Intensive and Extensive Margin Labor Supply Responses to Kinks in Disability Insurance Programs

Myhre, Andreas

9 June 2021

Intensive and Extensive Margin Labor Supply Responses to Kinks in Disability Insurance Programs*

Andreas S. Myhre†

This version: June 9, 2021

Abstract: While kinks are prevalent in tax and transfer systems, the fiscal revenue and behavioral responses are not fully understood. In disability insurance (DI) programs, for instance, kinks help balance the moral hazard effects from the induced entry with the provision of work incentives for recipients who regain their ability to work. Using quasi-random variation in kink points in the benefit schedule for Norwegian DI recipients, I identify intensive and extensive margin earnings responses to the implicit tax on earnings as DI benefits are phased out above the kink. To identify the intensive margin responses, I implement a non-parametric bunching design that does not require functional form assumptions or deciding an excluded region around the kink. Responses correspond to an earnings elasticity with respect to the implicit net-of-tax rate of about 0.18. Using a regression discontinuity design, I further show that the kink in the benefit schedule induces significant responses at the extensive margin. I use the estimated earnings responses to evaluate how the benefit offset affects program costs, and find that relaxing the benefit offset reduces public expenditures only if program entry is very inelastic. My findings speak to recent policy-proposals aiming to improve work incentives of DI recipients.

Keywords: labor supply, disability insurance, policy evaluation, bunching
JEL codes: H53, H55, I38, J21

*This project received financial support from the Norwegian Welfare Administration. I would also like to thank Kristoffer Berg, Andreas Kostøl, Herman Kruse and Edwin Leuven for valuable input.

†University of Oslo, Department of Economics; Statistics Norway; The Norwegian Labor Force Administration. E-mail: a.s.myhre@econ.uio.no

1 Introduction

Disability Insurance (DI) programs are among the largest public transfer programs in developed countries. In the OECD, total spending on DI programs amounts to approximately 2.5% of GDP (OECD, 2010). In an attempt to reduce fiscal costs, many countries have already implemented or are considering implementing policies designed to improve incentives to work for the programs' beneficiaries.[1] While such policies allow beneficiaries to keep a larger share of their benefits if they engage in employment, they typically involve a high implicit tax on earnings above an exemption threshold, creating a large kink in the budget constraint. If labor supply responses to the kink are large, policymakers could relax the exemption threshold or reduce the implicit tax rate to increase work effort among DI recipients. This would improve welfare among current recipients and increase revenue from income taxes. On the other hand, a more lenient policy might also increase expenditures on DI benefits and induce more potential applicants to apply for DI. Knowledge about recipients' labor supply responses in such policies is therefore crucial for optimal policy design.

The main contribution of this paper is to assess this policy trade-off by investigating intensive and extensive margin labor supply responses of DI recipients to kinked budget sets. Two key features in the Norwegian DI program allow me to do this. First, the kink in the benefit schedule is salient and large in magnitude. DI beneficiaries in Norway can earn up to approximately \$4,937 per annum without losing benefits. If earnings exceed this threshold, DI benefits are reduced by approximately \$2 for every \$3 in earnings above the threshold. Second, a sharp discontinuity in benefit schedules provides a particularly attractive setting to analyze behavioral responses. As a transitional policy from prior work incentives, recipients awarded DI before 1st of January 2015 were subject to a relaxed exemption threshold of \$8,000 until 2018.[2] This sharp discontinuity in benefit schedules therefore allows me to analyze intensive and extensive margin responses separately. I identify the intensive margin elasticity using a non-parametric bunching design, where the sharp discontinuity allows me to observe bunching at the kink for one group of individuals and at the same time observe the earnings distribution of a comparison group where the implicit tax on earnings does not change. To identify extensive margin responses, I implement a regression discontinuity design, and estimate an extensive margin elasticity with respect to participation tax rates. In combination, the intensive and extensive margin elasticities allow me to analyze how different benefit schedules affect overall public expenditures.

My main empirical strategy implements the theoretical framework of Blomquist *et al.* (2019) who show that the earnings elasticity can be identified non-parametrically by inverting the cumulative earnings distributions of two groups subject to different kinked budget sets. In my setting, identification relies on recipients with DI award on either side of the cut-off date (i.e. 1st of January 2015) being drawn from the same distribution of potential earnings, i.e. the earnings distributions of the two groups would have been comparable if they were subject to the same benefit schedule. Each point in the cumulative distributions of the two groups will then correspond to individuals with the same earnings potential. This allows me to estimate the intensive margin earnings elasticity by comparing earnings of recipients at the same point in the cumulative distributions where the marginal tax rate is different between the two

[1] One example is the "\$1 for \$2 offset" in the US that has been proposed for many years but never been implemented. Under this policy, DI benefits would be reduced by \$1 for every \$2 in earnings above about \$14,000 per annum. Other examples include i.e. the "Ticket to Work" program in the UK. Switzerland tested a conditional cash program that offered DI recipients cash payments if they expanded or stared working. Sweden introduced the so-called continuous deduction program back in 2009.

[2] Before 2015, the annual exemption threshold was \$12,000.

groups. An advantage of this method is that it does not require functional form assumptions or deciding an excluded region around the kink as in standard bunching applications. To shed light on these issues in the standard applications, I estimate bunching responses using common strategies in the literature and compare these with the non-parametric approach.[3]

A kink may also create responses along the extensive margin as it increases the average tax of participating in the labor force.[4] Therefore, if individuals have a fixed cost of labor force participation, it is possible that the high kink at $8,000 induces some recipients to start working. To identify this effect, I implement a regression discontinuity design which compares recipients with DI award on each side of the cut-off date, and pin down an extensive margin elasticity with respect to participation tax rates. Furthermore, it is possible that the extensive margin responses could affect the observed earnings distributions, and therefore induce a bias in the estimated intensive margin elasticity. To shed light on this matter, I perform Monte Carlo simulations generating data from a utility function with a fixed cost of labor force participation. This allows me to investigate how extensive margin responses affect the estimated elasticity in the bunching setting.

My main empirical findings can be summarized by the following conclusions. First, I find large and sharp bunching in recipients' earnings around each kink. If DI benefits were not reduced above the threshold, recipients who bunch at the baseline kink at $4,937 would have earned $944 or about 19 percent more. This response corresponds to an earnings elasticity with respect to the implicit net-of-tax rate of about 0.18. This elasticity is higher than estimates found in studies that examine bunching at kinks in the income tax schedule, but is in line with estimates found in studies that examine bunching in social security programs.[5] Second, I find that the benefit offset creates sizable responses at the extensive margin. Labor force participation is about 0.8 percentage points higher among recipients with a kink at $8,000 which is sizable considering that only about 11 percent of DI recipients participate in the labor force. The response corresponds to an elasticity of labor force nonparticipation with respect to participation taxes of about 0.11, and is in line with estimates found in other studies.[6]. Third, I find that extensive margin responses induce a large bias in the bunching elasticity in my setting. The Monte Carlo simulations reveal that the bias amounts to about 70 percent of the true elasticity. However, I show that the bias is negligible if one accounts for extensive margin responses in earnings distributions. Fourth, my findings indicate that relaxing the benefit offset increases disposable income and reduces costs for current recipients of the program. However, overall costs for the government are likely to increase if a more lenient benefit offset policy attracts more individuals to the DI program.

A caveat with this study is that it is not informative about the level of increased program inflow when recipients are allowed to keep a larger share of their benefits as they earn more. Because recipients were unable to manipulate the DI award date, and the more lenient benefit schedule was a transitional policy, I am unable to identify this effect. I do, however, calculate the size of induced entry that has to be generated by more generous benefit schedules in order to increase program costs. Based on findings in

[3]In particular, I estimate the bunching elasticity using the polynomial approach of Chetty *et al.* (2011) and the linear approximation approach suggested by Saez (2010).

[4]See e.g. Gelber *et al.* (2017b, 2020a).

[5]Saez (2010); Chetty *et al.* (2011); Bastani & Selin (2014); Paetzold (2019) find elasticities in the range of 0-0.05 among wage earners. Zaresani (2020) and Ruh & Staubli (2019) estimate structural earnings elasticities of 0.20 and 0.27 using a kink in the Canadian DI program and a notch in Austrian DI program, respectively. Gelber *et al.*, 2020b find an observed elasticity of 0.19 in the US social security program.

[6]Kostøl & Mogstad (2014) estimate an elasticity of labor force nonparticipation with respect to participation taxes of about 0.12 in the Norwegian DI program. Ruh & Staubli (2019) find an elasticity of about 0.10 in the Austrian DI program.

previous studies, I find that relaxing the benefit offset is likely to increase overall public expenditures.[7]

My paper primarily contributes to a small and inconclusive literature that assesses fiscal costs of the benefit schedule in DI programs. Kostøl & Mogstad (2014) find that replacing a notch at the exemption threshold with a kink in the Norwegian DI program was likely to reduce program costs. In contrast, Ruh & Staubli (2019) exploit a notch in the Austrian DI program, and conclude that abolishing the notch would increase program costs.[8] They find that most of the increased government revenue comes from extensive margin responses if the benefit schedule is relaxed. I contribute to this literature by assessing policy implications in a DI policy where recipients are subject to kinked incentives.

My paper is also closely related to the broader literature that examines labor supply effects of financial incentives to work for DI recipients. Zaresani (2020) estimates a structural earnings elasticity of 0.20 using kinks in the Canadian DI program. Weathers & Hemmeter (2011) examine effects of a pilot project that replaced a notch at the exemption threshold with a gradual reduction in DI benefits in the US. They find that the policy change significantly increased the number of recipients with earnings above the exemption threshold. Campolieti & Riddell (2012) find that the introduction of an exemption threshold significantly increased labor force participation of Canadian DI beneficiaries. They find no changes in program inflow or outflow. In contrast to the above studies, Schimmel *et al.* (2011) find that increasing the exemption threshold from \$500 to \$700 in the US only increased beneficiaries' earnings by a small amount. Butler *et al.* (2015) investigate employment effects of a conditional cash program that offered cash claims for DI recipients who expanded work in a randomized experiment. They only find small and negligible effects on employment. I contribute to this literature by identifying intensive margin and extensive margin responses separately. My paper is also related to the literature on the potential work capacity of DI recipients.[9]

Finally, my paper also relates to a large literature that studies behavioral responses to kinked incentives. In particular, Saez (2010) shows that bunching at kinks can identify a behavioral elasticity. Chetty *et al.* (2011) extend this framework to allow for adjustment costs. More recently, Blomquist *et al.* (2019) show how an elasticity can be identified from variation in budget sets. I contribute to this literature by implementing the conceptual framework of Blomquist *et al.* (2019) and estimate the bunching elasticity non-parametrically. In terms of understanding how standard empirical implementations of the bunching approach perform, I contribute by investigating how these approaches compare to the non-parametric approach. My findings indicate that common estimation approaches in the literature perform well in my context.[10] My paper is also related to several studies using bunching at kinks to identify earnings responses in social security programs (e.g. Le Barbanchon, 2016; Gelber *et al.*, 2020b; Zaresani, 2020). The bunching approach has also been used to identify responses in many other contexts.[11] Kleven (2016) provides a review of this literature.

[7] See e.g. Hoynes & Moffitt (1999); Gruber (2000); Campolieti & Riddell (2012); Mullen & Staubli (2016); Castello (2017).

[8] Ruh & Staubli (2019) estimate a structural elasticity driving the responses of about 0.27.

[9] A number of studies document that DI receipt significantly reduces earnings and labor force participation by using rejected applicants as a control group for DI recipients (e.g. Bound, 1989; Chen & van der Klaauw, 2008; Singleton, 2012; Maestas *et al.*, 2013; French & Song, 2014). There is also some evidence on the work capacity for DI recipients who have endured longer spells on DI (Borghans *et al.*, 2014; Moore, 2015).

[10] Specifically, fitting a flexible polynomial to the empirical distribution following Chetty *et al.* (2011) yields an elasticity of about 0.15-0.19 depending on the polynomial order, while the non-parametric approach yields an elasticity of about 0.18. Using the approach of Saez (2010), I estimate an elasticity of about 0.19.

[11] Some examples include Einav *et al.* (2017) who use a dynamic model to investigate responses to health insurance contracts, retirement decisions (Manoli & Weber, 2016), effects of minimum wages (Harasztosi & Lindner, 2019), transaction taxes in housing markets (Kopczuk & Munroe, 2015; Best & Kleven, 2017) and responses to speed controls (Traxler *et al.*, 2018).

The remainder of this paper is organized as follows. Section 2 describes the Norwegian DI program and the incentives to work. Section 3 describes the data and sample used in the empirical analysis. Section 4 outlines the methodology and the validity of the empirical design. Section 5 presents the main results of the estimated earnings elasticity. Section 6 analyzes the extensive margin responses. Section 7 performs an analysis of the elasticity using common bunching methods in the literature. Section 8 calculates fiscal effects of alternative policies and discusses their implications. Section 9 concludes.

2 Institutional Background

2.1 The Norwegian DI Program

The Norwegian DI program is designed to provide partial earnings replacement to working-age individuals whose work capacity is permanently reduced due to a medically verifiable physical or mental impairment. The program is part of the broader social security system and is financed by payroll taxes. Of the OECD countries, Norway has one of the highest proportions of the working age population on DI rolls. From 1961 to 2004, the percentage of the working age population on DI rolls increased consistently from 2.2 to 10.4 percent. Since 2004, the proportion decreased slightly and is around 9.5 percent as of 2016, with public spending around $10 billion, or 3 percent of GDP (NAV, 2018).

Pathways into DI and Determination Process In order to apply for DI benefits, individuals' ability to work must have been clarified by a primary medical doctor and appropriate vocational measures must have been completed. Only individuals in the working age population 18-67 years are eligible to apply. Individuals' ability to work must be permanently reduced by at least 50 percent, and illness or injury must be the main reason for the reduced work capacity.[12] In Norway, most applicants for permanent DI benefits are beneficiaries of a temporary DI program whose purpose is to evaluate and improve beneficiaries' ability to work. More than 80 percent of allowed cases for permanent DI benefits are individuals who have endured spells on this program. The temporary DI program in its current form was implemented in March 2010 when three different types of temporary DI or rehabilitation programs were replaced by the current program. As a general rule, temporary DI benefits are provided for up to 4 years for eligible individuals during the time period I consider in this paper.[13] Other pathways into DI include individuals on sickness benefits while some clear cut cases lead directly to award.

If the criteria for permanent disability are met, individuals must submit an application to the Norwegian Labor and Welfare Administration (NAV) whose disability examiners and medical doctors assess the medical evidence and verify the validity of the claim.[14] Processing an application usually takes between 4-9 months depending on the complexity of the application, workload and geographical location of the local DI office. If the disability examiner concludes that the applicant cannot engage in more than 50 percent of full-time employment because of the health impairment, and appropriate vocational measures have been completed, a disability award is made. Approximately 85 percent of applications are accepted. Of allowed cases, about 80 percent are allowed a full disability claim.[15]

[12] For individuals on the temporary DI program at the time of application, a 40 percent permanent reduction in earnings capacity is sufficient. For individuals whose disability is due to an approved occupational illness or injury, a 30 percent permanent reduction in earnings capacity is sufficient.

[13] From 2018, benefits are provided for up to 3 years as a general rule.

[14] Dahl *et al.* (2014) explain the disability determination process in detail.

[15] The remaining 20 percent of allowed DI cases are allowed a partial DI claim where disability rating depends on the perceived ability to work. I abstract from partially disabled recipients in this paper.

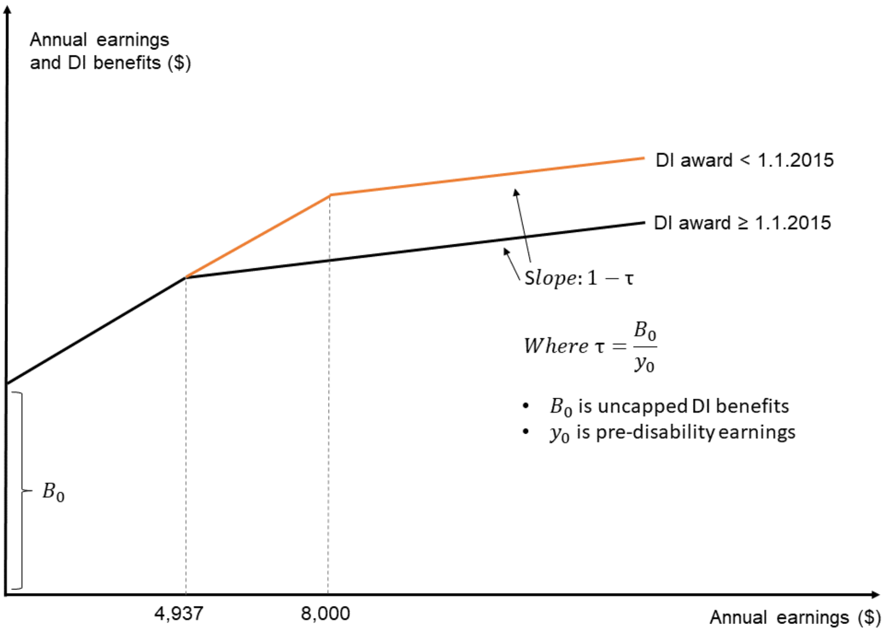
2.2 Benefit Phase-out and Levels

DI recipients deemed as totally disabled in Norway can earn up to an exemption threshold K each year without any reductions in DI benefits B.[16] If annual earnings z exceed this threshold, DI benefits are gradually reduced by a share τ for every dollar in earnings above the threshold. The relationship can be summarized as follows, where B_0 is the uncapped DI benefits before any reductions:

$$B = \begin{cases} B_0 & if \quad z \leq K \\ B_0 - \tau(z-K) & if \quad z > K \end{cases} \tag{1}$$

where the implicit tax-rate τ is equal to the replacement rate of DI benefits which is approximately 66 percent for most recipients. Specifically, $\tau = B_0/y_0$ where y_0 is pre-disability earnings adjusted for wage growth.[17] Both earnings and DI benefits are subject to regular income taxes. The general annual exemption threshold is $4,937 in 2016 dollars and is indexed annually according to the average wage growth. This is equivalent to about 4 hours of work per week for the representative DI recipient. Until 2018, a transitional policy applied to recipients awarded DI before 1st of January 2015 who were subject

Figure 1: Budget Sets



Notes: The figure shows the budget set for recipients awarded DI before and after 1.1.2015, respectively, during years 2015-2018. For illustrative purposes and with minimal loss of generality, I assume that recipients awarded DI before and after 1.1.2015 receive the same amount of (uncapped) DI benefits, and disregard dependent benefits and income taxation.

[16] Partially disabled recipients are subject to the same benefit phase-out rules as totally disabled, but with an individual exemption threshold K_i which depends on disability rating and past earnings.

[17] Pre-disability earnings are defined as each individual's earnings potential if engaged in full-time employment prior to disability onset. For recipients with little or no earnings history, pre-disability earnings are set to minimum levels. If annual earnings exceed 80 percent of pre-disability earnings, recipients are considered engaging in full-time employment and no benefits are provided for that year. DI recipients can exceed this threshold for up to five consecutive years and keep their DI receipt.

to a relaxed threshold of \$8,000, or about 6 hours of work per week.[18] Figure 1 represents the budget constraints for the two groups.

Benefit Levels DI benefits replace about 66 percent of past earnings up to a maximum amount of about \$49,000 per annum. If recipients have little or no earnings history, they receive a minimum amount of DI benefits of about \$31,000.[19] In 2015, the definition of past earnings which calculations of DI benefits is based upon was changed. For recipients awarded DI before 1st of January 2015, benefits were based on projected retirement savings as if the individual had continued to work until the general retirement age of 67 years. The years with the lowest income, or projected income in the future were excluded so that a maximum of 20 years of earnings history were used in the calculations of DI benefits. For recipients awarded DI on 1st of January 2015 or later, DI benefits were calculated using the highest average income in three of the last five years prior to disability onset. Because DI benefits were based on more recent earnings history, this group of recipients receive slightly higher (uncapped) DI benefits on average due to real wage growth, approximately 4% more on average. Minimum benefits were the same independently of award date.

3 Data and Sample Selection

This section describes the administrative data sources, key outcome variables and the main sample for the empirical analysis.

3.1 Data Sources

In the empirical analysis, I use data from three main sources that can be linked by unique and anonymized identifiers for every resident individual. The data on DI recipients is provided by the Norwegian Labor and Welfare Administration (NAV) and contains monthly records of all DI recipients who entered the permanent DI program until 31st of December 2015. It contains information about the level of DI benefits received, disability rating, month of DI award, month of disability onset, pre-disability earnings and a rich set of demographic and socioeconomic information including gender, age and cohabitant status. The earnings data is also provided by NAV and contains monthly records of wage earnings for each employer-employee relationship during years 2015-2017. Finally, I use administrative data provided by Statistics Norway which contains individual demographic and socio-economic information such as education, number of children and date of death.

The administrative nature of the data reduces the extent of measurement errors in disability variables and employment relationships. Because individual disability status and earnings are third-party reported (i.e. by NAV and the employers), the coverage and reliability are rated as exceptional by international quality assessments (see e.g. Atkinson *et al.*, 1995). Since administrative data are a matter of public record, there is no attrition due to non-response or non-consent by individuals or firms, and individuals can only exit these data sets due to natural attrition (i.e. death or out-migration).

[18] Before 2015, the exemption threshold was approximately \$12,000 and recipients were subject to a different set of work incentives. The exemption threshold at \$8,000 was therefore a transitional policy from the previous work incentives until 2018. From 2019 and onward, all recipients are subject to a common threshold of \$4,937.

[19] For cohabitant recipients, minimum benefits amounts to about \$28,000. In addition, recipients classed as "young disabled" get an additional amount of about \$5,000 if they were 26 years or younger at disability onset.

3.2 Variables

Because phase-out of DI benefits is determined at the annual level, the main outcome variable I consider is annual wage earnings.[20] As the earnings data comes from monthly records, I construct this variable by summarizing earnings from all employer-employee relationships (if more than one) for each month during the calendar year. The second key outcome variable I consider is whether recipients have any annual earnings. In Norway, employees are subject to holiday pay which is based on last year's earnings.[21] In order to distinguish between recipients who have engaged in employment and those whose earnings are based on last year's earnings, I therefore define this variable as positive earnings excluding holiday pay. The time period I consider is 2016-2017 as some recipients in the estimation sample were awarded DI in 2015, and 2017 is the last year of the data.

3.3 Estimation Sample

The main sample used in the empirical analysis consists of recipients awarded DI between 1st of April 2014 and 30th of September 2015, i.e. +/- 9 months around the cut-off date for being subject to the relaxed benefit phase-out. I restrict the sample to recipients who are deemed totally disabled by NAV due to a lack of information on the exact location of the exemption threshold among partially disabled recipients. Furthermore, I exclude recipients who turn 67 years during the calendar year due to eligibility for old-age pension beginning at age 67.

Table 1 reports summary statistics for all totally disabled DI recipients awarded DI before 2015 as well as the estimation sample of recipients with DI award on each side of the cut-off date, respectively. Compared to the average DI recipient, recipients in the estimation sample have lower earnings and fewer recipients have any labor market earnings. By construction, they have also spent fewer years on DI and are somewhat younger. Otherwise, individual characteristics are fairly similar. As for the two groups in the estimation sample, recipients awarded DI in 2015 receive about 5 percent more (uncapped) DI benefits compared to recipients awarded DI in 2014. While the two groups share fairly similar characteristics, there is in particular one notable difference. Recipients awarded DI in 2014 are far more likely to have been on a prior temporary DI program before the current temporary DI program was implemented in March 2010. As the general maximum spell on the current temporary DI program is 4 years, many recipients who were transferred from the prior programs were awarded DI in 2014. As a result, these individuals are slightly younger and have endured longer spells since disability onset before being awarded DI on average.

[20] While income from self-employment is also subject to a reduction in DI benefits, very few DI recipients have income from self-employment (less than 1 percent of recipients with some earnings).

[21] As a general rule, employees get 12 percent of last year's earnings as holiday pay in June the following year regardless of current employer-employee relationship(s).

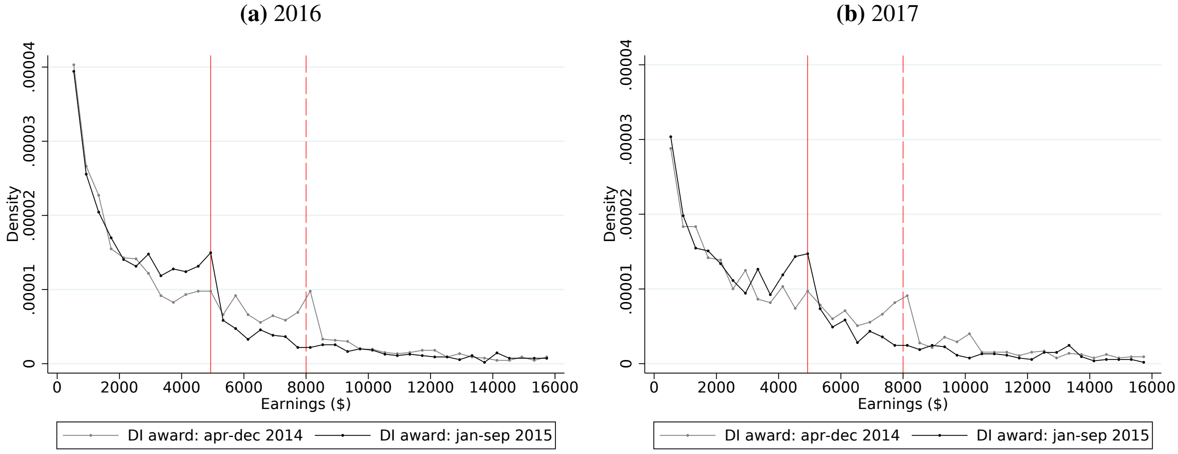
Table 1: Summary Statistics

	All DI recipients		Estimation sample			
DI award:	- dec 2014		apr-dec 2014		jan-sep 2015	
Kink point:	$8,000		$8,000		$4,937	
Outcome variables:	mean	sd	mean	sd	mean	sd
Annual earnings ($)	993	(4,309)	627	(3,229)	463	(2,450)
Any annual earnings (%)	15.6		12.1		11.0	
DI Information:						
Uncapped DI benefits ($)	35,566	(6,258)	34,485	(6,570)	36,207	(8,193)
Pre-disability earnings ($)	57,334	(15,197)	59,410	(15,233)	60,224	(16,395)
Benefit replacement rate	.62	(.09)	.59	(.10)	.61	(.11)
DI award date	2003	(11)	2014	(0)	2015	(0)
Disability onset date	1999	(9)	2007	(4)	2009	(4)
Fraction from TDI program	.41		.92		.89	
Fraction from prior TDI programs	.27		.57		.30	
Individual characteristics:						
Age	52.3	(11.0)	47.6	(12.1)	48.0	(13.1)
Age at DI award	40.3	(13.8)	46.3	(12.1)	47.4	(13.1)
Years of schooling	10.7	(2.1)	10.9	(2.2)	11.0	(2.2)
Number of children	1.6	(1.4)	1.7	(1.4)	1.6	(1.4)
Fraction females	.56		.56		.54	
Fraction cohabitants	.48		.51		.51	
Number of recipients	236,568		16,620		13,697	

Notes: All samples consist of totally disabled DI recipients with DI receipt 31.12.2015. Outcome variables are measured in 2016 and 2017. Age, years of schooling, cohabitant status and number of children are measured in 2015. All other covariates are either pre-determined or constant over time. Earnings and DI benefits are measured in 2016 dollars (NOK/USD = 7.5).

Next, I examine the earnings distributions for each group in the estimation sample around the annual kinks. Figure 2 shows the raw earnings distributions for each group in 2016 and 2017 grouped into $400 bins. Specifically, the black solid line indicates the density for recipients awarded DI between January and September 2015, and the vertical red solid line indicates the kink point for these recipients at $4,937. The gray solid line indicates the density for recipients awarded DI between April and December 2014, with the vertical red dashed line indicating the kink point for this group at $8,000. I use the full sample of recipients (i.e. including recipients with zero earnings) to calculate the density for each group. Notably, there is large bunching around each kink. Otherwise, the densities appear to track each other very closely in regions outside of the two kinks. The similarities in densities below the first kink point is particularly striking, indicating that the earnings potential between the two groups is similar.

Figure 2: Earnings Distributions Around the Annual Kinks



Notes: Panel (a) and (b) show the earnings distributions in $400 bins for DI recipients awarded DI between April 1st and December 31st 2014 (gray line) and recipients awarded DI between January 1st and September 30th 2015 (black line) in 2016 and 2017. The red dashed line and the red solid line indicate the kink point in the budget constraint for each group, respectively. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

4 Methodology

This section outlines the conceptual framework for my empirical strategy. In the empirical application, my goal is to estimate the earnings elasticity with respect to the implicit tax rate implied by the phase-out of DI benefits. Identification of the elasticity relies on the fact that DI recipients are subject to different budget sets depending on award date.

4.1 Theoretical Framework

My framework follows Blomquist *et al.* (2019) but focus on the kink in the consumption-leisure space for DI recipients as opposed to a kink in the income tax schedule. I assume that recipients maximize the following quasi-linear utility function:

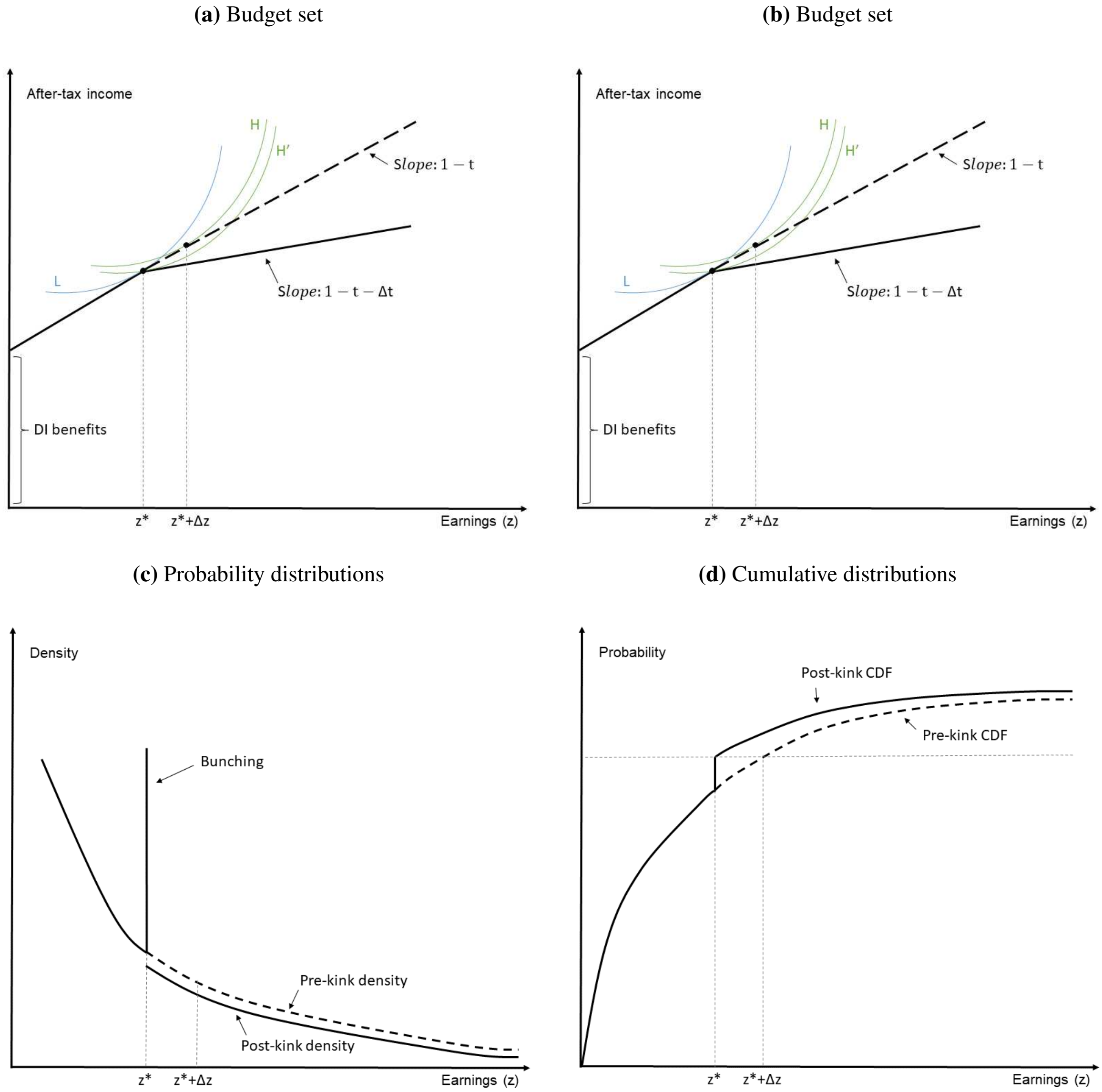
$$U(c,z) = c - \frac{n}{1+\frac{1}{e}}\left(\frac{z}{n}\right)^{1+\frac{1}{e}} \tag{2}$$

subject to the budget constraint $c = B + z - T(z;B)$ where c is consumption, B is DI benefits and z is before-tax earnings. $T(z;B)$ is the implicit tax liability, and depends on earnings and DI benefits. n is an ability parameter and e is the earnings elasticity with respect to the marginal net-of-tax rate $1-t$. A key identifying assumption is that the distribution of ability n is smooth in the population. This assumption implies that, given a linear tax system $T(z;B) = t\cdot(z+B)$, the smooth ability distribution translates into a smooth after-tax earnings distribution. Maximization of $U(c,z)$ subject to the linear budget constraint yields $z = n(1-t)^e$. Note that $z = n$ if $t = 0$, i.e. n can be interpreted as potential earnings if there were no implicit taxes on earnings. Because of the quasi-linearity assumption, the model rules out income effects of tax changes on earnings.[22]

[22] If income effects are present, the estimated elasticity will be downward biased as the income effect induces individuals to work more (assuming leisure is a normal good). Hence, the compensated elasticity (accounting only for substitution effects) will be larger than the uncompensated elasticity (accounting for income and substitution effects). However, Bastani & Selin (2014) show that even large income effects induced by a kink in the budget set have little impact on the compensated elasticity.

Now, suppose that a kink is introduced at some threshold z^*, with the tax rate increasing from t to $t+\Delta t$. The tax schedule can now be expressed as $T(z;B) = t\cdot(z+B) + \Delta t\cdot(z-z^*)\cdot 1(z \geq z^*)$. Figure 3 (a) and (b) illustrate how the slope in the budget set changes at z^*. Individual H is the individual with the highest earnings before the kink is introduced who would locate at z^* with the kink, illustrated by the slope in the indifference curve H' being exactly equal to the slope in the budget set above the kink $1-t-\triangle t$. This is the marginal bunching individual. The individual would locate at $z^*+\triangle z$ before the kink is introduced and locate at z^* with the kink in the budget set. Individual L would locate at z^* in both cases. Individuals with earnings in the interval $[z^*, z^*+\triangle z]$ before the kink is introduced would locate at z^* with the kink. This is illustrated in Figure 3 (c) which shows the probability density distributions pre- and post the introduction of the kink. The earnings distribution is smooth in the population before

Figure 3: Budget Set and Density Distributions



Notes: Panel (a) and (b) show after-tax income as a function of annual earnings with a linear tax t (dashed line) and with the tax increasing from t to $t+\triangle t$ at the kink point z^* (solid line). $z^*+\Delta z$ denotes the earnings of the marginal buncher, i.e. the individual with the highest earnings without the implicit tax on DI benefits who will locate at z^* with the implicit tax on DI benefits above z^* . Panel (c) shows the probability density distribution with and without the kink at z^*. Panel (d) shows the corresponding cumulative distributions.

the kink is introduced, while there is substantial bunching at the kink after the kink is introduced. Figure 3 (d) shows the corresponding cumulative distributions. Because the pre-kink earnings distribution is smooth in the population, the point in the pre-kink CDF at $z^* + \triangle z$ corresponds to the point in the post-kink CDF of the marginal bunching individual who is now locating at z^*. This is illustrated by the horizontal dashed line.

Next, consider two groups $j = 1,2$ drawn from the same distribution of ability n. Group 1 is subject a constant marginal tax rate t for the whole budget set, while for group 2 the marginal tax rate increases from t to $t + \Delta t$ at z^*. Let $F_1(Z)$ and $F_2(Z)$ be the corresponding cumulative distribution functions for each group, i.e. $F_i(Z) = Pr(Z_i \leq z_i)$ for $j = 1,2$. Then from theorem 3 in Blomquist *et al.* (2019), it follows that

$$e = \frac{\ln(\frac{z^*}{z^*+\Delta z})}{\ln(\frac{1-t-\Delta t}{1-t})} \tag{3}$$

where e is the earnings elasticity with respect to the net-of-tax rate t, and Δz is the earnings response of the marginal bunching individual defined as the point where $F_1(z^* + \Delta z) = F_2(z^*)$.[23] Intuitively, this corresponds to the same point in the CDF between the two groups at the kink point for group 2 where the marginal tax rate differs between the two groups. Because the distribution of ability is smooth, each point in the CDFs of the two groups corresponds to individuals with the same ability, or potential earnings.

4.2 Empirical Implementation

My approach to estimate the earnings elasticity e relies on the fact that two groups of DI recipients in the estimation sample are subject to different budget sets. In particular, the kink point for recipients awarded DI in 2015 is $z^* = \$4,937$ where the marginal tax rate on earnings increases from t_0 to t_1. Here, t_0 is the regular marginal tax rate on income and $t_1 = t_0 + \tau(1 - t_0)$, where τ is the benefit phase-out rate.[24] At the same point in the budget set, the marginal tax rate is t_0 for recipients awarded DI before 2015. Using the fact that $t_0 \neq t_1$ at z^*, I estimate the earnings elasticity e using the following formula:

$$\hat{e} = \frac{\ln(\frac{z^*}{z^*+\Delta\hat{z}})}{\ln(\frac{1-t_1}{1-t_0})} \tag{4}$$

where $\Delta\hat{z}$ is the estimated response of the marginal buncher, and is given by $\hat{F}_0(z^* + \Delta\hat{z}) = \hat{F}_1(z^*)$, where $F_i(Z)$ is the cumulative distribution function of DI recipients with treatment status $i = 0,1$ where treatment status $i = 0$ indicates recipients awarded DI before 2015 and $i = 1$ indicates recipients awarded DI in 2015 or later. Intuitively, the estimated earnings response of the marginal buncher, or the last individual with earnings at the kink $z^* = \$4,937$ in the treated group corresponds to the individual at the same point in the CDF in the non-treated group. The crucial assumption in my setting is that potential earnings, i.e. earnings without the phase-out of DI benefits for individuals with DI award on either side of the cut-off date are drawn from the same distribution, and that this distribution is smooth. Additionally, I assume that the earnings distribution of recipients awarded DI before 2015 is unaffected by the kink at \$8,000 in the interval $[0, z^* + \Delta z]$, i.e. recipients who would locate above this region without the kink would not locate in this region with the kink.

[23] A formal proof of this result is provided in Blomquist *et al.* (2019).

[24] For most recipients, the marginal tax rate on income t_0 is about 35 percent. At the kink, the marginal tax rate therefore increases from about 35 percent to about .35 + .66·(1-.35) = 78 percent.

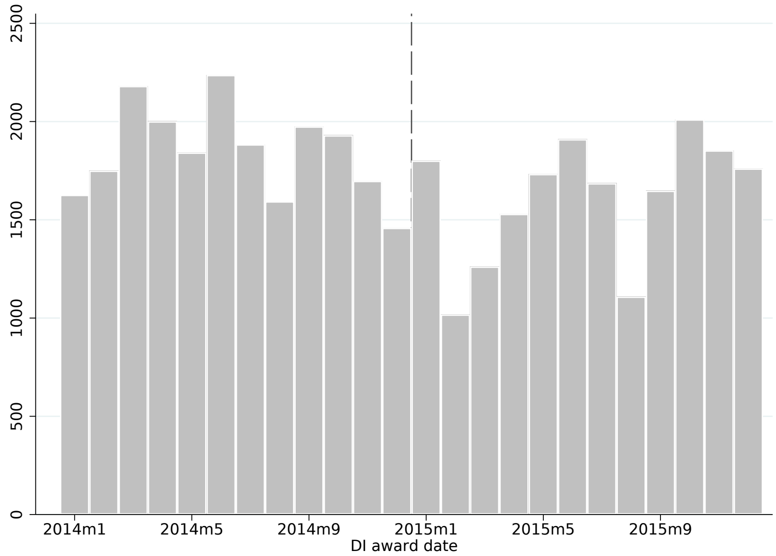
Inference To calculate the standard error of the estimated elasticity, I use bootstrap methods. First, I generate many earnings distributions by random resampling with replacement. As the estimation sample may include the same individuals more than once, I use a pairs-cluster bootstrap that accounts for clustering at the individual level, keeping all observations of each individual that I resample. Second, I re-estimate the elasticity within each sample, and define the standard error as the standard deviation of the distribution of the elasticity. In all estimations, I use 500 repetitions.

4.3 Threats to Identification

The validity of my empirical design hinges on the assumption that recipients with DI award on either side of the cut-off at 1st of January 2015 are drawn from the same distribution of potential earnings. In other words, the earnings distributions should be comparable if these recipients were subject to the same benefit phase-out policy. The validity of my design therefore requires that recipients are unable to precisely manipulate the DI award date. Crucially, there were no changes to eligibility for DI around the cut-off date. Although some institutional details were formalized as early as 2011, the policy change was announced as late as October 2014. Because processing times of applications usually take between 4-9 months, recipients were unable to gain entry before the cut-off date of January 1st, 2015. Even if potential applicants would have some influence over when to apply for DI, recipients would be unable to manipulate the award date precisely because of uncertainty in the processing time of applications. Therefore, the variation in treatment should be randomized close to the cut-off.[25]

Figure 4 shows the distribution of DI award date around the cut-off. Because I only have monthly data on DI award, the assignment variable in this context is discrete. I therefore follow Frandsen (2017) and perform a formal statistical test for bunching on either side of the cut-off. While this test rejects the null hypothesis of no bunching, the test also rejects the null in more than half of hypothetical placebo

Figure 4: Distribution of DI Award Date



Notes: The figure shows the distribution of DI award between January 2014 and December 2015. The sample consists of totally disabled recipients on DI receipt 31.12.2015 aged 18-66 years.

[25] See Lee & Lemieux (2010).

cut-off points during the time period considered in the empirical analysis.[26] This suggests a natural high variation in the number of allowed applications may explain why the test of no bunching fails in this context rather than manipulation of the DI award date.

While recipients awarded DI before the cut-off date were subject to a more lenient benefit phase-out, the policy was announced as a transitional policy which would revert to a common policy in 2019. Therefore, the gain from being subject to the more lenient benefit phase-out was limited. If one would worry about potential manipulation of the DI award date, a bigger concern is the fact that calculation of DI benefits was also changed at the cut-off date. Because DI benefits were calculated using more recent years of earnings history for awards after 1st of January 2015, most recipients would receive slightly higher levels of DI benefits if awarded DI after this date due to real wage growth. While this would require detailed information about earnings history and institutional details among recipients, I cannot rule out this possibility. In order to shed light on this concern, I calculate the DI benefits that recipients hypothetically would receive if awarded DI after the cut-off date.[27] If DI recipients were able to manipulate the award date, one would expect recipients with high potential DI benefits post the cut-off date to locate to the right of the cut-off, and vice versa for recipients with low potential DI benefits who would locate before the cut-off date. As a formal test, I run a regression of projected DI benefits (if recipients were awarded DI after the cut-off date) on a dummy which equal to 1 if the observed award date is after the cut-off date using 2 months of bandwidth on each side of the cut-off and a triangular kernel. Reassuringly, the coefficient on the dummy is highly insignificant with a p-value of 0.47, lending support to the claim that recipients were unable to manipulate the DI award date to get maximum DI benefits.

To shed further light on possible manipulation of the DI award date, I extend the above exercise to include a rich set of individual characteristics. If recipients indeed were unable to manipulate the DI award date, any pre-determined covariate should have the same distribution on each side, close to the cut-off. In Appendix Table A.1, I report coefficients and standard errors of each pre-determined covariate running the same regression as described above. While most covariates appear smooth around the cut-off and are insignificant at conventional levels, one exception is years of education which is significant at the 1% level. However, based on the large number of covariates I consider, the probability of observing changes in one covariate around the cut-off is quite large. If I perform a joint test for all covariates, I cannot reject the null of no manipulation at conventional levels of significance as reported in Appendix Table A.1. The p-value of the joint test is 0.18.[28]

Income Effects and Weighting Strategy As recipients awarded DI before the cut-off date receive a slightly lower DI benefit, this might induce this group of recipients to work more than recipients awarded DI after the cut-off date through an income effect.[29] In that case, this would shift the pre-kink CDF in Figure 3 (d) to the right and bias the elasticity upwards. Ideally, one would want to identify the

[26] The test rejects 12 out of the 22 placebo cut-off points between each pair of months of DI award between January 2014 and December 2015.

[27] Unfortunately, I am unable to calculate DI benefits using the definition before 2015 due to data limitations.

[28] Using the sample with only 1 month of bandwidth on each side yields the same conclusions for each covariate separately as well as jointly, with the p-value of the joint test being 0.13.

[29] Several studies have shown that increasing (reducing) the generosity of DI benefits reduces (increases) beneficiaries labor supply through an income effect. See e.g. Gruber, 2000; Marie & Castello, 2012; Gelber *et al.*, 2017a; Deuchert & Eugster, 2019.

income and substitution effects separately. Unfortunately, separating the two effects is not possible in this context as I only have one instrument that is the DI award date. In order to address the issue of DI benefits not being directly comparable between the two groups, I apply the weighting approach proposed by Kline (2011).[30] This approach accounts for differences in pre-determined covariates between the two groups, and in particular the level of DI benefits. Intuitively, recipients awarded DI before the cut-off with high DI benefits are assigned a higher relative weight in estimations. While the approach does not fully account for the difference in DI benefits between the two groups, it reduces the difference from about 4% to 1%. In Section 5, I show that the bias in the compensated elasticity is small if income effects induces recipients to work more due to a lower DI benefit.

Another advantage of the weighting approach is that it accounts for differences in (other) pre-determined covariates between the two samples, including years of education which was significant in the balancing tests. While my estimation sample ideally would only include observations very close to the cut-off, deciding the bandwidth, i.e. the sample of DI recipients on each side of the cut-off date is a trade-off between bias and variance. As one includes observations further away from the cut-off, differences in pre-determined covariates increase. In particular, a larger share of recipients awarded DI early in 2014 had endured spells on a prior temporary DI program before being awarded DI. These recipients were slightly younger and had endured longer spells between DI award and disability onset, and could therefore have slightly different earnings potential than other recipients if i.e. health improves or worsens over time.[31] The weighting approach assigns lower weights to these recipients as they are less likely to be awarded DI after the cut-off date. Appendix Table A.2 shows that the difference between the two groups in the estimation sample is insignificant at all conventional levels when using the weighting approach. This result holds for all covariates separately as well as jointly. The p-value of the joint test is 0.43.

As the goal of my main estimation strategy is not to identify average effects of the different incentives to work, it is not entirely clear how to decide the bandwidth in this context. In my baseline specification, I use 9 months of bandwidth on each side of the cut-off which is the optimal bandwidth suggested by Calonico *et al.* (2014).[32] A potential worry using observations further away from the cut-off is trends in earnings potential if i.e. health improves or worsens over time. Because of this, I use triangular weights in my baseline specifications. To examine the validity of my findings, I perform several robustness checks. In particular, I show that the estimated earnings elasticity is relatively robust to bandwidth selection. I also show that average effects are practically indistinguishable if I include linear or quadratic trends in the DI award date in a standard regression discontinuity design.

[30] I implement this adjustment by estimating the probability of each recipients being awarded DI after the cut-off date $P(I_i = 1|x_i)$ using a logistic regression. As the level of DI benefits may be correlated with other covariates such as age, education and pre-DI earnings, only re-weighting the level of DI benefits might induce imbalance in covariates that are correlated with DI benefits between the two samples. Therefore, I include the full set of covariates included in Table A.2 along with uncapped DI benefits as control variables. Recipients awarded DI before the cut-off date are then re-weighted using the propensity score weight $w(x_i) = \frac{1-P(I=1)}{P(I=1)} \frac{P(I_i=1|x_i)}{1-P(I_i=1|x_i)}$ where $P(I = 1)$ denotes the probability of being awarded DI after the cut-off date.

[31] As a robustness check, I exclude recipients who have endured spells on a prior DI program. As opposed to the full (unweighted) sample of DI recipients, this alternative sample is balanced in terms of pre-determined covariates as reported in Appendix Table A.1. Although less precise, the estimated elasticity is practically indistinguishable from the estimated elasticity using the full sample with the weighting approach.

[32] The optimal bandwidth is calculated using the weighted sample and a triangular kernel with no (linear) trends in the assignment variable.

Extensive Margin Responses While my baseline model does not incorporate responses at the extensive margin, it is possible that the lower kink induces some individuals to stop working altogether. Theoretically, a higher average tax on earnings will induce some individuals to stop working if individuals have a fixed cost of labor force participation. In Section 6, I document that the fraction of recipients with some earnings is lower among recipients with the kink at $4,937 compared to recipients with the kink at $8,000. I also show that the elasticity is substantially upward biased if I do not account for extensive margin responses by calibrating a model with a fixed cost of labor force participation to the empirical distribution. Intuitively, the recipients who stop working would have earned above the kink under the more lenient benefit phase-out policy. Because of this, there would be missing mass in the upper part of the earnings distribution which would shift the post-kink CDF in Figure 3 (d) to the left. If one does not account for this response, the response of the marginal bunching individual will be overstated and the estimated elasticity upward biased. In order to adjust for extensive margin responses, I follow Ruh & Staubli (2019) and assume that the distribution of recipients who stop working is the same as the observed earnings distribution above the kink.[33] In Section 6, I show that the estimated elasticity is very close to the theoretical elasticity when incorporating this adjustment procedure in a simulation exercise.

5 Main Results

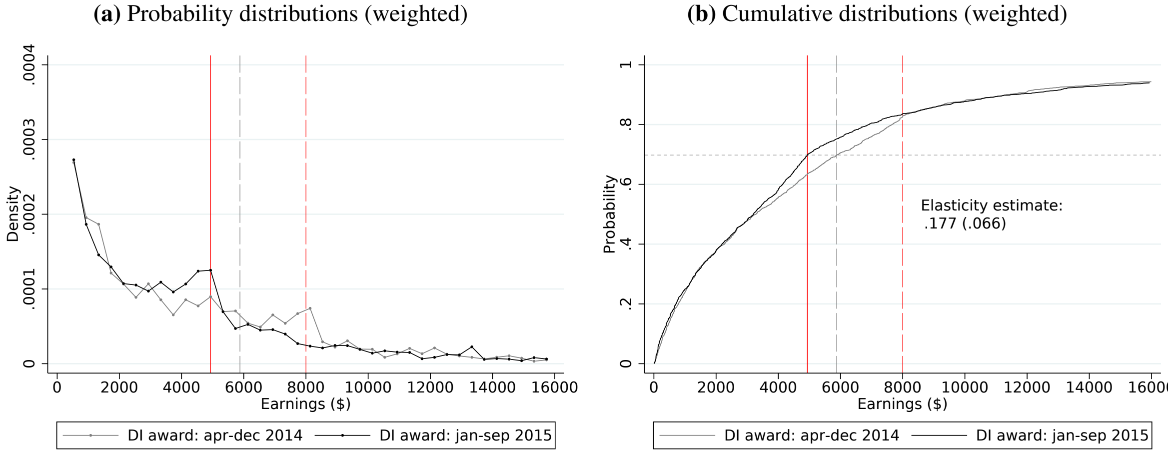
This section presents the main results and begins with a graphical representation of the estimation strategy. I then proceed by presenting the main analytical results before challenging the empirical specification in several ways.

5.1 Graphical Evidence of Behavioral Responses

I begin my analysis by providing a graphical representation of the estimation procedure. Figure 5 (a) shows the weighted earnings distributions for the pooled sample of recipients with some earnings for each group in the estimation sample grouped into $400 bins. First, recipients are weighted by a triangular weight so that recipients close to the cut-off are assigned higher relative weights. Second, recipients awarded DI before the cut-off are weighted by propensity score weights that accounts for differences in pre-determined covariates between the two groups, and in particular the level of uncapped DI benefits. Third, I incorporate the adjustment procedure for extensive margin responses outlined in Section 4 by adding recipients awarded DI after the cut-off to the right of the kink until the fraction of recipients with some earnings is the same between the two groups. Responses should therefore be interpreted as intensive margin responses to the implicit tax on earnings as DI benefits are phased out above the kink.

[33] Specifically, I add individuals to the right of the kink for recipients subject to the lower exemption threshold until the fraction of working individuals is the same for both groups.

Figure 5: Graphical Representation of Earnings Elasticity Estimation: Pooled Sample



Notes: Panel (a) shows the (weighted) pooled earnings distributions for 2016 and 2017 in \$400 bins for recipients with positive earnings awarded DI between April 1st and December 31st 2014 (gray line) and recipients awarded DI between January 1st and September 30th 2015 (black line). The red dashed line and the red solid line indicate the kink point in the budget constraint for each group, respectively. Both groups are weighted by a triangular kernel weight in estimations. For recipients awarded DI in 2015, I add recipients to the right of the kink until the fraction of recipients with positive earnings is the same as for recipients awarded DI in 2014 in order to adjust for extensive margin responses. Recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i=1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Panel (b) shows the corresponding cumulative distributions. The horizontal dashed line indicates the CDF at the kink for recipients awarded DI in 2015. The vertical gray dashed line indicates earnings of the marginal buncher $z^* + \triangle z$. Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

Figure 5 (b) shows the corresponding cumulative earnings distributions for each group. The distributions appear to track each other very closely until about \$3,000 from which the cumulative distribution for recipients with DI award after the cut-off increases more steeply due to recipients bunching around the kink at \$4,937 (indicated by the red solid line). The same pattern is observed for recipients with DI award before the cut-off who bunch around the kink at \$8,000 (indicated by the red dashed line). Reassuringly, the cumulative distributions appear to track each other very closely above the second kink, indicating that the distributions would have been comparable if the two groups were subject to the same benefit phase-out policy. The horizontal dashed line indicates the point in the CDF for the last individual who bunches at the kink at \$4,937 for the sample of recipients awarded DI after the cut-off. The vertical dashed line indicates the earnings of the individual at the same point in the CDF for the sample of recipients with DI award before the cut-off. This is the estimated earnings of the marginal bunching individual. Then, I plug this estimate into the formula for the elasticity given by Equation 4 and estimate an earnings elasticity of about 0.18.

5.2 Earnings Elasticity Estimates

In this section, I present the main estimation results. Table 2 reports estimates of the earnings elasticity (e), the earnings response of the marginal buncher (Δz) and the average intensive margin response for the main estimation sample. I present estimates using the weighting approach that accounts for differences in pre-determined covariates between the two groups and unweighted estimates for comparison. Both specifications use triangular weights and incorporates the adjustment procedure for extensive margin responses outlined in Section 4.

Table 2: Earnings Elasticities for Pooled Sample and by Year

	Elasticity (e)		Earnings response (\$) Marginal buncher (Δz)		Average response		Observations <individuals>
Full sample	.177***	.198***	944**	1,067***	157	531*	59,753
(2016-2017)	(.066)	(.059)	(392)	(357)	(308)	(307)	<30,317>
By year:							
2016	.132**	.172***	689*	914***	151	455	30,317
	(.067)	(.058)	(368)	(341)	(322)	(316)	
2017	.217***	.218***	1,186***	1,192***	163	532	29,436
	(.064)	(.063)	(401)	(392)	(396)	(410)	
Weighted	Yes	No	Yes	No	Yes	No	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports intensive margin estimates of the elasticity, the earnings response of the marginal buncher defined as the earnings response at the first kink and the average response of recipients for the full sample and by each year (2016 and 2017). The sample consist of totally DI recipients awarded DI between 1st of April 2014 and 30th of September 2015. For recipients awarded DI in 2015, I add recipients to the right of the kink until the fraction of recipients with positive earnings is the same as for recipients awarded DI in 2014 in order to adjust for extensive margin responses. For the weighted estimates, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i=1|x_i)$ is estimated with a logit model using the individual characteristics in Table A.2 and uncapped DI benefits as control variables. Both groups are weighted by a triangular kernel weight in all estimations. Standard errors (in parentheses) are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

For the weighted pooled sample, the estimated earnings response for the marginal buncher is large and statistically significant. I estimate that without the kink, the marginal bunching individual who bunches at the kink at \$4,937 would have earned \$944 or about 19 percent more. This response corresponds to an earnings elasticity with respect to the implicit net-of-tax rate of about 0.18. In contrast, the average earnings response is \$157 and is only 17 percent as large as the response of the marginal buncher. This estimate can be interpreted as the average intensive margin response of increasing the kink point from \$4,937 to \$8,000. The fact that this estimate is small and insignificant suggests that increasing the kink point only affects recipients with earnings in a narrow region. The unweighted estimates are slightly larger, but qualitatively similar to the weighted estimates. This suggests that differences in predetermined covariates, and in particular DI benefits do not change my main conclusions. Notably, the estimated earnings response for the marginal buncher and the corresponding elasticity is larger for 2017 than for 2016. While this might be explained by responses increasing over time as recipients overcome frictions such as changing hours worked and learn the tax schedule, the difference between the two years is not significant at conventional levels. The average response is almost indistinguishable between the two years.

In comparison to other studies, the estimated elasticity is significantly higher than in studies that exploit kinks in the income tax schedule. These studies typically find elasticities in the range of 0-0.05 among wage earners (Saez, 2010; Chetty *et al.*, 2011; Bastani & Selin, 2014; Paetzold, 2019). Compared to similar studies on DI beneficiaries, Ruh & Staubli (2019) and Zaresani (2020) estimate earnings elasticities of 0.27 and 0.20 in Austria and Canada, respectively. However, both these studies estimate a structural elasticity as opposed to my study. Although the kink in this setting is large and salient implying that recipients are more likely to overcome adjustment costs, my estimate might be attenuated by i.e. lumpy hours or imperfect information about the benefit phase-out. In that case, the estimate represents a lower bound of the long-run elasticity. Also, it is important to keep in mind that my estimation approach relies on local moments around the kink at $4,937. Compared to other countries, the exemption threshold in Norway is quite low.[34] The earnings elasticity might differ in other countries where the exemption threshold is higher, as this subgroup of beneficiaries have higher earnings capacity and might also be different in other dimensions. In that case, it is likely that my estimate represents a lower bound compared to recipients around the exemption threshold in other countries.

Heterogeneity To shed further light on my main findings, I explore heterogeneity in earnings responses to the implicit tax on earnings. Appendix Table A.3 reports the estimated earnings responses for different subgroups in the population. Somewhat surprisingly, I am unable to detect any statistically significant differences in effects across the different subgroups. Point estimates are slightly higher for older recipients, and slightly higher for recipients with high (uncapped) DI benefits. However, I lack statistical precision to draw any firm conclusions. I am unable to detect any notable differences in responses between genders, recipients with different levels of education and recipients with different levels of earnings prior to disability onset, with point estimates being very similar across subgroups.

5.3 Robustness Analysis

In order to verify the validity of my main results, I do a series of robustness checks reported in Table 3. The first row reports estimates using the baseline specification with triangular weights and 9 months of bandwidth on each side of the cut-off. Next, I estimate responses using rectangular weights implying that all recipients in my sample are assigned the same relative weight in the initial estimation procedure. Although slightly lower, the estimated effects are well within one standard error of the baseline specification. In the third specification, I exclude recipients who had endured spells on a prior temporary DI program as this group of recipients were more likely to be awarded DI early in 2014 (i.e. before the cut-off). This alternative sample is well balanced in terms of pre-determined covariates as reported in Appendix Table A.1. Using this alternative sample, estimates are remarkably similar as to the full sample of recipients. Next, I perform a placebo test by pretending that the cut-off date for being subject to the different phase-out policies in DI benefits was 1st of January 2014 instead of 1st of January 2015. Reassuringly, the point estimates are small and insignificant. Appendix Figure A.1 shows the probability distributions and cumulative distributions for each group in the placebo sample, respectively. The distributions of the two groups in the placebo sample appear remarkably similar. This lends some support to the assumption of earnings potential being comparable for recipients with slightly different DI award dates.

Next, I examine how the estimated earnings responses change as I deviate from the baseline bandwidth selection of 9 months. While the estimated elasticity is somewhat lower if I use only 1 month of bandwidth on each side of the cut-off, the estimated elasticity is within one standard error of my main specification. For specifications using 2 months of bandwidth or more, the estimated earnings responses appear stable and are very similar to my main specification which is reassuring. In Appendix Figure A.2 (a), I show how the estimated elasticity vary with bandwidth selection by plotting point estimates with 95 percent confidence intervals for each bandwidth between 1 and 12 months. The figure yields the same conclusion of point estimates being relatively stable to bandwidth selection.

[34] In the US and the UK, the exemption thresholds are about $14,000 and $8,000, respectively.

Table 3: Robustness Checks for Earnings Responses

	Elasticity (e)		Earnings response ($) Marginal buncher ($\Delta z$)		Average response		Observations <individuals>
Baseline	.177***	.198***	944**	1,067***	157	531*	59,753
specification	(.066)	(.059)	(392)	(357)	(308)	(307)	<30,317>
Rectangular	.141***	.177***	741***	943***	107	559**	59,753
weights	(.049)	(.037)	(278)	(218)	(241)	(245)	<30,317>
Alternative sample:	.181**	.173**	951**	907**	60	60	32,782
Not on prior TDI programs	(.073)	(.074)	(426)	(426)	(350)	(347)	<16,762>
Placebo sample:	-.026	-.043	-116	-187	115	-70	57,189
DI award 2013-2014	(.033)	(.033)	(142)	(139)	(315)	(298)	<29,046>
Alternative bandwidths:							
1 month	.117	.117	585	585	-797	-709	6,422
	(.126)	(.115)	(655)	(607)	(931)	(918)	<3,260>
2 months	.182	.182*	947	947	-151	-67	11,757
	(.112)	(.106)	(627)	(604)	(723)	(722)	<5,976>
4 months	.190**	.208***	999**	1,106**	46	277	24,972
	(.083)	(.079)	(478)	(465)	(507)	(521)	<12,673>
6 months	.172**	.181***	906**	958**	53	341	39,006
	(.071)	(.066)	(412)	(387)	(404)	(405)	<19,794>
12 months	.154***	.186***	809**	996***	140	552**	81,764
	(.056)	(.048)	(322)	(285)	(281)	(273)	<41,500>
Weighted	Yes	No	Yes	No	Yes	No	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table presents intensive margin estimates of the elasticity, the earnings response of the marginal buncher defined as the earnings response at the first kink and the average response of recipients for the baseline specification and each alternative specification. For recipients awarded DI in 2015, I add recipients to the right of the kink until the fraction of recipients with positive earnings is the same as for recipients awarded DI in 2014 in order to adjust for extensive margin responses. For the weighted estimates, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i=1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Both groups are weighted by a triangular kernel weight in estimations. Standard errors (in parentheses) are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

As I am unable to impose functional form assumptions in the assignment variable (i.e. the DI award date) due to the nature of the empirical design, a potential worry using observations further away from the cut-off arises if earnings potential is correlated with DI award due to e.g. health improving or worsening over time. To further investigate the validity of my findings, I examine whether earnings are correlated with the assignment variable. Appendix Table A.4 reports regression discontinuity estimates using no trend in the assignment variable, a linear trend and a quadratic trend, respectively.[35] This is the average effect of being subject to the more lenient benefit phase-out policy (i.e. kink at $8,000 versus $4,937) in the population. I use a triangular kernel and 9 months of bandwidth on each side of the cut-off as in the main specification. Reassuringly, the point estimates (with control variables) are very similar across the different specifications. The point estimate is $117 using no functional form in the assignment

[35] The regression can be expressed as $y_{it} = \alpha + f(x_i) + \beta I_{x_i<c} + \delta X_i + \varepsilon_{it}$ where y is earnings, x is the assignment variable (i.e. the DI award date) and c is the cut-off date at 1st of January 2015. $I_{x_i<c}$ is a dummy equal to 1 if being awarded DI before the cut-off date. X is a vector of covariates and ε is the error term. $f(x)$ takes the functional form for each specification as explained in text.

variable, \$113 with linear trends and \$108 for the quadratic specification, respectively.

Finally, I investigate how income effects would affect the elasticity estimate. As recipients awarded DI in 2015 or later receive slightly higher levels of DI benefits on average due to institutional changes, this might induce recipients awarded DI before the cut-off date to work more compared to recipients awarded DI after the cut-off through an income effect. While evidence on intensive margin responses to the benefit generosity is scarce, Gelber *et al.* (2017a) estimate an income elasticity of earnings with respect to DI benefits of about 1.[36] Using the weighting approach, the difference in DI benefits between the two groups in the estimation sample is about 1 percent. Assuming an income elasticity of 1, i.e. recipients reduce earnings by 1 percent if the level of DI benefits increase by 1 percent, a back-of-the-envelope calculation yields an elasticity of .166 or about 6 percent lower than the baseline estimate.[37] This suggests that the upward bias in the elasticity due to income effects is small in this context.[38]

6 Extensive Margin Responses

Since a lower kink point increases the average tax for recipients with earnings above the kink, it is possible that the lower kink induces some recipients to stop working altogether. In this section, I estimate the magnitude of the extensive margin response in my setting. I then investigate how extensive margin responses affect the estimate of the intensive margin earnings elasticity.

6.1 Empirical Analysis

To assess the extensive margin responses of being subject to the different benefit phase-out policies, I implement a simple regression discontinuity (RD) design. Specifically, I run the following regression:

$$y_{it} = \alpha + f(x_i) + \beta I_{x_i<c} + \delta X_i + \varepsilon_{it} \tag{5}$$

where y is the outcome variable (such as a dummy for having positive earnings), x is the assignment variable (i.e. the DI award date) and c is the cut-off date at 1st of January 2015. $f(x)$ is an unknown functional form of the assignment variable and $I_{x_i<c}$ is a dummy equal to 1 if individual i is awarded DI before the cut-off date. X is a vector of covariates and ε is the error term. β is the coefficient of interest, and measures the average effect of being subject to the more lenient benefit phase-out policy (i.e. kink at \$8,000 versus \$4,937) on labor force participation in the population. The validity of my RD design hinges on recipients not being able to manipulate the assignment variable, which I outlined in Section 4.3. I use the same baseline specifications as for the main empirical strategy using a triangular kernel, 9 months of bandwidth on each side of the cut-off and no trend in the assignment variable (i.e. $f(x) = 0$). For consistency, I also incorporate the same weighting approach as outlined in Section 4.

Table 4 reports estimates of Equation 5 for the full estimation sample. Column 1 and 2 shows that the more lenient exemption threshold increased labor force participation by about 0.8 percentage points or about 7 percent compared to recipients with the kink at \$4,937. This estimate is robust to including trends in the assignment variable, yielding a point estimate of .010 for a linear trend and .009

[36] Most evidence on the effect of benefit generosity on labor supply for DI beneficiaries investigate extensive margin responses. See e.g. Gruber (2000); Marie & Castello (2012); Deuchert & Eugster (2019).

[37] If recipients awarded DI after the cut-off decrease earnings by 1 percent, earnings of the marginal bunching individual would be $.99 \cdot (z^* + \triangle z) = .99 \cdot (4937 + 944) = \$5,822$. The estimates earnings response of the marginal bunching individual would then be $5822 - 4937 = \$885$. Plugging this into Equation 4 yields an elasticity of .166.

[38] As the general level of DI benefits in Norway is higher on average than in the US, it is possible that the income effect for Norwegian DI recipients is smaller than for US recipients . In that case, the bias represents an upper bound.

for a quadratic trend, respectively. Column 3 and 4 shows that the more lenient exemption threshold significantly decreased the average tax of participating in the labor market by about 3 percentage points, or 8 percent. In order to shed light on the magnitude of this response, I follow Kostøl & Mogstad (2014) and calculate the elasticity of labor force nonparticipation with respect to the participation tax rate.[39] The results suggest an elasticity of about 0.11 which is comparable to similar studies on DI recipients.[40] Figure A.2 (b) shows how the elasticity of labor force nonparticipation vary with bandwidth selection. Although the estimate is somewhat higher if I use 1 or 2 months of bandwidth on each side of the cut-off, the estimated elasticity appears relatively robust to bandwidth selection.

Table 4: Extensive Margin Responses and Implied Elasticity of Labor Force Nonparticipation

	Labor force participation		Participation tax rate		Nonparticipation elasticity (ε)		Observations <individuals>
Full sample	.008*	.007*	-.032***	-.032***	.106*	.104*	59,753
(2016-2017)	(.004)	(.004)	(.001)	(.001)	(.058)	(.056)	<30,317>
	[.109]	[.109]	[.401]	[.400]			
Weighted	Yes	No	Yes	No	Yes	No	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports estimates of labor force participation, the participation tax rate and the implied elasticity of non-participation using a regression where the outcome variable is regressed on a dummy which is equal to 1 if recipients are awarded DI in 2014 using a triangular kernel. The sample consists of totally disabled recipients awarded DI between 1st of April 2014 and 31st of September 2015. For the weighted estimates, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i=1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Labor force participation is defined positive earnings excluding holiday pay. The participation tax rate is defined as the implied tax of participating in the labor force including income taxes and the implicit tax of DI benefits. The elasticity of labor force nonparticipation is defined as $\varepsilon = \frac{\Delta(1-LFP)/(1-LFP)}{\Delta PTR/PTR}$ where $\Delta(1-LFP) = -\Delta LFP$ is the estimated effect on labor force nonparticipation. LFP and PTR are the mean labor force participation and participation tax rate of recipients awarded DI in 2015 (in brackets). ΔPTR is the difference in participation tax rates between the different benefit phase-out policies evaluated for the earnings distribution of recipients awarded DI in 2015. Standard errors (in parentheses) are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications.

6.2 Elasticity Estimation with Extensive Margin Responses

I now investigate how extensive margin responses would affect the estimated intensive margin earnings elasticity. As shown in Section 6.1, the lower kink induces some recipients to stop working altogether because of a higher participation tax rate for earnings above the kink. Because this only affects individuals who would have earned above the kink, the density above the kink would shift downwards. This would have a knock-on effect and shift the density below the kink upwards and shift the cumulative distribution to the left. Because of this, I would overstate the response of the marginal bunching individual, and hence also the intensive margin earnings elasticity. To assess the magnitude of this bias, I do a simulation exercise following the same steps as Ruh & Staubli (2019). I base my simulations on the utility function in Equation 2 with the addition of individuals having a fixed cost of labor force participation q that is smoothly distributed across the population. Individuals choose earnings z to maximize

[39] The elasticity of labor force nonparticipation is defined as $\varepsilon = \frac{\Delta(1-LFP)/(1-LFP)}{\Delta PTR/PTR}$, where $\Delta(1-LFP) = -\Delta LFP$ is the estimated effect of labor force nonparticipation being subject to the more lenient benefit phase-out policy, LFP is the mean labor force participation for the recipients with the (low) kink at \$4,937, ΔPTR is the difference in the participation tax rate and PTR is the average participation tax rate for recipients with the (low) kink at \$4,937.

[40] Kostøl & Mogstad (2014) estimate an elasticity of labor force nonparticipation with respect to participation taxes of about 0.12. Ruh & Staubli (2019) find an elasticity of about 0.10.

the following utility function:

$$U(c,z) = c - \frac{n}{1+\frac{1}{e}}\left(\frac{z}{n}\right)^{1+\frac{1}{e}} - q \tag{6}$$

subject to the budget constraint $c = B + z - T(z;B)$. Individuals will only participate in the labor force if $q \leq u^*(c,z) - u_0$ for some $z > 0$, where u_0 denotes the utility from nonparticipation. If the tax of participating in the labor force increases, some individuals would stop working because of the fixed cost q. This allows me to investigate how taxes create extensive margin responses and how it affects the estimated intensive margin elasticity.

The simulation exercise proceeds as follows. First, I calibrate a vector of ability parameters n that best resembles the empirical ability distribution.[41] I assume that n follows a gamma distribution because this distribution most closely resembles the empirical distribution. Appendix Figure A.3 shows that the simulated probability and cumulative distributions closely resembles the empirical distributions. Second, I assign individuals to the two different tax systems considered in this paper, one with a kink at \$4,937 and the other with a kink at \$8,000. I then calculate each individual's optimal earnings z using the estimated elasticity $e = .177$ from Section 5. Third, I calibrate a vector of fixed costs q following Liebman (2002). Specifically, I draw a random fixed cost from a uniform distribution with a lower limit of zero and an upper limit equal to the difference between the individual's utility at the optimal z with no (implicit) taxes and the utility with zero earnings. These fixed costs are then divided by a scalar so that the extensive margin response is consistent with the response estimated in the empirical analysis.

Next, I estimate the intensive margin elasticity considering the three following scenarios: In the first scenario, I assume no fixed cost of labor force participation, i.e. $q = 0$. The second scenario considers individuals with $q \geq 0$ who work only if the utility from working is larger than the utility from nonparticipation. I then estimate the elasticity as in Section 5 but ignore the adjustment procedure for extensive margin responses. In the third scenario, I estimate the elasticity incorporating the adjustment procedure outlined in Section 4. Specifically, I add individuals to the right of the kink for treated individuals until the fraction of working individuals is the same as for non-treated individuals. I assume that the distribution of recipients who have stopped working because of the fixed cost of participation is the same as the observed earnings distribution above the kink.

The results from the simulation exercise for the different cases are shown in Table 5 (column 1) and the corresponding empirical estimates (column 2). With no extensive margin responses, the elasticity is precisely estimated. If I allow for a fixed cost of labor force participation, ignoring the extensive margin response induces a large bias in the estimated elasticity of 71% in this context. However, the bias is small when I incorporate the adjustment procedure explained above. As reported in the table, the adjusted elasticity estimate is slightly smaller than the theoretical elasticity, but the bias is only 2%. This suggests that the adjustment procedure works well in this context. Appendix Figure A.4 shows a graphical representation of the estimation procedure comparing the cumulative distributions for the simulation exercise and the empirical sample, respectively. From the figures, it is clear that I overestimate the response of the marginal buncher when I do not account for extensive margin responses

[41] Under the assumption that ability, or potential earnings being the same for the two groups in the estimation sample, non-treated recipients at the same point in the CDF as treated recipients should resemble the potential earnings of treated with earnings below the kink. To construct the empirical ability distribution, i.e. the earnings distribution if there were no (implicit) taxes on earnings, I assume that ability of treated is $n = F_0^{-1}(z)$ when $F_1 = F_0$ for treated with earnings $z \leq z^*$, where $F_j(z)$ is the cumulative distribution for treatment status $j = 0, 1$. For treated with earnings $z^* > z$, I use the estimated elasticity $e = .177$ to calculate the earnings response which yields ability $n = z(\frac{1-t_0}{1-t_1})^e$.

as the CDF for the treated group has shifted to the left when some individuals above the kink have stopped working.

Table 5: Simulation Exercise and Adjustment for Extensive Margin Responses

	Simulation	Empirical
Without extensive margin responses	.177 (0%)	
With extensive margin responses		
unadjusted	.303 (71%)	.315
adjusted	.173 (-2%)	.177

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table shows intensive margin elasticity estimates with and without adjustment for extensive margin responses for the simulated and empirical earnings distributions. For the simulated earnings distributions (see text for details), I assume an elasticity of $e = .177$. In the first scenario, I assume no extensive margin responses. In the second scenario, I calibrate a fixed cost of labor force participation that resembles the estimated empirical extensive margin response. For the unadjusted estimates, I calculate the intensive margin elasticity using observations (individuals) with positive earnings only. For estimates adjusted for extensive margin responses, I add observations (individuals) to the right of the kink until the fraction of observations (individuals) with positive earnings is the same between the two samples. In parentheses, I report the bias in the estimated elasticity relative to the theoretical elasticity.

7 Elasticities using Bunching Methods

In most studies that use bunching at kinks to identify a behavioral response, the counterfactual distribution, i.e. what the distribution would have looked like without the kink, is unobserved. As pointed out by Blomquist *et al.* (2019), the amount of bunching at the kink is not informative about responses unless one is willing to impose restrictions on the counterfactual distribution. The common way to deal with this issue in the literature is to use the observed density to estimate the counterfactual density using a flexible polynomial. As a result, identification of the behavioral response depends on the assumed shape of the counterfactual distribution. However, without information on the true counterfactual density, it is not clear whether the polynomial approach provides a valid estimate of the counterfactual distribution, and therefore whether it provides a valid estimate of the behavioral responses.

To shed light on this matter, I re-estimate the earnings elasticity for the sample of recipients with the kink at \$4,937 using bunching methods common in the literature. My first approach follows Chetty *et al.* (2011) and fits a flexible polynomial to the observed distribution to estimate the amount of bunching around the kink. Specifically, I group individuals into earnings bins of \$400 and estimate a regression of the following form:

$$c_j = \sum_{i=0}^{p} \beta_i (z_j)^i + \sum_{k=z^L}^{z^U} \gamma_k \mathbf{1}(z_j = k) + \varepsilon_j \tag{7}$$

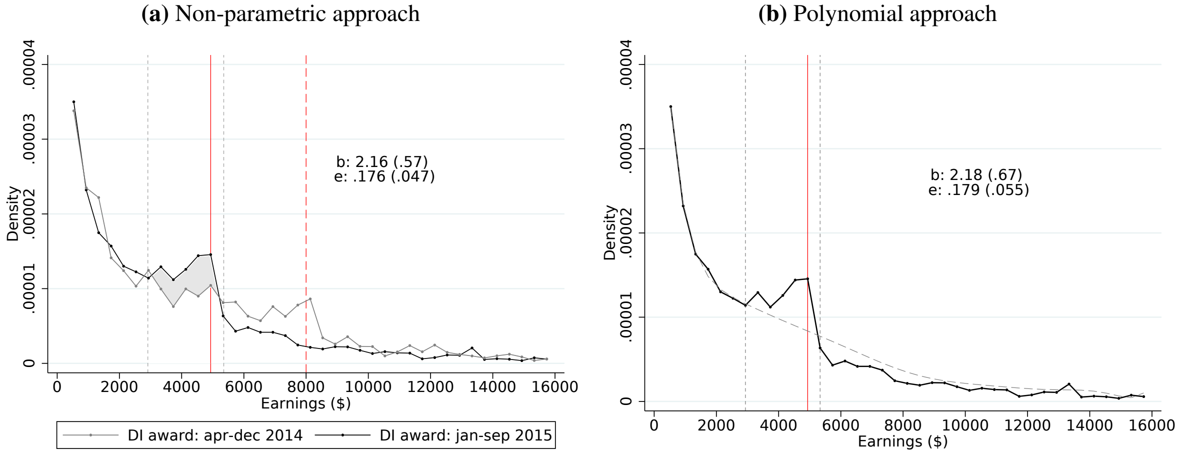
where c_j is the number of individuals in bin j, z_j is the earning level of bin j, z^L and z^U is the lower and upper limit of the excluded range around the kink and p is the order of the polynomial. The counterfactual density is obtained as the predicted values from Equation 7 omitting the contribution of the dummies in the excluded range, i.e. $\hat{c}_j = \sum_{i=0}^{p} \hat{\beta}_i (z_j)^i$. This density is then adjusted so that the estimated missing mass above the kink is equal to the estimated bunching mass around the kink.[42] The estimated amount of "bunching", or excess mass around the kink is then determined by the sum of the predicted values of the dummies, i.e. $\sum_{k=z^L}^{z^U} \hat{\gamma}_k \mathbf{1}(z_j = k)$. This estimate is then normalized by the estimated density

[42]This is achieved by upward shifts in the counterfactual distribution to the right of the kink, which is done in increments until the counterfactual satisfies the integration constraint.

at the kink. Multiplying this estimate with the binwidth obtains an estimate of the earnings response of the marginal buncher with the same interpretation as in my main estimation approach. Plugging this into Equation 4 yields the estimated earnings elasticity using the polynomial approach. As a comparison to the polynomial approach, I estimate the amount of “bunching” and the earnings elasticity using the same framework, but instead using the sample of recipients awarded DI before the cut-off as the counterfactual density. This should give an indication of how well the polynomial approach works in this context.

Figure 6 provides a graphical representation of the two approaches that I consider. In Panel (a), the gray area indicates the estimated amount of “bunching” using the non-parametric approach where recipients with DI award before the cut-off serve as the counterfactual density for recipients awarded DI after the cut-off. The vertical gray dashed lines indicate the lower and upper limits of the excluded region and are determined by visual inspection. The estimated amount of “bunching” (b) amounts to 2.16 and implies that the excess mass around the kink indicated by the gray area amounts to 2.16 of the counterfactual density at the kink. The estimated elasticity ($e = .176$) is almost indistinguishable from the estimated elasticity in my main empirical approach using the cumulative distributions of the two groups ($e = .177$). Panel (b) shows the estimated “bunching” and elasticity using the polynomial approach where I use the same excluded region as in the non-parametric approach. The dashed line indicates the estimated counterfactual density using a 10th degree polynomial fitted to the empirical distribution. The estimated “bunching” and elasticity ($e = .179$) are remarkably similar to the non-parametric approach suggesting that the polynomial approach provides a valid estimate of earnings responses in this context.

Figure 6: Graphical Representation of Bunching Estimates



Notes: Panel (a) illustrates the bunching estimate using the non-parametric approach, where bunching is estimated using the earnings distribution of recipients awarded DI between April 1st and December 31st 2014 (gray line) as a counterfactual density for recipients awarded DI between January 1st and September 30th 2015 (black line), both distributions using \$400 bins. The vertical dashed lines indicate the bunching region where the estimated bunching is illustrated by the gray area. The red dashed line and the red solid line indicate the kink point in the budget constraint for each group. Recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i = 1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Both groups are weighted by a triangular kernel weight. Panel (b) illustrates the bunching estimate using the polynomial approach, where a 10th degree polynomial is fitted to the empirical distribution of recipients awarded DI between January 1st and September 30th 2015, illustrated by the gray dashed line. Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

Table 6 reports estimates of the elasticity (e), bunching (b) and the earnings response of the marginal buncher $(\triangle z)$ for each approach. Note that the polynomial approach is unaffected by the weighting approach as it only uses data on recipients to the right of the cut-off. In addition to the approaches discussed in this section, I estimate the earnings elasticity following Saez (2010) which relies on a linear approximation of the counterfactual density around the kink.[43] While this elasticity estimate is slightly higher than in the main empirical approach, it is qualitatively similar. For the polynomial approach, the estimated responses are based on a 10th degree polynomial fitted to the empirical distribution. As in most bunching applications, it is not clear how to decide the order of the polynomial in this setting. I therefore perform a robustness analysis and estimate responses using alternative orders of the polynomial fitted to the empirical distribution. Appendix Table A.5 reports estimates of "bunching" and the earnings elasticity using polynomials of order 8 - 12, and Appendix Figure A.5 provides a graphical representation.[44] The estimated elasticity range from .150 to .194 suggesting that the estimates are fairly robust to alternative specifications of the counterfactual density.

Table 6: Earnings Elasticity Estimates from Bunching Methods

	Elasticity (e)		Bunching (b)		Earnings response ($) Marginal buncher (Δz)		Observations <individuals>
CDF method	.177***	.198***	2.36**	2.67***	944**	1,067***	59,753
	(.066)	(.059)	(.98)	(.89)	(392)	(357)	<30,317>
Bunching methods:							
Counterfactual:	.176***	.184***	2.16***	2.26***	863***	902***	59,753
Awarded DI in 2014	(.047)	(.045)	(.58)	(.55)	(230)	(218)	<30,317>
Counterfactual:	.179***	.179***	2.18***	2.18***	873***	873***	26,950
Fitted polynomial	(.055)	(.055)	(.67)	(.67)	(267)	(267)	<13,697>
Saez method	.190***	.190***	2.32***	2.32***	928***	928***	26,950
	(.051)	(.051)	(.63)	(.63)	(250)	(250)	<13,697>
Weighted	Yes	No	Yes	No	Yes	No	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports intensive margin estimates of the elasticity, bunching and the earnings response of the marginal buncher defined as the earnings response at the first kink for four different methods (see text for details). The sample consists of totally DI recipients awarded DI between 1st of April 2014 and 30th of September 2015. For the weighted estimates, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i = 1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Both groups are weighted by a triangular kernel weight in all estimations. For the specification using a fitted polynomial, I use a 10th degree polynomial fitted to the empirical density using the sample of recipients awarded DI in 2015. Standard errors (in parentheses) are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

A caveat with this exercise is that even though my analysis suggests that the polynomial approach works well in this context, the same may not be true in other settings. In general, the performance of the polynomial approach will depend on choices made by the researcher, and in particular the order of the polynomial and the upper and lower limit of the excluded region. Moreover, the approach will

[43] The elasticity is derived from Equation (5) in Saez, 2010 which can be solved explicitly for the elasticity e: $B = z^*[T-1]\frac{h(z^*)_- + h(z^*)_+/T}{2}$ where $T = (\frac{1-t_0}{1-t_1})^e$. B denotes the estimated fraction of "bunching" in the population, z^* is the kink point and $h(z^*)_-$ and $h(z^*)_+$ denote the estimated densities just below and just above the kink, respectively.

[44] Based on visual inspection, orders lower than 8 clearly underfits the empirical distribution in my setting.

depend on the size of the kink and the size of the earnings response. If bunching is less sharp, deciding the excluded region is often not clear. Additionally, one often relies on using observations further away from the kink to estimate the counterfactual density. The predicted density then may serve as a poor counterfactual density around the kink.

8 Fiscal Effects and Policy Implications

In this section, I assess fiscal effects of different benefit offset policies for the government and program's beneficiaries, and the associated policy implications. While a more lenient policy may improve welfare of recipients and increase tax revenues, it might increase expenditures on DI benefits. It could also increase program inflow as the program becomes more desirable for potential applicants. Therefore, it is not clear how such policies should be designed and how different policies affect total program costs.

To shed light on these matters, I examine implications of two alternative policies on recipients' disposable income, DI benefits paid, income taxes and net public expenditures. The alternative policies are compared to the baseline policy with the kink at \$4,937. The first policy I consider is relaxing the exemption threshold to \$8,000 which is the temporary policy for recipients awarded DI in 2014 or earlier. To calculate effects of this policy, I estimate Equation 5 for each outcome that I consider for the main estimation sample. The second policy I consider is abolishing the phase-out of DI benefits entirely. Under this policy, recipients would keep full DI benefits regardless of how much they earn. In both scenarios, recipients would still have to pay regular income taxes. To calculate effects of this policy, I decompose responses into intensive and extensive margin responses. For the intensive margin response, I estimate responses in the same way as in the main empirical approach for recipients with earnings below the kink.[45] For recipients with earnings above the kink, I calculate responses using the estimated intensive margin elasticity ($e = .177$).[46] To calculate the extensive margin response, I use the estimated elasticity of labor force nonparticipation from Section 6 and calculate the number of recipients who would start working under the alternative policy.[47] I assume that the earnings distribution of additional working recipients is the same as the observed earnings distribution above the kink. Lastly, I calculate the changes in recipients' disposable income, taxes and government expenditures based on the earnings responses.

The results from the two alternative policy changes are reported in Table 7. The two first columns show that relaxing the kink from \$4,937 to \$8,000 significantly increases recipients' disposable income with \$83 on average for the weighted approach. This effect is mainly driven by working recipients who increase labor supply, while a few recipients start working under the more lenient policy. Most recipients do not experience increased disposable income as they do not work under either policy. While estimated government expenditures on DI benefits increase with \$15 on average per recipient, this effect is offset by an increase in income taxes by \$42. Because of this, estimated net government expenditures decrease with \$27 per recipient. However, this effect is too imprecisely estimated to draw firm conclusions. Column 3 and 4 show that abolishing the phase-out of DI benefits entirely increases disposable income

[45] The earnings response can be expressed as $F_0(z + \triangle z) = F_1(z)$ where $\triangle z$ is the earnings response, F_0 is the CDF of recipients with DI award before the cut-off, and F_0 the CDF of recipients awarded DI after the cut-off.

[46] The earnings response can be expressed as $\triangle z = z(\frac{1-t_0}{1-t_1})^e - z$ where $\triangle z$ is the earnings response, z is current earnings, t_0 is the marginal tax rate below the kink and t_1 is the marginal tax rate above the kink.

[47] More specifically, the change in labor force participation is calculated as $\triangle LFP = -\varepsilon \frac{\triangle PTR}{PTR}(1 - LFP)$ where $\triangle PTR$ denotes the policy-induced reduction in participation tax rate.

by \$183 which is more than twice as much as the first alternative policy. DI benefits paid and income taxes also increase more. Although net government expenditures are estimated to be \$22 lower for each recipient than under the current policy, they are slightly higher than under the first alternative policy. Again, I lack precision to draw firm conclusions.

Table 7: Annual Fiscal Effects of Increased Incentives to Work

Outcome:	Relax kink		Abolish kink	
Disposable income (\$)	83***	100***	183***	191***
	(17)	(17)	(26)	(24)
DI benefits (\$)	15	1	50	51
	(11)	(11)	(122)	(109)
Payroll taxes (\$)	42***	52***	72***	76***
	(12)	(12)	(15)	(14)
Net expenses (\$)	-27	-51**	-22	-26
	(22)	(24)	(22)	(22)
Implied elasticity of induced entry	.07	.09	.01	.01
Weighted	Yes	No	Yes	No
Individuals	30,317		30,317	
Observations	59,753		59,753	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports estimates of alternative benefit phase-out policies on annual disposable income, DI benefits, payroll taxes and net public expenditures. The first alternative policy considers relaxing the annual kink from \$4,937 to \$8,000. Column 1 and 2 report estimates using a regression where outcome variables are regressed on a dummy which is equal to 1 if recipients are awarded DI in 2014. Columns 3 and 4 report estimates of abolishing the DI phase-put policy entirely (see text for details). In all estimations, I use the full sample of totally disabled recipients awarded DI between 1st of April 2014 and 31st of September 2015 and a triangular kernel. For the weighted estimates, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)}\frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i=1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Standard errors (in parentheses) are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. All variables are measured in 2016 dollars (NOK/USD = 7.5).

It is important to keep in mind that the estimated effects are specific to my estimation sample of recipients with only 1-4 years on DI receipt. Therefore, it is possible that the alternative policies would have different implications for the full sample of DI recipients. As shown in Table 1, the full sample of recipients differ in some dimensions compared to the estimation sample. In particular, they are slightly older, have spent more years on DI and have about 50% higher earnings on average. It is therefore likely that the alternative policies would have larger impacts for the full sample of recipients. Additionally, it is possible that the long-run earnings responses, which my fiscal calculations are based upon, are higher than the observed responses if responses are attenuated by e.g. lumpy hours or imperfect information about the benefit phase-out. In that case, my estimates represent a lower bound of the true effects.

While my exercise shows that the alternative policies might decrease program costs for current recipients, these estimates ignore the possibility that a more generous DI program could induce more program entry.[48] To shed light on this matter, I calculate how elastic program inflow would have to be in order to not increase program costs. Specifically, I calculate the elasticity of induced entry as in Kostøl

[48] A more generous DI program could also lead to fewer program exits by current beneficiaries in the long run. However, this effect is likely to be small because the exit rate from DI is already very low.

& Mogstad (2014), defined as the percentage increase in the number of DI recipients relative to the percentage increase in disposable income as a DI recipient.[49] Table 7 shows that relaxing the kink yields an induced entry elasticity of about 0.07, while abolishing the disincentives to work entirely yields a very low elasticity of about 0.01. These calculations suggest that both programs are likely to increase program costs.[50]

9 Conclusion

In this paper, I have examined recipients' labor supply responses to the financial incentives induced by the benefit phase-out. Using earnings distributions in a local experiment that assigned recipients to different benefit offset policies, I provide transparent and credible identification of labor supply responses of DI beneficiaries. I find evidence of large behavioral responses around the exemption threshold suggesting that working recipients would have earned considerably more if benefits were not phased out above the threshold. I also find that recipients subject to the higher exemption threshold are more likely to participate in the labor force. My framework is also useful for understanding responses to kinked budget sets. My findings suggest that common estimation strategies in the literature that identify behavioral responses using bunching at kinks in the budget set performed well in this context.

As my study investigates recipients of the Norwegian DI program, one needs to exercise caution in applying these findings to other countries. In particular, the exemption threshold in Norway is lower than in most other countries. This difference is important as my main estimation strategy exploits recipients who locate around the threshold. Furthermore, I advise readers to exercise the usual caution in interpreting findings from a local experiment. In this context, the study considers recipients who have entered the DI program fairly recently and have a lower earnings capacity compared to recipients with longer spells on the program. Therefore, it is likely that responses are larger for the full population of DI recipients.

The estimated labor supply responses are particularly useful for guiding policymakers in how different benefit offset policies will affect recipients' disposable income and program costs. My findings indicate that relaxing the exemption threshold increases disposable income and reduces costs for current recipients of the program. A caveat with this study is that it is not informative about the level of increased program inflow when recipients are allowed to keep a larger share of their benefits as they earn more. I do, however, calculate the size of induced entry that has to be generated by more generous benefit offset policies in order to increase program costs. Based on findings in other studies, I conclude that more generous policies will likely increase public expenditures.

[49] The elasticity is defined as $\varepsilon_{entry} = \frac{E(\triangle NE)/E(B|Award=1)}{P(Award=1)\cdot E(\triangle I|Award=1)}$ where $\triangle NE$ is the change in net government expenditures and $\triangle I$ is the change in disposable income between the current policy and the alternative policy.

B is DI benefits. I assume that new entries have the same earnings distribution and DI benefits as recipients with earnings above the kink as the alternative programs gives no further incentives for entries who would earn below the kink. Furthermore, I assume a probability of award of 0.85 which is roughly the award rate in the Norwegian DI program.

[50] While the literature on induced entry of DI recipients is somewhat inconclusive, Gruber (2000) reports induced entry elasticities in the range of 0.28-0.36 in Canada. Mullen & Staubli (2016) and Hoynes & Moffitt (1999) report an elasticity of about 1.2 in Austria and the US, respectively. In contrast, Campolieti & Riddell (2012) and Castello (2017) do not find any evidence of induced entry in Canada and Spain, respectively.

References

ATKINSON, A., RAINWATER, L., & SMEEDING, T.M. 1995. Income Distributions in OECD countries: Evidence from the Luxembourg Income Study. *OECD Publications and Information Center*.

BASTANI, SPENCER, & SELIN, HÅKAN. 2014. Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, **109**(Jan.), 36–49.

BEST, MICHAEL CARLOS, & KLEVEN, HENRIK JACOBSEN. 2017. Housing market responses to transaction taxes: Evidence from notches and stimulus in the UK. *The Review of Economic Studies*, **85**(1), 157–193.

BLOMQUIST, SÖREN, NEWEY, WHITNEY, KUMAR, ANIL, & LIANG, CHE-YUAN. 2019. On Bunching and Identification of the Taxable Income Elasticity. *NBER Working Paper 24136*.

BORGHANS, LEX, GIELEN, ANNE C., & LUTTMER, ERZO F. P. 2014. Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in a disability Insurance Reform. *American Economic Journal: Economic Policy*, **6**(4), 34–70.

BOUND, J. 1989. The Health and Earnings of Rejected Disability Insurance Applicants. *American Economic Review*, **79**(3), 482–503.

BUTLER, MONIKA, LECHNER, MICHAEL, DEUCHERT, EVA, STAUBLI, STEFAN, & THIEMANN, PETRA. 2015. Financial Work Incentives for Disability Benefit Recipients: Lessons from a Randomized Experiment. *IZA Journal of Labor Policy*, **4**(18), 1–18.

CALONICO, SEBASTIAN, CATTANEO, MATIAS D., & TITIUNIK, ROCIO. 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, **82**(6), 2295–2326.

CAMPOLIETI, MICHELE, & RIDDELL, CHRIS. 2012. Disability Policy and the labor market: Evidence from a natural experiment in Canada, 1998-2006. *Journal of Public Economics*, **96**, 306–316.

CASTELLO, JUDIT VALL. 2017. What happens to the employment of disabled individuals when all financial incentives to work are abolished? *Health Economics*, **26**(S2), 158–174.

CHEN, S., & VAN DER KLAAUW, W. 2008. The work disincentive effects of the disability insurance program in the 1990s. *Journal of Econometrics*, **142**(2), 757–784.

CHETTY, RAJ, FRIEDMAN, JOHN, OLSEN, TORE, & PISTAFERRI, LUIGI. 2011. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, **126**(2), 749–804.

DAHL, GORDON B., KOSTØL, ANDREAS RAVNDAL, & MOGSTAD, MAGNE. 2014. Family Welfare Cultures. *The Quarterly Journal of Economics*, **129**(4), 1711–1752.

DEUCHERT, EVA, & EUGSTER, BEATRIX. 2019. Income and substitution effects of a disability insurance reform. *Journal of Public Economics*, **170**(Feb.), 1–14.

EINAV, LIRIAN, FINKELSTEIN, AMY, & SCHRIMPF, PAUL. 2017. Bunching at the kink: Implications for spending responses to health insurance contracts. *Journal of Public Economics 146 (2017) 27-40*.

FRANDSEN, BRIGHAM R. 2017. *Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete*. Advances in Econometrics, vol. 38. Emerald Publishing Ltd. Chap. Regression Discontinuity Designs, pages 281–315.

FRENCH, ERIC, & SONG, JAE. 2014. The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy*, **6**(2), 291–337.

GELBER, ALEXANDER, MOORE, TIMOTHY J., & STRAND, ALEXANDER. 2017a. The Effect of Disability Insurance Payments on Beneficiaries Earnings. *American Economic Journal: Economic Policy*, **9**(3), 229–261.

GELBER, ALEXANDER M., JONES, DAMON, SACKS, DANIEL W., & SONG, JAE. 2017b. Using Non-Linear Budget Sets to Estimate Extensive Margin Responses: Method and Evidence from the Social Security Earnings Test. *NBER Working Papers 23362, National Bureau of Economic Research, Inc*.

GELBER, ALEXANDER M., JONES, DAMON, SACKS, DANIEL W., & SONG, JAE. 2020a. The Employment Effects of the Social Security Earnings Test. *Working Papers 2020-05, Becker Friedman Institute for Research in Economics*.

GELBER, ALEXANDER M., JONES, DAMON, & SACKS, DANIEL W. 2020b. Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test. *American Economic Journal: Applied Economics*, **12**(1), 1–31.

GRUBER, JONATHAN. 2000. Disability Insurance Benefits and Labor Supply. *Journal of Political Economy*, **108**(6), 1162–83.

HARASZTOSI, PÉTER, & LINDNER, ATTILA. 2019. Who Pays for the minimum Wage? *American Economic Review*, **109**(8), 2693–2727.

HOYNES, HILARY WILLIAMSON, & MOFFITT, ROBERT. 1999. Tax Rates and Work Incentives in the Social Security Disability Insurance Program: Current Law and Alternative Reforms. *National Tax Journal*, **52**(4), 623–54.

KLEVEN, HENRIK. 2016. Bunching. *Annual Review of Economics, Annual Reviews*, **8**(1), 435–464.

KLINE, PATRICK. 2011. Blinder-Oaxaca as a reweighting estimator. *American Economic Review: Papers and Proceedings*, **101**(3), 532–537.

KOPCZUK, WOJCIECH, & MUNROE, DAVID. 2015. Mansion Tax: The Effect of Transfer Taxes on the Residential Real Estate Market. *American Economic Journal: Economic Policy*, **7**(2), 214–257.

KOSTØL, ANDREAS RAVNDAL, & MOGSTAD, MAGNE. 2014. How Financial Incentives Induce Disability Insurance Recipients to Return to Work. *American Economic Review*, **104**(2), 624–55.

LE BARBANCHON, THOMAS. 2016. The effect of the potential duration of unemployment benefits on unemployment exits to work and match quality in France. *Labour Economics*, **42**(C), 16–29.

LEE, DAVID S., & LEMIEUX, THOMAS. 2010. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, **48**(June), 281–355.

LIEBMAN, JEFFREY B. 2002. The Optimal Design of the Earned Income Tax Credit. *Making Work Pay: The Earned Income Tax Credit and Its Impact on American Families*, **edited by Bruce D. Meyer and Douglas Holtz-Eakin**, 196–234. New York: Russel Sage Foundation.

MAESTAS, NICOLE, MULLEN, KATHLEEN J., & STRAND, ALEXANDER. 2013. Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt. *American Economic Review*, **103**(5), 1797–1829.

MANOLI, DAY, & WEBER, ANDREA. 2016. Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions. *American Economic Journal: Economic Policy*, **8**(4), 160–182.

MARIE, OLIVIER, & CASTELLO, JUDIT VALL. 2012. Measuring the (income) effect of disability insurance generosity on labour market participation. *Journal of Public Economics*, **96**(1-2), 198–210.

MOORE, TIMOTHY J. 2015. The Employment Effects of Terminating Disability Benefits. *Journal of Public Economics*, **124**, 30–43.

MULLEN, KATHLEEN J., & STAUBLI, STEFAN. 2016. Disability Benefit Generosity and Labor Force Withdrawal. *Journal of Public Economics*, 49–63.

NAV. 2018. Nav-ytelsene frem mot 2060 - En forenklet analyse av store penger. *NAV-rapport nr. 1-2018*.

OECD. 2010. *Sickness, Disability and Work: Breaking the Barriers. A synthesis of findings across OECD countries*. OECD Publishing, Paris.

PAETZOLD, JOERG. 2019. How do taxpayers respond to a large kink? Evidence on earnings and deduction behavior from Austria. *International Tax and Public Finance*, **26**(1), 167–197.

RUH, PHILIPPE, & STAUBLI, STEFAN. 2019. Financial Incentives and Earnings of Disability Insurance Recipients: Evidence from a Notch Design. *American Economic Journal: Economic Policy*, **11**(2), 269–300.

SAEZ, EMMANUEL. 2010. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, **2**(3), 180–212.

SCHIMMEL, JODY, STAPELTON, DAVID C., & SONG, JAE G. 2011. How Common is 'Parking' among Social Security Disability Insurance Beneficiaries? Evidence from the 1999 Change in the Earnings Level of Substantial Gainful Activity. *Social Security Bulletin*, **71**(4), 77–92.

SINGLETON, PERRY. 2012. Earnings of rejected applicants to the Social Security Disability Insurance program. *Economics Letters*, **116**(2), 147–150.

TRAXLER, CHRISTIAN, WESTERMAIER, FRANZ G., & WOHLSCHLEGEL. 2018. Bunching on the Autobahn? Speeding responses to a 'notched' penalty scheme. *Journal of Public Economics*, **157**(Jan.), 78–94.

WEATHERS, ROBERT R., & HEMMETER, JEFFREY. 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, AREZOU. 2020. Adjustment Costs and Incentives to Work: Evidence from a Disability Insurance Program. *Journal of Public Economics*, **188**(104223).

Appendix: Additional Tables and Figures

Table A.1: Balancing Tests of Pre-determined Covariates

Dependent variable:	Full sample (2 months around cut-off)			Alternative sample: Not from prior TDI programs (9 months around cut-off)		
	difference	std. error	p-value	difference	std. error	p-value
Number of recipients	-86	(307)	.778	-155	(122)	.206
Characteristics:						
Age at DI award	-.67*	(.37)	.072	-.27	(.25)	.282
Fraction females	.002	(.014)	.902	-.004	(.009)	.634
Years of schooling	-.18***	(.06)	.005	-.06	(.04)	.108
Pre-disability earnings ($)	-225	(468)	.631	89	(301)	.766
Projected DI benefits ($)	-169	(234)	.471	-5	(147)	.971
Years since onset date	-.12	(.12)	.320	-.08	(.07)	.232
Fraction cohabitants	-.005	(.014)	.745	-.012	(.009)	.168
Number of children	.01	(.04)	.755	.008	(.025)	.759
Fraction from TDI program	-.013	(.009)	.125	-.016**	(.007)	.013
Fraction from prior TDI programs	-.015	(.013)	.253	-	-	-
Joint test			.176			.194
Observations		5,976			16,762	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports results of a regression using a triangular kernel where each variable is regressed on a dummy which is equal to 1 if recipients are awarded DI in 2014. The full sample consist of totally disabled recipients awarded DI between 1st of November 2014 and 28th of February 2015. The alternative sample consists of recipients awarded DI between 1st of April 2014 to 30th of September 2015 and excludes recipients who received temporary DI benefits before a reform in the temporary DI program in March 2010. Projected DI benefits are defined as the level of DI benefits if recipients were awarded DI after 1st of January 2015. Years of schooling, cohabitant status and number of children are measured in 2015. All other covariates are fixed over time. Standard errors (in parentheses) are clustered at the individual level and are robust to heteroskedasticity. Monetary variables are measured in 2016 dollars (NOK/USD = 7.5).

Table A.2: Balancing tests of Pre-determined Covariates: Weighted Estimates

Dependent variable:	mean: DI award apr-dec 2014		mean: DI award jan-sep 2015	difference (std. error)				p-value	
Net uncapped DI benefits ($)	27,647	28,405	28,741	-1,095***	(68)	-337***	(69)	<.001	<.001
Characteristics:									
Age at DI award	46.68	47.34	47.02	-.34**	(.17)	.32*	(.19)	.048	.085
Fraction females	.550	.531	.531	.019***	(.007)	.000	(.007)	.004	.962
Years of schooling	10.93	11.01	11.01	-.08***	(.03)	.00	(.03)	.008	.998
Pre-disability earnings ($)	59,557	60,304	60,224	-668***	(214)	80	(254)	.002	.753
Years since DI onset date	6.89	6.12	6.12	.77***	(.06)	.00	(.06)	<.001	.948
Fraction cohabitants	.504	.502	.507	-.003	(.007)	-.005	(.007)	.626	.457
Number of children	1.65	1.60	1.62	.03	(.02)	-.02	(.02)	.119	.314
Fraction from TDI program	.913	.887	.890	.022***	(.004)	-.003	(.005)	<.001	.436
Fraction from prior TDI programs	.498	.303	.305	.193***	(.006)	-.002	(.006)	<.001	.709
Joint test (p-value)								<.001	.429
Weighted	No	Yes	No	No		Yes		No	Yes
Observations	16,620	16,620	13,697	30,317					

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports weighted and unweighted means of recipients awarded DI before/after 1st of January 2015, the difference between the weighted and unweighted means and the corresponding p-values. The sample consist of totally disabled recipients awarded DI between 1st of April 2014 and 31st of September 2015. Years of schooling, cohabitant status and number of children are measured in 2015. All other covariates are constant over time. Means and the corresponding differences are weighted by a triangular kernel. For the weighted means and differences, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i = 1|x_i)$ is estimated with a logit model using the covariates in Table (including uncapped DI benefits) as control variables. Standard errors (in parentheses) are clustered at the individual level and are robust to heteroskedasticity. Monetary variables are measured in 2016 dollars (NOK/USD = 7.5).

Table A.3: Subsample Analysis of Earnings Responses

			Earnings response ($)				Observations
	Elasticity (e)		Marginal buncher (Δz)		Average response		<individuals>
Full sample	.177***	.198***	944**	1,067***	157	531*	59,753
(2016-2017)	(.066)	(.059)	(392)	(357)	(308)	(307)	<30,317>
Age							
18-49	.115*	.182***	606	999***	242	720*	27,029
	(.069)	(.064)	(392)	(383)	(407)	(422)	<13,960>
50-66	.254***	.295***	1,379***	1,634***	0	458	21,845
	(.091)	(.076)	(536)	(460)	(492)	(461)	<11,896>
Gender							
Male	.181**	.194***	946**	1,021**	153	740	26,839
	(.078)	(.074)	(446)	(431)	(488)	(527)	<13,600>
Female	.185*	.199**	1,022*	1,106**	232	335	32,914
	(.097)	(.084)	(586)	(514)	(387)	(345)	<16,717>
Education							
High	.155	.176**	804	921*	331	759	24,133
	(.098)	(.084)	(551)	(482)	(483)	(501)	<12,244>
Low	.179**	.218***	972**	1,207***	-70	348	35,620
	(.076)	(.072)	(459)	(448)	(413)	(411)	<18,073>
DI benefits							
High	.257***	.245***	1,466***	1,385***	362	624	35,165
	(.077)	(.067)	(486)	(422)	(421)	(398)	<17,912>
Low	.151	.158*	766	804	214	371	24,588
	(.096)	(.090)	(526)	(496)	(490)	(480)	<12,405>
Pre-DI earnings							
High	.189**	.184**	931*	904**	299	865*	26,429
	(.093)	(.077)	(496)	(416)	(492)	(520)	<13,482>
Low	.179**	.232***	1,034**	1,384***	-107	209	33,324
	(.080)	(.065)	(494)	(418)	(387)	(358)	<16,835>
Weighted	Yes	No	Yes	No	Yes	No	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table presents intensive margin estimates of the elasticity, the earnings response of the marginal buncher defined as the earnings response at the first kink and the average response of recipients for the full sample and by each subgroup. The sample consist of totally DI recipients awarded DI between 1st of April 2014 and 30th of September 2015. For recipients awarded DI in 2015, I add recipients to the right of the kink until the fraction of recipients with positive earnings is the same as for recipients awarded DI in 2014 in order to adjust for extensive margin responses. For the weighted estimates, recipients awarded DI in 2014 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2015 and $P(I_i=1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Both groups are weighted by a triangular kernel weight in estimations. Low education is defined as not finishing high school or less. High education is defined as high school education or more. Low (uncapped) DI benefits is defined as receiving minimum benefit levels, and high DI benefits otherwise. Low (high) pre-DI earnings are defined as less than or equal to (larger than) the sample median. Standard errors (in parentheses) are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5)

Table A.4: Regression Discontinuity Estimates

Outcome:	Trends in running variable (DI award)						Observations
	None		Linear		Quadratic		<individuals>
Annual	117***	151***	113	117	108	112	59,753
earnings ($)	(39)	(39)	(85)	(86)	(85)	(85)	<30,317>
	[531]	[513]	[533]	[530]	[560]	[552]	
Controls	Yes	No	Yes	No	Yes	No	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports results of regression discontinuity estimates using a triangular kernel for different specifications of the running variable: No trend (differences in means), a common linear trend and a common quadratic trend. The sample consist of totally disabled recipients awarded DI between 1st of April 2014 and 31st of September 2015. Controls include the variables in Table A.2, including (uncapped) DI benefits and the implicit tax rate on DI benefits (equal to the benefit replacement rate). Results reports the coefficient of the dummy which is equal to 1 if recipients are awarded DI in 2014. Standard errors (in parentheses) are clustered at the individual level and are robust to heteroskedasticity. Dependent means in brackets. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

Table A.5: Parametric Bunching Estimates for Alternative Orders of Polynomial

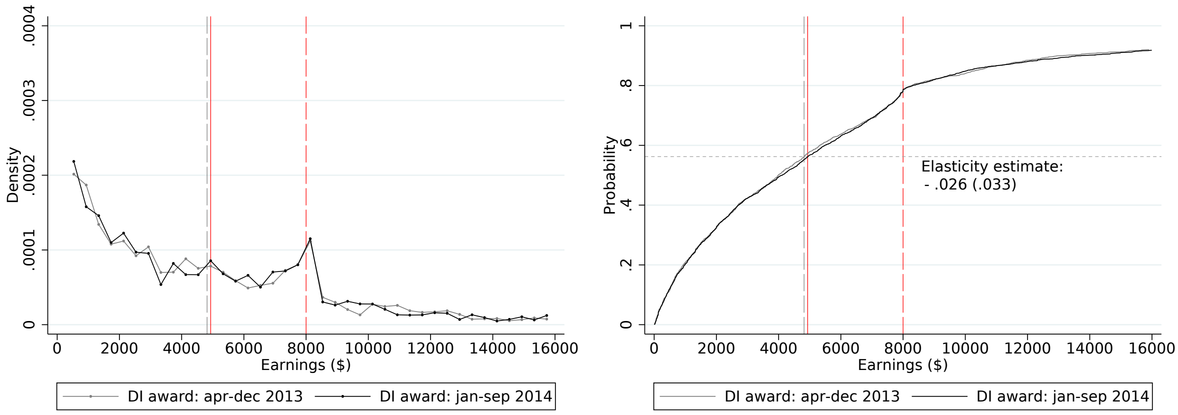
Order of polynomial:	8	9	10	11	12	Observations <individuals>
Bunching (b)	1.83***	1.97***	2.18***	2.17***	2.37***	59,753
	(.51)	(.53)	(.67)	(.64)	(.78)	<30,317>
Elasticity (e)	.150***	.161***	.179***	.177***	.194***	
	(.042)	(.044)	(.055)	(.053)	(.064)	

*** significant at 1% level, ** significant at 5% level, * significant at 10% level

Notes: The table reports bunching and elasticity estimates from the parametric bunching method (see Section 7 for details) using alternative orders of polynomial fitted to the empirical distribution. In the baseline specification I use a 10th order polynomial. The sample consist of recipients awarded DI between January 1st and September 30th 2015 and weighted by a triangular kernel (assigning more weight to the individuals awarded DI early in the year). Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications.

Figure A.1: Placebo Elasticity Estimates: Recipients Awarded DI 2013-2014

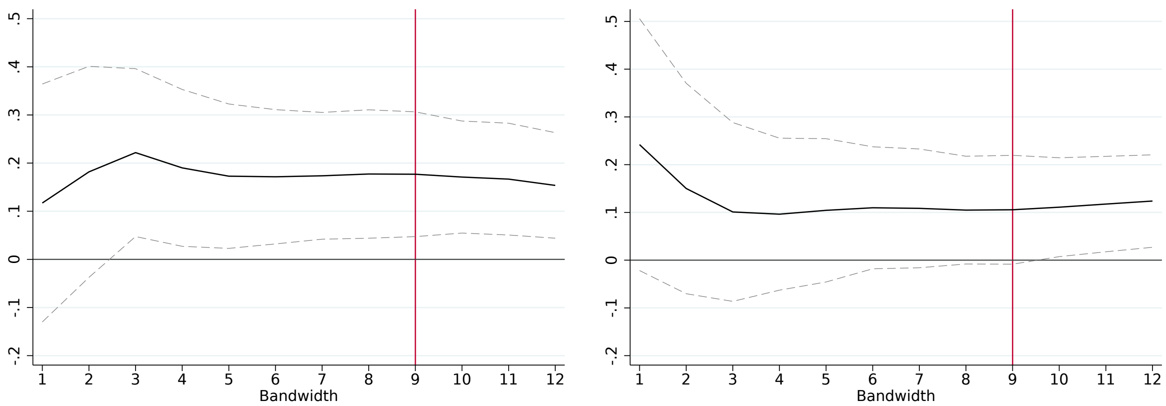
(a) Probability distributions **(b)** Cumulative distributions



Notes: Panel (a) shows the (weighted) pooled earnings distributions for 2016 and 2017 in $400 bins for recipients with positive earnings awarded DI between April 1st and December 31st 2013 (gray line) and recipients awarded DI between January 1st and September 30th 2014 (black line). The red solid line indicates the kink point in the budget constraint for both groups. Both groups are weighted by a triangular kernel weight in estimations. For recipients awarded DI in 2014, I add recipients to the right of the kink until the fraction of recipients with positive earnings is the same as for recipients awarded DI in 2013 in order to adjust for extensive margin responses. Recipients awarded DI in 2013 are weighted by propensity score weights $w(x_i) = \frac{P(I_i=1|x_i)}{P(I=1)} \frac{1-P(I=1)}{1-P(I_i=1|x_i)}$ where $P(I=1)$ denotes the probability of being awarded DI in 2014 and $P(I_i = 1|x_i)$ is estimated with a logit model using the covariates in Table A.2 (including uncapped DI benefits) as control variables. Panel (b) shows the corresponding cumulative distributions. The horizontal dashed line indicates the CDF at the kink for recipients awarded DI in 2014. The vertical gray dashed line indicates earnings of the marginal buncher $z^* + \triangle z$. Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).

Figure A.2: Earnings Elasticity Estimates for Different Bandwidths

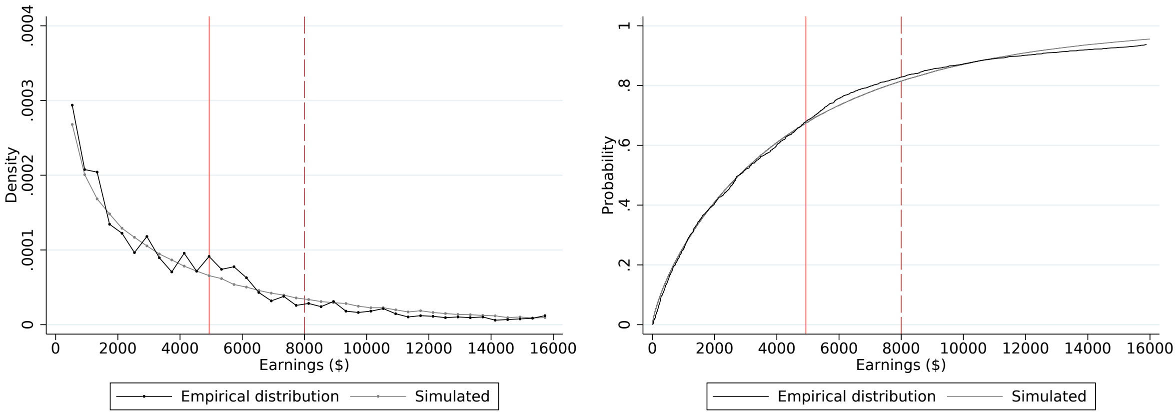
(a) Intensive margin **(b)** Extensive margin



Notes: Panels (a) and (b) displays the weighted estimates for different bandwidth choices of the intensive margin elasticity (see Section 5.2 for details) and the extensive margin elasticity (see Section 6 for details), respectively. The black solid line indicates the point estimates, and the dashed lines indicate 95% confidence intervals. Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. The red vertical line indicates the baseline specification of 9 months bandwidth. The sample consists of totally disabled recipients awarded DI between 1st of January 2014 and 31st of December 2015.

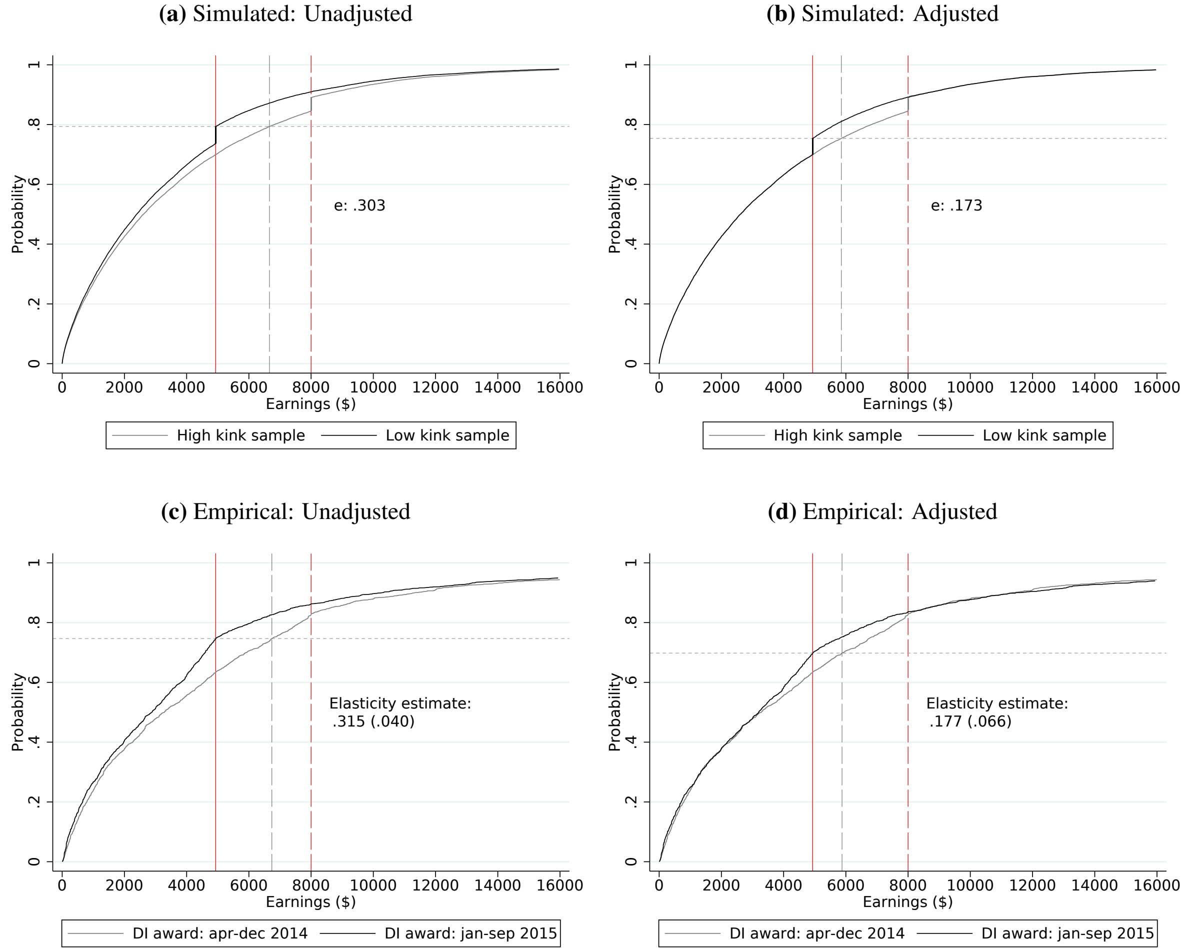
Figure A.3: Empirical and Simulated Ability Distributions

(a) Probability distributions **(b)** Cumulative distributions



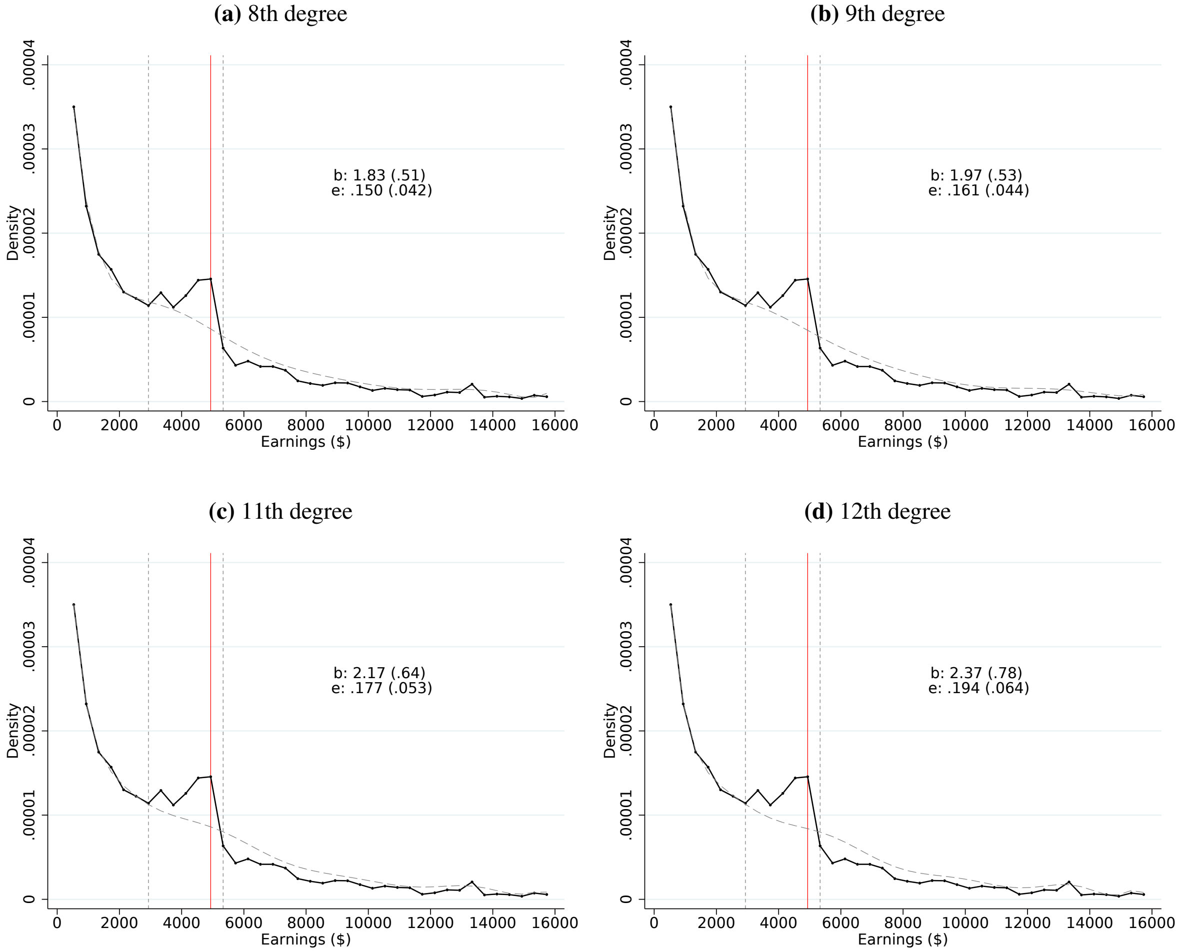
Notes: Panel (a) shows the empirical and simulated ability distributions (see Section 6 for details of how the empirical ability distribution is constructed), and Panel (b) shows the corresponding cumulative distributions. For the empirical distributions, the sample consists of recipients awarded DI between 1st of January and 30th of September 2015. The sample is weighted by a triangular kernel (putting more weight on recipients awarded DI earlier in the year). For the simulated distribution, I calibrate a gamma distribution that best resembles the empirical probability density distribution.

Figure A.4: Simulation-based Adjustment for Extensive Margin Responses



Notes: Panel (a) shows the cumulative distributions from the simulated earnings distributions without adjustment for extensive margin responses and the estimated intensive margin elasticity. The earnings distributions are constructed using the simulated ability distributions in Figure A.3, and are then split into two samples: One with a (low) kink at \$4,937 indicated by the red solid line, and the other with a (high) kink at \$8,000 indicated by the red dashed line. I then construct earnings responses assuming quasi-linear utility and an elasticity $e = .177$ and calibrate a fixed cost of labor force participation that resembles the estimated empirical extensive margin response. Only individuals with earnings larger than the fixed cost of labor force participation are used in the estimation. The horizontal dashed line indicates the CDF at the (low) kink at \$4,937. The vertical gray dashed line indicates earnings of the marginal buncher $z^* + \triangle z$. Panel (b) follows the same procedure where I add observations to the right of the kink for the sample with kink at \$4,937 until the fraction of observations participating in the labor force is the same between the two samples in order to adjust for extensive margin responses. I assume that the additional observations follow the same distribution as the earnings distribution to the right of the kink. Panel (c) shows the corresponding (weighted) cumulative distributions and the estimated intensive margin elasticity for the empirical sample, where the sample consists of recipients with positive earnings only. Panel (d) shows the corresponding (weighted) cumulative distributions and the estimated intensive margin elasticity for the empirical sample, where I add recipients to the right of the kink until the fraction of individuals with positive earnings is the same between the two samples. Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications.

Figure A.5: Parametric Bunching Estimates for Alternative Orders of Polynomial



(a) 8th degree

(b) 9th degree

(c) 11th degree

(d) 12th degree

Notes: The figures illustrate the parametric bunching approach using alternative specifications of the polynomial fitted to the empirical distribution which is grouped into $400 bins. The vertical dashed lines indicate the excluded region, and the red solid line indicate the kink point. The sample consist of recipients awarded DI between January 1st and September 30th 2015 and weighted by a triangular kernel (assigning more weight to the individuals awarded DI early in the year). Standard errors are calculated by a pairs cluster bootstrap which accounts for clustering at the individual level using 500 replications. Earnings are measured in 2016 dollars (NOK/USD = 7.5).