the empirical results. As our baseline estimation, we investigate the impact of taxation on the probability of remaining in the AWP by the end of the year, as well as the number of days an individual stays in the program throughout the year. Additionally, we expand our analysis to assess the causal effects on two key outcomes: the probability of returning to full DI status and the probability of re-entering the formal labor market. Last, we delve into a more detailed analysis by exploring potential heterogeneity in responses, considering factors such as gender, pathology, social status, and length of disability.

6.1. Impact of Taxation on the Probability of Remaining in the AWP

6.1.1. Graphical Evidence

Once decided the key parameters, first step when implementing a RKD is to assess the graphical evidence. We start analyzing the probability to stay in AWP by the end of the year depending on the salary, and hence on the taxation faced. Figure 7 shows the relationship between the probability of an individual remaining in the AWP in December 2013 and the assignment variable at the kink. Using a bandwidth of 8 euros and bins of 1 euro, and displaying a linear trend at each side of the kink, we can evaluate the graphical evidence. The sharp reversal in the slope of the kink suggests a significant behavioral response. The probability of remaining in the program increases with salary up to the kink. However, after the marginal tax rate increases, the probability of remaining in the AWP by the end of the year decreases as salary rises.

The figure suggests a sharp change in the kink, where the tax increases from 20% to 50%. This provides supportive evidence for identifying the effect of higher taxation, which depends on salary, on the labor supply of DI recipients through the AWP. To confirm this graphical evidence, we empirically estimate treatment effects and report the results in the next subsection.

6.1.2. Baseline Empirical Results

Columns 1 and 2 of Table 3 present the results for the baseline estimation of equation (7).[15] The table begins by reporting the value of the deterministic denominator in the first stage, followed by the treatment effects. In each line, we report the estimated coefficient β in the RKD analysis, which is a result of dividing the estimated treatment effect by the known denominator, the robust confidence interval, the calculated mean of the dependent variable within the defined bandwidth around the kink, and the number of observations. Together with the estimated coefficients, we also report the associated significance levels and standard errors.

Panel A reports the first-stage deterministic coefficient, which is 0.3 as enforced by the policy rule. This coefficient indicates the increase in the marginal taxation rate. Panel B reports the treatment effects of the kink. Column 2 does not include covariates in the regression, while column 2 introduces age, gender, region, and social state as additional control variables. Since we only have information on pathology for individuals in long-term disability, we lose 11,576 observations when including pathology as a control variable. However, the results for this regression do not differ significantly depending on this variable. Thus, in the regressions presented here, the set of covariates does not include pathology, but the results controlling for this variable are reported in Appendix Table A2. In the next subsection, we analyze separately the samples for short- and long-term disability recipients and are able to evaluate differential effects by pathology.

Our main coefficient of interest is β, which indicates that an implicit increase of €1 per working day in the implicit marginal tax rate leads to a significant reduction of 4.0 to 4.4 percentage points in the probability of still being in the AWP at the end of the year. In the cases where the marginal tax is 20%, individuals do not seem to be discouraged from continuing to offer some labor capacity while retaining part of their disability benefits; in fact, one can even detect an increase in labor supply as salary increases. This finding changes when approaching the kink: at a marginal tax of 50%, every extra euro earned results in half of it being deducted from benefits. At this point, both the graphical representation and the estimated coefficients show strongly significant evidence of a reversal in this trend. From here, labor supply through the AWP decreases with every extra euro earned.

To improve the robustness of our analysis, we employ an alternative variable to assess the participation in AWP: the total number of days an individual works through the program in the whole year. The findings using this variable align with those previously obtained and are detailed in the appendix. Figure A2 illustrates a notable decline in the average number of days worked through the program after crossing the kink. Furthermore, Table A3 shows a 1€ per working day increase in the implicit marginal taxation around the kink is associated with a 8.8

[15] To obtain the value in equation (7), we parametrically estimate equation (8) and divide by the deterministic value in the first stage (0.3).

percentage points decrease in the number of days, on average, that an individual remains in the AWP throughout the whole year, which corresponds to around 32 fewer days in the program.

The empirical results confirm the graphical evidence and align with previous literature findings in the literature: financial disincentives reduce participation in return-to-work programs. However, our findings raise two important questions about the behavioral responses affecting labor supply: Where do these individuals who exit the AWP go? Do they return to full DI or to a complete work resumption?

6.2. Impact on Exit Paths

We extend the analysis by investigating the impact of marginal tax rate changes on the probability of going back to full DI and of resuming full-time work. These results extend the existing literature, providing novel insights on the effect of financial (dis)incentives on effective work resumption and benefits dependency.

As in Section 6.1.1, we begin by providing graphical evidence. Panel A of Figure 8 illustrates the relationship between the assignment variable and the probability of an individual in the AWP returning to full DI status due to lower financial incentives. The probability of returning to full DI decreases with salary, and the trend line illustrates a significant negative relationship between these two variables. However, as the kink is approached, the line flattens, indicating that after the increase in marginal taxation, the probability of returning to full DI rises significantly. Panel B illustrates the relationship between the assignment variable and the probability of completely resuming work. There is no change in the slope at the kink, suggesting that an increase in marginal taxation does not affect labor resumption.

To confirm the graphical evidence and quantify which of these probabilities exhibits a greater effect, we estimate treatment effects empirically, as reported in columns 3–6 of Table 3. Columns 3 and 4 present the estimated causal effect of the kink on the probability of returning to full DI status, both without and with covariates. Columns 5 and 6 show the estimated impact on the probability of resuming full-time employment. Our main coefficient of interest is β, which reveals that a 1€ per working day increase in the implicit marginal taxation rate at the kink leads to a significant 3.5 percentage points increase in the probability of returning to full DI status by the end of the year (3.3 p.p. when covariates are included). These coefficients are statistically significant at high levels with narrow robust confidence intervals. In contrast, for the probability of resuming full-time work, as suggested by the graphical evidence, the estimated coefficients are very close to zero and not statistically significant.

Our findings suggest that an increase in marginal taxation for individuals participating in the AWP effectively pushes them out of the program. While this results in a notable increase in the likelihood of returning to full DI status, there is no observable effect on the probability of resuming full-time employment—even if the program's primary objective is to foster reintegration into the labor force. Other potential exit paths from the AWP, though less common, include pensions, program exclusion, unemployment, and mortality.

Our results also align with the labor-leisure decision model. The increase in taxation acts as a reduction of the budget constraint, and for a representative agent with convex preferences—who normally prefers a combination of both goods—a higher taxation rate provides more utility from returning to DI rather than fully abandoning the benefit dependency and relying solely on labor earnings.

Additionally, these findings may be of considerable interest to policymakers, as they highlight the unintended consequences of a policy designed to facilitate work resumption. When labor earnings are highly taxed, the policy fails to achieve its goal and instead has the opposite effect, pushing individuals who were already participating in the labor force back into full benefit dependency.

6.3. Heterogeneous Effects

In this subsection, we examine the heterogeneous effects on the likelihood of individuals responding to a higher implicit marginal taxation rate by either returning to full-time employment or to reverting to full DI after crossing the threshold where the marginal taxation rate increases by 30% in the AWP. We conduct separate regression analyses, segregating the data by gender, age bracket, pathology, social status, and prior disability length. Table 4 presents the significant differences based on these characteristics, while Appendix Figures A3 to A7 report the graphical evidence. Below, we describe the results and discuss the intuition behind them.

6.3.1. Effects by Gender

To analyze the impact of gender, we conduct six different regression analyses. Panel A of Table 4 presents the results for both male and female subgroups. Our findings indicate that men are more responsive to the increase in marginal taxation, with a 4.5 percentage point higher probability of leaving the program by the end of the year. Additionally, their probability of returning to full DI status increases by 5.8 percentage points after the taxation rises, while the probability for women increases by 2.4 percentage points. For both genders, the probability of returning to full-time employment at the kink does not shows significant values. These results align with the trends observed in the overall sample reported in Table 4, but are higher in magnitude and significance for men than for women.

One plausible explanation for these findings is that, on average, women tend to have higher rates of part-time employment compared to men (OECD, 2022). Consequently, they may exhibit lower responsiveness to changes in taxation rates, as they are more likely to view their alternative to work resumption as part-time work without additional benefits. As a result, they may find it optimal to continue working part time despite the higher implicit taxation. In contrast, men are generally less inclined toward part-time employment, which makes them more sensitive to changes in taxation rates and more likely to return to full DI when financial incentives diminish.

Furthermore, due to gender norms and domestic responsibilities, women may prefer continuing to work part time, even if financial incentives decline. This preference aligns with existing literature highlighting how social norms and domestic roles shape women's labor supply decisions (Kleven et al., 2019; Alesina et al., 2013). Additionally, some studies suggest that women are generally less reactive to economic shocks than men. For instance, during the COVID-19 pandemic, women's likelihood of working outside the home decreased more significantly compared to men, without any corresponding increase in unpaid work (Farré et al., 2022). Thus, women might be less inclined to exit a combined DI and salary scheme entirely, while men, faced with reduced incentives, are more likely to return to full DI status. This behavior could be influenced by the perceived utility gains that women derive from balancing work with domestic responsibilities, even if it involves financial sacrifices (Akerlof & Kranton, 2000).

6.3.2. Effects by Age

To analyze the impact of age, we divide the sample into three age groups: young individuals aged 18–35, middle-aged individuals aged 36–49, and older individuals aged 50–59. Notably, the youngest group represents the smallest proportion of AWP participants, with only 7,779 observations, compared to the middle-aged and older groups, which have 20,484 and 19,512 observations, respectively. This suggests that the program is more popular among individuals nearing retirement, with 59 being the most common age among all participants. Appendix Figure A8 illustrates the age distribution of AWP participants.

When analyzing the first outcome—the probability of remaining in the program in response to an increase in the marginal taxation rate—we find that the youngest group has little to no response. In contrast, the middle-aged group exhibits the strongest response, with a 4.2 percentage points decrease, followed by a 3.6 percentage points decrease for the older group (50 years and above). None of the age groups show significant coefficients regarding the probability of resuming full-time employment at the kink. However, all demonstrate a significant impact on the probability of returning to full DI status, with the older age group showing the most statistically significant results. These findings are consistent with the baseline outcomes, reinforcing the broader trends observed across the entire sample.

6.3.3. Effects by Pathology

When analyzing the impact of pathology type, we focus on mental health disorders and musculoskeletal disorders, as these are the two most common pathologies among DI recipients and collectively constitute a significant portion, approximately 60%, of the entire sample (see Table 2). Our findings, shown in Panel C of Table 4, reveal distinct tendencies based on the pathology type. Individuals with mental health conditions show a slightly stronger response to a higher implicit marginal taxation. They are more inclined to return to full DI rather than fully resume work when faced with an increase in taxation. Specifically, those with mental health disorders are 4.8 percentage points more likely to return to full DI status at the taxation kink, and no significant impact is observed on the probability of returning to full-time employment for this group.

In contrast, individuals with musculoskeletal disorders show a higher propensity to resume full-time work at the taxation increase. For this group, the probability of returning to full-time employment shows a modestly significant coefficient, increasing by 2 percentage points. Unlike other subgroups, there is no significant impact on the probability of reverting to full DI status for this group. These findings suggest that our overall results are mostly driven by individuals with mental health conditions. For individuals with musculoskeletal disorders, however, the increase in marginal taxation does not lead to a full return to DI, and in this case, it may even achieve the intended goal of encouraging them to resume employment.

The existing body of literature has already established a clear, negative link between mental health and employment. Shen (2023), analyzing Canadian data, demonstrates the significant impact of poor mental health on employment outcomes. Similarly, Ringdal & Rootjes (2022), using data from the Netherlands, arrive at the same conclusion. Insights from this literature suggest that the primary challenges faced by individuals with mental health issues in the labor market stem from a decline in cognitive abilities, including difficulties in concentration, motivation, and social interactions. Furthermore, Noordik et al. (2011) present a qualitative study showing the barriers to a full return to work for individuals with mental health conditions. They explain that although most of patients aim for a full return to work, there are critical intention-behavioral gaps between proposed solutions and actual behavior for these type of patients. Our study aligns with this literature, showing that in the context of the AWP, individuals with mental health-related conditions are the least likely to fully return to work after a decrease in financial incentives.

This finding is particularly relevant for policymakers, as it reveals that diagnosis is a key factor when designing return-to-work strategies for DI recipients. While financial incentives seem somewhat effective for individuals with musculoskeletal disorders, they appear uneffective for those with mental health conditions. Evidence from studies such as Fontenay & Tojerow (2022) shows that alternative interventions, like Supported Employment programs, which focus on job coaching and follow-up support, may be more effective for individuals with mental health conditions, fostering higher employment rates and reducing reliance on DI benefits.

6.3.4. Effects by Social Status

In our exploration of heterogeneous effects based on social status, we segment our sample into blue- and white-collar workers. While we find no statistically significant results for white-collar workers, significant findings for blue-collar workers indicate a greater responsiveness to changes in taxation among this group. It is worth noting that the prevalence of blue-collar workers is higher in the sample around the kink, as illustrated in Table 1. In contrast, white-collar workers are more commonly found toward the right tall of the salary distribution. Consequently, the sample size of white-collar workers around the kink may not be sufficient to thoroughly analyze the impact of the taxation kink on this group. For blue-collar workers, our results indicate a 3.3 percentage points greater likelihood of returning to full DI, and, consistently with the rest of our results, any significant impact is found when analyzing the probability to resume full-time employment at the kink.

6.3.5. Effects by Disability Type

To examine the effects based on disability duration, we distinguish between individuals in the 'primary incapacity' program and those in the 'invalidity' program. As explained in Section 2, the key difference between these groups is that the former includes individuals incapacitated for one year or less, while the latter consists of those who have been on DI for over a year. Our findings show significant results for both groups.

Individuals with a longer history of disability have a decreasing likelihood of exiting DI over time, making policies aimed at facilitating their return to work particularly valuable. In this analysis, individuals on long-term DI demonstrate some responsiveness to a decrease in financial incentives into job resumption. However, they are still more likely to return to full DI (an increase of 2.6 percentage points) than to resume full-time work, with a modest but significant coefficient of less than 1 percentage point (see Panel E, Table 4). Similarly, individuals on short-term DI show a significant probability of returning to full benefit dependency at the kink with no significant effect on resuming full-time employment.

Additionally, we find that individuals on long-term DI exhibit greater labor supply elasticity in response to changes in financial incentives compared to those on short-term DI. This suggests that allowing long-term DI recipients to retain a larger portion of their benefits while working could be crucial in encouraging sustained labor force participation. Therefore, incorporating this flexibility into return-to-work policies could be essential for supporting this group.

7. Robustness Checks

To ensure the reliability and validity of our findings, we conduct a series of robustness checks. In this section, we detail the specific tests, which are standard in the RKD literature, and discuss how they contribute to the overall rigor of our analysis.

7.1. Sensitivity analysis for the density function

As explained in Section 5, the results from the RKD analysis are sensitive to several parameter choices, including the selection of the density function. To address this, we analyze the effect of varying this parameter, and Appendix Table A4 presents the results obtained using triangular and Epanechnikov density functions. Notably, the results remain robust across these two alternative density functions: the probability of staying in the AWP and returning to full DI remain significant across the different density functions, while the coefficient related to the probability of returning to full-time employment remains insignificant. These results indicate that all three outcomes are consistent and not significantly affected by the choice of density function.

7. 2. Sensitivity analysis for the choice of bandwidth

As discussed in Section 5.4, the choice of bandwidth in the RKD framework represents a critical trade-off between variance and bias. A larger bandwidth reduces variance by including more observations but may introduce greater bias into the estimator, while a smaller bandwidth can reduce bias but may lead to increased variance. If the selected bandwidth is too large, it may result in a considerable bias when estimating the conditional expectation function. In our specific context, it is crucial to avoid overlaps between different kinks, as it could confound the results. To prevent this, we establish an upper limit of 9.36 for the bandwidths. To determine the optimal bandwidth for each outcome, we use the data-driven procedure of Calonico et al. (2014), specifically designed for a fuzzy RKD. The bandwidths chosen by this method, typically fall between 8 and 10 euros around the kink, varying slightly depending on the outcome and the specific selection method.

Appendix Table A5 presents the results obtained using various bandwidths. For the outcome measuring whether individuals remain in the AWP by the end of the year, the coefficient remains significant for bandwidths between 8€ and 10€, including lower (7€ and 7.5€) and larger values (10€ and 20€). The optimal bandwidth suggested by the Calonico et al. (2014) selector for this outcome is 8.394€, which yields a significant effect of 5.9 percentage point at a p-value of 0.000. For the outcome measuring the probability of returning to DI, the coefficients are also significant across all bandwidths tested, with the optimal bandwidth set at 9.112 €, resulting in a coefficient of 3.1 percentage point and a p-value of 0.000.

On the other hand, for the "back to job" outcome, the coefficients are significant from bandwidths of 8.5€ or higher, but the p-values are typically around 0.050. However, many of these results fall beyond our established bandwidth limit of 9.36, which, as already explained, is not convenient in our context. The significance of this outcome is not stable throughout all the specifications.

In general, we observe that as the bandwidth increases, the coefficients tend to decrease slightly. However, for the "back to AWP" outcome, they range between 3.8 and 5.9 percentage points, and for "back to DI," they vary between 2.6 and 3.5 percentage points, which are still substantially significant in magnitude.

7.3. Test on the functional dependence

In Section 5, we discussed that the RKD relies on the assumption that the relationship between the assignment variable and outcomes (when there is no kink) should exhibit a degree of smoothness. A valid concern may be whether the effects detected by the RKD arise due to nonlinearities in this relationship. While the visual evidence presented earlier helps to alleviate this potential concern, one way to test it is by including different controls in our regressions to account for possible nonlinearities. Fontenay (2022) takes a similar approach in his analysis and explains that controlling for relevant covariates should strengthen the credibility of the RKD.

Tables 4 and A2 include baseline results with and without covariates, showing no significant differences in the outcomes when using these two specifications. The selected covariates (pathology, region, gender, age, and social state) are likely correlated with both the assignment variable and the outcomes of interest. In fact, they may be determinants as they influence an individual's income level. In Table 4, we have already shown that the results remain consistent whether these covariates are included or excluded.

In summary, after conducting a comprehensive series of tests, we can confidently assert that our findings are robust. Our analysis reveals that crossing the salary kink—where the marginal implicit tax rate increases in 30%—significantly decreases participation in the AWP and increases the likelihood of returning to full benefit dependency.

7.4. Sensitivity analysis for the functional form

The results can also be sensitive to the choice of the functional form. We begin testing this with a graphical analysis presented in Figure A9 in the Appendix, where we illustrate the outcomes of interest using a quadratic function of the assignment variable. In Panel A, the trend indicates a similar discontinuity at the kink for the probability of returning to full Disability Insurance (DI), whereas Panel B shows that for the probability of resuming full employment, there appears to be no significant discontinuity at the kink. The lines displayed in these graphics are aligned to the ones displayed when using the linear functional form in Figure 8.

To further validate these observations, we present the estimates for the treatment effects using a quadratic specification in Table A6 of the Appendix. Column 1 demonstrates that the probability of returning to full DI remains statistically significant, albeit at a somewhat lower statistical level. In contrast, Column 2 reveals that the probability of resuming full-time employment remains statistically insignificant. The increasing standard errors associated with the quadratic model, compared to those of the linear model, suggest that the linear polynomial is a more appropriate specification. This finding should not be surprising, as the policy rule is a linear function.

These findings corroborate the claim by Gelman and Imbens (2019) that the use of higher-order polynomials in regression discontinuity analysis can lead to inadequate coverage of confidence intervals. Consequently, our preferred specification is linear, consistent with most studies employing RKD (Ganong and Jäger, 2018).

8. Conclusion

DI represents a significant portion of social expenditure in modern economies. Addressing the rising number of individuals on DI through effective return-to-work policies is crucial, with financial incentives playing a key role in their design. In Belgium, the AWP is intended to serve as a policy tool to facilitate the reintegration of DI recipients into the labor market. However, our findings suggest that reducing financial incentives to remain in the AWP does not enhance full-time employment reintegration; rather, it leads to an increased dependency on full benefit.

Our study leverages a kink in the AWP scheme and applies an RKD to infer the causal impact of a 30% increase in marginal taxation rates on the likelihood of individuals remaining in or leaving the program. We find that a 1€ per working day increase in implicit marginal taxation rate reduces the probability of individuals remaining in the program by 4.4 percentage points. Further analysis of exit paths indicates a 3.5 percentage points increase in the probability of returning to full DI, with no significant impact on the probability of resuming full-time employment. This result is of highly relevance as it points out how return-to-work policies not always truly achieve their main objective of full employment reintegration, and highlights how the role of financial incentives is a determinant of this.

Exploring the heterogeneity in these results, we find that men are more sensitive to a taxation increase than women, but both genders are more likely to return to full DI status than full-time employment. Men, blue-collar workers, and individuals with mental health disorders are the ones more likely to resume the complete DI status while the length of disability and the age group does not appear to have major influence on the decision. Importantly, we demonstrate that individuals with mental health conditions tend to return to full DI, while those with musculoskeletal conditions are somewhat more likely to return to full-time employment. Indicating the relevance of consider diagnosis when designing return-to-work policies for DI recipients.

This study expands the literature on the impact of financial incentives in DI on labor supply and RTW policies by exploring a previously unexplored area: the extensive margin responses to reduced financial incentives for part-time work among DI recipients who are already participating in the labor force. Specifically, it examines the effects of an increase in implicit marginal taxation on effective full-time employment reintegration and on benefit dependency. It also sheds light on gender and pathology differences, which is novel in this context, thus also contributing to both the extensive literature on gender differences in the labor market and the scarce literature on the relationship between mental health and returning to work.

Moreover, this paper contributes to a broader literature on unemployment traps and exiting policies from other social welfare programs. The implications derived from our results research holds significant relevance for the design of such policies and can be extrapolated to other EU countries that have implemented or are contemplating return-to-work initiatives for DI recipients. In particular, the findings stress the importance of targeted policies and highlight that men, individuals with mental health conditions, and blue-collar workers face the greatest challenges when resuming full-time employment. Our empirical findings are supported by a theoretical framework relying on the labor-leisure consumption model, and a wide set of robustness checks validates their robustness.

Our findings are highly relevant for policymakers, as they suggest that increasing labor income taxation for DI recipients could hinder their gradual reintegration into employment, pushing those already working part time back into full benefit dependency. While this outcome challenges the sustainability of welfare states due to increased expenditures on DIs, it also negatively affects individuals by reducing their purchasing power and hindering their gradual return to employment.

References

Akerlof, G., & Kranton, R. E. (2000). Economics and Identity. *The Quarterly Journal of Economics*, 115(3): 715-753.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, 128(2): 469-530.

Autor, D. H., & Duggan, M. G. (2006). The growth in the social security disability rolls: A fiscal crisis unfolding. Journal of Economic Perspectives, 20(3), 71-96.

Bastiaans, M., Dur, R., & Gielen, A.C. (2024). Activating the long-term inactive: Labor market and mental health effects. Labour Economics, 90: 102593

Boeri, T., & van Ours, J. (2013). *The Economics of Imperfect Labor Markets: Second Edition* (REV-Revised, 2). Princeton University Press.

Böckerman, P., Kanninen, O., & Suonemi, I. (2018). A kink that makes you sick: The effect of sick pay on absence. *Journal of Applied Econometrics*, 33(4): 568-579.

Calonico, S., Cattaneo, MD., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression discontinuity designs. *Econometrica*, 82(6), 2295-2326.

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

Card, D., Lee, D. S., Pei, Z., & Weber, A. (2012). Nonlinear policy rules and the identification and estimation of causal effects in a generalized regression kink design. *National Bureau of Economic Research*. Working Paper 18564.

Card, D., Lee, D. S., Pei, Z., & Weber, A. (2015). Inference on causal effects in a generalized regression kink design. *Econometrica*, 83(6), 2453-2483.

Card, D., Lee, D. S., Pei, Z., & Weber, A. (2017). Regression Kink Design: Theory and practice. *Advances in Econometrics*, 38, 341-382.

Cattaneo, M.D., Jansson, M. & Ma, X. (2020). Simple local polynomial density estimations. *Journal of the American Statistical Association*, 115(531): 1449-1455.

Cheng, P.E., Johnson, E.G. & Liou, M. (1997). Standard Errors of the Kernel Equating Methods Under the Common-Item Design. *Applied Psychological measurement,* 21(4).

Chetty, R., Friedman. J.J., & Saez, E. (2013). Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review,* 103(7): 2683-2721.

Farré, L., Fawaz, Y., González, L., & Graves, J. (2022). Gender Inequality in Paid and Unpaid Work During Covid-19 Times. *The Review of Income and Wealth*, 68(2): 323-347.

Fernández, R., Fogli, A., & Olivetti, C. (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *The Quarterly Journal of Economics*, 119(4): 1249–1299.

Fontenay, S., & Tojerow, I. (2022). How Does Supported Employment Help Disability Insurance Recipients Work While on Claim? - IZA Discussion Paper 15386

Fontenyay, S. (2022). How Does Maternity Leave Allowance Affect Fertility and Career Decisions? *Working paper.*

Haller, A., Staubli. S., & Zweimüller, J. (2004). Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits?. Econometrica, 92 (1): 79-110

Imbens, G.W. & Lemieux, T. (2008). Regression discontinuity design: A guide to practice. *Journal of Econometrics*, 142(2): 625-635.

INAMI. (2013). Rapport annuel 2013.

https://www.inami.fgov.be/SiteCollectionDocuments/rapport-annuel-2013.pdf (Accessed October 16, 2024)

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4): 181-209.

Kools, L., & Koning, P. (2019). Graded return-to-work as a stepping stone to full work resumption. *Journal of Health Economics*, 65: 189-209.

Kostol, A.R., & Mogstad, M. (2014). How financial incentives induce disability insurance recipients to return to work. *American Economic Review*, 104(2): 624-655.

Kostol, A.R., & Myhre, A.S. (2021). Labor supply responses to learning the tax and benefit schedule. *American Economic Review*, 111(11): 3733-66.

Krekó, J., Prinz, D., & Weber, A. (2024). Take-up and labor supply responses to disability insurance earnings limits. Labour Economics, 89: 102583.

Landais, C. (2015). Assessing the welfare effects of unemployment benefits using the regression kink design. *American Economic Journal*, 7(4): 243-278.

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.

Marie, O., & Vall Castello, J. (2012). Measuring the (income) effect of disability insurance generosity on labour market participation. Journal of Public Economics, 96(1-2), 198-210.

Markussen, S., Røed, K., & Schreiner, R.C. (2017). Can Compulsory Dialogues Nudge Sick-listed Workers Back to Work? The Economic Journal, 128 (610): 1276-1303

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142: 698-714.

Meyer, B., & Rosenbaum, D. (2001). Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers. The Quarterly Journal of Economics, 116 (3) 1063-1114

Noordik, E., Nieuwenhuijsen, K., Varekamp, I., van der Klink, J. J., & van Dijk, F. J. (2011). Exploring the return-to-work process for workers partially returned to work and partially on long-term sick leave due to common mental disorders: a qualitative study. *Disability and Rehabilitation*, 33(17-18): 1625-1635.

OECD. (2021). Public expenditure on disability and sickness cash benefits in % GDP. OECD. https://www.oecd.org/en/data/indicators/public-spending-on-incapacity.html?oecdcontrol-89cf33ff83-var1=OECD%7CEU%7CBEL&oecdcontrol-9202e3bf52-var3=2021 (Accessed May 13, 2025)

OECD (2022). LMF1.6: Gender Differences in Employment – Definitions and methodology. *OECD Publishing*.

Ringdal, C., & Rootjes, F. (2022). Depression and labor supply: Evidence from the Netherlands. *Economics & Human Biology*, 45(C).

Ruh, P. & Staubli, S. (2019). Financial incentives and earnings of disability insurance recipients: evidence from a notch design. *American Economic Journal: Economic Policy*, 11(2): 269-300.

Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3): 180-212.

Shen, Y. (2023). Mental health and labor supply: Evidence from Canada. *SSM - Population Health*, 22: 101414.

Vall-Castelló, J. (2017). What happens to the employment of disabled individuals when all financial disincentives to work are abolished? *Health Economics,* 26(S2): 158-174.

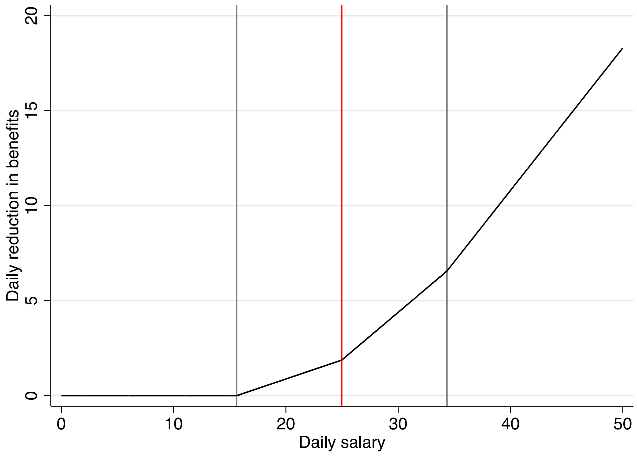
Weathers, R-R., & Hemmeter, J. (2011) The impact of changing financial work incentives on the earnings of social security disability insurance (SSDI) beneficiaries. *Journal of Policy Analysis and Management*, 30(4): 708-728.

Zaresani, A. (2018). Return-to-Work Policies and Labor Supply in Disability Insurance Programs. *American Economic Review Papers and Proceedings,* 272-276.

Zaresani, A., & Olivo-Villabrille, M. (2022). Return-to-work policies' clawback regime and labor supply in disability insurance programs. *Labour Economics, 78, 102215.*

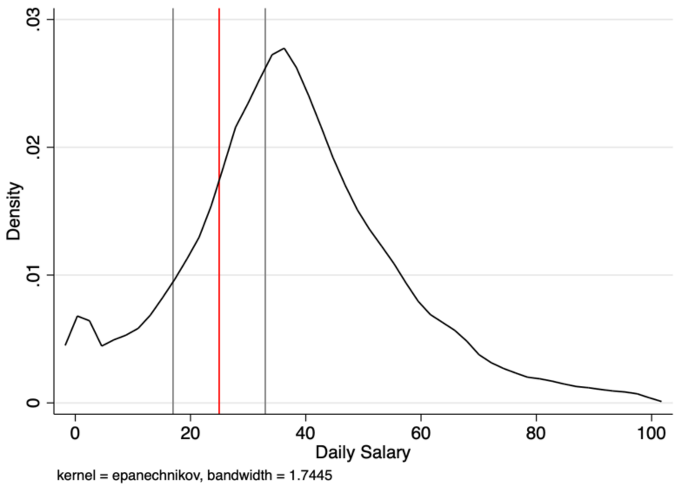
FIGURES & TABLES

Figure 1. Kinks in the reduction in benefits.



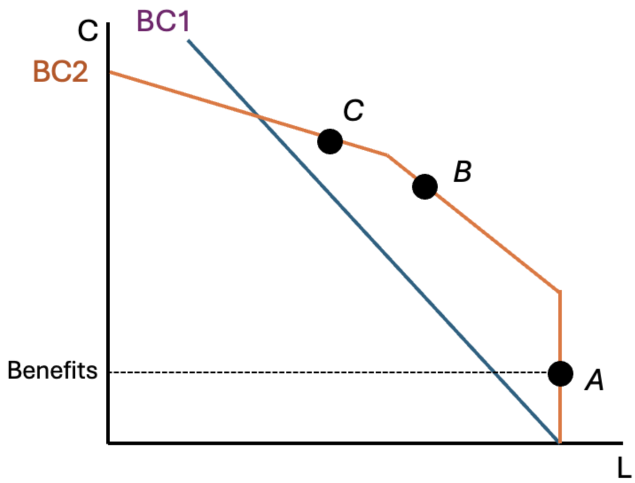
Notes: This graphic simulates the reduction in benefits based on the rules set by the NIHDI, which is the result of applying a marginal tax rate on salary of 20% after crossing the threshold of 15.6068€/day, 50% after the threshold of 24.9709€/day, and 75% from 34.335€/day.

Figure 2. Density distribution of the average daily salary.



Notes: This graphic assess the validity of the assumption 1 in RKD. Blue line represents the kernel density function distribution of the average daily salary. The red line indicates where the kink in the policy rule is located. Grey lines indicate the bandwidth of 8€. The absence of bunching around the kink visually supports the assumption of no sorting.

Figure 3. Representation of the Budget Constraints.

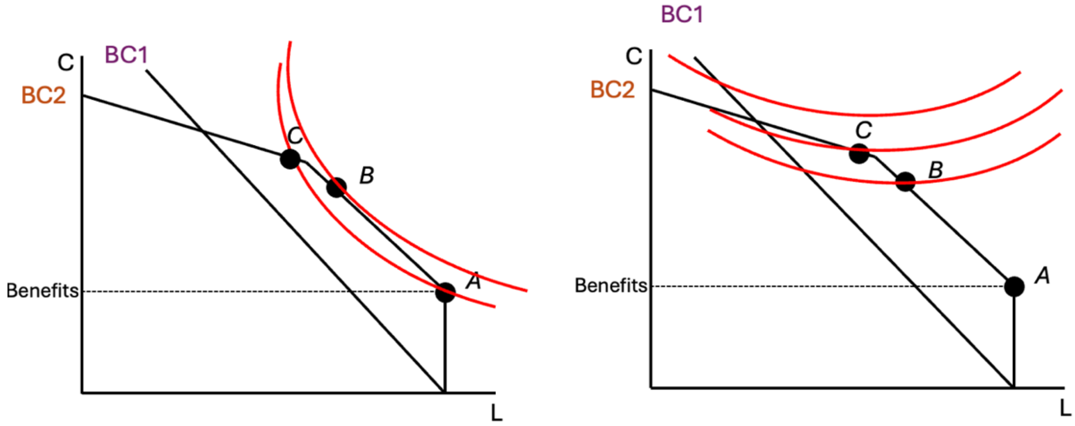


Notes: X-axis reports leisure level, y-axis reports the consumption level affordable at determined income. BC1 represents the budget constraint of being out of the DI program. BC2 represents the budget constraint of being on DI, which includes being completely disabled to work and offering some labour force through the AWP, this line represents the kink where the marginal tax rate increases in 30%. Individual in point A is fully disabled to work and all the affordable consumption is paid with DI benefits. Individual in point B is working through the AWP with the low marginal tax rate. Individual in point C is working through the AWP facing a higher marginal taxation rate.

Figure 4. Representation of agent's preferences

Panel A: Example of agent type I: Stronger preference for leisure

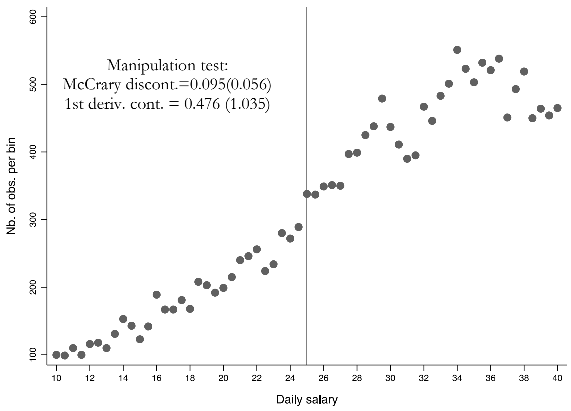
Panel B: Example of agent type II: Stronger preference for consumption



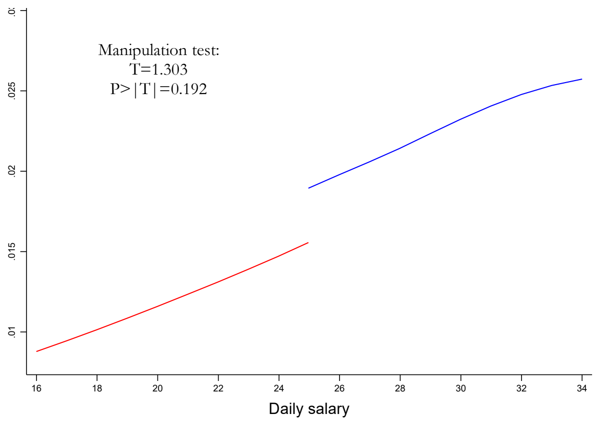
Notes: X-axis reports leisure level, y-axis reports the consumption level affordable at determined income. BC1 represents the budget constraint of being out of the DI program. BC2 represents the budget constraint of being on DI, which includes being completely disabled to work and offering some labor force through the AWP, this line represents the kink where the marginal tax rate increases in 30%. Panel A represents the indifference curves of agent type I, with stricter preferences for leisure. Panel B represents the indifference curves of agent type II, with stricter preferences for consumption.

Figure 5. Manipulation tests.

Panel A: Frequency distribution of assignment variable and McCrary's (2008) test.

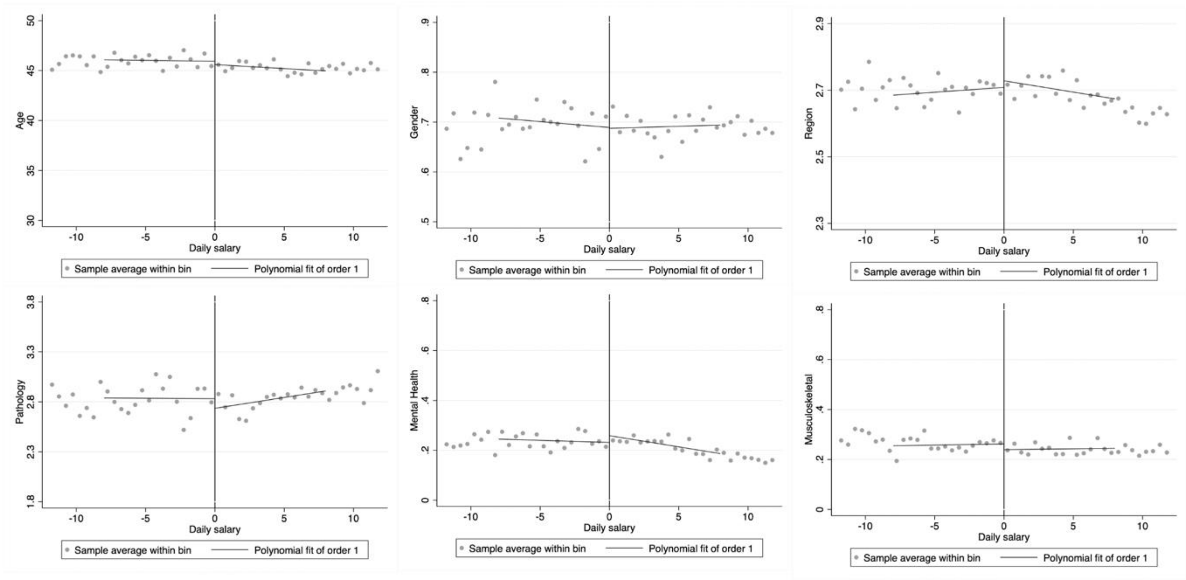


Panel B: Manipulation testing using local-polynomial density estimation.



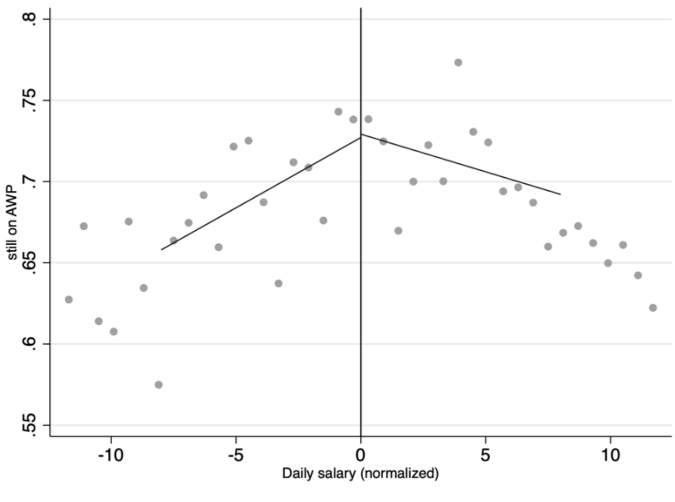
Notes: The graphics assess the validation of RKD assumption 1 of no sorting. Panel A shows the frequency distribution of the daily salary in 50 euros cents bins. The graph also displays two manipulation tests: the standard McCrary test (McCrary, 2008) that checks for a "jump" in the probability density function of the assignment variable, and the extension proposed by Card et al. (2015) to test that the first derivative of the p.d.f. is also continuous at the kink. We report the coefficients for both tests, as well as the corresponding standard errors in parentheses. The McCrary discontinuity test coefficient is 0.095 (0.056), dividing these values gives a t-value of 1.70, similarly, the first derivative continuity test coefficient is 0.476 (1.035), which results in 0.46, indicating in both cases that we do not reject the null hypothesis of continuity. Panel B represent the p.d.f of the daily salary around the kink, estimated using local polynomials, as proposed by Cattaneo et al., (2020), it includes the manipulation T- test coefficient and their corresponding p-value. The p-value > 0.05 indicates that there is not a statistically significant manipulation. Overall, graphical evidence and the formal tests suggest no manipulation of the assignment variable.

Figure 6. Distribution of control variables around the second kink.



Notes: The horizontal axis plots daily salary in euros in bins of around 1 euro, with a bandwidth of 8 euros. Vertical line indicates where the second kink is (24.9709€). The straight lines at both sides of the kink display the underlying linear relationship between the assignment variable and each of the control variables analysed here, they are estimated using local nonparametric regressions. The control variables represented are: age (continuous variable), gender (dichotomic variable), region (categorical variable: Brussels, Wallonia, Flanders), pathology (categorical variable), and a binary variable indicating the prevalence of both mental health and musculoskeletal conditions, which are the two more frequent pathologies in our sample.

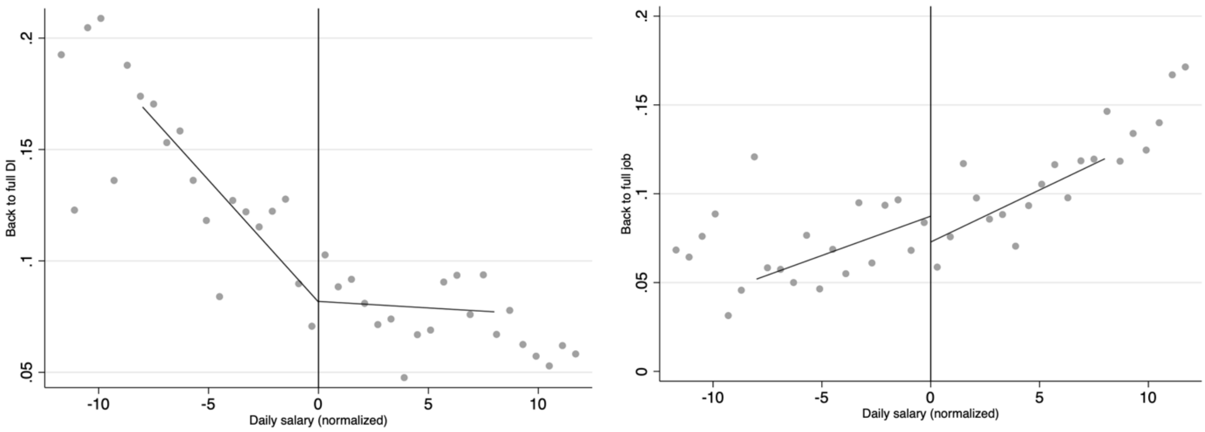
Figure 7. Relationship between the outcome and salary level.



Notes: Horizontal axis plots normalized salary from the adapted work (relative to the kink), using 1 euro per bin and a bandwidth of 8 euros. The vertical axis plots the mean of the outcome variable (the probability to stay in AWP by the end of the year, which is a dummy variable taking value 1 if the individual is still on AWP in December 2013 and 0 otherwise) in each bin. The straight lines display the underlying relationship on each side of the kink. Vertical line indicates the salary kink.

Figure 8. Relationship between the outcome and salary level.

Panel A: Probability of going back to full job Panel B: Probability of going back to full DI



Notes: Horizontal axis plots normalized salary from the adapted work (relative to the kink) in 1-euro bins at each side of the cut-off. The vertical axis plots the mean of the outcome variable (the probability to go back to full job, or the probability to go back to full DI) in each bin. The straight lines display the underlying relationship on each side of the kink. Vertical line indicates the salary kink.

Table 1. Main descriptive statistics for the sample in 2013.

	Full sample Mean (SD)	**Kink Sample Mean (SD)**
Gender (women=1)	0.65 (0.48)	0.69 (0.46)
Age	46.00 (9.09)	45.53 (9.30)
Social State (blue collar=1)	0.54 (0.50)	0.72 (0.45)
Days in AWP	170.24 (108.28)	185.25 (102.35)
Daily salary	39.36 (31.46)	26.34 (4.40)
Daily benefits	36.72 (20.34)	37.06 (15.07)
Observations	36,778	10,136

Notes: The table report the mean value and, in parentheses, the standard deviation of all the salaried individuals in the AWP in 2013. The kinks samples include the same statistics around the three kinks, using a bandwidth of 8 euros.

.

Table 2. Other descriptive statistics for the sample in 2013.

	Full Sample Total (%)	Kink Sample Total (%)
Pathology		
Musculoskeletal disorders	6,995 (30.17%)	2,356 (32.54%)
Mental health	6,289 (27.13%)	2,205 (30.45%)
Cancer	3,123 (13.47%)	681 (9.40%)
Nervous system	1,522 (6.57%)	453 (6.26%)
Circulatory system	1,495 (6.45%)	392 (5.41%)
Trauma and poisoning	1,327 (5.72%)	394 (5.44%)
Other	2,423 (10.54%)	759 (10.48%)
Unknown	8 (0.03%)	1 (0.01%)
Reason of exit AWP		
Back full-time work	5,554 (40.81%)	892 (29.79%)
Back full DI	3,227 (23.71%)	958 (32.00%)
Unemployment	88 (0.65%)	26 (0.87%)
Death	67 (0.49%)	12 (0.40%)
Exclusion	1,411 (10.37%)	396 (13.23%)
Other	1,255 (9.22%)	313 (10.45%)
Unknown	2,007 (14.75%)	397 (13.26%)
Region		
Brussels	1,637 (4.48%)	343 (3.40%)
Wallonia	9,847 (26.98%)	2,347 (23.29%)
Flanders	25,019 (68.54%)	7,388 (73.31%)

Notes: The table reports some additional descriptive statistics for all the salaried individuals in the AWP in 2013. The kinks samples include the same statistics around the three kinks, using a bandwidth of 8 euros. The total value and the percentage of the total sample in parentheses.

Table 3. Regressions Results

	(1)	(2)	(3)	(4)	(5)	(6)
	Remain AWP	Remain AWP	Back to full DI	Back to full DI	Back to full job	Back to full job
Panel A:						
First stage	0.3		0.3		0.3	
Panel B:						
Treatment effects						
β	-0.044***	-0.040**	0.035***	0.033***	0.005	0.005
	(0.014)	(0.014)	(0.010)	(0.010)	(0.008)	(0.008)
Robust CI	[-0.121, 0.015]	[-0.108, 0.004]	[0.018,0.092]	[0.015, 0.089]	[-0.041, 0.022]	[-0.042, 0.021]
Mean	0.705	0.705	0.095	0.095	0.088	0.088
Covariates	NO	YES	NO	YES	NO	YES
Observations	36,778	36,503	36,778	36,503	36,778	36,503
Eff observations	10,137	10,078	10,137	10,078	10,137	10,078

*Notes: all the coefficients reported are calculated from independent local polynomial nonparametric regressions, using a linear trend (polynomial of order 1), a defined symmetric bandwidth of 8 euros and uniform kernel functional form. Panel A reports first stage coefficient, which is the deterministic marginal tax change, 0.3 (0.5-0.2) for the kink located at the salary level of 24.97€, which is the same for all the outcomes. Panel B reports Treatment effects coefficients, the result of equation (7), which is the result of dividing the coefficient estimated by RKD of equation (8) by the deterministic first stage value. The coefficients show the estimated effect of a 1 euro per working day increase in the implicit marginal tax rate on three different outcomes (columns 1 and 2: the probability to stay in AWP by the end of the year; columns 3 and 4: the probability of returning to DI; columns 5 and 6: the probability of resuming full-time employment). Standard errors are in parentheses. wE report robust confidence intervals. The column "mean" reports the average of the dependent variable within the defined bandwidth. Covariates include: region, age, gender and social state. Significance levels: ***p<0.01, **p<0.05, *p<0.1.*

Table 4. Heterogeneous effects

Panel A) By Gender

	Men		Women	
	Back full DI	Back full job	Back full DI	Back full job
Panel A: First stage	0.3		0.3	
Panel B: Treatment effects	(1)	(2)	(3)	(4)
β	0.058*** (0.016)	-0.003 (0.012)	0.024** (0.012)	0.008 (0.010)
Robust CI	[0.045, 0.175]	[-0.078, 0.018]	[-0.011, 0.077]	[-0.044, 0.036]
Mean	0.095	0.058	0.094	0.101
Covariates	NO	NO	NO	NO
Eff. Observations	3,017	3,017	7,025	7,025

Panel B) By Age Group

	Youngers (19-35)		Medians (36-49)		Older (50-59)	
	Back full DI	Back full job	Back full DI	Back full job	Back full DI	Back full job
Panel A: First stage	0.3		0.3		0.3	
Panel B: Treatment effects	(1)	(2)	(3)	(4)		
β	0.050* (0.028)	-0.014 (0.027)	0.026* (0.014)	0.156 (0.012)	0.035** (0.015)	-0.004 (0.011)
Robust CI	[-0.013, 0.199]	[-0.179, 0.027]	[-0.021, 0.087]	[-0.023, 0.069]	[0.002, 0.114]	[-0.069, 0.018]
Mean	0.145	0.157	0.086	0.087	0.082	0.059
Covariates	NO	NO	NO	NO	NO	NO
Eff. Observations	1,724	1,724	4,409	4,409	4,004	4,004

Panel C) By Pathology

	Mental Health		Musculoskeletal	
	Back full DI	Back full job	Back full DI	Back full job
Panel A: First stage	0.3		0.3	
Panel B: Treatment effects	(1)	(2)	(3)	(4)
β	0.049** (0.021)	0.010 (0.009)	0.013 (0.019)	0.020* (0.011)
Robust CI	[-0.023, 0.144]	[-0.030, 0.046]	[-0.044, 0.978]	[0.001, 0.085]
Mean	0.110	0.023	0.091	0.038
Covariates	NO	NO	NO	NO
Eff. Observations	2,301	2,301	2,517	2,517

…

Panel D) By Social Status

	White Collar		Blue Collar	
	Back full DI	Back full job	Back full DI	Back full job
Panel A: First stage	0.3		0.3	
Panel B: Treatment effects	(1)	(2)	(3)	(4)
β	0.030	-0.016	0.033**	0.010
	(0.020)	(0.018)	(0.011)	(0.009)
Robust CI	[-0.024, 0.125]	[-0.135, 0.001]	[0.007, 0.093]	[-0.028, 0.043]
Mean	0.112	0.108	0.087	0.080
Covariates	NO	NO	NO	NO
Eff. Observations	2,872	2,872	7,260	7,260

Panel E) By Type of disability (short- vs. long-term)

	Short-term DI		Long-term DI	
	Back full DI	Back full job	Back full DI	Back full job
Panel A: First stage	0.3		0.3	
Panel B: Treatment effects	(1)	(2)	(3)	(4)
β	0.043**	-0.019	0.026**	0.009*
	(0.020)	(0.026)	(0.011)	(0.005)
Robust CI	[0.001, 0.157]	[-0.187, 0.012]	[-0.013, 0.076]	[-0.008, 0.032]
Mean	0.111	0.263	0.085	0.020
Covariates	NO	NO	NO	NO
Eff. Observations	2,767	2,767	6,231	6,231

*Notes: all the coefficients reported are calculated from independent local polynomial nonparametric regressions, using a linear trend (polynomial of order 1), a defined symmetric bandwidth of 8 euros, and uniform kernel functional form. The column first stage reports the deterministic marginal tax change, 0.3 for the kink, located at 24.97€. Treatment effects coefficient report result of equation (7). The coefficients show the estimated effect of a 1 euro per working day increase in the implicit marginal tax rate on the different outcomes depending on the indicated characteristic. Standard errors are in parentheses. We report robust confidence intervals. Significance levels: ***p<0.01, **p<0.05, *p<0.1.*