Intensive and Extensive Labor Supply Responses of Disability Insurance Recipients

Philippe Ruh, University of Zurich
Stefan Staubli, RAND, University of Zurich, and IZA

October 31, 2013

Preliminary and Incomplete. Comments Welcome.

Abstract

We study the impact of financial work incentives on the labor supply of disability insurance (DI) recipients along the intensive margin (intensity of work on the job) and along the extensive margin (participation into the labor force). We estimate the intensive margin elasticity nonparametrically using administrative data from Austria and a notch in the budget set of DI recipients. This notch creates excess bunching on the low-earnings side and missing mass on the high-earnings side of the cutoff. To identify the extensive labor supply response, we exploit the fact that screening stringency varies across states in Austria allowing us implement a spatial regression discontinuity design. We find evidence for significant behavioral responses along the intensive and the extensive margin. We complement our empirical analysis by developping a theoretical framework that can be used to analyze the implications of our estimates for the optimal design of financial work incentives.

Keywords: Disability insurance, earnings test, benefit notch, bunching

JEL Classification Numbers: J14, J26

1 Introduction

Disability Insurance (DI) programs are among the largest social insurance programs. In OECD countries, total expenditures on disability benefits account for approximately 2.5% of GDP on average (OECD, 2010). DI programs are designed to provide income replacement in the case of a permanent loss of earnings capacity due to poor or deteriorating health, but there have been concerns that DI discourages work. More precisely, in many countries DI recipients lose benefits if earnings exceed a substantial gainful activity (SGA) amount. The benefits loss at the SGA amount induces a high implicit tax on work and creates an incentive for beneficiaries to keep their earnings below the SGA level or to stop working completely.

This paper examines whether and to what extent the implicit tax on work affects the labor supply of DI recipients along the intensive margin (intensity of work on the job) and along the extensive margin (participation into the labor force). The intensive labor supply response is estimated using a discontinuous changes in the implicit tax on work in the Austria DI program. In Austria, DI recipients can earn up to a SGA threshold of 380 euros per month (around $500) without losing benefits. Above the SGA threshold, individuals lose a certain fraction of benefits, which depends on the sum of DI benefits and earnings. The loss in benefits is substantial in magnitude; the average DI recipients loses 15% of income if he or she earns 1 euro above the SGA threshold. We exploit this discontinuous change in the implicit tax on work (a notch) at the SGA threshold, to examine whether DI recipients bunch below the low-earnings side of the SGA threshold. The amount of excess mass at the SGA threshold can be used to estimate the counterfactual earnings of DI recipients in the absence of the notch in the budget set. Additionaly, by relating the counterfacutal earnings to the change in the implict tax rate at the notch point allows us to recover an estimate for the intensive labor supply elasticity.

To examine how the implicit tax on work affects the extensive labor supply decision, we exploit spatial variation in acceptance rates into the DI program accross different states in Austria. This variation arises because certain DI offices are more stringent in the screening of DI applicants than other offices. Since individuals need to apply for benefits in their state of residence, we can use a spatial regression discontinuity design that compares program entrants in states with relaxed screening with very similar applicants who were denied benefits because they live in states with a strict screening policy. The idea behind this approach is that labor supply participation rates of denied applicants provide a good counterfactual for the labor supply participation rates of program entrants in the absence of the high implicit tax on work. By relating the difference in labor supply participation rates to the difference in implicit tax rates, we can estimate an exentsive labor supply elasticity.

Our empirical analysis yields the following results. First, the distribution of earnings of DI recipients exhibits significant "bunching" at earnings levels just below the SGA threshold. Using an approach similar to Kleven and Waseem (2013) to infer behavioral parameters from notches, we estimate that the excess number of DI recipients at the notch point equals the total number that should be observed with earnings up to 180 euros above the notch, suggesting that average earnings would be up to 60% larger in the absence of the notch. Relating the earnings response to the change in the implicit tax rate implies an intensive labor supply elasticity of around 0.17, which is similar to previous estimates by Gelber et al. (2013) for Social Security claimants in the U.S. and Chetty et al. (2011) using kinks in the Danish tax schedule. Second, we find substantial heterogeneity in the bunching response across different subgroups, with women and younger age groups exhibiting more excess mass at the notch point than men and older age groups. Third, we find significant discontinuities in acceptance rates to the DI program at borders between strict and lenient DI screening states. The average applicant is 9.1 percentage points more likely to be accpted to the program if he or she lives in a lenient state as opposed to a strict state. We also find significant discontinuities in the labor force participation rates at state borders. More specifically, labor supply participation rates are 2.5 percentage points higher in lenient states relative to strict states.

We think that our paper is of general interest for several reasons. First, while differing in the exact amount, DI programs in many countries specify an SGA amount above which DI recipients lose some or all of their benefits. For example, in the U.S. DI recipients lose all their benefits if monthly earnings exceed $1,040 ($1,740 for blind recipients) over a longer time period. We therefore believe that our paper sheds light on an important work disincentive that is present in many DI programs. Second, many countries consider implementing reforms that encourage return-to-work of DI beneficiaries by relaxing the strict earnings restrictions. The effectiveness of these policy measures depends on the responsiveness of DI beneficiaries labor supply to financial incentives, but research on this topic has been hampered by the lack of quasi-experimental variation in financial incentives. This project circumvents this problem by exploiting exogenous spatial and institutional discontinuities in Austria.

Third, Saez (2002) shows that optimal income transfers for low income individuals depend on the intensive and extensive labor supply elasticity. Using his framework, our estimates for the intensive and extensive labor supply responses can be used to shed light on the optimal design of return-to-work policies for DI recipients. Fourth, we rely on exceptionally rich administrative data from the Austrian Social Security Administration database (ASSD), containing the complete labor market history of all employees in Austria over the period 1972 to 2012. The data can be linked with tax register data, which allow us to precisely measure earnings and DI benefits.

The previous literature has relied on kinks in the budget constraint to estimate labor supply elasticities with respect to the tax rate (Saez (2010);Chetty et al. (2011)), while the current project exploits a notch in the budget constraint. As discussed in Kleven and Waseem (2013), one difference between notches and kinks is that notches, by introducing a discrete jump in tax liability, may create an extensive labor supply response, while kinks only affect labor supply through the intensive margin. A second difference is that earnings elasticities from notches will be a mix of compensated and uncompensated elasticities, while earnings elasticities from kinks are compensated elasticities because there is no income effect. This paper is also related to the growing literature that examines the work disincentive effects of DI programs (e.g., Bound (1989); von Wachter et al. (2011); Maestas et al. (2013)). Most of this literature focuses on the impact of the DI program on labor force participation, while ignoring responses along the intensive margin. This paper helsp to fill this gap by exploiting the notch in the budget set at the SGA thereshold, which exists in many DI programs in industrialized countries.

This paper proceeds as follows. Section 2 describes Austria's DI program. Section 3 outlines our identification strategies to estimate intensive and extensive labor supply responses of DI recipients using bunching at the SGA threshold and variation in screeing stringency across regional DI offices. Section 4 summarizes the data and presents descriptive statistics. Section 5 presents the empirical results. Section 6 summarizes and draws policy conclusions.

2 Institutional Background

2.1 The Austrian DI Program

Austria's DI program covers all active labor market participants and provides partial earnings replacement to workers below the full retirement age who are unable to engage in substantial gainful activity because of a medically determinable health impairment that will last for at least six months.

To apply for DI benefits, an individual must submit an application to the DI office of the state of residence. Austria has nine states and hence there are nine such offices. Employees at the DI office first check the non-medical eligiblity criteria for DI benefits. Only individuals who have contributed to the program for at least 5 years in the past 10 years and are not yet eligible for retirement benefits can apply for DI benefits. In contrast to the U.S. system, non-employment is not a precondition for applying for DI benefits. If an applicant meets the nonmedical criteria, a team of disability examiners and physicians assesses the applicant's overall ability to work and the medical severity of the applicant's disability. A disability award is made if the medical examination finds that a medically determinable impairment causes more than 50% of a reduction in ability to work relative to that of a healthy person with comparable education.[1] If the health impairment is expected to be temporary, DI benefits are granted for a limited time period of typically two years. DI benefits are awarded for an indefinite time period in case of permanent health impairments. Applicants who disagree with the decision of the DI office can appeal within 3 months and have several levels of appeal.

Once DI benefits are awarded, there are three main pahtways out of the program. First, DI claimants may no longer meet the medical or non-medical eligiblity criteria for disability benefits. For example, the health status may improve such that the DI recipient is no longer disabled. In 2011, medical improvements and return to work accounted for roughly 2 percent of program exits. Second, DI claimants may reach the full retirement age, at which point they can ask to be transferred to the old-age pension program. However, few beneciaries do so because in most cases the corresponding old-age pension would be lower than the disability pension.[2] Third, DI recipients may die, which accounted for XXX percent of program exits in 2011.

[1] Medical criteria for disability classification are relaxed starting at age 57. See Staubli (2011) for the impact of this relaxation on labor force participation of older workers.

[2] Only xxxx percent of beneficiaries who reached the full retirement age in xxxx transferred to the old-age pension program.

DI benefits are subject to income taxation and mandatory health insurance contributions. The replacement rate after income and payroll taxes is on average 75% of the pre-disability net earnings up to a maximum of approximately 2,800 euros per month (around $3,500). The level of benefits depends on a assessment basis and a pension coefficient. The assessment basis corresponds to the average earnings over the best 20 years after applying a cap to earnings in each year. The pension coefficient is the percentage of the assessment basis that is received in the pension. The pension coefficient increases with the number of contribution years up to a maximum of 80% (roughly 45 contribution years). Applicants under age 60 qualify for a special increment if their pension coefficient is below 60%.

2.2 The Notch at the Substantial Gainful Activity Amount

Similar to the U.S. Social Security DI program, disability beneficiaries in Austria can earn up to a Substantial Gainful Acitivity (SGA) amount whitout loosing any benefits. In 2013, the monthly SGA amount was 397 euros (around $530 using an exchange rate of $1.37 per euro), which is about half of the SGA amount for non-blind DI recipients in the U.S.[3] However, in contrast to the U.S. program, DI recipients will only lose part of their benefits if their earnings exceed the SGA threshold. The amount of the reduction DI benefits in a given month depends on the sum of DI benefits (b) and earnings (w) in that month as follows:

$$\triangle b = \begin{cases} 0 & \text{if } w + b \leq K_1 \\ 0.3(w + b - K_1) & \text{if } K_1 < w + b \leq K_2 \\ 0.3(K_2 - K_1) - 0.4(w + b - K_2) & \text{if } K_2 < w + b \leq K_3 \\ 0.3(K_2 - K_1) + 0.4(K_3 - K_2) + 0.5(w + b - K_3) & \text{if } w + b > K_3, \end{cases} \tag{1}$$

where the reduction in benefits $\triangle b$ cannot exceed $\min(0.5b, w)$. The values K_1, K_2, and K_3 are adjusted each year to account for inflation; in 2013 the corresponding values were 1,108, 1,662, and 2,216 euros. In addition to loosing part of the DI benefits, there is also a change in the average payroll tax rate on earnings at the SGA threshold from 4% to around 11.5%. Together, these rules create a discrete change in the implicit tax on work – a notch – at the SGA threshold. For example,

[3] In the U.S. in 2013 the SGA amount was $1,040 for non-blind DI recipients and $1,740 for blind DI recipients.

the average DI recipient with 1,000 euros monthly disability benefits loses 141 euros in income on the first euro of earnings above the SGA threshold. Thus, the SGA notch creates an incentive for DI recipients to bunch just below the notch point in order to avoid the high implicit tax on work on the high-earnings side of the SGA threshold. Our empirical approach uses the amount of bunching at the notch point to estimate the counterfactual earnings in the absence of the benefit notch. Relating the counterfactual earnings to the change in the implicit tax at the notch point allows us to calculate an intensive labor supply elasticity for DI recipients.

3 Theory and Empirical Methodology

3.1 Empirical Strategy for Estimating the Intensive Margin Response

Our approach to exmine the responsivness of DI recipients to the notch in the budget constraint at the SGA level is based on the framework by Kleven and Waseem (2013) who study the intensive labor supply response to notches in the Pakistani income tax schedule. The idea behind the approach is that a notch creates excess bunching on the low-earnings side and missing mass on high-earnings side of the cutoff. The amount of excess mass at the notch point is an indicator for the responsiveness of DI recipients to financial incentives along the intensive margin.

Figure 1 illustrates the empirical approach graphically. Earnings below the SGA level z^* are taxed with a marginal tax rate t which is determined by the income tax schedule. Gross earnings net of payroll taxes are subject to the income tax with marginal tax rates in the different tax brackets of 0%, 21%, 31%, 41% and 50%. There is a discontinuous change in tax liability if the earnings exceed the SGA threshold, which comes from two sources. First, DI recipients lose a fraction of their DI benefits as illustrated in equation (1). Second, there is also a change in the the average tax rate on earnings because individuals have to start paying social security payroll taxes (including health insurance contributions) on *all* earnings.[4]

The notch in the budget set at z^* induces some DI recipients who in the absence of the notch would have earned more than z^* to instead "bunch" at z. By relating the bunching mass at z^* to the change in the implicit tax on work allows us to calculate

[4]It is possible that the average tax rate may decline for some DI recipients at the notch because they fall into a lower income tax bracket. However, because tax brackets in Austria are very wide, this effect is not very important.

an intensive labor supply elasticity. Because the level of DI benefits depends on the past contribution and earnings history, the change in tax liability at the notch varies across pension recipients. Assume that the elasticity is homogeneous across DI recipients, then the amount of bunching is defined by:

$$B = \int_b \int_{z^*}^{z^* + \triangle z_b^*} \tilde{h_0}(z,b) dz db \approx h_0(z^*) E[\triangle z_b^*], \tag{2}$$

where $(z^*, z^* + \triangle z_b^*)$ is the bunching segment associated with DI benefits b, $\tilde{h_0}(z,b)$ is the joint distribution of earnings and pension benefits in the baseline without a notch, and $h_0(z) = \int_b \tilde{h_0}(z,b) db$ is the unconditional earnings distribution in the baseline. Hence, with heterogeneity in DI benefits bunching can be used to estimate the average earnings response $E[\triangle z_b^*]$.

Calculating the excess mass B requires a prediction for the counterfactualy earnings distribution in the absence of the SGA notch. To do so, we follow the approach by Chetty et al. (2011) and Kleven and Waseem (2013). More specifically, we group individuals into earnings bins of 10 euros indexed by j and run the following regression:

$$c_j = \sum_{i=0}^{p} \beta_j (z_j)^i + \sum_{i=z_L}^{z_U} \gamma_i I(z_j = i) + \varepsilon_j, \tag{3}$$

where c_j is the number of individuals in bin j, z_j is the earnings level in bin j, and p is the order of the polynomial. The counterfactual earnings distribution is calculated as the predicted values from equation (3) omitting the contribution of the dummies in the excluded range ($\hat{c}_j = \sum_{i=0}^{p} \hat{\beta}_j (z_j)^i$). Excess mass is estimated as the difference between the observed and counterfactual earnings distribution in the range $[z_L, z^*]$ where $z_L$is determined by visual inspection. To determine the upper bound z_U, we exploit the fact that missing mass created by the bunching response $\hat{M} = \sum_{j>z^*}^{z^U} (c_j - \hat{c}_j)$ must be equal to the bunching mass at the SGA threshold $\hat{B} = \sum_{j=z_l}^{z^*} (c_j - \hat{c}_j)$. Thus, we choose z_U such that $\hat{B} = \hat{M}$.

Given an estimate for the bunching mass B and the counterfactual density h_0 in the absence of the notch, we can calculate an elasticity of earnings with respect to the implicit net-of-tax rate as follows:

$$e = \frac{\triangle z^* / z^*}{\triangle t^* / (1 - t^*)} = \frac{\hat{B} / \hat{h}_0(z^*) z^*}{\triangle t^* / (1 - t^*)}, \tag{4}$$

where $\triangle t^*$denotes the change in the implicit marginal tax rate from t to t^* when the notch is approximated as a hypothetical kink (see Figure 1). We calculate standard errors using a bootstrap procedure in which we generate a large number of earnings distributions (and associated estimates of each variable) by random resampling of individuals from the population with replacement. The standard error of each variable is defined as the standard deviation in the distribution of estimates of the given variable.

3.2 Empirical Strategy for Estimating the Extensive Margin Response

Our approach to exmine the impact of work disincentives on the responsivness of DI recipients along the extensive margin exploits regional variation in screening stringency of applications for disability benefits. Each state in Austria has its own DI office which is responsible to examines applications of individuals with residence in that state. There are nine such offices in total and among them acceptance rates vary by as much as 20%. To the extent that this variation is not correlated with observable and unobservable characteristics of applicants, it can be used to examine how program entry affects labor force participation of applicants. More specifically, our identification strategy is a spatial regression discontinuity design that exploits the discontinuous changes in the screening stringency at the border between lenient and strict screening states. This jump in screening stringency implies that applicants who are very similar in observable characteristics have very different acceptance rates depending on the screening stringency of the state of residence.

This comparison can be implemented by estimating regressions of the following type:

$$y_i = \alpha + \beta D_i + \gamma_0(1 - D_i)f(Z_i) + \gamma_1 D_i f(Z_i) + X_i'\delta + \varepsilon_i \tag{5}$$

where i denotes individual; D_i is a dummy that is equal to 1 if an applicant's place of residence is in a strict state and 0 if the place of residence is in a lenient state, Z_i denotes the driving distance (in minutes) to the border between strict and lenient screening states, and f is a function of the driving distance to the border between strict and lenient screening states. The vector X_i contains individual and region-specific characteristics. The coefficient of interest is β which measures the impact of living in a lenient screening state relative to a strict screening state on the outcome variable y_i. We will first examine the difference in the acceptance rates between lenient and strict screening states. In this case the outcome variable is an indicator that is 1 if an applicant got awarded benefits and 0 otherwise. In a second step we will estimate the impact of being awarded benefits on labor force participation rates. Here the outcome variable y_i is an indicator for whether individual i is doing any work after applying for benefits. By relating the labor force participation response to the change in the implicit tax on work gives us an estimate for the extensive labor supply response of DI recipients with respect to the implicit tax rate.

The first key identifying assumption underlying our spatial regression discontinuity approach is that applicants on either side of the geographical border are screened differently but are otherwise comparable in all observable (and unobservable) characteristics. To shed light on this assumption, we will examine whether individual characteristcs such as age, last wage, and past work experience are smooth around the state border. The second key identifiying assumption is that applicants do not manipulate the location of residence. In our case this assumption would violated if, for example, individuals moved to states with a lenient screening policy before applying for benefits. Since we have several years of data prior to the application date, we can shed light on the validity of this assumption by examining whether applicants tend to move between states before submitting an application. We will also run a series of placebo tests in which we will estimate equation (5) for states that are very similar in their screening stringency.

4 Data and Descriptive Evidence

4.1 Data and Sample Selection

We combine register data from three different sources. First, the Austrian Social Security Database (ASSD) contains very detailed longitudinal information for all private sector workers and self employed in Austria between 1972 and 2012. At the individual level the data include gender, nationality, month and year of birth, blue-collar or white-collar status, and labor market history. Labor histories are summarized in spells. Specifically, all employment, unemployment, disability, sick leave, and retirement spells are recorded. The data contain several firm-specific variables: geographical location, industry affiliation, and firm identifiers that allow us to link both individuals and firms. Second, we use administrative tax records for the years 1994 to 2012 which contains detailed information on benefits from the various social insurance programs, earnings, and other sources of income that are not subject to payroll taxes. These data allow us to precisely measure earnings which is crucial to capture bunching at the SGA threshold. Also, the data contain information on the community of residence, which is crucial for our spatial regression discontinuity design. Third, we have information on all applications for DI benefits between 2004 and 2010. These data contain very detailed information whether the application was accepted or rejected, the application date, the decision date, the reason for a rejection, and the underlying health impairment (for the years 2008 to 2010). The tax records and the applications data can be linked with the ASSD via an identifyer variable.

We construct two analysis samples; one to examine bunching at the SGA level and one to examine spatial variation in DI acceptance rates. The sample to examine bunching consists of all DI spells that started in the years 2001 to 2012 by workers who are between age 20 and 56 at the start of the DI spell. We exclude spells that started prior to 2001 because earnings restrictions were not uniformly regulated for these spells. We exclude individuals who start claiming benefits after age 57 because they face stricter earnings restrictions. More specifically, they lose all DI benefits if earnings exceed the SGA level and they are not allowed to work in the same occupation as before the disability onset. We observe individuals in the sample at a monthly frequency. The analysis sample to examine the extensive labor supply response consists of all individuals in the age group 20 to 56 who applied for benefits between 2004 and 2010. Given the long time span covered in the ASSD, we can follow these individuals up to 40 years prior and 8 years after their application date.

Table 1 provides summary statistics for our analysis samples. Column 1 shows summary statistics for all individuals in the bunching sample, while column 2 shows summary statistics for those individuals who "bunch" at the SGA level. A comparison across the two columns shows that both groups are very similar in their background characteristics. Columns 3 and 4 provide summary statistics for accepted and rejected DI applicants. There are more apparent differences visible between these two groups. In particular, accepted applicants earn higher wages prior to DI entry than rejected applicants, which is consistent with the findings from previous studies (see, e.g., Maestas et al. (2013)).

Table 1

5 Empirical Results

5.1 Intensive Labor Supply Response

We start our analysis by examining graphically whether there is evidence for bunching at the SGA level. To do so, we pool all available years of data and calculate the difference between earnings and the SGA level in the relevant year. We then group individuals into 10 euro bins and plot the number of individuals per bin in the bins around the SGA level, as shown in Figure 2. The solid black line beneath the empirical distribution is a sixth-degree polynomial fitted to the empirical distribution using equation (3). In the Appendix we show that the counterfactual earnings distribution is not very sensitive to the choice of the degree of the polynomial.

Several things can be observed from the figure. First, there is large and sharp bunching at the SGA threshold. We estimate that excess bunching is 12.9 times the height of the average counterfactual distribution and this estimate is strongly significant. Second, the earnings distribution exhibits significant missing mass given that earnings fall discretely above the notch. Third, there are no holes in the earnings distribution as the distribution of earnings is relatively flat above the notch, suggesting that optimization frictions might be important.

Figure 2

The SGA level is adjusted each year to account for inflation and real wage growth. It is therfore interesting to examine whether the excess mass follows the movement of the SGA level. Figure 3 plots the distribution of earnings for the years 2002, 2004, 2006, 2008, 2010, and 2012. The vertical red line denotes the corresponding SGA level in a given year, which adjuste each year by around 10 euros. Clearly, the excess mass follows the movement of the SGA threshold very closesly.

Figure 3

Next, we investigate how the excess mass at the SGA threshold differs across subgroups of the population. Figure 4 shows that the amount of bunching is relatively similar for men (b=14.42) and for women (b=13.45). Notice that this does not necessarily imply that men and women are equally responsive to financial incentives given that level of DI benefits varies by gender and hence the implicit tax on work is potentially quite different. Figure 5 plots the distribution of earnings for three the age groups 20-34, 35-44, and 45-54. For each age group the distribution of earnings exhibits very large and sharp bunching at the SGA level. The estimated excess mass is very similar, but as before the responsiveness to financial incentives might be different given that the implicit tax rate may vary across age groups.

Figure 4

Figure 5

In the next step we translate the observed excess mass at the SGA threshold into a counterfactual earnings estimate in the absence of the notch using two different approaches. First, we assume that there are no adjustment frictions such that addition earnings in the absence of the notch $\triangle z^*$ is given by $\hat{B}/\hat{h}_0(z^*)z^*$. The second approach accounts for adjustment frictions following the approach by Kleven and Waseem (2013). More specifically, we calculate the number of individuals who are located in the range $[z^*, z^* + \triangle z^*]$ and whose total income is lower than the hypothetical total income at the notch point. We divide this number by the total number of individuals under the counterfactual density in the range $[z^*, z^* + \triangle z^*]$ to get an estimate for the share of constrained individuals $\hat{a}$. Dividing the bunching estimate $\hat{B}$ by the fraction of constrained individuals $\hat{a}$ yields an estimate for the counterfactual earnings that is not attenuated by adjustment frictions. Table 5 shows the results for the full sample and for different subgroups using a polynomial of degree six for the counterfactual earnings distribution.

Table 5

5.2 Extensive Labor Supply Response

In this section we examine the impact of work disincentive effects using a spatial regression discontinuity design that exploits variation in acceptance rates across local DI offices. We start our analysis by examining graphically whether there is a discontinuous jump in acceptance rates between strict and lenient screening states. For this exercise we focus on the border between the states Lower-Austria and Burgenland for two reasons: First, Burgenland is one of the most lenient screening states with a mean acceptance rate of almost 50% while Lower-Austria is rather strict with an acceptance rate of only 38% on average. Second, Burgenland and

Lower-Austria share a long common border with a relatively large population living on either side of the border.[5]

Table 6

Figure 6 plots acceptances rates in 5 minute intervals from the border between Lower-Austria and Burgenland (as indicated by a red line). Acceptance rates trend relatively smoothly on the left side (Lower-Austria) and the right side (Burgenland) of the border. However, there is a discontinuous jump right at the border of around 9 percentage points, providing clear evidence for differences in screening stringency of applicants. Does this discontinuity in program entry also translate into different labor force participation rates on either side of the border? Figure provides an answer to this question by examining labor force participation rates of applicants on either side of the border. There is a small jump in labor force participation rates of -2.9 percentage points.

Table 7

As discussed in section 3.2, the spatial regression discontinuity design provides causal estimates for the extensive labor supply response of disability programs if applicants for benefits are comparable in observable and unobservable characteristics on either side of the border. To examine the validity of this assumption, we examine whether the distribution of important covariates such as age and last wage are smooth around the border threshold. Figure 8 shows that there are no significant jumps for these covariates at the border threshold, supporting the validity of the identification assumption.

Table 8

6 Conclusion

In this paper, we examine how work disincentive in disabiliy insurance programs affect the labor supply of benefit recipients along the intensive margin (intensity of work on the job) and extensive margin (participation into the labor force). To identify the intensive labor supply response, we exploit a notch in the budget set

[5]Many borders between states in Austria are not useful for this exercise given that only few people live near the border due to the Alps.

of DI recipients in Austria at the substantial gainful activity (SGA) amount. Many countries specifiy a SGA level in their DI program and if earnings exceed the SGA amount for an extended period of time DI recipients lose typically part or all of their benefits. This discontinuous change in the implicit tax rate – a notch – at the SGA level creates excess bunching on the low-earnings side and missing mass on the high-earnings side of the cutoff. To identify the extensive labor supply response, we exploit the fact that screening stringency varies across states in Austria allowing us to implement a spatial regression discontinuity design. More specifically, we compare labor force participation rates of accepted DI applicants who are subject to the high implicit tax on work with labor force participation rates of similar applicants who got denied benefits because their application was screened by a very strict DI office.

We find evidence for large and sharp bunching at the SGA earnings threshold, suggesting that many DI recipients would earn considerably more in the absence of the notch at the SGA level. Our estimate suggests that the excess number of DI recipients at the notch point equals the total number that should be observed with earnings up to 150 euros or 50% more than the SGA level. The bunching response is heterogeneous across different subgroups of the population, with women and younger age groups being more responsive than men and older age groups. These results suggest that the SGA threshold in DI programs is a major work disincentive that reduces labor supply of DI beneficiaries at the intensive margin.

Our empirical strategy to examine the impact of the DI program on labor force participation rates shows that there are significant discontinuities in acceptance rates to the program at borders between strict and lenient disability screening states. The average applicant is 10 percentage points more likely to be accepted into the program if he or she lives in a lenient state as opposed to a strict state. We also find significant discontinuities in the labor force participation rates at state borders. More specifically, labor supply participation rates are xxx percentage points higher in lenient states relative to strict states, implying an extensive labor supply elasticity of xxx.

Our empirical estimates help to shed light on the optimal design in financial work incentives for DI beneficiaries. From a policy perspective, our study suggests that policy reforms aiming at increasing labor supply of DI recipients via financial incentives may be more successful if they target the intensive margin rather than the extensive margin. However, ths policy conclusion depends crucially on the impact of financial incentives on program entry. By allowing DI recipients to work more without loosing benefits may induce more people to seek and ultimately receive benefits. There is little evidence on the impact of financial incentives on program entry and more research on this topic would be extremely valuable.

References

Bound, John. 1989. "The Health and Earnings of Rejected Disability Insurance Applicants." *American Economic Review*, *79* (3): 482–503. 1

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics*, *126* (2): 749–804. 1, 3.1

Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2013. "Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test." *NBENBER Working Paper. 19491*: 1–91. 1

Kleven, Henrik J. and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics*, *128* (2): 669–723. 1, 3.1, 3.1, 5.1

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review*, *103* (5): 1797–1829. 1, 4.1

OECD. 2010. *Sickness, Disability and Work: Breaking the Barriers.* OECD Publishing. 1

Saez, Emmanuel. 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *Quarterly Journal of Economics*, *117* (3): 1039–1073. 1

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, *2* (August): 180–212. 1

von Wachter, Till, Jae Song, and Joyce Manchester. 2011. "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program." *American Economic Review*, *101* (7): 3308–3329. 1

Table 1: Descriptive Statistics

	Intensive Margin Sample		Extensive Margin Sample	
	Not Working	Working	Rejected	Accepted
Age	47	46	49	51
	(7.56)	(8.37)	(9.16)	(9.09)
Last Wage	48.3	56.5	31.7*	41.5*
	(38.7)	(16.8)	(23.4)	(27.0)
Insurance Years	22.9	24.5		
	(9.52)	(9.07)		
UI Duration	1.28	.890		
	(1.54)	(1.26)		
Experience	9.05	11.3		
	(4.79)	(3.90)		
Blue Collar	.676	.723	0.725	0.692
	(0.47)	(0.45)	(0.446)	(0.462)
Self-Employed	.102	.062		
	(0.30)	(0.24)		
Gender (women=1)	0.480	0.445	0.519	0.330
	(0.500)	(0.497)	(0.500)	(0.470)
Individuals	106,464	22,352	121,793	76,815
Observations		412,658		

Notes: The Intensive Margin Sample consists of all individuals starting a disability pension between 2001 and 2012. One observation refers to one individual. All variables are determined at the start of the disability spell. Individuals Working consists of individuals that work at least once during the disability spell. The Extensive Margin Sample consists of all individuals filing a disability claim between 2004 and 2010. Last Wage in the Extensive Margin Sample is the average daily wage in the last twenty years.

Table 2: Intensive labor supply response for full sample and different subgroups

	Excess Mass (b)	Frictions (a)	Δz^*	$\frac{\Delta z^*}{1-a}$
A. Baseline				
Polynomial Order 6	12.88	0.27	129	176
	(0.34)	(0.01)	(3.43)	(3.93)
B. Gender				
Men	14.42	0.04	144	150
	(1.05)	(0.00)	(13.16)	(13.13)
Women	12.48	0.01	125	126
	(0.93)	(0.00)	(4.78)	(4.78)
C. Age Groups				
Below Age 35	11.43	0.03	114	117
	(0.58)	(0.00)	(5.49)	(5.53)
Between Age 35 and 45	17.98	0.07	180	192
	(1.90)	(0.01)	(23.92)	(23.98)
Above Age 45	12.83	0.14	128	150
	(0.83)	(0.01)	(10.27)	(10.43)

Notes: The sample consists of all individuals starting a disability pension between 2001 and 2012. Frequency is monthly and only observations with positive wages are included. Standard errors in parenthesis are obtained using a bootstrap procedure where we sample directly from the population with replacement. The standard deviation of the distribution is shown in brackets. All estimations are provided with a polynomial of order 6.

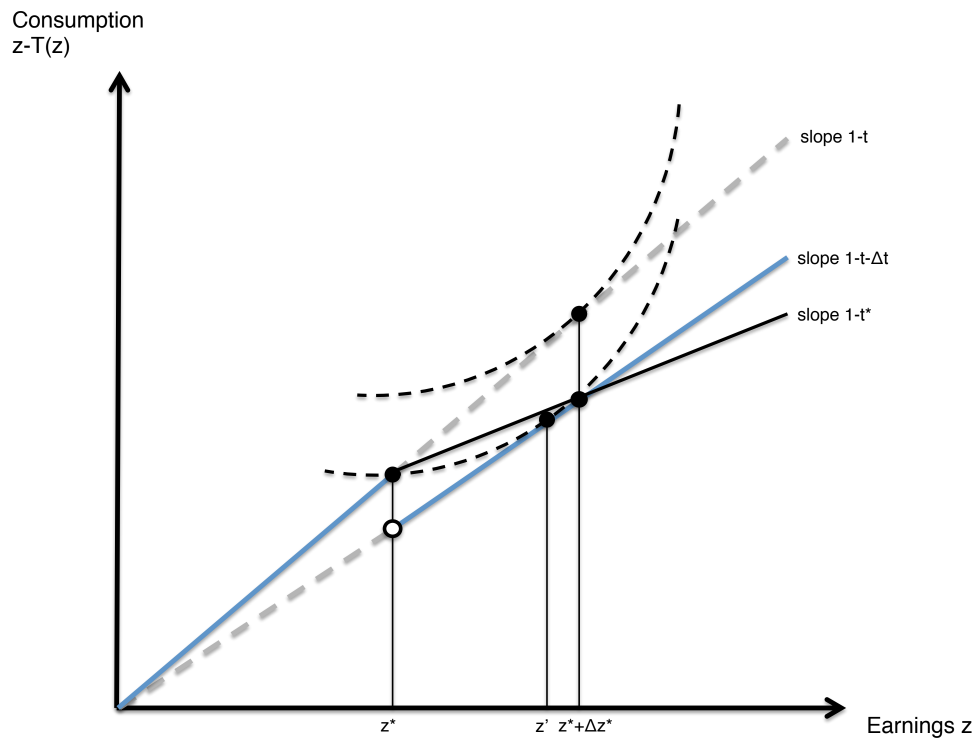

Figure 1: Reduced-form approximateion of intensive labor supply elasticity

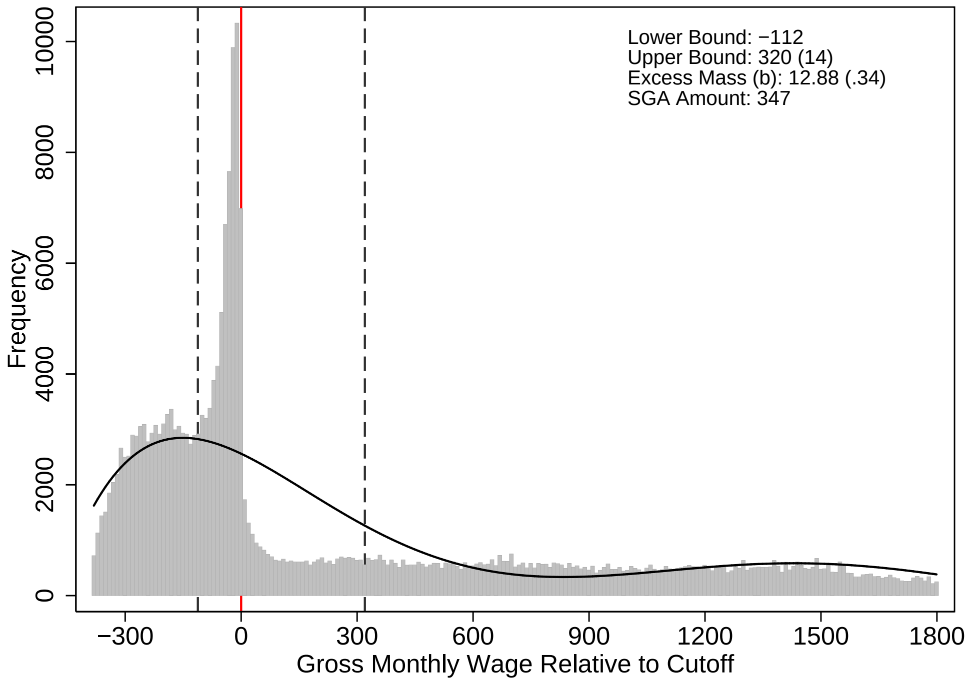

Figure 2: Earnings distribution around the SGA threshold, all years

Notes: This figure shows the earnings distribution around the SGA cutoff (demarcated by the vertical red line at 0) for DI beneficiaries between 2001-2012. The histogram bins are taxable earnings relative to the SGA level in the relevant year. The bin width is 10 euros. The solid line beneath the empirical distribution is a fifth-degree polynomial fitted to the empirical distribution using equation (3).

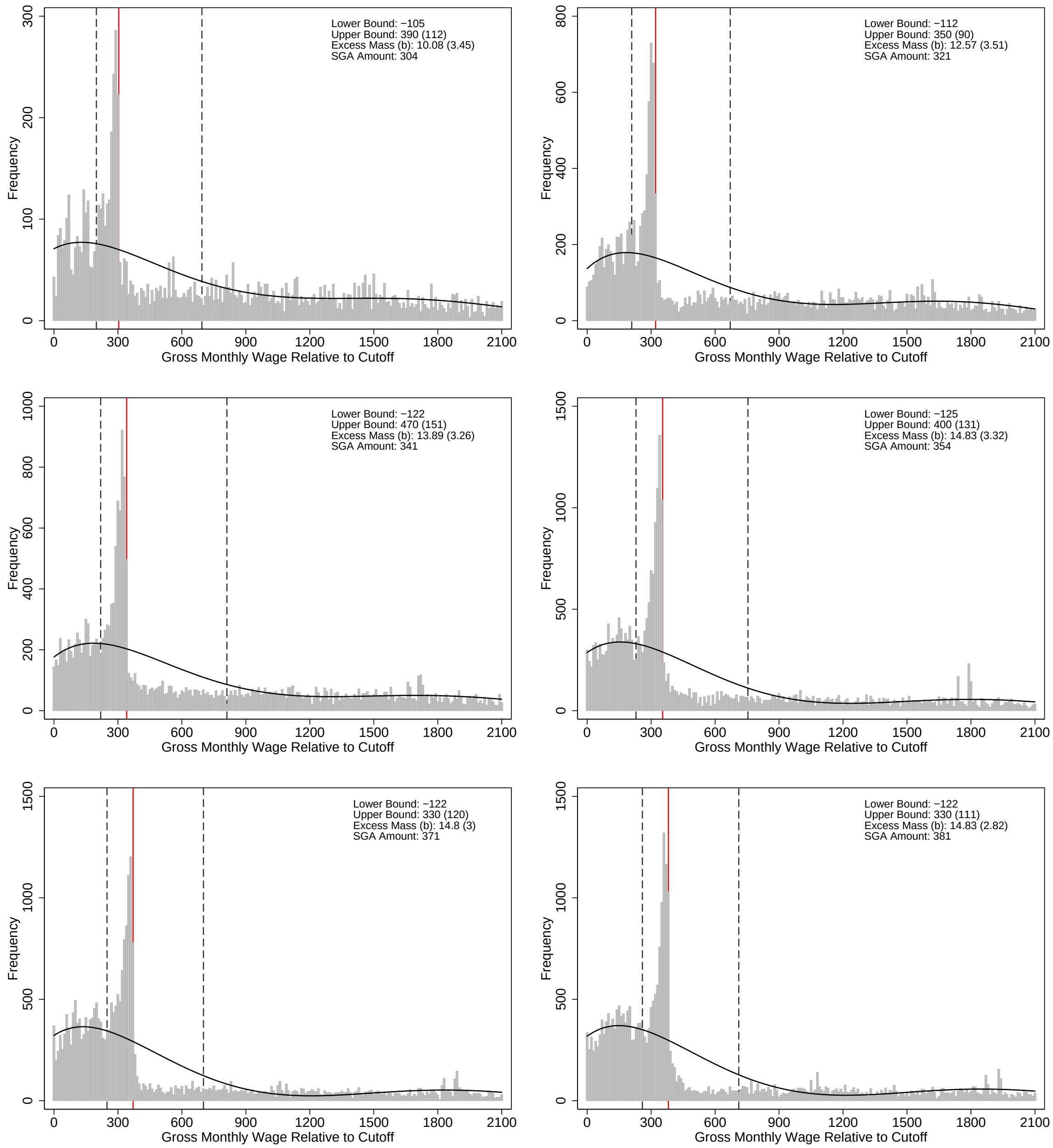

Figure 3: Earnings distribution around the SGA threshold, 2002-2012

Notes: These figures show the earnings distribution around the SGA cutoff (demarcated by the vertical red line at 0) for DI beneficiaries in the years 2002, 2004, 2006, 2008, 2010, and 2012. The histogram bins are taxable earnings relative to the SGA level in the relevant year. The bin width is 10 euros. The solid line beneath the empirical distribution is a sixth-degree polynomial fitted to the empirical distribution using equation (3).

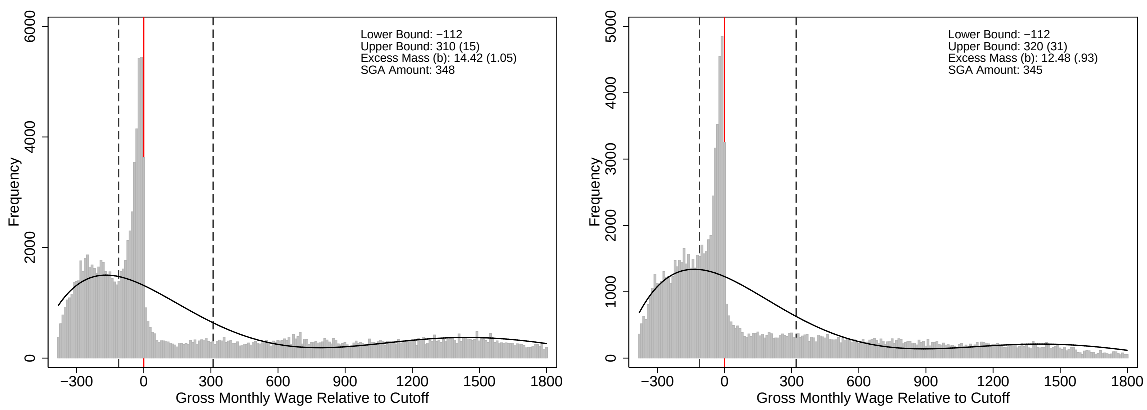

Figure 4: Earnings distribution around the SGA threshold for men (left panel) and women (right panel), all years.

Notes: This figure shows the earnings distribution around the SGA cutoff (demarcated by the vertical red line at 0) for DI beneficiaries between 2001-2012. The histogram bins are taxable earnings relative to the SGA level in the relevant year. The bin width is 10 euros. The solid line beneath the empirical distribution is a sixth-degree polynomial fitted to the empirical distribution using equation (3).

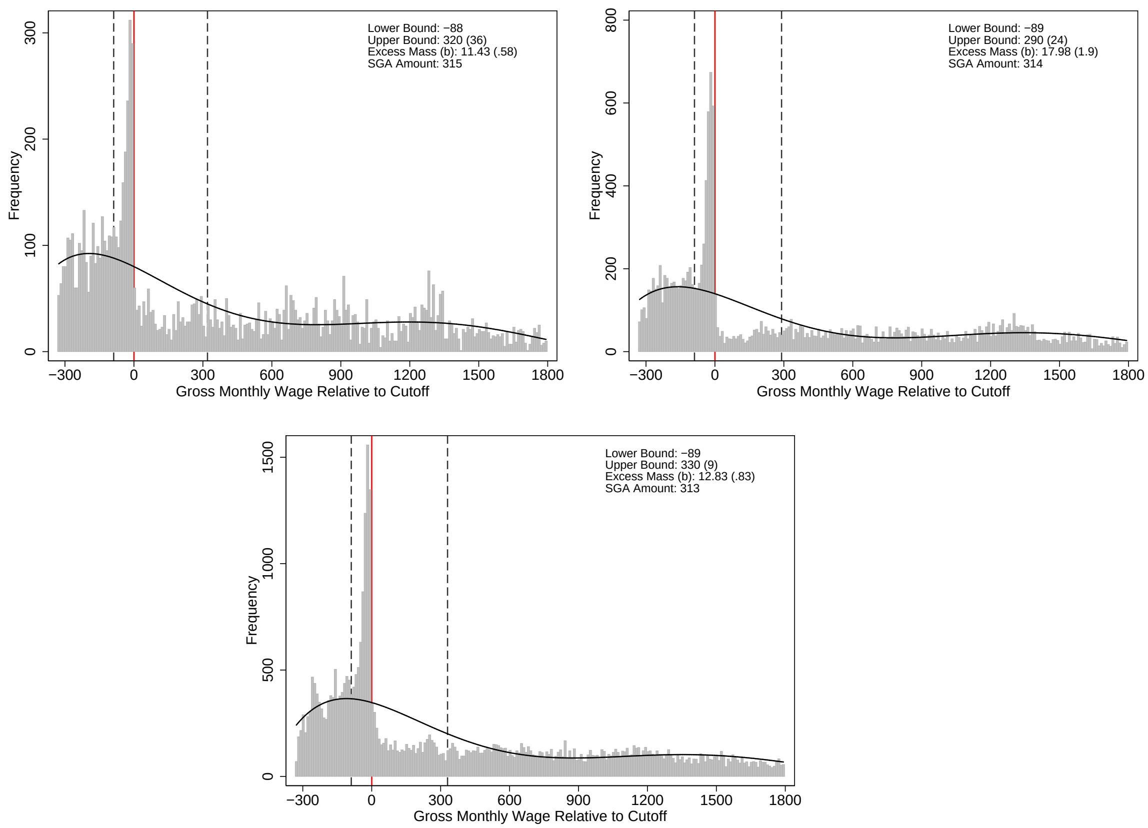

Figure 5: Earnings distribution around the SGA threshold for the age group 20-34 (top left panel), age group 35-44 (top right panel), and age 45-54 (bottom panel).

Notes: These figures show the earnings distribution around the SGA cutoff (demarcated by the vertical red line at 0) for DI beneficiaries between 2001-2012. The histogram bins are taxable earnings relative to the SGA level in the relevant year. The bin width is 10 euros. The solid line beneath the empirical distribution is a sixth-degree polynomial fitted to the empirical distribution using equation (3).

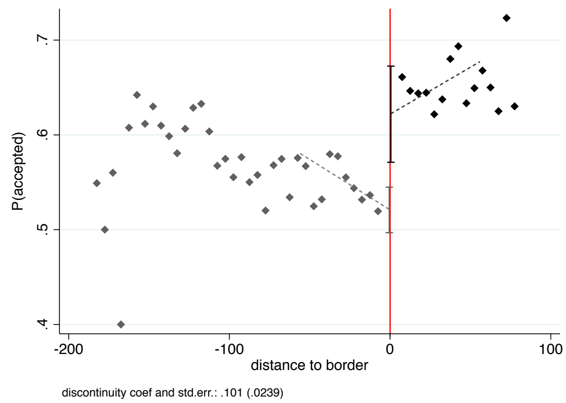

Figure 6: Acceptance rates by distance to border between Lower-Austria and Burgenland

Notes: This figure shows acceptance rates into the DI program by driving distance to the border (5 minute intervals) of Lower Austria (negative distance) and Burgenland (positive distance).

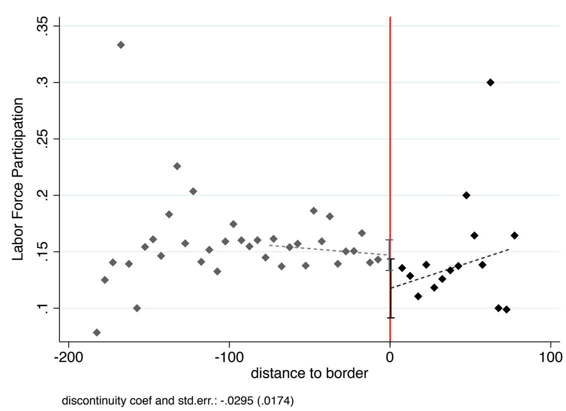

Figure 7: Labor force participation by distance to border between Lower-Austria and Burgenland

Notes: This figure shows the labor force participation rate into the DI program by driving distance to the border (5 minute intervals) of Lower Austria (negative distance) and Burgenland (positive distance).

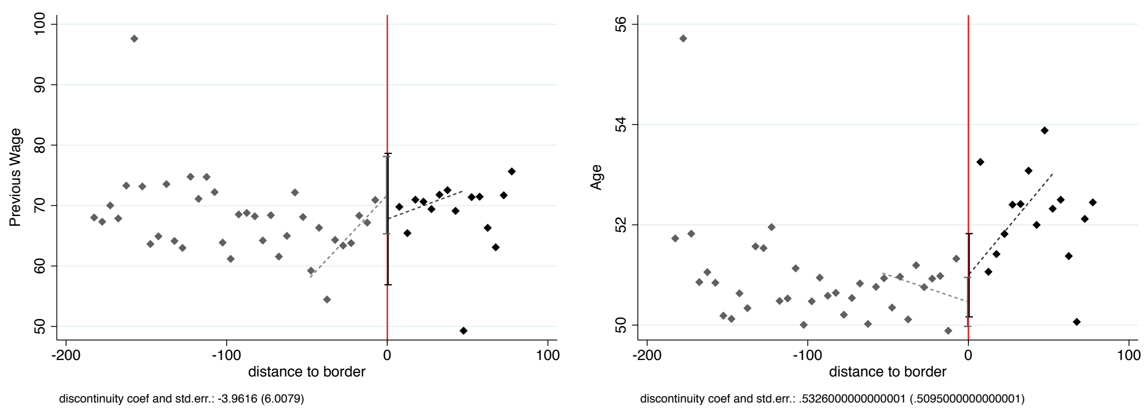

Figure 8: Different covariates by distance to border between Lower-Austria and Burgenland

Notes: This figure shows different covariates (age and wage year prior to application) into the DI program by driving distance to the border (5 minute intervals) of Lower Austria (negative distance) and Burgenland (positive distance).