I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13539

Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits?

Andreas Haller
Stefan Staubli
Josef Zweimüller

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13539

Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits?

Andreas Haller
University of Zurich and IZA

Stefan Staubli
University of Calgary, NBER, CEPR and IZA

Josef Zweimüller
University of Zurich, CEPR, CESifo and IZA

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits?*

We study the welfare effects of disability insurance (DI) and derive social-optimality conditions for the two main DI policy parameters: (i) DI eligibility rules and (ii) DI benefits. Causal evidence from two DI reforms in Austria generate fiscal multipliers (total over mechanical cost reductions) of 2.0-2.5 for stricter DI eligibility rules and of 1.3-1.4 for lower DI benefits. Stricter DI eligibility rules generate lower income losses (earnings + transfers), particularly at the lower end of the income distribution. Hence, to roll back the Austrian DI program, policy makers should implement tighter DI eligibility rules rather than lower DI benefits. An application of our framework to the DI system of the U.S. suggests that DI eligibility rules are too strict and DI benefits are too low.

JEL Classification: H53, H55, J14, J21, J65

Keywords: disability insurance, screening, benefits, policy reform

Corresponding author:
Andreas Haller
University of Zurich
Mühlebachstrasse 86
CH-8008 Zurich
Switzerland
E-mail: andreas.haller@econ.uzh.ch

* We thank Tabea Bucher-Koenen, Raj Chetty, Richard Disney, Thomas Hoe, Lucija Muehlenbachs, Timothy Moore, Matthew Notowidigdo, Luigi Pistaferri, Philippe Ruh, Florian Scheuer, Johannes Spinnewijn, Alexander Strand, Conny Wunsch and seminar participants at Erasmus University Rotterdam, University of Amsterdam, University of Bonn, University of Manitoba, University of Melbourne, Universitat Pompeu Fabra, University of Salzburg, University of Zurich, CEPR Labour Economics Symposium, CEPR Public Economics Symposium, CEPR/NBER Aging/Health workshop, NBER Summer Institute, workshop on "Family, Aging, Social Insurance" in Bergen, and the VfS Population Economics meeting in Basel for helpful comments. This research was supported by the U.S. Social Security Administration through grant #1DRC12000002-04 to the National Bureau of Economic Research as part of the SSA Disability Research Consortium. The findings and conclusions expressed are solely those of the author and do not represent the views of SSA, any agency of the Federal Government, or the NBER. All remaining errors are our own.

1 Introduction

The number of disability insurance (DI) recipients has risen rapidly over the past decades in most OECD countries despite generally improving health, higher material living standards, and less physically demanding working conditions.[1] The increasing financial burden of DI programs for taxpayers has led many governments to implement DI reforms aiming explicitly at reducing the DI program inflow and DI expenditures. While restrictive DI reforms reduce the fiscal burden for taxpayers, they also impose utility losses on individuals suffering from a disability. The welfare consequences ultimately depend on how DI reforms address this incentive-insurance trade-off.

In this paper, we pursue the "sufficient-statistics" approach of policy evaluation to shed light on the welfare effects of DI reforms; an approach that has been extensively applied to optimal unemployment insurance (UI). DI programs differ from UI programs in two key dimensions. DI receipt is typically permanent, implying that a key behavioral margin are DI applications driving program *inflow*, and the assessment of DI eligibility requires a medical test, implying that DI reforms need to address the incentive-insurance trade-off of *two* policy instruments: DI eligibility rules and DI benefits.[2] Indeed, proponents of rolling back the DI program argue that DI benefits are not only too generous but also too easy to obtain. Our paper is the first to analyze in a comprehensive way the welfare effects of both DI policy instruments using sufficient statistics.

Our analysis comes in three steps. *First*, we set up a general theoretical framework to study how DI affects individuals' choices. We derive social-optimality formulas that characterize the incentive-insurance trade-off in DI under general economic environments. The incentive costs – both for DI eligibility rules and DI benefits – can be expressed in terms of a *fiscal multiplier*. The fiscal multiplier measures the total cost savings of a DI reform relative to the "mechanical" fiscal effect, the fiscal cost savings in the absence of any behavioral responses. The fiscal multiplier is a key benchmark for welfare analysis: A DI reform is welfare enhancing if the fiscal multiplier is larger than the associated insurance losses. Put differently, a DI reform generating a fiscal multiplier of x is welfare enhancing if one dollar in the hands of a marginal DI recipient has a lower social value than x dollars in public funds.

In the *second step* of our analysis, we provide a causal analysis of two DI reforms that were implemented in Austria in 2003 and 2013. The 2003 reform implemented changes to the pension formula reducing DI

[1] In the U.S., 2.6 percent of individuals in the age group of 20 to 64 were receiving DI benefits in 1992, but by 2012 this fraction had risen to 5.3 percent. Many European countries have also experienced significant growth of their DI programs, particularly Scandinavian countries. In 2015, public spending on disabilility insurance ("incapacity") programs amounted to an OECD-average of 1.9% of GDP, with substantial variation across countries: 1.3% in the U.S., 1.7% in France, 1.9% in the UK, 2.1% in Austria and Germany, and more than 4% in Sweden, Norway and Denmark (OECD, 2020).

[2] In contrast, UI receipt is typcially temporary and assessing UI eligibility is straightforward. Moreover, the relevant behavioral margin in UI (job-search effort) affects the program *outflow,* which is primarily driven by the generosity of a single policy instrument: the UI benefit level. Baily (1978) has pioneered the theoretical analysis of optimal UI and Chetty (2006) has developed the sufficient-statistics approach in the UI context. More recent applications of the sufficient statistic approach for optimal UI include Shimer and Werning (2007), Chetty (2008), Kroft (2008), Schmieder et al. (2012), Landais (2015), Kroft and Notowidigdo (2016), Landais et al. (2018) and Kolsrud et al. (2018). See the article by Chetty and Finkelstein (2013) for a detailed discussion of this literature.

benefit levels substantially for some individuals but less so for others. The quasi-experimental variation in DI benefits over time and across individuals allows us to identify the causal effect of DI benefits. The 2013 reform implemented stricter DI eligibility rules by increasing the "relaxed screening age" (RSA), the age at which vocational factors in the DI determination process increase DI award rates substantially. Because of a staggered increase in the RSA, we can compare "adjacent" cohorts to identify the causal effect of stricter DI eligibility rules. Using population data from the Austrian social security register (ASSD) merged with the universe of DI applications (provided by the Austrian Ministry of Social Affairs, BMASK), we find that stricter DI eligibility rules and lower DI benefits in the Austrian DI reforms generated behavioral responses, which lowered DI program costs substantially.[3]

The *third step* of our analysis explores the welfare effects of the Austrian DI reforms. To estimate the fiscal multiplier, we can draw on our reduced-form estimate of the total fiscal effect (the numerator), but we still have to estimate the mechanical fiscal cost savings in the absence of behavioral responses (the denominator). In the case of *lower DI benefits,* the mechanical fiscal effect of a one-percent reduction in DI benefits is simply one percent of the pre-reform mean of DI expenditures, which yields a fiscal multiplier between 1.3 and 1.4. Estimating the mechanical fiscal effect of *stricter DI eligibility rules* is less straightforward, because we need to know who is an always applicant (who does not abstain from applying even under stricter DI eligibility rules). Unfortunately, we cannot observe in the data who is an always applicant and who is a marginal applicant (who abstains from applying under stricter eligibility rules).[4] We argue (and provide supportive evidence) that the mechanical fiscal effect can be inferred from the re-application behavior of previously rejected DI applicants. Based on this strategy, we estimate a fiscal multiplier between 2.0 and 2.5.

Taken together, the relative size of fiscal multipliers suggests that stricter DI eligibility rules are more effective than lower DI benefits in reducing program expenditures. But to assess the relative welfare effects, we also need to compare the insurance losses of the two policy instruments. While we lack the necessary data (on health, wealth, and consumption) to estimate the insurance losses directly, we find that income losses (earnings plus transfers) associated with stricter DI eligibility rules are smaller than those associated with lower DI benefits, particularly in the lowest quintile of the income distribution. Through the lens of our theoretical framework, this pattern suggests lower insurance losses of stricter DI eligibility rules compared to lower DI benefits. Hence,

[3]The linked social security and DI applications data provide us with a "unique" data set in the sense that we observe not only all DI applications but also all workers (applicants *and* non-applicants) covered in the ASSD (about 80% of the Austrian population). Observing applicants and non-applicnts allow us to study in detail individuals' DI application behavior. Staubli (2011) studies the labor market effects of an earlier increase in the RSA in 1996, but he has no application data and cannot study application behavior, which is important in the present context, since individuals' application behavior is a key driver of the welfare effects of DI reforms.

[4]If always applicants and marginal applicants had identical (observed and unobserved) characteristics, the mechanical fiscal effect of stricter DI eligibility rules could be estimated simply by applying lower DI award rates to the *average* pre-reform DI applicant. We conduct a complier analysis (Imbens and Rubin 1997; Abadie 2003) which shows that always applicants and marginal applicants differ significantly with respect to a number of observed characteristics. Hence, approximating the mechanical effect (the fiscal cost savings of always applicants) by the cost savings of average applicants (a mixture of always applicants and marginal applicants) is misleading.

we conclude that stricter DI eligibility rules dominate lower DI benefits as a policy tool for rolling back the Austrian DI program.[5]

While we think our paper makes progress using the sufficient-statistics approach for optimal DI, we need to keep in mind the limitations of this approach. First, the welfare implications are drawn from reduced-form estimates, which apply only "locally" to the particular Austrian context. However, our analysis is of more general interest, since many DI programs feature eligibility rules similar to Austria that are based on vocational factors such as age or work history. For example, in the U.S. DI system applicants older than age 55 are evaluated based on more lenient eligibility standards than applicants between ages 50 and 55, who are subject to more lenient standards than applicants below age 50.[6] Different from what we find for Austria, Chen and van der Klaauw (2008) and Deshpande et al. (forthcoming) do not find sorting of applications around these age cutoffs, suggesting no behavioral response and a fiscal multiplier of DI eligibility rules of 1.[7] Previous U.S. estimates on the behavioral responses to changes in DI benefit levels result in a fiscal multiplier very similar to the one we estimate for Austria. For example, Low and Pistaferri (2015) estimate an elasticity of DI applications with respect to benefit levels of 0.62. Together with a DI award rate of 0.67 (French and Song, 2014), this elasticity suggests a fiscal multiplier of lower DI benefits of 1.3-1.4 for the U.S.

A second limitation of the sufficient-statistics approach is that it applies only to marginal (infinitesimally small) policy changes, while in reality we are interested in non-marginal policy changes (Kleven, forthcoming). We address this issue by deriving social optimality conditions for non-marginal DI policy changes and show that our analysis of the fiscal multiplier, a core concept of our framework, is also valid with non-marginal policy changes. We further show, for non-marginal policy changes, how income losses (along the income distribution) can be used to bound insurance losses and how these bounds are useful for ranking the two DI policy instruments.

A third limitation of the sufficient-statistics approach is that, without restrictions on preferences and the economic environment, one typically ends up with a large number of elasticities to be estimated. Our concept of the fiscal multiplier (with its focus on overall program costs) is useful, because it permits welfare analysis without making specific restrictions to reduce the number of elasticities. In this respect, our framework is similar to Lee et al. (forthcoming) and Hendren and Sprung-Keyser (2020).[8]

[5] While our analysis cannot put a number on the absolute insurance losses, a tentative analysis (based on CRRA preferences and hand-to-mouth consumers) suggests that insurance losses and fiscal gains of lower DI benefits are of similar size (i.e. DI benefits are optimal), while insurance losses fall short of fiscal gains of tighter DI eligibility rules (i.e. tightening eligibility rules is welfare-enhancing).

[6] There are consideration of revising the U.S. vocational factors. In 2012 the Congressional Budget Office proposed to increase the age cutoffs of the relaxed eligibility rules (Mann et al., 2014).

[7] Appendix Figure D.25 contrasts the U.S. application behavior to the Austrian application behavior. In Austria, there is a large spike in DI applications exactly at the RSA. In the U.S., there is no spike in DI applications at the age cutoffs but a discontinuous jump in award rates.

[8] Lee et al. (forthcoming) estimate the fiscal externality of UI benefit reforms. The fiscal externality is the behavioral fiscal effect relative to the mechanical fiscal effect. Hence, what we refer to as the fiscal multiplier is 1+fiscal externality. Hendren and Sprung-Keyser (2020) use the concept of the Marginal Value of Public Funds ("MVPF") to evaluate 133 historical policy changes in the U.S. The MVPF is the willingness to pay for a policy divided by the net cost to the government. In our application, the MVPF corresponds to the insurance value divided by the fiscal multiplier. We separate the two effects. In the case of DI, determining the insurance value is not straightforward (and to some degree a judgment call), while the fiscal multiplier can be estimated with

An interesting aspect of our approach relates more specifically to the impact of DI eligibility rules. One might have thought that a welfare analysis of tighter DI rules requires pinning down type-I and type-II errors (false rejections and false acceptances). However, we show that it suffices to estimate the mechanical fiscal effect, which is an advantage, because it substantially reduces the data requirements and the complexity of the analysis. The disadvantage is that we cannot address the accuracy of the screening process – an important open question in DI research. Instead, a more structural approach, such as the one of Low and Pistaferri (2015), is able to directly estimate the type-I and type-II errors. Benitez-Silva et al. (2004) and Low and Pistaferri (2019) provide further evidence on the classification errors of the DI screening process.

This paper contributes to an active literature on the labor market and welfare effects of disability insurance programs (for reviews, see Bound and Burkhauser, 1999; Low and Pistaferri, forthcoming). Our paper complements the strand of literature that evaluates the incentive-insurance trade-off in DI programs using structural models. Most closely related are the U.S. studies of Bound et al. (2004) and Low and Pistaferri (2015). Bound et al. (2004) simulate the benefits and costs of changes in disability benefit levels and find that the implicit price of providing an additional dollar of income to DI recipients – what we call the fiscal multiplier of DI benefits – is 1.5, very similar to our estimate. Low and Pistaferri (2015) assess the optimal DI benefit levels and eligibility criteria and find that eligibility criteria are too strict and DI benefits too low in the U.S. We reach the same conclusion by applying our framework to existing U.S. estimates. Bound et al. (2010) specify a structural model to study the interplay between health and labor force participation. They find that removing the DI program entirely would have little effect on individuals in good health but would hurt individuals in bad health. Autor et al. (2019) use a judge leniency instrumental variable design and a structural model to estimate the consumption and welfare effects of DI in Norway. They show that DI increases household income and consumption for singles but not for married individuals. Their results point, like ours, to the importance of benefit substitution.

Our paper also contributes to the theoretical literature on optimal disability insurance. We extend the model of Diamond and Sheshinski (1995) by expressing the social optimality conditions of DI eligibility rules and DI benefit levels as a function of sufficient statistics, which we can estimate empirically using program evaluation methods. We also generalize their setup to a dynamic environment with rich heterogeneity across agents. Also related are the U.S. studies of Meyer and Mok (2019) and Deshpande et al. (forthcoming) who apply the Bailey-Chetty formula for optimal UI benefits to DI and estimate the effect of receiving DI on consumption and financial outcomes. Similarly, Ball and Low (2014) estimate the effect of DI on consumption in the UK to infer the insurance value of DI benefits. We go beyond these papers by studying the welfare effects of DI eligibility rules and comparing them to the welfare effects of DI benefits. Understanding the effects of eligibility rules is

reduced-form methods.

important as the discussion about DI reforms focuses on whether individuals are truly eligible for DI benefits.

Our paper also relates to the strand of literature that estimates the impact of DI on applications, DI take-up, and labor supply using reduced-form methods without considering welfare effects.[9] Autor and Duggan (2003) find that relaxed eligibility rules and increases in the DI replacement rates explain the stark growth of DI rolls in the U.S. and lead to a lower unemployment rate.[10] Parsons (1991) and Gruber and Kubik (1997) exploit variation in DI rejection rates across U.S. states over time and find that an increase rejection rates reduces DI applications and increases labor force participation. We find similar evidence for self-screening in response to stricter eligibility rules, i.e. a decline in applications, and also show that stricter eligibility rules target healthier individuals via our complier analysis.[11] Our paper also relates to studies that explore the effects of DI benefit levels for application behavior and labor supply (Gruber, 2000; Campolieti, 2004; Mullen and Staubli, 2016). We build on these papers by estimating the effects of DI benefits on benefit substitution and fiscal costs, which are key for assessing the welfare effects of lower DI benefits.

The paper is organized as follows. The next section presents a model of disability insurance and formulas for optimal disability eligibility and benefits. Section 3 describes the data and institutional background in Austria. Sections 4 and 5 present the empirical results on stricter DI eligibility rules and lower DI benefit levels, respectively. Section 6 estimates the fiscal multipliers of these two policy instruments and discusses how our estimates can be used for welfare evaluation. Section 7 applies our framework to the U.S. disability system. Section 8 concludes.

2 Theoretical Framework

In this section, we explore how the two main DI policy parameters – the strictness of DI eligibility rules and the level of DI benefits – affect social welfare, as well as labor supply and application behavior of potential DI claimants.[12] Section 2.1 starts with the static framework of Diamond and Sheshinski (1995) and Section 2.2

[9] Another important strand of the literature studies the impact of DI receipt on labor force participation by comparing accepted and rejected DI applicants (Bound, 1989; von Wachter et al., 2011), by exploting variation in eligibility rules (Chen and van der Klaauw, 2008), and by exploiting the random assignment of DI applicants to examiners and administrative law judges (Maestas et al., 2013; French and Song, 2014). We do not directly contribute to this literature, but changes in labor force participation are reflected in our program cost estimates. We also do not study outflow from DI (Campolieti and Riddell, 2012, Borghans et al., 2014, Moore, 2015) or earnings of DI recipients (Kostol and Mogstad, 2014, Gelber et al., 2017, Ruh and Staubli, 2019 and Kostøl and Myhre, 2020), but these responses enter the fiscal multiplier as well.

[10] Autor and Duggan (2006) discuss potential DI reforms to counteract the cost explosion of the DI program in the U.S. They point out that there are three ways to reduce the size of DI programs: (i) tightening the screening process (eligibility rules), (ii) reducing the incentives to seek benefits (lower DI benefits) and (iii) encouraging faster exit. Our framework sheds light on the welfare effects of options (i) and (ii).

[11] de Jong et al. (2011) and Godard et al. (2019) find that another aspect of the application process, more intense screening of applicants, also reduces DI applications and improves targeting to more deserving applicants. In contrast, Deshpande and Li (2019) show that higher application costs have adverse targeting effects by inducing individuals who would have qualified for DI to no longer apply.

[12] By increasing the "strictness of disability eligibility rules" we mean any policy making it more difficult that a DI application – with a given degree of disability – gets accepted. This is what Low and Pistaferri (2015) and Diamond and Sheshinski (1995) call, respectively, "strictness of screening" and "disability standard". The terms disability rules, disability standard, and disability screening are used interchangeably. The formal definition of strictness is discussed in detail in section 2.1.

extends the analysis to a dynamic setting.

2.1 A Static Model of Optimal DI

Setup. Consider an agent living for two periods. In the first period, she works, earns a wage w, pays a lump-sum tax τ (which finances the DI program) and enjoys utility $u(w-\tau)$. There are no savings nor any other choices in the first period.[13] In the second period, the agent suffers a disability shock θ, modelled as a random draw from a continuous distribution $F(\theta)$. If θ is small (= the disability not very severe), the agent continues working and enjoys second-period utility $u(w)-\theta$. If θ is sufficiently large (= the disability severe), the agent applies for DI benefits. A DI application causes disutility ψ, capturing the extensive medical checks, the bureaucratic hassle, etc. associated with the DI assessment process. The fixed application cost ψ is important in the present context as it ensures that DI application choices depend on the eligibility rules of the DI system.[14] With probability $p(\theta)$ the application is accepted, where $p'(\theta) > 0$.[15] When the application is accepted, the agent withdraws from work, claims DI benefits b and gets second-period utility $v(b)-\psi$. When the application is rejected, the applicant either resumes work and gets second-period utility $u(w)-\theta-\psi$; or claims social welfare $z < b$ and gets second-period utility $v(z)-\psi$. (No disutility or uncertainty are associated with claiming social welfare.) Appendix Figure A.1 illustrates the sequence of events and agent's choices in the second period.

DI Applications and Labor Supply. Let us now look at the DI application choice and the labor supply decision. An individual prefers working over claiming social welfare benefits if her disability is $\theta < \theta^R \equiv u(w) - v(z) > 0$, i.e. if the utility of claiming social welfare falls short of the utility of working. Hence, θ^R denotes the "marginal social welfare claimant". Consider an agent whose disability is not extremely severe, $\theta < \theta^R$. (This implies she goes back to work in case her DI application gets rejected.) Her application choice compares the utility when staying employed, $u(w)-\theta$, to the expected utility when applying for DI, $p(\theta)\,v(b) + [1-p(\theta)]\,(u(w)-\theta)-\psi$. The "marginal applicant," the agent who is indifferent between filing a DI application and remaining employed, has disability

$$\theta^A = u(w) - v(b) + \frac{\psi}{p(\theta^A)}. \tag{1}$$

It follows that agents with disability $\theta \geq \theta^A$ apply for DI, while agents with disability $\theta < \theta^A$ remain employed. Figure 1, Panel (a) characterizes the outcome of agents' DI application choices. It draws the probability of

[13] The setup follows Chetty (2006) who reconsiders Baily's (1978) formula of optimal unemployment insurance (UI). The stylized two-period framework - tax payments but no DI application choices in the first period, while no tax payments but DI application choices in the second period - simplifies the formula without changing the substance of the argument.

[14] Here we deviate from Diamond and Sheshinski (1995) who do not consider application costs. Recent empirical studies support the idea that application costs are important drivers of DI applications, e.g. Deshpande and Li (2019) and Godard et al. (2019).

[15] Below, we will analyze a situation where the government has control over the $p(\theta)$-function. By adopting stricter eligibility rules, the $p(\theta)$-function shifts down, so that p takes a lower value for any given θ (and vice versa).

DI award $p(\theta)$ against θ and indicates the disability cutoff-levels θ^A and θ^R. Agents with a disability $\theta \geq \theta^A$ apply for DI; if rejected, those with disability $\theta \in [\theta^A, \theta^R)$ return to work, while those with $\theta \geq \theta^R$ go on social welfare.

Equation (1), and its graphical representation in Figure 1, applies when $\theta^A < \theta^R$, i.e. a marginal applicant returns to work in case her DI application is rejected. We discuss in Appendix A.1 the formal conditions under which $\theta^A < \theta^R$ holds. Note that this is not a critical assumption and we do not impose it in the general model. Moreover, it is worth emphasizing that this is a natural assumption in the present context. With $\theta^A < \theta^R$ the model predicts that DI policy parameters affect labor supply decisions. Distortionary labor supply effects of DI programs are supported by a large body of empirical evidence.[16]

DI Policy Instruments. We now assess the welfare effects of two policy instruments that characterize any DI system: the level of DI benefits and the strictness of DI eligibility rules. While the role of DI benefits b is straightforward and poses no major conceptual problems, the role of DI eligibility rules θ^* needs further discussion. The inherent problem of the DI assessment process is that the true disability θ is the agent's private information. For this reason, a DI applicant has to undergo a disability assessment process, which delivers an estimate of her disability to the government. Formally, the government observes $s = \theta + e(\theta)$, where s is a noisy signal, θ is the applicant's true disability and $e(\theta)$ is the noise.[17] The strictness of DI eligibility rules – the policy parameter under direct control of the government – can be captured by a critical value of s, call it θ^*, such that a DI application with $s \geq \theta^*$ is accepted, while an application with $s < \theta^*$ is rejected. The acceptance probability can then be written as $p(\theta; \theta^*)$.[18] In what follows, we consider the case where the government can change θ^* but takes the signal as given. This is the context of our empirical analysis below, which exploits quasi-experimental variation in the "relaxed screening age" (RSA) at which DI eligibility rules become more lenient. In our notation, the strictness of DI eligibility equals $\theta^* = \theta^H$ before the RSA and falls to $\theta^L < \theta^H$ after the RSA. An increase in the RSA from age R to some higher age $R + \Delta$, implies that, during the age window $[R, R + \Delta]$, the treated cohort is subject to the strict DI eligibility standard θ^H, while the control cohort is subject to the lenient standard θ^L. If cohorts are otherwise similar (in productivity, health, preferences, etc.), a plausible assumption for "adjacent" cohorts, comparing treated to control cohorts identifies the causal effect

[16] A number of papers provide direct evidence on the work behavior of rejected DI applicants. These findings are perfectly consistent with the predictions of the model with $\theta^A < \theta^R$. Bound (1989); von Wachter et al. (2011); Maestas et al. (2013); French and Song (2014) use rejected DI applicants as a control group for accepted applicants to study the impact of DI on labor supply. For instance,von Wachter et al. (2011) report that, in 69.6% of rejected DI applicants aged 30-44 in the U.S. report positive yearly earnings two years after the DI application and 57.4% report earnings higher than three months of full-time employment at the minimum wage in 2000. The corresponding numbers are 52.6% and 42.7% for rejected DI applicants aged 45-64. In the Norwegian study by Kostol and Mogstad (2014), about 30 percent of rejected DI applicants aged 18-49 are participating on the labor market.

[17] The variance of the noise is likely to vary with the severity of the disability as very severe and perhaps also very weak disabilities are more easy to assess than intermediate cases.

[18] In the following we assume that the DI assessment process is informative, i.e. we assume $\partial p(\theta; \theta^*)/\partial \theta \geq 0$. This implies that in an applicant pool with a more severe disability a smaller fraction of DI assessments fall short of an arbitrary cutoff θ^* and will ensure that on average the award probability is increasing in the severity of the disability.

of an increase in θ^* on the outcomes of interest.[19]

Welfare Effects of DI Reforms. We follow the literature assuming society's objective can be represented by a utilitarian social welfare function. Assuming a population of mass unity and abstracting from discounting, the social welfare function is given by

$$\begin{aligned} W(\theta^*, b) = \quad & u(w-\tau) + \int_0^{\theta^A} (u(w)-\theta)dF(\theta) + \int_{\theta^A}^{\theta^R} (1-p(\theta;\theta^*))(u(w)-\theta)dF(\theta) + \\ & + \int_{\theta^A}^{\infty} p(\theta;\theta^*)v(b)dF(\theta) + \int_{\theta^R}^{\infty} (1-p(\theta;\theta^*))v(z)dF(\theta) - \int_{\theta^A}^{\infty} \psi dF(\theta). \end{aligned} \tag{2}$$

The right-hand-side terms sum up the welfare levels of the various agents: first-period workers, all of whom are working and paying taxes (first term on the right-hand-side); the working healthy (second term), the rejected DI applicants resuming work (third term); the DI recipients (fourth term); and the social-welfare recipients (fifth term). The last term takes account of the aggregate welfare losses associated with DI application costs. When designing the optimal DI program, the government needs to take into account agents' behavioral responses to changes in DI policy parameters. Furthermore, the social planner is constrained by a balanced-budget requirement: DI and social welfare benefit payments have to be covered by the taxes raised in the first period,

$$\tau = b \int_{\theta^A}^{\infty} p(\theta;\theta^*)dF(\theta) + z \int_{\theta^R}^{\infty} (1-p(\theta;\theta^*))dF(\theta). \tag{3}$$

In what follows, we discuss the welfare effects of DI reforms. We first look at the effects of implementing more stringent DI eligibility rules, before we turn to the effects of reducing DI benefits. The discussion is framed in terms of implementing a more restrictive DI system, because most policy debates center around reducing the financial burden of the DI program. Of course, analogous arguments hold for reforms that increase the generosity of the DI system.

Stricter DI Eligibility Rules: Marginal Increase in θ^*. The utilitarian government sets DI eligibility rules θ^* to maximize social welfare W, taking into account the balanced-budget requirement and agents' DI application responses. In Appendix A.1, we show that the welfare effect of increasing θ^* is

[19] It should be clear that the government could, in principle, take measures other than varying θ^* to manipulate the DI award probability $p(\theta;\theta^*)$. For instance, the goal of a DI reform could be to increase the precision of DI screening, to avoid type-I and type-II errors (= false acceptances and false rejections) of an imperfectly functioning DI assessment system. This could be done through more extensive medical checks, better equipment, monitoring of DI applicants, etc.. Such measures would reduce the variance of the noise $e(\theta)$. However, unlike changing θ^*, changing the precision of the signal requires resources and welfare calculations need to take into account society's willingness to pay for improved DI screening. While such policies are clearly relevant in practice, we do not analyze their welfare implications here, mainly because we cannot address them empirically with our data. However, we consider this a potentially interesting direction for future research. Low and Pistaferri (2015) and Low and Pistaferri (2019) make progress in this direction by estimating type-I and type-II errors in the U.S. award process.

$$\frac{\partial W}{\partial \theta^*} = u'(w-\tau) \underbrace{[B(\theta^*)+M(\theta^*)]}_{\text{fiscal cost reduction}} - \underbrace{\Big[[v(b)-(u(w)-\tilde{\theta})]M_W + [v(b)-v(z)]\,M_Z\Big]}_{\text{insurance losses}}. \quad (4)$$

Condition (4) highlights the two opposing effects of stricter DI eligibility rules θ^* on social welfare. On the one hand, a higher θ^* raises social welfare because it saves taxpayers' money (fiscal cost reduction). On the other hand, a higher θ^* reduces social welfare, because fewer agents are awarded DI when hit by a severe disability shock (insurance losses).

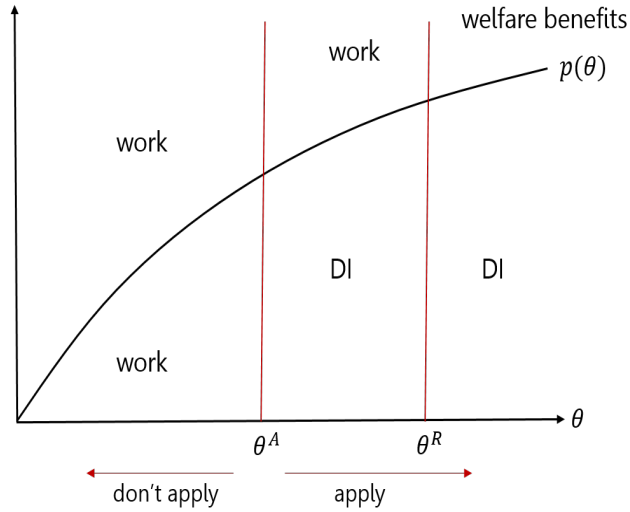
The fiscal cost reduction consist of two components: the behavioral fiscal effect $B(\theta^*)$ and the mechanical fiscal effect $M(\theta^*)$. The behavioral fiscal effect measures the reduction in DI expenditures due to fewer DI applications. The mechanical fiscal effect $M(\theta^*)$ comes from fewer DI applications getting accepted. To see the behavioral and mechanical effects more clearly, note that the DI inflow probability is the product of two factors: the probability of filing an application times the probability that the application gets accepted, $Pr(\text{DI}) = Pr(\text{Apply}) * Pr(\text{Accept}|\text{Apply})$. In the above notation, the application probability is $Pr(\text{Apply}) = 1 - F(\theta^A)$, while the acceptance probability is $Pr(\text{Accept}|\text{Apply}) = \left[\int_{\theta^A}^{\infty} p(\theta;\theta^*)dF(\theta)\right] / \left[1-F(\theta^A)\right]$. The derivative of the application probability with respect to θ^* yields the average agent's change in application behavior, $(\partial\theta^A/\partial\theta^*)p(\theta^A;\theta^*)f(\theta^A)$, which is the red area in Panel (b) of Figure 1. Multiplying with the DI benefit b yields the behavioral fiscal effect $B(\theta^*) = (\partial\theta^A/\partial\theta^*)p(\theta^A;\theta^*)f(\theta^A)\cdot b$. The derivative of the acceptance probability with respect to θ^* equals $-\int_{\theta^A}^{\infty}(\partial p(\theta;\theta^*)/\partial\theta^*)dF(\theta)$, which is the sum of the gray and the blue area in Panel (b) of Figure 1. The gray area captures the rejected working applicants $M_W \equiv -\int_{\theta^A}^{\theta^R}(\partial p(\theta;\theta^*)/\partial\theta^*)dF(\theta)$; the blue area are the rejected applicants on social welfare $M_Z \equiv -\int_{\theta^R}^{\infty}(\partial p(\theta;\theta^*)/\partial\theta^*)dF(\theta)$. Each rejected applicant resuming work saves the amount b to the taxpayer (recall that, in the second period, workers do not pay taxes), while each rejected applicant substituting DI for social welfare saves $b-z>0$ to the taxpayer. The mechanical *fiscal* effect is therefore $M(\theta^*) \equiv M_W\cdot b + M_Z\cdot(b-z)$. Since fiscal savings are used to reduce taxes, the total fiscal gain, $B(\theta^*)+M(\theta^*)$, is valued at the marginal utility of consumption of the taxpayer $u'(w-\tau)$ in equation (4).

Adopting stricter DI eligibility rules θ^* does not only save money to taxpayers, it also reduces the insurance value of the DI system. The lower DI acceptance probability corresponds to a higher probability that a DI applicant eventually has to resume work, M_W, or has to claim social welfare, M_Z. The average utility loss of the former is $v(b)-(u(w)-\tilde{\theta})>0$, where $\tilde{\theta}$ is the average disability level of rejected applicants who go back to work.[20] The utility loss of the latter is $v(b)-v(z)>0$. Note the reduction in the insurance value depends only

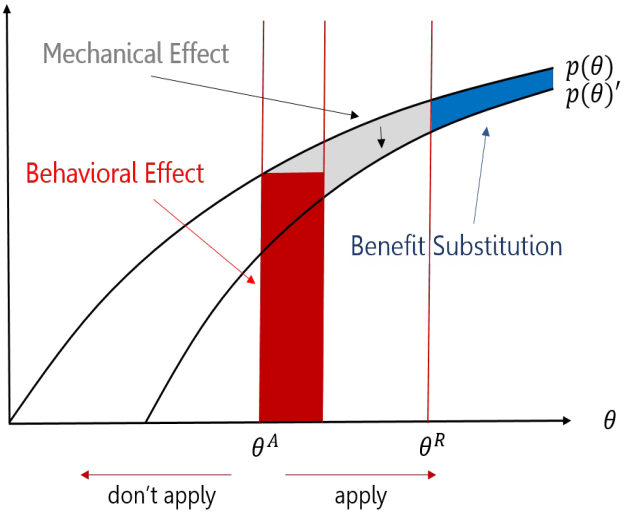
[20] Formally, $\tilde{\theta}$ is the average disability level of agents with a disability shock in the range $[\theta^A,\theta^R)$, so that $\tilde{\theta} \equiv \int_{\theta^A}^{\theta^R}(\partial p(\theta;\theta^*)/\partial\theta^*)\theta dF(\theta) / \int_{\theta^A}^{\theta^R}(\partial p(\theta;\theta^*)/\partial\theta^*)dF(\theta)$.

Figure 1: Illustration of Static Model and Effects of Stricter Eligibility Rules

(a) Illustration of Model



(b) Effects of Stricter Eligibility Rules



Notes: Panel (a) illustrates the basic setup. Individuals are characterized by disability level θ and can choose whether to work, apply to DI or leave the labor force and consume social welfare benefits. The award process to DI is noisy and individuals are awarded DI with probability $p(\theta)$. We assume that $p(\theta)$ is weakly increasing in θ. This captures that (i) it is difficult to assess the true disability level of an individual and (ii) the assessment contains nonetheless some valuable information on the true disability level. The marginal DI applicant is denoted by θ^A and individuals with $\theta \geq \theta^A$ apply to DI. The marginal welfare benefits type is denoted by θ^R and individuals with $\theta \geq \theta^R$ will go on welfare benefits if they are rejected. Panel (b) illustrates the effects of stricter eligibility criteria. Stricter criteria shift down the award probability curve. The area between the two award probability curves is the mechanical effect. A fraction of the mechanically rejected applicants returns to work (gray area). The other fraction substitutes DI benefits with welfare benefits (blue area). Stricter eligibility criteria also shift the marginal applicant to the right. The change in the marginal applicant times the award probability of the marginal applicant is the behavioral effect (red area).

on the mechanical effect but not on the behavioral effect. This is a direct implication of the Envelope theorem.[21] Intuitively, only marginal applicants react to a marginal change in the strictness of eligibility rules. Marginal applicants are indifferent between applying and not applying. Hence, if a marginal increase in θ^* induces them not to file an application, their welfare is not directly affected. However, fewer applications reduce the financial burden of the DI system, thus they generate a positive fiscal effect that benefits taxpayers.

The optimal strictness of eligibility rules θ^* balances the trade-off between insurance loss and fiscal gain, where (4) is set to zero. For later use, we rewrite this condition as

$$\frac{\partial W}{\partial \theta^*} \gtreqless 0 \iff 1 + \frac{B(\theta^*)}{M(\theta^*)} \gtreqless \frac{L_W + L_Z}{u'(w-\tau)M(\theta^*)}, \tag{5}$$

where $L_W \equiv [v(b) - (u(w) - \tilde{\theta})]M_W > 0$ and $L_Z \equiv [v(b) - v(z)]\, M_Z > 0$ are the aggregate utility losses suffered by the additionally rejected applicants resuming work (L_W) and claiming social welfare (L_Z), respectively. The two sides of the inequality have an intuitive interpretation. The left-hand-side is the fiscal multiplier, $1 + B(\theta^*)/M(\theta^*)$, and measures the reduction in the financial burden for the taxpayer per *mechanically* saved dollar (= hypothetical fiscal gain when application behavior remains unchanged). The right-hand-side is the

[21] While the decision to apply is discrete the envelope theorem applies because we have a marginal change in the policy parameter θ^*. In Appendix A.1 we show this formally and also discuss how the welfare evaluation changes in case of a discrete (non-marginal) change of θ^*.

corresponding reduction of the insurance value in monetary units. Dividing by the marginal utility of consumption of the taxpayer $u'(w-\tau)M(\theta^*)$ yields the insurance loss (in monetary terms) per mechanically saved dollar.

Stricter DI Eligibility Rules: Discrete Increase in θ^*. The welfare implications of stricter DI eligibility rules, as summarized in condition (5) hold true for a marginal increase in θ^*. In our empirical implementation, however, we study an RSA increase from age R to some higher age $R+\Delta$, which implies a discrete increase in the DI eligibility standard from an initially lenient standard θ^L to a strict standard θ^H during the age window $[R, R+\Delta]$. In Appendix A.1 we show that, for a discrete increase in θ^*, condition (5) needs to be rewritten as

$$\Delta W \gtreqless 0 \iff 1+\frac{B_\Delta(\theta^*)}{M_\Delta(\theta^*)} \gtreqless \frac{L_{W\Delta}+L_{Z\Delta}}{u'(w-\tau_\Delta)M_\Delta}+\frac{L_{MA}}{u'(w-\tau_\Delta)M_\Delta},$$

where the subscript Δ highlights that the corresponding effect has been generated by a discrete change in θ^*.[22] The main takeaway is that fiscal cost reductions (the left-hand-side of the social welfare condition) can be measured by the fiscal multiplier $1+B/M$. However, the insurance losses of marginal applicants (those abstaining from a DI application under the now stricter rules) can no longer be ignored. These welfare losses are captured by the second term on the right-hand-side of the adjusted welfare condition.

Lower DI Benefits. The second key DI policy parameter is the level of DI benefits b. It is straightforward to show (see Appendix A.1) that the condition for a socially optimal DI benefit level is

$$\frac{\partial W}{\partial(-b)} \gtreqless 0 \iff 1+\frac{B(b)}{M(b)} \gtreqless \frac{v'(b)}{u'(w-\tau)}. \tag{6}$$

Similar to condition (5) above, condition (6) tells us that a reduction in DI benefits b is welfare-improving if the fiscal gains to taxpayers exceeds the insurance loss suffered by disabled workers. On the one hand, a lower b reduces the financial burden of the DI system because fewer agents apply for DI. In condition (6), this is captured by the fiscal multiplier, $1+B(b)/M(b)$. On the other hand, a lower b reduces the consumption smoothing benefit, because it reduces the consumption possibilities when hit by a disability shock. This is captured by the ratio of the marginal utility of a DI benefit recipient relative to the marginal utility of a taxpayer.

The behavioral fiscal effect is $B(b) \equiv -\left(\partial\theta^A/\partial b\right) p(\theta^A) f(\theta^A) \cdot b$ and the mechanical fiscal effects is $M(b) \equiv \int_{\theta^A}^{\infty} p(\theta) dF(\theta)$. The ratio of behavioral over mechanical fiscal effect corresponds to the DI inflow elasticity, i.e. $\xi = (\partial DI/\partial b)(b/DI) = -\left(\partial\theta^A/\partial b\right) p(\theta^A) f(\theta^A) b / \int_{\theta^A}^{\infty} p(\theta) dF(\theta) = B(b)/M(b)$. This yields an interesting analogy of the optimal DI formula to the famous Baily (1978) formula for optimal unemployment insurance

[22] τ_Δ is the (discrete) reduction in taxes made possible by the (discrete) increase in strictness of DI eligibility rules, so as to keep the government DI (+ social welfare) budget balanced.

(UI). Both in the case of UI and in the case of DI, the condition for the socially optimal benefit level can be written as $1+\eta = v'(b)/u'(w-\tau)$. In the Baily (1978) model of optimal UI, η is the elasticity of unemployment duration with respect to the UI benefit level; in the above model of optimal DI, $\eta = \xi$, the elasticity of the DI inflow with respect to the DI benefit level. In other words, the relevant moral-hazard margin in the case of DI is the program *inflow*, while the relevant margin in the case of UI is the program *outflow*.[23]

So far, we have derived conditions for social optimality for each single DI policy parameter, holding the other policy parameter fixed. A natural question is how a DI reform should optimally combine these two policy parameters. More precisely: how strongly – and in which direction – should DI eligibility rules θ^* be changed per unit change of DI benefits b? In Appendix A.1 we discuss how the formulas for optimal benefits and eligibility rules are informative for the optimal policy mix.

2.2 The General Model

The above model highlights the basic trade-offs of DI policy reforms but misses two ingredients that are crucial in designing and evaluating DI reforms: heterogeneity across individuals and intertemporal choices. In the model of section 2.1, agents differ only in θ and all actions happen within one period. In what follows, we allow for multiple sources of heterogeneity (such as wages and other factors) and we extend the model to multiple periods. This latter extension allows us to capture the intertemporal nature of the DI application choice. In the context of our empirical analysis below – which exploits an RSA increase from R to some higher age $R+\triangle$ – it is obvious, that the question "When should I apply?" becomes crucial. To address the DI application timing in a meaningful way, a dynamic framework is needed.

Agents' Choices and Social Welfare. Assume that the agent's time horizon consists of T periods, indexed by $t = 0, \ldots, T-1$. Denote by $\theta_{i,t}$ the disability shock, by $\chi_{i,t}$ a vector of other shocks (such as wages/productivity and other factors) influencing the DI application choice, and by $A_{i,t}$ the level of financial assets available at the beginning of period t. Once the state vector $X_{i,t} = (\theta_{i,t}, A_{i,t}, \chi_{i,t})$ is revealed, agent i decides whether to apply for DI, and if rejected, whether to resume work or claim social welfare. The application and work decisions are based on knowledge of $X_{i,t}$ and expectations about future realizations of $X_{i,t+s}$, $s = t+1, \ldots, T-1$. Simultaneously with the DI application choice, the agent decides how much to consume and save in period t.[24] The decisions in period t determine $A_{i,t+1}$ and, together with realizations $\theta_{i,t+1}$ and $\chi_{i,t+1}$, form the state vector X_{t+1}, on the basis of which the agent makes her $t+1$ choices, and so on.

[23] The implicit assumption here is that DI generosity does neither affect the intensive margin of labor supply (DI recipients do not work on the labor market) nor the outflow from DI (DI is an absorbing state, no DI spell ever terminates to a regular job or any other destination).

[24] The within-period sequence of work and DI-application choices is just like the one of the static model, captured in Figure A.1. However, the general model also admits the possibility that $\theta^A \geq \theta^R$, so that equation (A.1) is violated. This might occur for agents with low wage realization and low DI acceptance probabilities.

The utilitarian government can freely choose DI policy parameters $P = (\theta_0^*, \ldots, \theta_{T-1}^*; b_0, \ldots, b_{T-1})$ and seeks to maximize the objective

$$\max_P W(P) = \int_i V_i(P) di + \lambda \left(G(P) - \bar{G} \right), \tag{7}$$

where $W(P)$ denotes social welfare under policy P; $V_i(P)$ is the (expected) indirect lifetime utility of agent i (who responds optimally to policy P), λ is the Lagrange multiplier on the government's budget constraint, $G(P)$ is the net fiscal revenue, and $\bar{G}$ is an exogenous revenue constraint. $G(P)$ is given by

$$G(P) = \int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} \left(W_{i,t} \cdot \tau_{i,t} - D_{i,t} \cdot b_{i,t} - Z_{i,t} \cdot z_{i,t} \right) \right] di, \tag{8}$$

where $(D_{i,t}, W_{i,t}, Z_{i,t})$ denote the probabilities that in period t agent i is on DI, at work or on social welfare. In Appendix A.2 we show that agent i's indirect (expected) lifetime utility can be written as

$$\begin{aligned} V_i(P) \quad = \quad & \max E \left[\sum_{t=0}^{T-1} \beta^t \left(v(c_{i,t}^D) \cdot D_{i,t} + v(c_{i,t}^Z) \cdot Z_{i,t} + \left(u(c_{i,t}^W) - \theta_{i,t} \right) \cdot W_{i,t} - \Lambda_{i,t} \cdot \psi \right) \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \left((1+r_t) A_{i,t} + b_{i,t} - c_{i,t}^D - A_{i,t+1} \right) D_{i,t} \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^W \left((1+r_t) A_{i,t} + w_{i,t} - \tau_{i,t} - c_{i,t}^W - A_{i,t+1} \right) W_{i,t} \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^Z \left((1+r_t) A_{i,t} + z_{i,t} - c_{i,t}^Z - A_{i,t+1} \right) Z_{i,t} \right], \end{aligned} \tag{9}$$

where the first line summarizes agent $i's$ period utilities, with $(c_{i,t}^D, c_{i,t}^W, c_{i,t}^Z)$ as the consumption levels in the various states, and $\Lambda_{i,t}$ as the DI application indicator. The remaining lines are agent i's budget constraints associated with being on DI (second line), at work (third line), and on social welfare (fourth line). The corresponding Lagrangian multipliers are denoted by $(\mu_{i,t}^D, \mu_{i,t}^W, \mu_{i,t}^Z)$.

Stricter DI Eligibility Rules. We now explore the welfare effects of marginally changing the strictness of DI eligibility rules θ_s^*, while leaving all other elements of the DI policy vector $P = (\theta_0^*, \ldots, \theta_{T-1}^*; b_0, \ldots, b_{T-1})$ unchanged. Notice that this thought experiment is equivalent to an RSA increase, the policy change we exploit below to empirically estimate the effect of stricter DI eligibility rules. An RSA policy implies that θ_t^* takes high values up until age $R-1$ and falls to lower values from age R onward. If the relaxed screening age is increased from age $R = s$ to $R = s+1$, this is equivalent to an increase in θ_s^* but unchanged values of $\theta_{t \neq s}^*$.[25] In Appendix

[25] Notice further that our analysis in the text studies the welfare effects of a marginal increase θ_s^* while an RSA policy typically implies a discrete change in θ_t^* at the RSA. Assume that $\theta_t^* = \theta^H$ for ages $t = 0, \ldots, R-1$ and $\theta_t^* = \theta^L < \theta^H$ for ages $t = R, \ldots, T-1$. Then an increase in the RSA from $R = s$ to $R = s+1$ is associated with a discrete change in θ_s^* equal to $\triangle \theta_s^* = \theta^H - \theta^L$. We discuss the welfare effects of a discrete change in θ^* in Appendix A.1 for the static model and in Appendix A.2 for the general

A.2 we show that $\partial W(P)/\partial\theta_s^* \gtreqless 0$ is equivalent to

$$1 + \frac{\mathbb{E}\left[B(\theta_s^*)\right]}{\mathbb{E}\left[M(\theta_s^*)\right]} \gtreqless \frac{\mathbb{E}\left[L_W\right] + \mathbb{E}\left[L_Z\right]}{\lambda \cdot \mathbb{E}\left[M(\theta_s^*)\right]}, \tag{10}$$

where the operator $\mathbb{E}[Y]$ encompasses aggregation of the variable $Y_{i,t}$ across individuals, time and states of nature.[26] The left-hand-side is the fiscal multiplier of increasing θ_s^* where $\mathbb{E}\left[M(\theta_s^*)\right]$ denotes the mechanical fiscal effect and $\mathbb{E}\left[B(\theta_s^*)\right]$ is the behavioral fiscal effect. The right-hand side, $\mathbb{E}\left[L_W\right] + \mathbb{E}\left[L_Z\right]$ are the dynamic insurance losses arising from fewer agents being admitted to the DI program in period s. Normalizing by the Lagrange multiplier λ (= the value to society of relaxing the government budget constraint), yields the money-metric of these utility losses. In Appendix A.2, we make explicit how $\mathbb{E}\left[B(\theta_s^*)\right]$, $\mathbb{E}\left[M(\theta_s^*)\right]$, $\mathbb{E}\left[L_W\right]$ and $\mathbb{E}\left[L_Z\right]$ are determined.

Notice the similarity of the social optimality condition (10) to the social optimality condition (5) of the simple static framework of section 2.1.[27] A key difference between the static and the general model is that an increase in θ_s^* – stricter DI eligibility rules at some age s – does not only affect the DI inflow at that age s, but also at other ages. The behavioral fiscal effect of an increase in θ_s^*, $\mathbb{E}\left[B(\theta_s^*)\right]$, can occur in all periods, even before age s, as forward-looking individuals might change their behavior already at younger ages. The mechanical fiscal effect, $\mathbb{E}\left[M(\theta_s^*)\right]$, persists at older ages because DI is an absorbing state. If many applicants are screened out today, more applicants will reapply tomorrow. As a result, the mechanical effect $\mathbb{E}\left[M(\theta_s^*)\right]$ spreads out over the age window $[s, T-1]$. In Section 6 we will decompose the estimated fiscal cost reductions into its behavioral and mechanical components, and provide direct evidence on the persistence of the mechanical fiscal effect.

Discrete versus Marginal Increase in θ^*. As mentioned above, the policy change that allows us to study the impact of an increase in the strictness of DI eligibility rules is an RSA increase, which is associated with a discrete (rather than marginal) increase in θ^*. In Appendix A.2 we show this leaves the left-hand-side of social optimality condition unchanged. However, the right-hand-side needs to take into account the insurance losses of the marginal applicants. In other words, the exact same logic that we discussed in the static framework above, applies to the general model.

Lower DI Benefits. Alternatively, the DI reform may implement lower DI benefits. So, let us consider the welfare effects of a reduction in the DI benefit b_s (while leaving DI benefits unchanged at all other ages). In

model. Our empirical implementation of the fiscal multiplier is robust to non-marginal changes. Kleven (forthcoming) discusses the issues when studying discrete rather than marginal changes in benefit levels.

[26] Formally $\mathbb{E}[Y] = \int_i \sum_{t=0}^{T-1} E(Y_{i,t})di$ with $E(Y_{i,t}) = \int_{X(i,t)} Y(i,t)dF(X_{i,t})$.

[27] The optimal DI formulas (10) and (11) let us calculate the welfare gains of DI reforms in a broad set of stochastic dynamic environments, such as investments in health or human capital (that might accommodate the disability and productivity shocks), borrowing constraints, spousal labor supply, home production, etc. In this respect, the analysis of optimal DI is analogous to the analysis of optimal UI studied in Chetty (2006a).

Appendix A.2 we show that that condition $-\partial W(P)/\partial(b_s) \gtreqless 0$ is equivalent to

$$1 + \frac{\mathbb{E}\left[B(b_s)\right]}{\mathbb{E}\left[M(b_s)\right]} \gtreqless \frac{\mathbb{E}\left[v'(c^D)\right]}{\lambda \cdot \mathbb{E}\left[M(b_s)\right]}, \tag{11}$$

where $\mathbb{E}\left[B(b_s)\right]$ and $\mathbb{E}\left[M(b_s)\right]$ are the behavioral and mechanical fiscal effects of a marginal reduction of b_s. Again, this looks very similar to the static model. Just like before, behavioral responses to a reduction of b_s occur in all periods. Mechanical responses occur at age s only (because we consider lower benefits paid out at age s but unchanged benefits at all other ages).

3 Institutional Background and Data

3.1 Institutional Background

Like in many developed countries, Austria has three transfer programs that provide income replacement for economic or health reasons: disability insurance (DI), sickness insurance (SI), and unemployment insurance (UI). The DI program is financed by a payroll tax on earned income and provides partial earnings replacement to workers below the full retirement age with at least 5 insurance years within the last 10 years.[28]DI applicants must submit their application to the local DI office. Employees at the DI office first check whether the applicant meets the formal requirements for DI receipt. Importantly, and different from the U.S, DI applicants are not required to stop working. Then a team of disability examiners and physicians assesses the severity of the medical impairment and the applicant's residual earnings capacity. An impairment is considered to be severe if it lasts at least six months and limits the applicant's mental or physical ability to engage in substantial gainful activity.

DI Eligibility Rules. The assessment of the applicant's residual earnings capacity depends on work experience and whether his or her age is below or above a *relaxed screening age (RSA)* threshold, currently set at age 60. Applicants below the RSA are awarded DI benefits if the earnings capacity has been reduced to less 50% of the earnings capacity of a healthy person in *any reasonable* occupation the individual could be expected to carry out.[29] Applicants above the RSA (who have worked for at least 10 years within the last 15 years) need to have an earnings capacity of less than 50% in a *similar* occupation.[30] The RSA was 57 until the end of 2012 and was increased in three one-year steps to age 60 by 2017. We exploit the variation in the RSA to identify

[28]Insurance years include both contribution years (periods of employment, including sick leave and maternity leave) and non-contribution years (periods of unemployment, military service, or secondary education). The required insurance years increase by one month for every two months above age 50 up to a maximum of 15 insurance years.The insurance years requirement does not apply if the disability is job-related; for each occupation there exists an explicit list of qualifying impairments.

[29]Eligibility standards are less strict for semi-skilled and skilled applicants below the RSA threshold, whose set of reasonable occupations is more limited. To be classified as semi-skilled or skilled, an applicant must have worked in a semi-skilled or skilled occupation for 7.5 years or more in the most recent 15 years.

[30]Access to disability insurance is also relaxed in other countries at older ages, including Australia, Canada, Denmark, Sweden until 1997 (Karlström et al., 2008), and the United States (Chen and van der Klaauw, 2008).

the labor market effects of stricter DI eligibility rules (section 4). Once benefits are awarded, DI beneficiaries receive monthly payments until their return to work, medical recovery or death. DI benefits can be granted for a temporary period, but less than 4 percent of claimants ever leave the DI rolls.

DI Benefits. DI benefits are subject to income and payroll taxation and replace approximately 70 percent of pre-disability net earnings up to a maximum of about €4,500 per month. The level of DI benefits is calculated by multiplying a pension coefficient, which varies by age and insurance years, with an assessment basis, which is the average indexed capped earnings over a given period of time (e.g., the best 16 years in 2004 at the beginning of our study period). Younger applicants with limited work experience qualify for a special increment to supplement their benefits. DI beneficiaries may continue work, but those earning more than an exempt threshold lose up to 50 percent of their benefits.[31] A pension reform in 2004 gradually decreased the benefit levels for most workers, providing exogenous variation we use to identify the labor market effects of changes in benefit levels (section 5).

SI and UI Benefits. In case of a temporary illness, employers continue to pay 100% of earnings for up to 12 weeks. Once the right to full benefits paid by the employer has expired, individuals may claim SI benefits which are taxed and replace approximately 65% of the last net wage up to the same maximum that applies to DI benefits. SI benefit duration is 52 (26) weeks for individuals who have worked at least (less than) 6 months in the previous 12 months. UI benefits replace 55 percent of the previous wage subject to a minimum and maximum. The maximum UI benefit duration 39 weeks of regular UI benefits for workers below 50 and 52 weeks for workers above 50 (provided they have paid UI contributions for at least 9 years in the last 15 years). Job losers who exhaust the regular UI benefits can apply for unemployment assistance. These means-tested transfers last for an indefinite period and are about 70 percent of regular UI benefits.

3.2 Data

We merge data from two administrative registers. First, the Austrian Social Security Database (ASSD) contains detailed longitudinal information for the universe of workers in Austria between 1972 and 2018. The ASSD records all employment, unemployment, disability, sick leave, and retirement spells as well as a limited set of background characteristics (gender, month and year of birth, blue- or white-collar status). Spells before 1972 are available for individuals who have claimed a public pension by the end of 2008. The ASSD also contains some firm-specific information: geographic region, industry affiliation, and firm identifiers that allow us to link both individuals and firms. See Zweimüller et al. (2009) for a detailed description of the data. Second, we

[31] Ruh and Staubli, 2019 show that this policy induces DI beneficiaries to keep their earnings below the exempt threshold in order to retain benefits.

use data on all DI applications, which cover the period 2004 to 2017 and contain detailed information on the date of the application, the date of the decision, the decision itself (i.e. reject or accept), the reported medical impairment of the applicant, and the stage of the application (i.e. first application, re-application, or appeal).

Starting from the population data set, we impose three restrictions. First, we exclude women because their eligibility age for an old age pension gradually increased from age 56 to age 60 during our observation window, making it difficult to disentangle the effect of DI reforms from the effect of increasing the retirement age.[32] Second, we exclude self-employed and civil service workers, because they are covered by a different pension system than private-sector workers. Third, we exclude observations in which individuals are over age 62, at which point many become eligible for an old age pension. Our sample covers more than three quarters of all active labor market participants in Austria. Since we observe complete work histories, we can precisely calculate how much DI benefits individuals would get at any point in time and whether individuals have sufficient work experience to apply for DI benefits under the relaxed eligibility criteria above the RSA.

Table B.1 in Appendix B shows summary statistics for the sample we use to study the effects of stricter DI eligibility rules. To capture changes in labor market behavior around the RSA, we limit the sample to men between age 54 and age 62 with at least 10 employment years in the past 15 years (measured at age 56). These men are considered eligible for relaxed DI eligibility, while men with less than 10 employment years in the past 15 years are considered ineligible.[33] We will use the sample of ineligibles for placebo tests. Since our empirical strategy exploits increases in the RSA from 57 to 58 and from 58 to 59, we distinguish between three cohorts of men: RSA 57, RSA 58, and RSA 59 who qualify for relaxed DI eligibility at age 57, age 58, and age 59, respectively. We observe individuals on a quarterly basis.

Our first set of outcome variables focus on DI application behavior. *DI application ever* is an indicator for whether an individual has ever applied for DI benefits. *DI application yearly* is an indicator for whether an individual has applied for DI benefits at a particular age. We also distinguish yearly applications by the underlying health impairment (mental disorders, musculoskeletal disorders, and other disorders) and whether the applications is a re-application, meaning that the applicant has applied for DI before. Our second set of outcome variables focus on labor market outcomes. *DI benefit receip*t is an indicator for whether an individual is receiving DI benefits, *employment* is indicator for whether an individual is employed, and *other benefit receipt* is an indicator for whether the individual is receiving UI or SI benefits.[34] In the empirical analysis, we also calculate the benefit and earnings streams associated with each labor market status, allowing us to study the

[32] Staubli and Zweimüller (2013) show that this increase had sizeable employment and unemployment effects.

[33] Note that only individuals who worked in a *similar occupation* for 10 of the last 15 years are eligible for relaxed DI eligibility, while our definition is based on whether somebody has worked in *any occupation* for 10 years of the last 15 years because we can only observe industry affiliation and not occupation. This implies that the eligible sample will include some individuals who are in fact not eligible for relaxed screening, but this number is likely small because what constitutes a similar occupation is defined broadly.

[34] DI spells are back-dated in the ASSD to the date the claim was filed, so an individual who applied for DI benefits late in the calendar year and was awarded benefits in the next calendar year is observed to claim benefits in the calendar year when the application was filed.

fiscal effects of stricter DI eligibility rules.

Table B.2 in Appendix B shows summary statistics for the sample, we use to study the effects of changes in DI benefit levels. Following Mullen and Staubli (2016), we define a reference date, January 1, and obtain all information to compute potential DI benefits and other relevant individuals characteristics as of this date for each year an individual is not receiving DI benefits. We estimate the effects separately for the age groups 30-56 and 57-60, which is also the age group of interest when studying the effects of stricter DI eligibility rules. Our main outcome variables of interest are indicators for whether, within a year, individuals apply for DI (*DI application*), are awarded DI benefits (*DI inflow*), exit employment (*employment outflow*), or stop receiving UI or SI benefits (*other benefit outflo*w).

4 The Effect of Tighter DI Eligibility Rules

4.1 The 2013 DI Reform

In April 2012, the Austrian government announced the 2. Stability Act (2. Stabilitätsgesetz), which became effective on January 1, 2013. The Act had two objectives: reduce expenditures in the public pension systems and foster employment among older workers. The only change to the DI program was a stepwise increase in the RSA threshold from age 57 to age 60. Up until December 2012 the RSA was age 57. The RSA was increased to age 58 in January 2013, followed by further increases to age 59 in January 2015 and age 60 in January 2017. Individuals who had not worked in a similar occupation for 10 years in the last 15 years were not affected by the increases as they were not eligible for relaxed DI eligiblity rules. We focus on the increases in the RSA to 58 and 59, because the available data preclude the analysis of the increase in the RSA to 60.

The RSA increases create variation in the tightness of DI eligibility rules at certain ages across birth cohorts. For example, the RSA is 58 for men who turn 57 between December 2012 and November 2013 (those born after November 1955 and before December 1956).[35] We label this birth cohort the RSA-58 cohort. Conversely, the RSA is 57 for men born before December 1955 and we label this cohort the RSA-57 cohort. Men in the RSA-58 cohort, compared to men in the RSA-57 cohort, face stricter disability screening at age 57. The RSA is 59 for men born after November 1956 and we label this cohort the RSA-59 cohort. Men in the RSA-59 cohort, compared to men in the RSA-57 cohort, face stricter DI eligibility rules at ages 57 and 58. Figure B.5 in Appendix B illustrates the step-wise increase graphically.

Figure B.6 in Appendix B provides descriptive evidence on the labor market effects of the RSA increases. Trends in labor market outcomes across birth cohorts are remarkably similar until age 57 – the relaxed screening age for the RSA-57 cohort. At this age, the DI recipient rate rises sharply in the RSA-57 cohort. The percent

[35] Applications are assessed using the rules in the month after filing. Therefore, if someone turns 57 in December 2012 and applies to DI his application is evaluated in January 2013, when the new RSA of 58 applies.

of DI applicants also increases, suggesting that individuals are aware of the RSA and time their DI application to this age. Conversely, the percent of men who are employed or receive other benefits drops at age 57, pointing to the role of DI as a substitute for UI or SI. We observe similar breaks in trends when the RSA-58 and RSA-59 cohorts reach their RSA. Interestingly, cohorts with a higher RSA never catch up to cohorts with a lower RSA. To capture these persistence (and also potential anticipation) effects, our empirical strategy is designed to identify the entire age profile of the RSA increase.

4.2 Estimation Strategy

We exploit the exogenous variation in the RSA threshold across birth cohorts in a difference-in-differences design. Control (= older) birth cohorts are eligible to the more lenient DI eligibility rules already at age 57 (RSA=57), while treated (= younger) birth cohorts are eligible only at age 58 or age 59 (RSA=58 or 59) . Thus, we can identify the effect of stricter DI eligibility rules by comparing the age profiles of younger and older birth cohorts. This comparison can be implemented by estimating regressions of the following type:

$$y_{ict} = \alpha + \theta_a + \pi_c + \lambda_t + \sum_{k=54\backslash 56}^{61} \beta_k I[age = k] + X'_{ict}\delta + \varepsilon_{ict}, \tag{12}$$

where i denotes individual, c denotes birth cohort, and t denotes year-quarter; y_{ict} is the outcome variable of interest (such as an indicator for receiving DI benefits), θ_a are dummies for age in years to control for age-specific levels in the outcome variable, π_c are dummies for year-month of birth to capture time-constant differences across birth cohorts, λ_t are dummies for year and quarter to capture common time shocks and seasonal effects, and X_{ict} represent individual or region specific characteristics to control for any observable differences that might confound the analysis.[36] We cluster standard errors at the year-month of birth.

The key variables of interest are the indicators $I[age = k]$, which are equal to one if an individual's age is equal to k, where k runs from 54 to 61 using $k = 56$ as the reference age. Each β_k-coefficient measures the average causal effect of an RSA increase at age k. To obtain the average effect of an RSA increase over a wider age interval, we can simply take the average of different β_k-coefficients. For example, $\sum_{k=57}^{61} \beta_k/5$ measures the average change in the outcome variable at each age in the age interval 57 to 61.

We estimate the effects of the RSA-58 and RSA-59 change separately, using always the RSA-57 cohort as the control group. This way we can directly compare the effects of a one-year and a two-year RSA increase. Another reason to estimate the effects separately is that, compared to the RSA-58 cohort, men in the RSA-59 cohort have more time to adjust to the reform. They just turned 55 years old when the reform was announced, while men in the RSA-58 cohort were almost 57 years old. Having more time to adjust increases the scope for

[36] The fixed effects θ_a, π_c, and λ_t are not collinear, because each age-in-year and year-quarter cell contains cohorts with different year-month of birth.

anticipation effect: changes in behavior even before age 57.

The identification assumption is that, absent the increase in the RSA, the change in y_{ict} at a certain age would have been comparable between treated birth cohorts (RSA=58 or 59) and control cohorts (RSA=57). A potential concern is that age-specific trends in the outcome variable could change across birth cohorts for reasons unrelated to the RSA increases. The estimated β_k-coefficients for $k < 57$ provide placebo checks for spurious trends. They should not be statistically significant if the identification assumption holds, although they could also pick up anticipation effects. As an additional placebo check, we estimate equation (12) for men who never become eligible for to the lenient DI eligibility rules because they have worked less than 10 years in the past 15 years. They should not respond to the changes in the RSA.

4.3 Empirical Results

Figure 2 shows the estimated β_k-coefficients from equation (12) for the RSA-58 and the RSA-59 increases for four key outcomes: DI benefit receipt, DI application ever, employment, and other benefits. The shaded area denotes the 95 percent confidence interval. In all graphs, we see that the estimates before age 57, the pre-reform RSA, are close to zero and statistically insignificant, providing evidence that the estimates are not confounded by differential trends across birth cohorts.

As panel (a) shows, because of the RSA increases, fewer men receive DI benefits between ages 57 and age 61. DI recipiency rates drop by about 4 percentage points at age 57. For the RSA-58 cohort, DI recipiency rate remains lower after age 57, even though DI eligibility rules have become more lenient. For the RSA-59 cohort, the DI recipiency rate declines further at age 58 and is still lower at age 59 when this cohort qualifies for relaxed DI eligibility rules. If applying for DI imposes utility costs, we would expect that fewer people apply when eligibility criteria are strict. Indeed, panel (b) shows that DI application rates for the RSA-58 and RSA-59 cohorts drop at all ages above 56.[37] Panels (c) and (d) show that stricter DI eligibility rules increase employment and other benefit receipt above age 56.[38] The expansion in employment persists until the last age we can observe in the data, and is about twice as large for the RSA-59 cohort compared to the RSA-58 cohort. While the rise in other benefit receipt is temporary for the RSA-58 cohort, it persists up to the last age for the RSA-59 cohort.

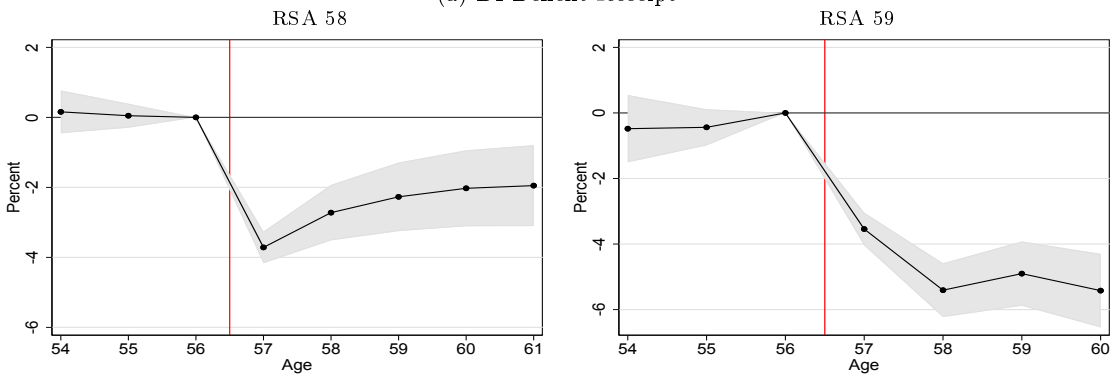
It is interesting to look at the timing and dynamics of the estimates in Figure 2. First, we find no evidence for anticipation effects, which is less surprising for the RSA-58 cohort, because they learned about the reform

[37] Appendix Figure B.11 decomposes DI application ever by impairment type. It shows that fewer individuals apply with musculoskeletal and other impairments, but the same number of individuals apply with mental impairments.

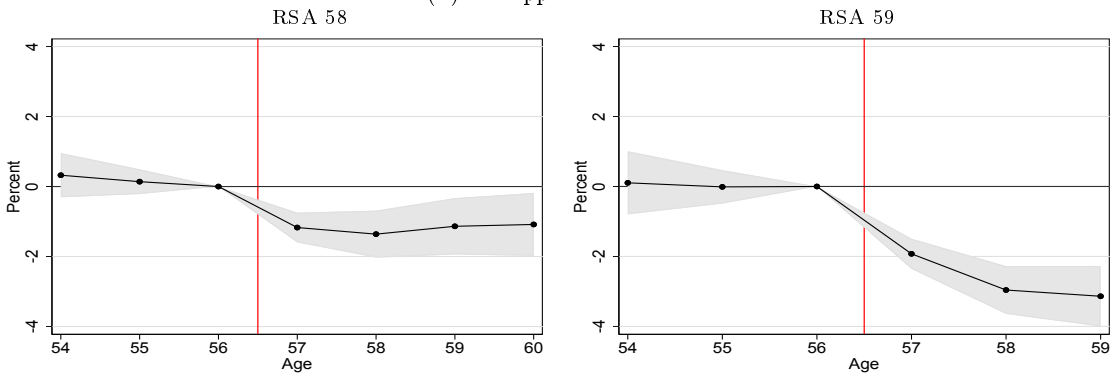
[38] Appendix Figure B.12 shows that the expansion in employment is primarily driven by individuals who are already employed and who stay employed longer, rather than by individuals who are on other benefi receipt and who start working when the RSA increases. Similarly, Appendix Figure B.13 shows that the increase in other benefit receipt is primarily driven by individuals who are already reciving other benefits and now receive other benefits longer until they reach the new RSA, at which point many transition to the DI program.

Figure 2: Effects of RSA on Labor Market States and DI Application Ever by Age

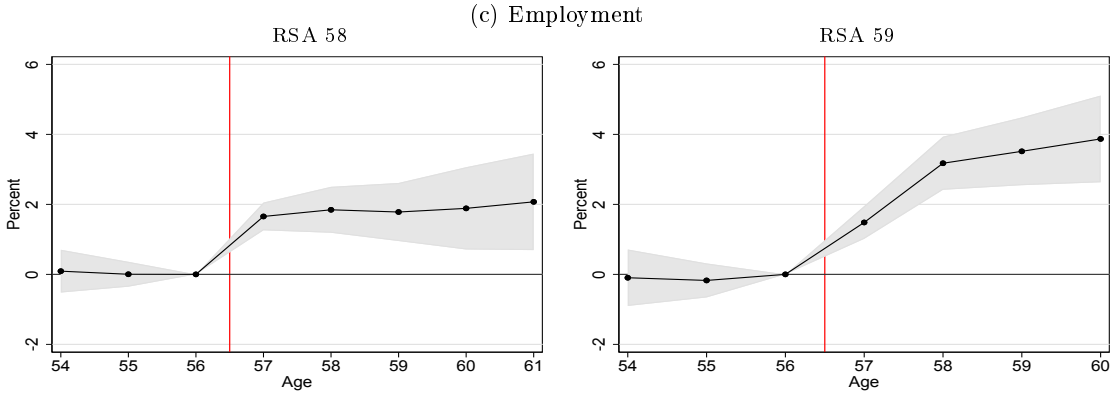
(a) DI Benefit Receipt



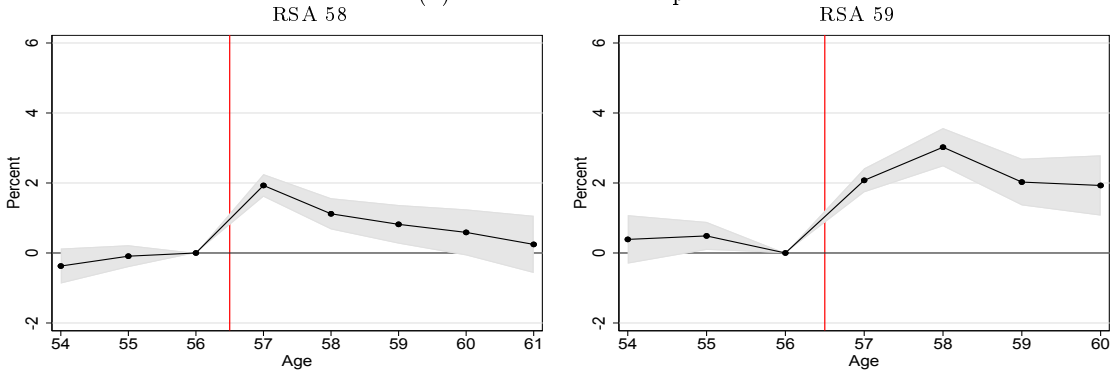
(b) DI Application Ever



(c) Employment



(d) Other Benefit Receipt



Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA-58 and RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

just a couple months before turning 57. The RSA-59 cohort knew about the reform two years before turning 57 and had time to adjust, but all estimates before age 57 are close to zero and insignificant. Second, the DI application rate falls at age 57, implying that individuals are aware of the RSA and adjust their behavior. If the estimated effects were purely mechanical, applications at age 57 should not react. Third, the estimated effects are highly persistent and show up at ages beyond the RSA. This is consistent with persistent mechanical effects as discussed in section 2.2 above. The strength of mechanical and behavioral effects is of crucial importance as their relative size determines the effect of tightening DI eligibility rules on social welfare. We discuss welfare effects in Section 6, where we propose an empirical strategy to directly estimate the mechanical effect, allowing us to split up the total effect of the interesting outcomes into its behavioral and mechanical component.

In Figure B.7 in Appendix B, we plot the estimated β_k-coefficients from equation (12) for men with too little work experience to be eligible for the lenient DI eligibility rules. For this "placebo" groups, we find that DI benefit receipt, DI application ever, employment and other benefit receipt do not differ significantly across birth cohorts, even after age 56. This provides strong support that our main estimates are not confounded by differential trends across birth cohorts.

A useful way to summarize the effects of tighter DI eligibility rules is by taking the average of the β_k-coefficients after age 56 (since point estimates are insignificant before age 57). We report these estimates in Table 1, distinguishing between men who are and those who are not eligible for relaxed DI eligibility rules. The estimates capture the average effect between age 57 and age 61 for the RSA-58 increase and between age 57 and age 60 for the RSA-59 increase.[39] The exception are DI application ever, which we observe for one year less.

Concerning the labor market effects (Panel A), we find that the share of men in the RSA-58 cohort receiving DI benefits declines by 2.54 percentage points, or about 14 percent relative to the mean above the RSA. Men in the RSA-58 cohort are also less likely to apply for DI, but the decline is only half as large as the decline in DI benefit receipt. Most men who do not qualify for DI anymore continue to work: the average employment rate increases by 1.85 percentage points. But benefit substitution is also important: other benefit receipt increases by 0.94 percentage points. The labor market effects for the RSA-59 cohort are qualitatively similar but about twice as large compared to the RSA-58 cohort. On the other hand, men not eligible to the RSA (the placebo group) barely change their labor market behavior. The DI recipiency rate declines for non-eligible men in the RSA-59 cohort, but the reduction is about five times smaller relative to eligible men and likely reflects that some non-eligible men qualify for relaxed DI eligibility as they get older.

Panel B reports the fiscal effect of the RSA increases, which are crucial for assessing the welfare effects of stricter DI eligibility rules. We focus on four outcomes: DI benefits, tax revenue, other benefits, and the total

[39] While the effects are still visible at age 61 for RSA 58 and age 60 for RSA 59, these effects should disappear at age 62 when most men in Austria retire (see Figure B.8 in the Appendix). We can estimate the effect of RSA 58 and RSA 59 up to age 62 if we assume that the RSA increases have the same effect until age 62 as at the last age, we currently observe in the data. This assumption is reasonable because the effects stabilize after age 58 as Figure 2 shows. Appendix Table B.3 shows the corresponding average effects up to age 62. They are statistically indistinguishable from the estimates in Table 1.

Table 1: Average Effect of Stricter DI Eligibility Rules

	Eligible				Non-eligible (placebo)			
	RSA 58		RSA 59		RSA 58		RSA 59	
	Estimate	Mean	Estimate	Mean	Estimate	Mean	Estimate	Mean
A. Labor market effects (%)								
DI benefit receipt	-2.54***	18.56	-4.82***	17.3	-0.40	38.17	-0.91**	37.52
	(0.44)		(0.41)		(0.38)		(0.39)	
DI application ever	-1.19***	21.81	-2.67***	20.29	-0.16	38.61	-0.1	37.89
	(0.34)		(0.32)		(0.32)		(0.34)	
Employment	1.85***	68.36	3.01***	71.59	0.32	14.34	0.2	14.73
	(0.39)		(0.39)		(0.3)		(0.33)	
Other benefit receipt	0.94***	7.55	2.26***	7.30	-0.01	19.8	0.49	20.08
	(0.25)		(0.27)		(0.38)		(0.39)	
B. Fiscal effects (Euro)								
DI benefits	-884***	6756	-1727***	6245	-115	11012	-395***	10721
(A)	(161)		(150)		(120)		(113)	
Tax revenue	263***	11185	407***	11625	16	1582	-10	1608
(B)	(56)		(59)		(33)		(35)	
Other benefits	172***	1217	448***	1182	-5	2233	89	2277
(C)	(46)		(57)		(55)		(62)	
Total fiscal effect	-976***	-3213	-1686***	-4199	-135	11663	-297***	11389
(A-B+C)	(185)		(176)		(115)		(113)	
No. Observations	2,444,975		2,176,311		916,207		806,100	

Notes: The table reports the average effect of the RSA for the ages above age 56. The estimates are constructed by taking the average of the β_k-coefficients from equation (12) for $k \geq 57$. The *eligible*-sample comprises men who, at age 56, have worked for at least 10 years in the last 15 years. The *non-eligible*-sample comprises men who, at age 56, have worked less than 10 years in the last 15 years. *Mean* denotes the mean above the RSA for the RSA-57 cohort. Fiscal effects are reported in 2018 Euro. Standard errors clustered at the year-month of birth level are reported in parentheses. Levels of significance: *1%, **5%, and ***1%.

fiscal effect, which is the sum of benefits received minus taxes paid. We calculate the outcomes on an individual basis, multiplying at each age the number of days an individual spends in a given labor market state times the daily benefit received or taxes paid in that state. We then estimate equation (12) for each outcome separately and average the β_k-coefficients above age 56. Appendix Figure B.14 plots the estimated β_k-coefficients. They are close to zero and statistically insignificant before age 57 and statistically different from zero after age 56.

Tighter DI eligibility rules (RSA=58 or 59) lessen spending on DI benefits and raise tax revenues from increased work activity, but they also raise spending on other benefits because of benefit substitution. The reduction in DI benefits in the RSA-58 cohort is 884 Euro per individual and year, which in absolute value is about three times larger than the increase in tax revenue (263 euro) and five times larger than the increase in other benefits (172 euro). Overall, total fiscal costs at each age above 56 declines by 976 Euro per individual and year. The estimates for eligible men in the RSA-59 cohort are about twice as large compared to the RSA-58 cohort. The estimates for men who are not eligible for an RSA (placebo group) are small and mostly insignificant.

5 Impact of Benefit Generosity

The ideal experiment to analyze the impact of a change DI benefits would be to randomize the level of DI benefits across individuals. We emulate this ideal experiment with a quasi-experimental research design that exploits variation in DI benefits from a large pension reform. Our approach follows Mullen and Staubli (2016) who estimate the elasticity of DI claiming with respect to benefit generosity using variation in DI benefits in Austria from several reforms between 1987 and 2010. We differ from their study in two aspects. First, we update their estimates for a more recent time period (2004 to 2017). This period is characterized by lower replacement rates and stricter disability screening compared to the 1980s and 1990s, which could affect the responsiveness of DI claiming and applications to benefit levels. Second, we study the effect of benefit generosity on a novel set of outcomes, including employment, other benefit receipt, and fiscal costs, which are key for assessing the welfare effects of a change in benefit generosity.

5.1 The 2003 Pension Reform

In January 2004, the Austrian government implemented several changes to the calculation of DI benefits as part of a larger reform (Pensionsreform 2003). These changes reduced the potential benefit level for most individuals, although individuals with limited work history experienced an increase in the potential benefit level. Before the reform, they would qualify for a special supplement to their benefits if they were below age 56. The reform gradually increased the age limit for the special supplement to age 60 between 2004 and 2010. Over the same time, the reform phased in a reduction in the pension coefficient and an increase in the penalty for claiming benefits before the normal retirement age (age 65 for men and age 60 for women).[40] The reform also gradually increased the length of the assessment basis from 16 years to 40 years by 2028. The large scale reduction in benefits was heavily criticized by the public. In response to the backlash, the Austrian government passed legislation in 2005, limiting the maximum benefit reduction to five percent of the projected pre-reform benefits. The maximum benefit reduction was then increased by 0.25 percent each year; in 2017 it was equal to 8.25 percent of pre-reform benefits.

Figure B.9 in Appendix B illustrates the effect of the reform by showing the distribution of changes in potential DI benefits between 2004 and 2017. We plot separate figures for men ages 30 to 56 and men ages 57 to 60, which is the age group we focus on when studying stricter eligibility criteria. The reform produced potential winners and losers. About 90 percent of 57-60 year old men experienced a loss in potential DI benefits of up to 10 percent. The remaining 10 percent gained from the increase in age limit for the special increment

[40] Before the reform each insurance year increased the pension coefficient by 2 percentage points, while each year of claiming before the normal retirement age reduced the pension coefficient by 3 percentage points (capped at a maximum of 10.5 percentage points or 15 percent of the pre-penalty pension coefficient, whichever is lower). The reform gradually reduced the pension coefficient adjustment for each insurance year from 2 to 1.78 percentage points between 2004 and 2009 and changed the penalty for each year of early claiming to 4.2 percent of the pension coefficient (capped at 15 percent of the full pension).

and experienced a rise in potential DI benefits. The share of losers and winners are similar among 30-56 year old men, although the losses are more unevenly distributed. About 40 percent experienced a loss in potential DI benefits of at least 5 percent, while 20 percent experienced almost no loss.

5.2 Estimation Strategy

We exploit the variation in DI benefit levels stemming from the 2003 pension reform to estimate the causal impact of changes in benefit levels on labor market and fiscal outcomes. We are interested in estimating the following regression:

$$y_{it} = \alpha + X'_{it}\beta + \gamma b_{it}(Z_{it}) + \lambda_t + \varepsilon_{it}, \tag{13}$$

where i denotes individual, t denotes year, y_{it} is the outcome variable of interest such as applying for DI, X_{it} is a vector of demographic and labor market characteristics, $b_t(Z_{it})$ are log potential DI benefits which are a function of labor market characteristics $Z_{it} \in X_{it}$ (age, insurance years, and the assessment basis), λ_t are year fixed effects, and ε_{it} are any unobserved factors affecting the outcome such as taste for work. The parameter of interest is γ, which measures the average effect of a change in benefit levels on the outcome variable.

As Mullen and Staubli (2016) discuss, if b is a linear function of Z_{it}, we cannot separately identify γ and β because no variation is left in b after controlling for Z_{it}. If γ is a non-linear function of Z_{it}, we can identify γ as long as sufficient residual variation is left in b after controlling for Z_{it}.[41] A drawback of this identification strategy is that it relies heavily on functional form, creating bias in γ if the direct effect Z_{it} of y_{it} is incorrectly specified (Bound, 1989). This problem can be solved by exploiting the 2003 reform, because it creates variation in b that is independent from Z_{it}. Intuitively, with the policy reform we observe individuals with similar Z_{it} but different potential benefits b. This approach is akin to a difference-in-differences estimation strategy, where identification is obtained by relating individuals' differential response to their differential change in benefit levels stemming from the policy reform.

Mullen and Staubli (2016) show that the policy-induced variation in b can be isolated by including the individual-specific (log) hypothetical benefits under each policy regime as additional controls in equation (13).[42] Because of the phased-in nature of the 2003 policy reform, we have 14 different hypothetical benefits for each year from 2004 to 2017:

[41] For example, if we control for Z_{it} in a very flexible way by including polynomials or other transformations of Z_{it},γ may not be identified because potential benefits are collinear with Z_{it}.

[42] This approach has also been used by Fevang et al. (2017) to estimate the effect of temporary disability insurance benefits on the duration of temporary disability insurance spells using policy variation in Norway and by Nielsen et al. (2010) to estimate the response of college enrollment to changes in student aid using a Danish reform.

$$y_{it} = \alpha + X'_{it}\beta + \gamma b_{it}(Z_{it}) + \sum_{r=2005}^{2017} \delta_r b_r(Z_{it}) + \lambda_t + \varepsilon_{it}, \tag{14}$$

where $b_r(Z_{it})$ denotes hypothetical DI benefits under the policy regime r. By controlling for hypothetical DI benefits, we ensure that actual potential benefits are uncorrelated with any unobservable factors affecting the outcome variable, so that γ identifies the causal effect of DI benefits. We assess the quality of our prediction of hypothetical DI benefits under different policy regimes by comparing predicted DI benefits to actual DI benefits for the subsample of beneficiaries who received benefits in 2004 or who began receiving benefits after 2004. Appendix Figure B.15 plots mean matched DI benefits against mean predicted DI benefits. Actual benefits track our predicted benefits very closely. We cluster standard errors at the year-month of birth.

The identification assumption necessary for consistency of our estimates is the standard common trends assumption, which in this case requires that absent the 2003 reform the outcome variable would have evolved similarly across groups with differential change in benefit levels. To test the appropriateness of our identification strategy, we estimate 1,000 placebo regressions in which we randomly assign individuals within each cell defined by year, insurance-year decile, and assessment decile potential benefits $b_r(Z_{it})$ from a different year. If our empirical strategy isolates the policy-induced variation in DI benefits, then we expect the placebo estimates to be clustered around zero.

5.3 Empirical Results

Table 2 summarizes our main results with Panel A providing estimates of equation (14) for labor market outcomes and Panel B providing analogous estimates for fiscal outcomes, which serve as inputs for the fiscal multiplier. We find that a point percent increase in DI benefits increase the propensity to apply for DI benefits by 0.171 percentage points for the age group 57-60 and by 0.014 percentage points for the age group 30-56. These estimates correspond to a 0.64 percent and 0.84 percent increase in the application level. We also observe DI inflow increases by 0.093 percentage points for the age group 57-60 and by 0.003 percentage points for the age group 30-56. Taken together, these estimates imply an award rate of 54 percent (=0.093/0.171) for the marginal applicant in the age group 57-60, and a lower award rate of 21 percent for a marginal applicant in the age group 30-56. The higher award rate for 57-60 year old individuals highlights the significant relaxation in eligibility criteria above the RSA. An increase in benefit levels has no effect on employment but significantly increases outflow from other benefits, suggesting that marginal enrollees were receiving other benefits before being awarded DI benefits.

Concerning the fiscal effects, we observe that one percent increase in DI benefits expands spending on DI benefits and lowers tax revenue, but also lessens spending on other benefits because of benefit substitution. The increase in annual DI spending for a one percent increase in benefits amounts to 36.95 Euro for a 57-60 year

Table 2: Average Effect of Benefit Generosity

	Ages 57-60		Ages 30-56	
	Estimate	Mean	Estimate	Mean
A. Labor market effects (%)				
DI application ever	0.171***	26.71	0.014***	1.66
	(0.019)		(0.003)	
DI inflow	0.093***	18.68	0.003***	1.22
	(0.015)		(0.001)	
Employment outflow	-0.004	71.43	<0.001	89.24
	(0.011)		(0.001)	
Other benefit outflow	0.097***	9.89	0.003***	9.54
	(0.012)		(0.001)	
B. Fiscal effects (Euro)				
DI benefits	36.95***	4,516	2.26***	324
(A)	(3.16)		(0.26)	
Payroll taxes	-2.37***	9,915	-0.19***	10,322
(B)	(1.12)		(0.07)	
Other benefits	-20.62***	1,944	-1.27***	1,630
(C)	(2.33)		(0.24)	
Behavioral fiscal effect	18.69***	-3,455	1.18***	-8,368
(D=A-B+C)	(3.14)		(0.18)	
Observations	1,453,448		15,968,003	

Notes: The table reports estimates for γ from the econometric specification in (14). Fiscal effects are reported in annual 2018 Euro. *Mean* denotes the mean in levels for the year 2004. Standard errors clustered at the year-month of birth level are reported in parentheses. Levels of significance: *1%, **5%, and ***1%.

old individual and to 2.26 Euro for a 30-56 year old individual. About half of the additional DI spending is compensated with lower spending on other benefits, while the loss in tax revenue is quantitatively small. Overall, we find that the behavioral responses to a one percent increase in DI benefits raise annual fiscal spending by 18.69 Euro per 57-60 year individual and by 1.18 Euro per 30-56 year old individual.

Appendix Figure B.10 presents the results of our placebo regressions. The figure plots the empirical cumulative distribution of the 1,000 placebo estimates together with the true estimate for each outcome in Panel A of Table 2. The figures confirm that true increases in benefit generosity lead to unusually large increases in DI inflow, DI applications, and other benefit outflow (but have no effect on employment outflow). In contrast, the placebo increases in benefit generosity lead to estimates that are close to zero.

6 Estimating the Fiscal Multiplier of DI Reforms

The main purpose of this section is to estimate fiscal multipliers of DI policy reforms. There are two main reasons why this is of primary interest. First, fiscal multipliers provide us with an important benchmark for welfare analysis: A DI reform generating a fiscal multiplier of, say, 2 is welfare enhancing, if taking away one dollar from DI recipients yields an insurance loss of less than two dollars, or, put differently, if one dollar in the

hands of a marginal DI recipient has a lower social value as two dollars in the hands of the government.[43] Fiscal multipliers of stricter DI eligibility rules and lower DI benefits provide us with estimates of the left-hand-side of the social optimality conditions (10) and (11). Because we cannot estimate the right-hand-side of these conditions – due to lack of data on the dynamics of consumption, assets and health – we can only provide speculative estimates for the involved insurance losses.[44] The second reason for the crucial importance of fiscal multipliers is that they allow us to compare the effectiveness of alternative DI policy instruments. For instance, if stricter DI eligibility rules generate a fiscal multiplier substantially larger than the multiplier of reducing DI benefits, then a DI reform affecting the same group of agents should implement stricter DI eligibility rules.

In what follows, we first estimate the fiscal multiplier generated by the Austrian DI reforms, separately for tightening DI eligibility rules and reducing DI benefits. We then compare the effectiveness of the two DI policy instruments, taking account of who is mainly affected by them. We also provide a tentative assessment of the welfare effects of DI policies in the U.S., using previous empirical estimates of the impact on the DI inflow of (age discontinuities in) the DI eligibility criteria and of DI benefit changes.

6.1 The Fiscal Multiplier of Stricter DI Eligibility Rules

In the notation of our theoretical model, an increase in the RSA from age R to age $R+1$ corresponds to stricter DI eligibility rules during the age window $[R, R+1]$, which leads to total fiscal cost savings $\mathbb{E}\left[\triangle G(\theta_R^*)\right]$.[45] In Section 4 we have explored the effects of increasing the RSA and came up with an estimate of total fiscal cost savings. However, this does not yet allow us to put a number on the fiscal multiplier. To be able to do that, we need to decompose the total fiscal cost savings into its behavioral and mechanical components, $\mathbb{E}[B(\theta_R^*)]$ and $\mathbb{E}[M(\theta_R^*)]$. This decomposition poses a challenge to the researcher, because these two quantities cannot be directly estimated from the data. The problem is that we cannot directly observe who is a marginal applicant and who is an always applicant. However, this information is essential since the behavioral effect is driven by marginal applicants and the mechanical effect is driven by always applicants.

To make progress, we proceed in two steps. First, we characterize marginal and always applicants using the complier analysis method for difference-in-differences settings (Imbens and Rubin, 1997; Abadie, 2003; De Chaisemartin and D'Haultfoeuille, 2018; Jäger et al., 2019). The complier analysis shows that marginal and always applicants are different on a number of characteristics. Hence, decomposing the total fiscal effects

[43] "Taking away one dollar from DI recipients" here means "taking *mechanically* away one dollar", i.e. not taking into account possible behavioral responses. This corresponds to the thought experiment underlying the optimal DI policy formulas (10) and (11) which are normalized by mechanically saved dollars.

[44] For instance, insurance losses can be calculated, if one is willing assume that DI recipients (i) are hand-to-mouth consumers and (ii) have identical CRRA preferences. Such a procedure is often implemented when data for a direct estimation of insurance losses are not available. In Appendix D we show that, under those assumptions, we can calculate upper and lower bounds of the insurance losses resulting from stricter DI eligibility rules. We also show that insurance losses from reduced DI benefit generosity can be calculated in a straightforward way. Notice that the assumption of hand-to-mouth consumers assumes away self-insurance. Hence, these tentative estimates should be considered an upper bound of the true insurance losses.

[45] In the above empirical analysis, we studied the increase in the RSA from age 57 to both age 58 and age 59. Of course, the exact same logic applies if the RSA is increased by two years rather than only one year.

into its mechanical and behavioral component based on the average DI applicant (a mix of marginal- and always-applicants) is misleading.[46] In a second step, we therefore propose an empirical strategy estimating the mechanical fiscal effect based on a group that is arguably similar to always applicants in the whole population. We argue – and provide supporting evidence – that the subpopulation of previously rejected DI applicants provides such a group. The mechanical fiscal effect in the whole population can then be calculated as $\mathbb{E}[M(\theta_R^*)] = \pi^{AA} * \mathbb{E}[M(\theta_R^*)\,|\text{pre-57}]$, where π^{AA} is the share of always applicants in the whole population, estimated in the reduced-form/complier analysis; and $\mathbb{E}[M(\theta_R^*)\,|\text{pre-57}]$ is the mechanical fiscal effect of the typical always applicant, estimated from the subsample of previously rejected ("pre-57") DI applicants. With an estimate of the mechanical fiscal effect, $\mathbb{E}[M(\theta_R^*)]$, the behavioral fiscal effect is then just the residual of total fiscal costs savings and the mechanical fiscal effect, $\mathbb{E}[B(\theta_R^*)] = \mathbb{E}\left[\triangle G(\theta_R^*)\right] - \mathbb{E}[M(\theta_R^*)]$.

Always- versus Marginal Applicants: A Complier Analysis. Using the complier-analysis method for difference-in-differences settings, we compare the characteristics of DI applicants when DI eligibility rules are lenient to the characteristics of DI applicants when rules are strict. Any differences in these characteristics uncover how marginal applicants (who apply only under lenient rules) differ from always applicants (who apply even when rules are strict), and also from those of never applicants. The exact same contrast as for DI applicants can also be done for DI enrollees (those DI applicants whose application gets eventually accepted).[47]

Table 3 shows the population shares and average characteristics of marginal applicants and enrollees, always applicants and enrollees, and never applicants and enrollees for the RSA-58 change. We estimate a share always applicants $\pi^{AA} = 0.070$ (among individuals aged 57). The shares of marginal and never applicants are $\pi^{MA} = 0.014$ and $\pi^{NA} = 0.916$. Marginal applicants are less likely to be on sick leave at age 56 than always applicants. This is important in the present context because being on sick leave is a good proxy for underlying health problems. Marginal and always applicants have similar average earnings in the best 15 years, though at age 56 the labor market attachment of marginal applicants is stronger than the one of always applicants: 73 % of marginal applicants are employed at age 56, compared to 60 % of always applicants and 87 % of never applicants. Marginal applicants are more likely to be blue-collar workers and are more likely to apply with a musculoskeletal impairment, consistent with low-skilled/manual workers experiencing the largest relaxation in disability eligibility when reaching the RSA. Table 3 also reports the same contrast as for DI *applicants* for DI *enrollees* (accepted DI applicants) and shows similar patterns. The only major difference occurs at age 56 when marginal enrollees are less likely to be employed. This highlights that relaxed DI eligibility increases the

[46] Notice that the complication does not arise in the case of lower DI benefits, because a benefit reduction affects *all* DI recipients. The mechanical fiscal effect (= the fiscal cost reduction with DI application behavior unchanged) can be readily calculated from DI benefit reductions that accrue in the pre-reform control group. In contrast, the mechanical fiscal effect of tightening DI eligibility rules is driven only by the subgroup of always applicants: individuals who – despite the stricter rules – do not abstain from applying to DI.

[47] See Appendix C for a detailed discussion of how we implement the complier analysis in our setting.

probability of a DI award for applicants who are in better health but have poor labor market prospects.[48]

Pre-57 Applicants: Representative for Always Applicants? We argue that the mechanical fiscal effect can be estimated from the subpopulation of individuals who filed a DI application during ages 50-56. In what follows, we refer to this subpopulation as *pre-57 applicants.*[49] This raises the question whether pre-57 applicants – who re-apply at age 57 – are indeed representative for the always applicants of the whole population. Notice first that pre-57 applicants qualify as always applicants in the sense that they have applied under the strict DI eligibility rules (which apply to applications below the RSA). Using pre-57 applicants as our study group, we can perform the exact same evaluation exercise that we performed in Section 4 on the whole population. Just like the whole population, pre-57 applicants can be divided into a treated group (subject to RSA=58) and a control group (RSA=57), and the comparison of the two groups is informative on the causal impact of the RSA increase on pre-57 applicants. Clearly, any effect that emerges at age 57 (or later) can only be identified from agents who have not yet entered DI before they turned 57, i.e. individuals whose previous DI application got rejected, and who re-apply at age 57.[50]

The comparison of treated and control groups among pre-57 applicants delivers an unbiased estimate for the mechanical fiscal effect for the always applicants in the whole population, if the distribution of health shocks (and other relevant characteristics) among pre-57 applicants is identical to the distribution of these characteristics among always-applicants in the whole population. While this identifying assumption is per se fundamentally untestable, we provide several pieces of evidence consistent with that assumption.

First, we look at the (re-)application behavior of pre-57 applicants in the treated and control groups: (i) before age 57, (ii) at age 57, and (iii) after age 57. *Before age 57*, the two groups are subject to the same (strict) DI eligibility rules and we would therefore not expect any (mechanical) difference among them.[51] *At age 57,* the treated group is subject to strict eligibility rules (RSA=58), while the control group is already subject to the lenient rules (RSA=57). However, if the application behavior of pre-57 applicants is indeed representative for a group of always-applicants, a change in the strictness of DI eligibility rules should *not* change their application behavior. *After age 57*, however, we expect re-applications to be different between treated and control groups. As discussed in Section 2.2, the mechanical effect persists at older age, because DI is an absorbing state. If many applicants are screened out today, more applicants will reapply tomorrow. As a result, the mechanical

[48] The analogous results for the RSA-59 change resemble qualitatively the results for the RSA-58 change. This is shown in Appendix Table C.5.

[49] Appendix Table D.6 shows summary statistics for the pre-57 applicant sample.

[50] When the study group is the whole population (as in Section 4), individuals contributing to the identification of an RSA-effect at age 57 (or later) are those who did not yet apply for DI; and those who have previously applied but whose application got rejected. In contrast, the subpopulation of pre-57 applicants, individuals contributing to identification comprise, by construction, only of previously rejected DI applicants.

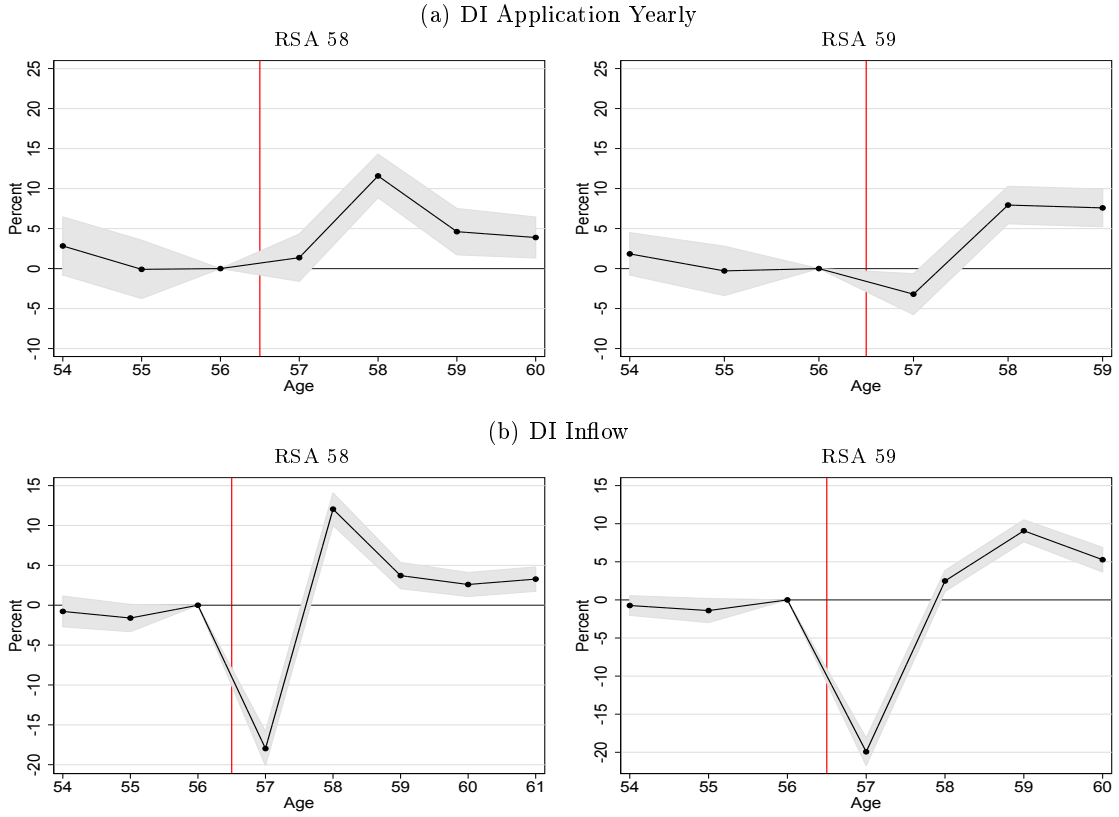
[51] In priniciple, pre-57 applicants in the treated group could change their application behavior *before age 57* – anticipating that their RSA will be larger than the one of the control groups and hence apply more frequently than the control group. If we see such behavioral effects, this would invalidate our assumption that pre-57 applicants are a study group that consists of always applicants only.

Table 3: Applicant and enrollee characteristics, RSA 58

	Marginal (M)	Always (A)	Difference M-A	Never (N)	Difference M-N
A. Applicants					
Share in population	0.014***	0.070***	-0.056***	0.916***	-0.902***
	(0.001)	(0.011)	(0.002)	(0.001)	(0.002)
Sick Leave at age 56 (%)	1.00	9.63***	-8.63***	1.03***	-0.03
	(1.87)	(0.30)	(2.11)	(0.02)	(1.87)
Unemployed at age 56 (%)	21.02***	26.02***	-5.00	4.91***	16.11***
	(3.38)	(0.59)	(3.89)	(0.04)	(3.38)
Employed at age 56 (%)	72.94***	60.29***	12.65***	86.57***	-13.64***
	(3.85)	(0.65)	(4.39)	(0.07)	(3.85)
Avg. annual earnings best 15 years (Euro)	41,183***	40,894***	289	46,074***	-4,891***
	(791)	(146)	(918)	(27)	(792)
Blue-collar (%)	93.26***	81.35***	11.91***	55.29***	37.98***
	(3.59)	(0.63)	(4.12)	(0.12)	(3.60)
Musculoskeletal impairment (%)	59.52***	43.89***	15.63***		
	(4.57)	(0.77)	(5.23)		
Mental impairment (%)	15.27***	14.28***	0.99		
	(3.44)	(0.64)	(4.02)		
Other impairment (%)	25.21***	41.83***	-16.62***		
	(4.66)	(0.77)	(5.33)		
B. Enrollees					
Share in population	0.038***	0.017***	0.022***	0.945***	-0.907***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Sick Leave at age 56 (%)	10.78***	15.9***	-5.18***	1.01***	9.77***
	(0.40)	(0.63)	(0.92)	(0.01)	(0.40)
Unemployed at age 56 (%)	36.13***	23.05***	13.07***	5.13***	31.00***
	(0.71)	(0.88)	(1.39)	(0.04)	(0.71)
Employed at age 56 (%)	49.41***	57.37***	-7.96***	86.44***	-37.03***
	(0.74)	(0.99)	(1.54)	(0.07)	(0.74)
Avg. annual earnings best 15 years (Euro)	40,639***	41,433***	-794*	45,919***	-5,280***
	(177)	(321)	(467)	(25)	(179)
Blue-collar (%)	88.12***	77.20***	10.91***	56.07***	32.04***
	(0.81)	(1.40)	(2.07)	(0.11)	(0.81)
Musculoskeletal impairment (%)	56.40***	28.83***	27.57***		
	(0.98)	(1.43)	(2.17)		
Mental impairment (%)	6.96***	23.46***	-16.50***		
	(0.78)	(1.47)	(2.14)		
Other impairment (%)	35.43***	46.26***	-10.83***		
	(1.02)	(1.69)	(2.53)		

Notes: The table reports the population shares and average characteristics of marginal applicants and enrollees, always applicants and enrollees, and never applicants and enrollees for the RSA-58 increase. We derive these estimates using the complier analysis for difference-in-differences settings described in Appendix C. Earnings are reported in 2018 Euro. Levels of significance: *1%, **5%, and ***1%.

Figure 3: Effect of RSA on DI Application Yearly and DI Inflow by Age for Pre-57 Applicants



Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA-58 and RSA-59 increases using the sample of pre-57 applicants. Pre-57 applicants comprise individuals who have applied for DI between age 50 and age 56. The shaded area denotes the 95 percent confidence interval.

effect $\mathbb{E}\left[M(\theta_R^*)\right]$ needs to include fiscal cost savings accruing after age $R = 57$.

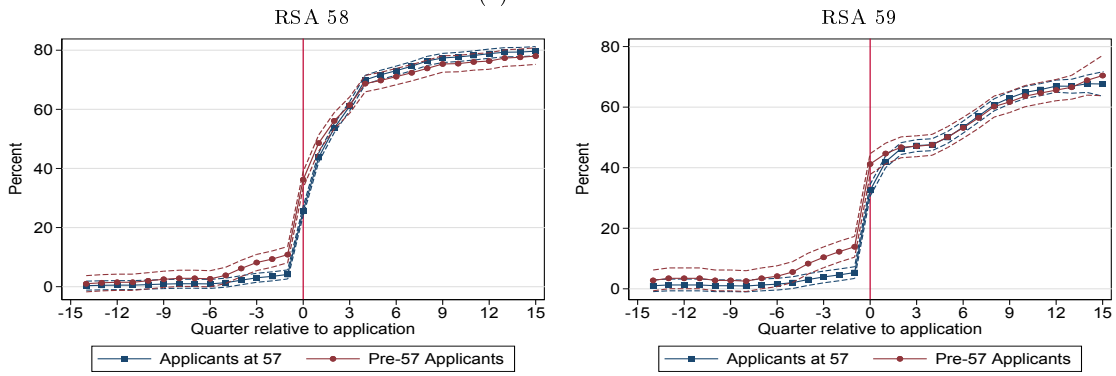
Figure 3 plots the coefficients of the same diff-in-diff strategy as in Section 4 among the subpopulation of pre-57 applicants, with panels (a) and (b) focusing on DI application yearly and DI inflow, respectively. The differences in DI application yearly and DI inflow between treated and control groups line up almost perfectly with those expected from always applicants, both for treated cohorts with RSA=58 (left figure) and RSA=59 (right figure). Before age 57, there are no differences between treated and control groups. At age 57 – when DI application yearly of the treated (but not those of controls) are evaluated under strict DI eligibility rules – there is no (or only a weak) differences in DI application yearly but a large downward spike in DI inflow. This is exactly what one would expect among always applicants: no differences in applications but more rejected applications among the treated. After age 57, we see a significant increase both in DI application yearly and in DI inflow among the treated cohorts.[52]

While the DI application behavior of pre-57 applicants is very similar to what one would expect from a group of always applicants, this does necessarily mean that the mechanical fiscal effect estimated from the

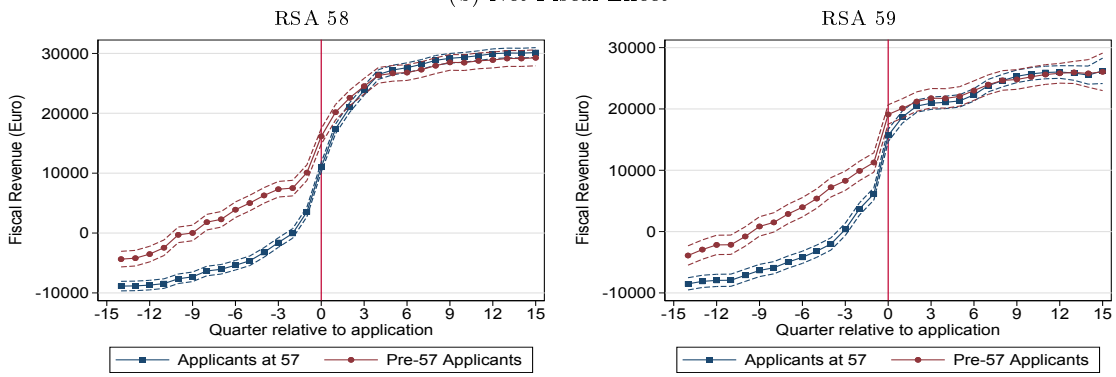
[52]Similar to Section 4, DI application and DI inflow rates are measured in percent of the entire pre-57 subpopulation (rather as hazard rates, where we would condition the application- and inflow-rates at age t on the subpopulation not yet on on DI at age $t-1$).

Figure 4: Comparison of Applicants at 57 and Pre-57 Applicants in Treatment Group

(a) DI Benefits



(b) Net Fiscal Effect



Notes: The figure compares trends in DI benefit receipt (panel a) and the net fiscal effect (panel b) for applicants at age 57 (always applicants) and pre-57 applicants. Always applicants are individuals who apply for DI at age 57 in the treatment group under the strict rules. Pre-57 applicants comprise individuals who applied for DI between age 50 and age 56 and re-apply for DI at age 57. The comparison shows that the two groups are very similar in outcomes after their application at age 57.

subpopulation of pre-57 is representative for the mechanical fiscal effect generated by the always applicants in the whole population. This is because the distribution of characteristics among pre-57 applicants could be very different from the always applicants in the whole population.[53] In Figure 4, panels (a) and (b), we compare, for treated cohorts (RSA=58 and RSA=59), DI benefit receipt and net fiscal expenditures of (i) age-57 applicants in the whole population (blue line) to (ii) age-57 re-applicants in the subpopulation of pre-57 applicants (red line). After date 0, the application date at age 57 (measured in quarters), DI benefit receipt and net fiscal expenditures of the two groups are very similar and 3 quarters after the application the two lines are no longer statistically significantly different from each other. Notice also that, before date 0, the subpopulation of pre-57 applicants looks different from the whole population. This should come to no surprise because pre-57 applicants,

[53] For instance, imagine that rejected pre-57 applicants re-appyling at age 57 have, on average, low incomes, while always applicants in the whole population applying at age 57 have, on average, higher incomes. Then we will estimate a small mechanical fiscal effect among the subpopulation of pre-57 applicants, while indeed the true mechanical fiscal effect – the response of always applicants in the whole population – is large. In that case the inference of the mechanical fiscal effect estimated from the subpopulation of pre-57 applicants to the whole population is invalid. The inference would also be invalid if the distribution of health shocks θ is different between the two populations. For example, it could be that pre-57 applicants are, on average, more healthy than always applicants in the whole population. This would mean that, while θ of each individual is above θ^A (meaning the individuals in both populations are always-applicants), the distribution of the θ's is more skewed to the right among always applicants in the whole population as compared the distribution of the θ's among pre-57 applicants. In that case, the inference is invalid, because based on the subpopulation of pre-57 applicants we underestimate the DI award probability of always applicants in the whole population. As a consequence, we underestimate the DI inflow and hence the mechanical fiscal effect.

by construction, already filed a DI application before date 0, while age-57 applicants in the whole population did not necessarily. Figure 4 supports our identifying assumption that pre-57 applicants are indeed representative for always applicants in the whole population.[54]

Estimating the Fiscal Multiplier. We perform the same diff-in-diff evaluation analysis among pre-57 applicants that we performed in Section 4 among the whole population. Figure 5 plots the diff-in-diff estimates by age for labor market outcomes and the net fiscal effect. Figure 5 provides empirical evidence on the persistence of the mechanical effect (as theoretically discussed in Section 2.2). For the RSA-58 cohort, DI benefit receipt significantly drops at age 57 and then steadily catches up and is at age 59/60 back to the level of the cohort with lenient DI eligibility at age 57. Interestingly, always applicants have a small employment effect at age 57 that vanishes afterwards, the benefit substitution effect is large. This implies that the permanent changes in employment and disability receipt in the population in Figure 2 must be driven by behavioral changes and are not due to a persistent mechanical effect. We see similar patterns for the RSA-59 cohort, where the mechanical effect persists for two years and then starts to disappear, as one would expect.

We are now ready to decompose the total fiscal cost savings into its mechanical and behavioral component and calculate the fiscal multiplier of tightening DI eligibility rules. Table 4 presents the decomposition of the fiscal effect into behavioral and mechanical fiscal effect. A one year increase in the RSA generates a net fiscal effect for always applicants of $\mathbb{E}\left[\Delta G(\theta_R^*)|\text{pre-57}\right] = 5,585\,\text{Euro}$.[55] If our identifying assumption is satisfied, we have $\mathbb{E}\left[\Delta G(\theta_R^*)|\text{pre-57}\right] = \mathbb{E}\left[M(\theta_R^*)|\text{pre-57}\right] = \mathbb{E}\left[M(\theta_R^*)\right]/\pi^{AA}$, where the first equality says that the total fiscal effect of pre-57 applicants is purely mechanical and the second equality says that pre-57 applicants are representative for always applicants in the whole population.[56] The mechanical fiscal effect in the whole population can then be calculated as $\mathbb{E}\left[M(\theta_R^*)\right] = 5,585\,\text{Euro} * 0.070 = 391\,\text{Euro}$ where $\pi^{AA} = 0.070$ is the share of always applicants from Table 3. The behavioral fiscal effect is calculated as the difference between the fiscal cost effect from Table 1 and the mechanical fiscal effect, $\mathbb{E}\left[B(\theta_R^*)\right] = 976 - 391 = 585\,\text{Euro}$. This decomposition implies a fiscal multiplier of 2.50. For the RSA increase from 57 to 59 we find a multiplier of 2.05 as displayed in Table 4.

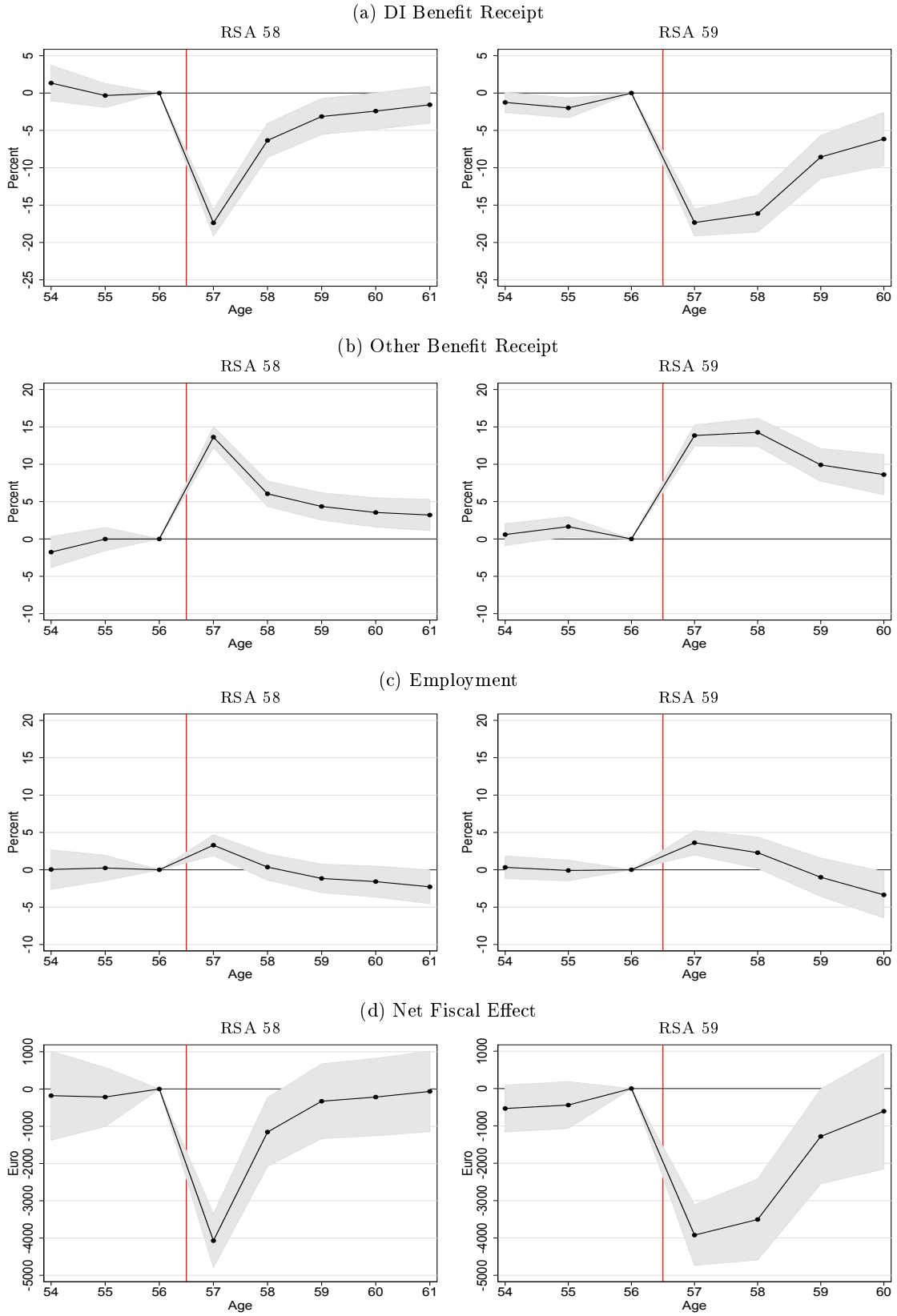
The multiplier has to be compared to the insurance value to assess the welfare effect of the reform. The insurance value measures the social value of one dollar in the hands of DI applicants who are mechanically screened out under the stricter DI eligibility rules. Hence, increasing the RSA by one year (two years) is welfare

[54] In Appendix Figures D.17-D.18 we explore the respresentativeness of pre-57 further. In Figure D.17 we show that pre-57 applicants award rates and application behavior (at and after age 57) that are very similar the those of all applicants at age 57. Appendix Figure D.18 shows that pre-57 and applicants at age 57 are also comparable in terms of employment and other benefit receipt after their age-57 application.

[55] Notice that $\mathbb{E}\left[\Delta G|\text{pre-57}\right]$ is defined as the net fiscal savings among pre-57 applicants *who reapply at age 57*. To obtain this effect we proceed as follows. We first estimate the net fiscal effect among all individuals belonging to the subpopulation of pre-57 applicants, irrespective of whether they filed a re-application at age 57. This yields an estimate of $\mathbb{E}\left[\Delta G|\text{pre-57}\right] * Pr(\text{re-apply at 57}) = 1,167\,\text{Euro}$. We then need to rescale by the probability that a pre-57 applicant reapplies at age 57, $P(\text{reapply at 57}) = 0.209$ to obtain $\mathbb{E}\left[\Delta G|\text{pre-57}\right] = 5,585\,\text{Euro}$.

[56] Notice that, by definition, the mechanical fiscal effect $\mathbb{E}\left[M(\theta_R^*)\right]$ comprises the total (mechanical) cost savings per individual in the whole population, i.e. unconditional on a DI application, while $\mathbb{E}\left[\Delta G|\text{pre-57}\right]$ conditions on an application at age 57.

Figure 5: Mechanical Effect of RSA on Labor Market States and Net Fiscal Revenue by Age



Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA-58 and RSA-59 increases using the sample of always applicants. Always applicants comprise individuals who have applied for DI between age 54 and age 56. The shaded area denotes the 95 percent confidence interval.

Table 4: Fiscal Multiplier for Eligibility Rules and Benefit Generosity

	Eligibility Rules				Benefit Generosity			
	RSA 58		RSA 59		Ages 57-60		Ages 30-56	
	Estimate	%	Estimate	%	Estimate	%	Estimate	%
Total fiscal effect	976		1,686		63.85		4.42	
Mechanical fiscal effect (M)	391	40%	823	49%	45.16	71%	3.24	73%
Behavioral fiscal effect (B)	585	60%	863	51%	18.69	29%	1.18	27%
Fiscal multiplier (1+B/M)	2.50		2.05		1.41		1.36	

Notes: Table presents estimates of the fiscal multiplier for stricter eligibility rules and more generous DI benefits. The fiscal multiplier of stricter eligibility is constructed as follows. The total fiscal effect is taken from Table 1 (with opposite sign). The mechanical fiscal effect is estimated using the sample of pre 57 applicants and then re-scaled by the population share of always applicants (see text for details). The behavioral fiscal effect is the total fiscal effect minus the mechanical fiscal effect.
The fiscal multiplier of benefit generosity is constructed as follows. The behavioral fiscal effect is taken from Table 2. The mechanical fiscal effect captures a 1%-increase in DI benefits for all DI beneficiaries in 2004. It is obtained by multiplying the mean DI benefits in Table 2 with 0.01. The total fiscal effect is the sum of the mechanical and behavioral fiscal effects.

increasing if 1 dollar in the hands of affected DI recipients has a social value of less than 2.50 dollars (2.05 dollars). The fiscal multiplier therefore is the evaluation benchmark for the insurance value. Due to lack of data on consumption and assets, we cannot directly estimate the insurance losses, the right-hand-side of the social optimality condition (10). Nevertheless, we also discuss in Appendix D.2 how we can put a number on the insurance value by assuming hand-to-mouth consumers with identical CRRA preferences. We find that increasing the RSA by one (two) years is welfare improving if risk aversion is below 2.8 (2.2).[57]

6.2 The Fiscal Multiplier of Lower DI Benefits

While estimating the fiscal multiplier of stricter DI eligibility rules is complicated, estimating the fiscal multiplier of the second important DI policy instrument, DI benefits, is rather straightforward. In section 5, we have directly estimated $\mathbb{E}[B(b_s)]$, the behavioral fiscal effect of a DI benefit reduction. According to our estimates in Table 2, the behavioral fiscal effect is 18.69 Euros per year for a 1% DI benefit cut during ages 57-60 (and 1.18 Euros per year for a 1% DI benefit cut during ages 30-56). To determine the fiscal multiplier, we additionally need the mechanical fiscal effect, $\mathbb{E}[M(b_s)]$ – which is simply one percent of the pre-reform mean of DI benefit expenditures. From Table 2, panel B column 2, a DI benefit cut for the 57-60 year old population yields a mechanical fiscal effect of $0.01 \cdot 4,516 = 45.16$ Euro per year (and $0.01 \cdot 324 = 3.24$ Euro per year for a DI benefit cut during ages 30-56, see column 4). The total fiscal effect – the sum of behavioral and mechanical fiscal effects – is then 63.85 Euro (benefit cut during ages 57-60) and 4.42 Euro per year (ages 30-56). The fiscal multipliers of reducing DI benefits is 1.41 for a DI benefit cut during the age group 57-60 (and 1.36 for the age group 30-56). Table 4 summarizes our results.

To assess the overall welfare consequences of a DI benefit cut, the fiscal multiplier needs to be compared to the associated insurance losses. As mentioned above, we lack the necessary data to perform a comprehensive

[57] To the extent that DI recipients are self-insured through own savings, the insurance loss of a DI reform is smaller. The critical level of risk aversion above which makes a DI reform welfare-reducing is then correspondingly higher.

welfare analysis.[58] Nevertheless, the fiscal multiplier provides us with an interesting benchmark: Cutting DI benefits by 1 % is welfare increasing if 1 dollar in the hands of DI recipients has a social value of less than 1.41 dollars.

6.3 Tightening DI Eligibility Rules or Reducing DI Benefits?

Using our results from Austrian DI reforms, we finally explore the *relative* performance of the two DI policy instruments in terms of economic welfare. The above analysis has shown that the fiscal multiplier of stricter DI eligibility rules is significantly larger than the fiscal multiplier of reducing DI benefits. A mechanical one-dollar reduction of the DI budget reduces the overall expenditures 1.8 times (= 2.50/1.41) more strongly when the reduction is due to stricter DI eligibility rules (RSA increase to age 58) rather than due to lower DI benefits. Put differently, if the insurance losses associated with stricter DI eligibility rules and lower DI benefits were equally large, policy makers should tighten eligibility rather than cutting benefits.[59]

Clearly, insurance losses of the two policy instruments are not identical. Comparing the right-hand-sides of the social optimality conditions (10) and (11) shows that stricter DI eligibility rules affect only the additionally screened-out DI applicants, while lower DI benefits affect all DI recipients alike. While we lack the necessary data (on consumption, wealth and health conditions) to put a number on the absolute value of the insurance losses, we can make progress comparing the insurance losses associated with the two DI policy instruments.

We proceed in two steps. We first use our theoretical framework developed in section 2.2 to derive a condition comparing the insurance loss of tighter DI rules with the loss of lower DI benefits. Then we implement this condition empirically and show that the insurance loss associated with tighter DI eligibility rules is, in all likelihood, smaller than the insurance loss associated with lower DI benefits.

In Appendix D.5, we show that a sufficient condition for the insurance loss of stricter DI eligibility rules being smaller than the insurance loss of lower DI benefits is

$$\mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\cdot\frac{v_i'(c_{i,t}^D)}{\lambda}\cdot\underbrace{\frac{\Delta D_{i,t}\left(b_{i,t}-z_{i,t}\right)}{\mathbb{E}\left[M_\Delta(\theta_s^*)\right]}}_{\text{upper bound income loss}}\right]\leq\mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\cdot\frac{v_i'(c_{i,t}^D)}{\lambda}\cdot\underbrace{\frac{D_{i,t}\left(b_{i,t}^H-b_{i,t}^L\right)}{\mathbb{E}\left[M_\Delta(b)\right]}}_{\text{lower bound income loss}}\right]. \tag{15}$$

The left-hand-side of condition (15) is an upper bound of the insurance loss associated with tighter DI eligibility

[58] Just like in the case of stricter DI eligibility rules, we quantify the possible insurance losses associated with DI benefit cut under the assumption of hand-to-mouth consumers with identical CRRA preferences. This yields a critical value of relative risk aversion of 1.1 (for a DI benefit cut during ages 57-60), above which the reduction of DI benefits was actually welfare-*reducing*. We show this in Appendix D.3.

[59] This holds if insurance losses are *smaller* than the fiscal gains for both policy instruments, i.e. it is optimal to cut the generosity of the DI program. If, instead, the insurance loss was *larger* than the fiscal multiplier (for both instruments), a DI reform should implement higher DI benefits rather than relaxed DI eligibility rules (given the insurance gains are equal) – the lower multiplier of DI benefits means that increasing benefits is less costly to taxpayers. We focus here on the relative performance of the two policy instruments if a policy maker wants to cut DI program costs.

rules which comes from additionally rejected DI applicants (indicated by $\Delta D_{i,t} = 1$).[60] The right-hand-side of the condition shows a lower bound of the insurance loss associated with cutting DI benefits. This insurance loss equals the utility loss of DI recipients (indicated by $D_{i,t} = 1$) all of whom are affected by lower DI benefits. Notice that condition (15) splits the insurance loss into a (bounded) income loss and a welfare weight, $v'(c^D_{i,t})/\lambda$. An obvious advantage is that income losses can be estimated directly from the data. Clearly, comparing income losses of the *average* individual is of limited interest, because the two DI policy instruments affect different individuals (who have different welfare weights that we cannot estimate). However, we can make progress by looking at income losses of individuals at different positions in the earnings distribution. If income losses associated with tighter DI eligibility rules were falling short of those associated with lower DI benefits *at all income levels* – roughly speaking, at all levels of c^D – then the above condition would be satisfied. In that case, we would conclude that tightening DI eligibility rules does not only generate a larger fiscal multiplier but also a smaller insurance loss than cutting DI benefits.

The second (empirical) step is to implement condition (15). For the left-hand-side of (15) we estimate the bounds on income losses, separately for each income quintile, using the empirical model (12) and individuals' income (= labor earnings net of payroll taxes + DI benefits + UI benefits + SI benefits) as a dependent variable. We fix the income quintile of an individual at age 55 and convert the estimated treatment-coefficients into an average yearly income loss during ages 57-61. For rejected DI applicants resuming work we implement the bound on income loss ($b_{i,t} - z_{i,t}$) by replacing their labor earnings with their potential social welfare benefits.[61] For the right-hand-side of (15) we can directly calculate the reduction in DI benefits ($b^H_{i,t} - b^L_{i,t}$) of DI recipients. Following condition (15), we normalize the absolute income losses by the mechanical fiscal effect of the respective DI policy instrument from Table 4. This "normalized" income loss measures, separately for each quintile, the average (bounded) income reduction associated with a 1 Euro mechanical reduction in fiscal spending of the respective policy instrument.[62]

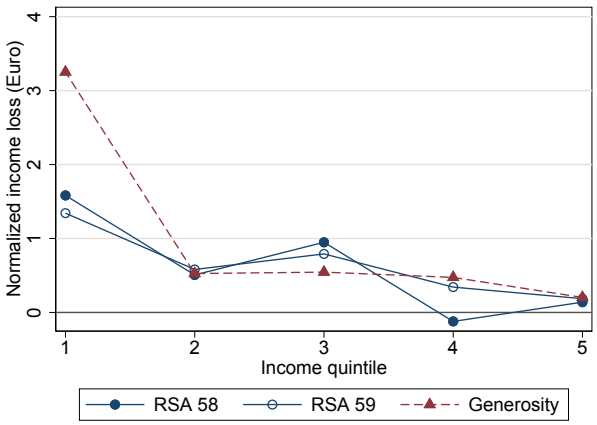
Figure 6 shows the normalized (bounded) income losses by income quintiles. In the lowest income quintile, stricter DI eligibility rules (blue lines in Figure 6) generate significantly lower income losses than lower DI benefits (red line in Figure 6), while the opposite is true for the third quintile. In contrast, the income losses are almost identical for the second, fourth and fifth quintile. In terms of magnitudes, the income-loss difference is high in the lowest income quintile – tighter DI rules generate normalized income losses that are about 1.50 Euros lower than the income losses generated by lower DI benefits. The income-loss difference is smaller in the middle of the income distribution – normalized income losses are 0.20-0.40 Euro higher with tighter DI rules

[60] The left-hand-side calculates the hypothetical insurance loss if all rejected DI applicants get social welfare benefits. In fact, many rejected DI applicants go back to work earning a higher income (but suffering from impaired health). The condition takes advantage of the fact that, for rejected DI applicants resuming work, we have $v(z_{i,t}) < u(w_{i,t}) - \theta$. Using $v(z_{i,t})$ as a lower bound for the utility of this group leads to the above condition.

[61] We construct the potential social welfare benefits as a weighted mix of UI and SI benefits (weighted by the relative share of UI and SI benefits recipients). This leads to an average replacement rate of 59 percent of previous earnings.

[62] Figures D.22 and D.23 show the income losses by income source (including labor earnings).

Figure 6: Bounds of Normalized Income Loss: Eligibility Rules vs. Benefit Generosity



Notes: The figure implements the relative comparison of the insurance losses of stricter eligibility rules and reduced DI benefits from condition (15). The Figure plots the bounds on the normalized income losses in each income quintile for the RSA increases and a reduction in benefit generosity. In the lowest income quintile, stricter DI eligibility rules (blue lines) generate significantly lower income losses than lower DI benefits (red line), while the opposite is true for the third quintile. The income losses are almost identical for the second, fourth and fifth quintile. In terms of magnitudes, the income-loss difference is high in the lowest income quintile – tighter DI rules generate normalized income losses that are about 1.50 Euros lower than the income losses generated by lower DI benefits. The income-loss difference is smaller in the middle of the income distribution – normalized income losses are 0.40 Euro higher with tighter DI rules compared with lower DI benefits. Roughly speaking, the insurance loss of stricter eligibility rules is smaller than the insurance loss of reduced DI benefits (=condition (15) holds) if a 40 cents reduction in income in the third income quintile does not generate a higher utility loss than a 1.50 Euro reduction in income in the lowest income quintile.

compared with lower DI benefits. Roughly speaking, condition (15) holds if a 40 cents reduction in income in the third income quintile does not generate a higher utility loss than a 1.50 Euro reduction in income in the lowest income quintile.[63] Naturally, we would expect the welfare weight of the lowest income quintile to be larger than the welfare weight of the third income quintile. We therefore conclude that the insurance loss associated with stricter DI eligibility is smaller than the insurance loss associated with reducing DI benefits.

In sum, our analysis of the Austrian DI reforms lead to the conclusion that fiscal multiplier of stricter DI eligibility rules are substantially larger than fiscal multipliers of cutting DI benefits. Moreover, we find that the overall insurance loss of tighter DI eligibility rules is, in all likelihood, smaller than the average insurance loss resulting from lower DI benefits. In the Austrian context it is therefore preferable to tighten DI eligibility criteria rather than cutting DI benefits. The next section applies our framework to the U.S. context, which has similar age specific eligibility criteria as the Austrian RSA.

7 Application to the U.S. Disability System

The U.S. DI eligibility criteria are also subject to vocational factors similar to the RSA in Austria. This medical-vocational grid introduces sharp discontinuities in initial award rates by age. Chen and van der Klaauw (2008) use these discontinuities to estimate the labor supply effects of DI benefit receipt. We use their estimates for our framework to discuss the welfare effects of abolishing/shifting these age cutoffs in the U.S. In contrast

[63] Of course, this statement holds only approximately, since there is heterogeneity within the income quintiles.

to Austria the U.S. age cutoffs do not seem to affect application behavior. There is no strategic bunching of applications at these ages as shown by Figure 6 in Chen and van der Klaauw (2008). Chen and van der Klaauw (2008) argue that the rules are not well-known among DI applicants and therefore there is no systematic sorting around the age cutoffs in the U.S. Appendix Figure D.25 contrasts the U.S. application behavior to the Austrian application behavior and reveals that the Austrian rules are well known as there is large spike of applications exactly at the RSA. In contrast the U.S. evidence suggests no behavioral response with respect to the age-dependent eligibility criteria ($\mathbb{E}[B(\theta^*_{US})] = 0$) and a fiscal multiplier of unity in the U.S. context. Therefore, tighter DI eligibility rules at these age cutoffs in the U.S. are welfare reducing – provided that one dollar in the hands of DI recipients has a social value of at least one dollar. Notice that this is almost certainly the case (to the extent that DI applicants are, on average, more deserving than the whole population).

There is no direct reduced-form estimate of the fiscal multiplier of DI benefit generosity in the U.S. In the simple static model from section 2.1 the benefit take-up elasticity ξ is a sufficient statistic (the fiscal multiplier is simply $1+\xi$). Bound and Burkhauser (1999) provide a literature review and report take-up elasticities in the range of 0.3-0.4. Low and Pistaferri (2015)'s structural model implies an application benefit elasticity of 0.62. Multiplied with the average award rate of 0.67 from French and Song (2014) this implies a take-up elasticity of $\xi = 0.41$. Hence, the U.S. evidence would imply a fiscal multiplier of 1.3-1.4. With this back of the envelope calculation we find a very similar fiscal multiplier in the U.S. as in Austria. The DI replacement rates in the U.S. are lower than in Austria and hence the insurance value in the U.S. should be higher than in Austria. This implies that benefits in the U.S. might to be too low. Bound et al. 2004 study the welfare effects of more generous DI benefits in the U.S. in a structural model. Interestingly, they estimate a fiscal multiplier of 1.5 for DI benefit generosity, which is similar in magnitude as our back of the envelope calculation based on elasticities only.[64]

It is worth noting that our findings are closely in line with the conclusions of Low and Pistaferri (2015) who find that U.S. DI reforms – which increase DI benefits and/or relax DI eligibility rules – are welfare improving. Low and Pistaferri (2015) reach this conclusion after setting up and estimating a structural life-cycle model. This is quite different from our methodology which is based on the sufficient-statistics approach and focuses on the local effect after age 55-60. Despite these methodological differences, conducting similar policy experiments lead to very similar conclusions.

[64] Bound et al. 2004 estimate the average implicit price of providing an additional dollar of income to recipients in the presence of moral hazard. This is the same concept as we refer to as the "fiscal multiplier".

8 Conclusion

In this paper, we provide a framework to analyze the welfare effects of stricter DI eligibility criteria versus lower DI benefits by developing sufficient-statistics formulas. We show that the fiscal multiplier is crucial for evaluation of the effectiveness of DI policies and estimate fiscal multipliers of stricter DI eligibility rules and lower DI benefits in the context of Austria.

To estimate the effects of stricter DI eligibility rules, we exploit variation in DI eligibility strictness that is generated by a policy reform. Prior to 2013 DI eligibility standards were significantly relaxed for workers above age 57 relative to those below age 57. A 2013 pension reform increased the relaxed screening age (RSA) threshold from age 57 to age 58, followed by further increases to age 59 in 2015 and age 60 in 2017. These step-wise increases generate quasi-experimental variation in the strictness of DI eligibility at a certain age by date of birth. To examine the impacts of changes in DI benefit levels, we exploit a large pension reform that reduced potential benefit levels for most individuals, although pension levels increased for some individuals with limited work experience.

We find that stricter eligibility creates fiscal multipliers of 2-2.5 and reducing benefit generosity has fiscal multipliers of 1.3-1.4. This implies that by tightening eligibility criteria the policy maker can induce larger behavioral changes and generate greater cost reductions compared to reducing DI benefits. Hence, on the cost side stricter eligibility criteria are more effective. Reducing benefit generosity is only preferable to stricter eligibility criteria if the insurance loss of reducing benefits was more than 1.1 dollars smaller than the insurance loss of stricter eligibility criteria. Our analysis of income losses along the income distribution suggests that the insurance loss of lower DI benefits is larger than the insurance loss of tighter eligibility rules.

References

Abadie, Alberto. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics*, *113* (2): 231–263.

Autor, David, Andreas Kostøl, Magne Mogstad, and Bradley Setzler. 2019. "Disability Benefits, Consumption Insurance, and Household Labor Supply." *American Economic Review*, *109* (7): 2613–54.

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

Autor, David H. and Mark G. Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." *Journal of Economic Perspectives*, *20* (3): 71–96.

Baily, Martin Neil. 1978. "Some Aspects of Optimal Unemployment Insurance." *Journal of Public Economics*, *10* (3): 379–402.

Ball, Steffan and Hamish Low. 2014. "Do Self-insurance and Disability Insurance Prevent Consumption Loss on Disability?" *Economica*, *81* (323): 468–490.

Benitez-Silva, Hugo, Moshe Buchinsky, and John Rust. 2004. "How Large are the Classification Errors in the Social Security Disability Award Process?" Working Paper 10219, National Bureau of Economic Research.

Borghans, Lex, Anne C. Gielen, and Erzo F.P. Luttmer. 2014. "Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform." *American Economic Journal: Economic Policy*, *6* (4): 34–70.

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

Bound, John and Richard V. Burkhauser. 1999. "Economic Analysis of Transfer Programs Targeted on People with Disabilities." In *Handbook of Labor Economics*, volume 3, chapter 51, Elsevier, pp. 3417–3528.

Bound, John, Julie Berry Cullen, Austin Nichols, and Lucie Schmidt. 2004. "The Welfare Implications of Increasing Disability Insurance Benefit Generosity." *Journal of Public Economics*, *88* (12): 2487–2514.

Bound, John, Todd Stinebrickner, and Timothy Waidmann. 2010. "Health, Economic Resources and the Work Decisions of Older Men." *Journal of Econometrics*, *156* (1): 106–129.

Campolieti, Michele. 2004. "Disability Insurance Benefits and Labor Supply: Some Additional Evidence." *Journal of Labor Economics*, *22* (4): 863–889.

Campolieti, Michele and Chris Riddell. 2012. "Disability Policy and the Labor Market: Evidence from a Natural Experiment in Canada, 1998–2006." *Journal of Public Economics*, *96* (3-4): 306–16.

Chen, Susan and Wilbert van der Klaauw. 2008. "The Work Disincentive Effects of the Disability Insurance Program in the 1990s." *Journal of Econometrics*, *142* (2): 757–784.

Chetty, Raj. 2006a. "A General Formula for the Optimal Level of Social Insurance." *Journal of Public Economics*, *90* (10-11): 1879–1901.

Chetty, Raj. 2006b. "A New Method of Estimating Risk Aversion." *American Economic Review*, *96* (5): 1821–1834.

Chetty, Raj. 2008. "Moral Hazard versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy*, *116* (2): 173–234.

Chetty, Raj and Amy Finkelstein. 2013. "Social Insurance: Connecting Theory to Data." In Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, eds., *Handbook of Public Economics: Volume 5*, chapter 3, Elsevier, pp. 111–193.

Clausen, Andrew and Carlo Strub. 2020. "Reverse Calculus and Nested Optimization." *Journal of Economic Theory*, *187*: 105019.

De Chaisemartin, Clément and Xavier D'Haultfoeuille. 2018. "Fuzzy Differences-in-Differences." *The Review of Economic Studies*, *85* (2): 999–1028.

de Jong, Philip, Maarten Lindeboom, and Bas van der Klaauw. 2011. "Screening Disability Insurance Applications." *Journal of the European Economic Association*, *9* (1): 106–129.

Deshpande, Manasi, Tal Gross, and Yalun Su. forthcoming. "Disability and Distress: The Effect of Disability Programs on Financial Outcomes." *American Economic Journal: Applied Economics*.

Deshpande, Manasi and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy*, *11* (4): 213–48.

Diamond, Peter A. and Eytan Sheshinski. 1995. "Economic Aspects of Optimal Disability Benefits." *Journal of Public Economics*, *57* (1): 1–23.

Fevang, Elisabeth, Inés Hardoy, and Knut Røed. 2017. "Temporary Disability Insurance and Economic Disincentives." *The Economic Journal*, *127* (603): 1410–1432.

French, Eric and Jae Song. 2014. "The Effect of Disability Insurance Receipt on Labor Supply." *American Economic Journal: Economic Policy*, *6* (2): 291–337.

Gelber, Alexander, Timothy J. Moore, and Alexander Strand. 2017. "The Effect of Disability Insurance Payments on Beneficiaries' Earnings." *American Economic Journal: Economic Policy*, *9* (3): 229–61.

Godard, M.C.M., P.W.C. Koning, and M. Lindeboom. 2019. "Targeting Disability Insurance Applications with Screening." WorkingPaper 036/V, Tinbergen Institute.

Gruber, Jonathan. 2000. "Disability Insurance Benefits and Labor Supply." *Journal of Political Economy*, *108* (6): 1162–1183.

Gruber, Jonathan and Jeffrey D. Kubik. 1997. "Disability Insurance Rejection Rates and the Labor Supply of Older Workers." *Journal of Public Economics*, *64*: 1–23.

Hendren, Nathaniel and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *Quarterly Journal of Economics*, *135* (3): 1209–1318.

Imbens, Guido W. and Donald B. Rubin. 1997. "Estimating Outcome Distributions for Compliers in Instrumental Variables Models." *The Review of Economic Studies*, *64* (4): 555–574.

Jäger, Simon, Benjamin Schoefer, and Josef Zweimüller. 2019. "Marginal Jobs and Job Surplus: A Test of the Efficiency of Separations." *NBER Working Paper No. 25492*.

Karlström, Anders, Må rten Palme, and Ingemar Svensson. 2008. "The Employment Effect of Stricter Rules for Eligibility for DI: Evidence from a Natural Experiment in Sweden." *Journal of Public Economics*, *92* (10-11): 2071–2082.

Kleven, Henrik J. forthcoming. "Sufficient Statistics Revisited." *Annual Review of Economics*.

Kolsrud, Jonas, Camille Landais, Peter Nilsson, and Johannes Spinnewijn. 2018. "The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden." *American Economic Review*, *108* (4-5): 985–1033.

Kostøl, Andreas and Andreas Myhre. 2020. "Labor Supply Responses to Learning the Tax and Benefit Schedule." *Working Paper*.

Kostol, Andreas R. and Magne Mogstad. 2014. "How Financial Incentives Induce Disability Insurance Recipients to Return to Work." *American Economic Review*, *104* (2): 624–55.

Kroft, Kory. 2008. "Takeup, Social Multipliers and Optimal Social Insurance." *Journal of Public Economics*, *92* (3-4): 722–737.

Kroft, Kory and Matthew J. Notowidigdo. 2016. "Should Unemployment Insurance Vary with the Unemployment Rate? Theory and Evidence." *Review of Economic Studies*, *83* (3): 1092–1124.

Landais, Camille. 2015. "Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design." *American Economic Journal: Economic Policy*, *7* (4): 243–78.

Landais, Camille, Pascal Michaillat, and Emmanuel Saez. 2018. "A Macroeconomic Approach to Optimal Unemployment Insurance: Theory." *American Economic Journal: Economic Policy*, *10* (2): 152–81.

Lee, David S., Pauline Leung, Christopher J. O'Leary, Zhuan Pei, and Simon Quach. forthcoming. "Are Sufficient Statistics Necessary? Nonparametric Measurement of Deadweight Loss in Unemployment Insurance." *Journal of Labor Economics*.

Low, Hamish and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off." *American Economic Review*, *105* (10): 2986–3029.

Low, Hamish and Luigi Pistaferri. 2019. "Disability Insurance: Error Rates and Gender Differences." *NBER Working Paper No. 26513*.

Low, Hamish and Luigi Pistaferri. forthcoming. "Disability Insurance: Theoretical Trade-Offs and Empirical Evidence." *Fiscal Studies*.

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

Mann, David R, David C Stapleton, and Jeanette de Richemond. 2014. "Vocational Factors in the Social Security Disability Determination Process: A Literature Review." Technical report, Mathematica Policy Research.

Meyer, Bruce D. and Wallace K.C. Mok. 2019. "Disability, Earnings, Income and Consumption." *Journal of Public Economics*, *171*: 51–69.

Moore, Timothy J. 2015. "The Employment Effect of Terminating Disability Benefits." *Journal of Public Economics*, *124*: 30–43.

Mullen, Kathleen J. and Stefan Staubli. 2016. "Disability Benefit Generosity and Labor Force Withdrawl." *Journal of Public Economics*, *143*: 49–63.

Nielsen, Helena Skyt, Torben Sorensen, and Christopher Taber. 2010. "Estimating the Effect of Student Aid on College Enrollment: Evidence from a Government Grant Policy Reform." *American Economic Journal: Economic Policy*, *2*: 185–212.

OECD. 2020. "Public Spending on Incapacity (Indicator)." Technical report.

Parsons, Donald O. 1991. "Self-Screening in Targeted Public Transfer Programs." *Journal of Political Economy*, *99* (4): 859–876.

Ruh, Philippe and Stefan Staubli. 2019. "Financial Incentives and Earnings of Disability Insurance Recipients: Evidence from a Notch Design." *American Economic Journal: Economic Policy*, *11* (2): 269–300.

Schmieder, Johannes F., Till von Wachter, and Stefan Bender. 2012. "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years." *Quarterly Journal of Economics*, *127* (2): 701–752.

Shimer, Robert and Iván Werning. 2007. "Reservation Wages and Unemployment Insurance." *Quarterly Journal of Economics*, *122* (3): 1145–1185.

Staubli, Stefan. 2011. "The Impact of Stricter Criteria for Disability Insurance on Labor Force Participation." *Journal of Public Economics*, *95* (9-10): 1223–1235.

Staubli, Stefan and Josef Zweimüller. 2013. "Does Raising the Early Retirement Age Increase Employment of Older Workers?" *Journal of Public Economics*, *108*: 17–32.

Strand, Alexander. 2016. "Effects of Tightening the Medical-Vocational Grid: Comments on Haller, Staubli and Zweimüller." In *Disability Research Consortium Annual Conference, Washington DC, August 2016*.

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

Zweimüller, Josef, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philipe Wuellrich, Oliver Ruf, and Simon Büchi. 2009. "Austrian Social Security Database." *Working Paper University of Zurich*, *401*: 1–80.

Online Appendix (Not for Publication)

A Theory

A.1 Static Model

Condition Under Which $\theta^A < \theta^R$ Holds. The discussion of the model in the main text – and the decision sequence in Figure A.1 – assume that $\theta^A < \theta^R$. This means that the social welfare program is a safety net to which an agent only applies after a DI application is rejected. Here we derive the condition under which this holds true.

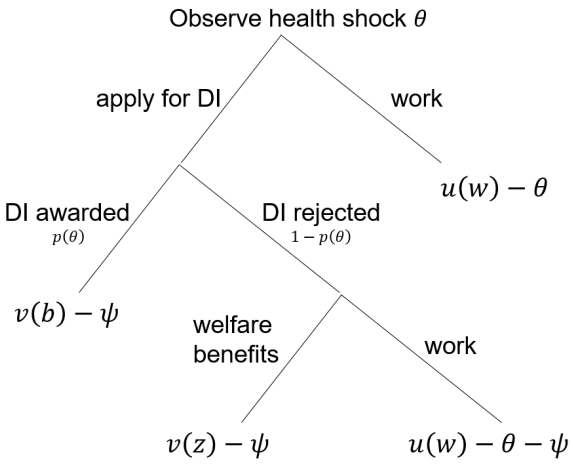
The utility of claiming social welfare falls short of the utility of working if the agent's disability is $\theta < \theta^R \equiv u(w) - v(z) > 0$. Hence, θ^R is the "marginal social welfare claimant". If $\theta \geq \theta^R$ the agent prefers social welfare over working and vice versa. An agent with $\theta \geq \theta^R$ applies to DI if $v(z) < p(\theta)\, v(b) + [1 - p(\theta)]\, v(z) - \psi$ (the utility on social welfare falls short of the expected utility of applying to DI). If this latter condition holds for $\theta = \theta^R$, it also holds for $\theta > \theta^R$, because $p'(\theta) > 0$ and $b > z$. Thus, no agent will claim social welfare benefits unless a previous DI application has been rejected, if

$$\psi < p(\theta^R)\,[(v(b) - v(z)]\,. \tag{A.1}$$

The condition is intuitive: if the DI program is generous (low ψ, high $p(\theta)$ and high b) and/or the social welfare program restrictive (low z), an agent with a severe disability first tries to get on DI, and claims social welfare only if her DI application gets denied. In the basic model, we assume condition (A.1) is satisfied.

Condition (A.1) also implies $\theta^A < \theta^R$, i.e. a marginal applicant returns to work in case her DI application is rejected. To see this, assume to the contrary, that $\theta^A \geq \theta^R$. Then the marginal applicant is indifferent between applying for DI and claiming social welfare, $p(\hat{\theta}^A)]v(b) + [1 - p(\hat{\theta}^A)]v(z) - \psi = v(z)$ or $\psi = p(\hat{\theta}^A)\,[v(b) - v(z)]$, where $\hat{\theta}^A$ is the corresponding threshold disability. However, this latter equality – together with $p(\hat{\theta}^A) \geq p(\theta^R)$ – implies that condition (A.1) is violated. In other words, while this alternative scenario is possible in principle, it is ruled out under the maintained parameter constellation. The intuition is similar as before: if the DI system is generous and/or social welfare restrictive, the rejected marginal applicant goes back to employment.

Figure A.1: DI Application Model: Decision Tree



Note: The figure shows the decision tree of the second period in the static model. If the disability level θ is small, the agent continues working and enjoys utility $u(w) - \theta$. If the disability is severe an agent applies to DI and is accepted with probability $p(\theta)$. Second period utility is $v(b) - \psi$in case of acceptance into DI. If the agent's application is rejected, she needs to decide whether to return to work (with utility $u(w) - \theta - \psi$) or consume other social welfare benefits (with utility $v(z) - \psi$).

Welfare Effect Strictness of DI Eligibility Rules. Starting from (2) the welfare effect of changing θ^* is given by

$$\begin{aligned}\frac{\partial W}{\partial \theta^*} &= -u'(w-\tau)\frac{\partial \tau}{\partial \theta^*} + \int_{\theta^A}^{\infty} \frac{\partial p(\theta;\theta^*)}{\partial \theta^*} v(b) dF(\theta) - \int_{\theta^A}^{\theta^R} \frac{\partial p(\theta;\theta^*)}{\partial \theta^*}(u(w)-\theta) dF(\theta) \\ &\quad - \int_{\theta^R}^{\infty} \frac{\partial p(\theta;\theta^*)}{\partial \theta^*} v(z) dF(\theta) + \frac{\partial \theta^A}{\partial \theta^*} p(\theta^A) f(\theta^A) \underbrace{\left(u(w) - v(b) + \frac{\psi}{p(\theta^A)} - \theta^A\right)}_{=0 \text{ by definition of } \theta^A \text{ in eq. (1)}}\end{aligned} \tag{A.2}$$

where

$$\frac{\partial \tau}{\partial \theta^*} = \int_{\theta^A}^{\infty} \frac{\partial p(\theta;\theta^*)}{\partial \theta^*} b dF(\theta) - b\frac{\partial \theta^A}{\partial \theta^*} p(\theta^A) f(\theta^A) - \int_{\theta^R}^{\infty} \frac{\partial p(\theta;\theta^*)}{\partial \theta^*} z dF(\theta) \tag{A.3}$$

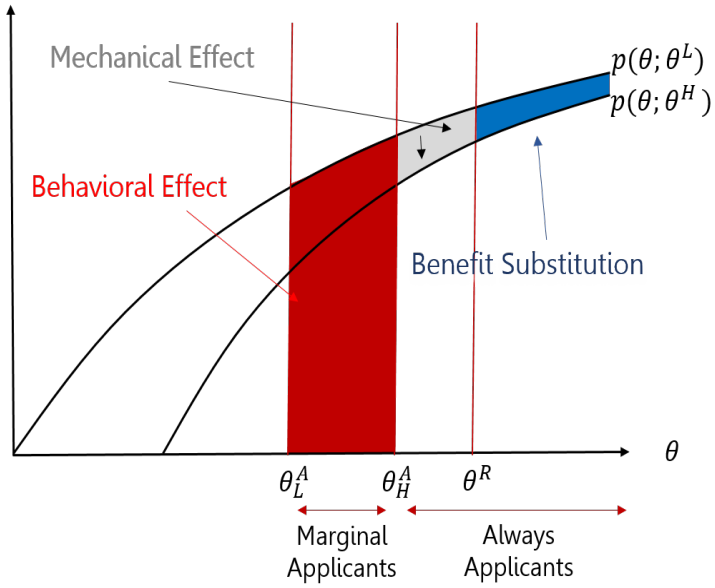
following from the planner's budget constraint. Defining $B(\theta^*) \equiv b\left(\partial \theta^A/\partial \theta^*\right) p(\theta^A) f(\theta^A)$, $M_W \equiv -\int_{\theta^A}^{\theta^R} \left(\partial p(\theta;\theta^*)/\partial \theta^*\right) dF(\theta)$, $M_Z \equiv -\int_{\theta^R}^{\infty} \left(\partial p(\theta;\theta^*)/\partial \theta^*\right) dF(\theta)$ and $M(\theta^*) \equiv M_W b + M_Z(b-z)$, we can rewrite $-\partial \tau/\partial \theta^* = B(\theta^*) + M(\theta^*)$. Plugging these terms into the above equation for $\partial W/\partial \theta^*$ yields condition (4) in the main text.

Welfare Effect Strictness of DI Eligibility Rules: Non-marginal Change. Condition (4) in the main text holds for a marginal change in θ^*. Here we consider the welfare effect of a discrete change. Suppose strictness

of eligibility rules increase from θ^L to $\theta^H > \theta^L$. This implies that the award probability falls $p(\theta;\theta^H) < p(\theta;\theta^L)$ and fewer individuals apply $\theta_H^A > \theta_L^A$ (θ_L^A denotes the marginal applicant with the lenient criteria θ^L and θ_H^A denotes the marginal applicant with the high standard θ^H). Note θ^R is still independent of the eligibility rules.

Figure A.2 illustrates the effects of a non-marginal change in strictness of eligibility rules. If rules become stricter, the award probability curve shifts down from $p(\theta;\theta^L)$ to $p(\theta;\theta^H)$. As a response fewer individuals apply. Individuals with $\theta < \theta_H^A$ no longer apply under the stricter rules. Individuals with $\theta \in \left[\theta_L^A, \theta_H^A\right]$ are therefore "marginal applicants" as they only apply under the lenient rules. The share of these marginal applicants is $\pi^{MA} = F(\theta_H^A) - F(\theta_L^A)$. The behavioral effect is the area under the old award curve of these marginal applicants. Individuals with a disability level above θ_H^A continue to apply. These are always applicants and their share is $\pi^{AA} = 1 - F(\theta_H^A)$. The difference between the old and new award curve for these always applicants corresponds to the mechanical effect. Some of these mechanically screened out individuals return to work (gray area) and some substitute to welfare benefits (blue area). Hence, we have the same effects as in the marginal case, but these effects are slightly differently defined.

Figure A.2: Illustration Non-Marginal Change



Note: This figure illustrates the effects of a non-marginal change in strictness of DI eligibility rules.

Let W_H and W_L denote welfare in the two regimes. Welfare in the two regimes $S \in \{H, L\}$ is

$$\begin{aligned} W_S = \quad & u(w-\tau) + \int_0^{\theta_S^A} (u(w)-\theta)dF(\theta) + \int_{\theta_S^A}^{\theta^R} (1-p(\theta;\theta^S))(u(w)-\theta)dF(\theta) + \\ & + \int_{\theta_S^A}^{\infty} p(\theta;\theta^S)v(b)dF(\theta) + \int_{\theta^R}^{\infty} (1-p(\theta;\theta^S))v(z)dF(\theta) - \int_{\theta_S^A}^{\infty} \psi dF(\theta). \end{aligned} \tag{A.4}$$

The welfare effect of this discrete change $\Delta W \equiv W_H - W_L$ is given by

$$\begin{aligned}\Delta W &= u(w-\tau^H) - u(w-\tau^L) \\ &\quad - \int_{\theta_H^A}^{\theta^R} \left[p(\theta;\theta^L) - p(\theta;\theta^H)\right] \left[v(b) - (u(w)-\theta)\right] dF(\theta) - \int_{\theta^R}^{\infty} \left[p(\theta;\theta^L) - p(\theta;\theta^H)\right] \left[v(b) - v(z)\right] dF(\theta) \\ &\quad - \int_{\theta_L^A}^{\theta_H^A} p(\theta;\theta^L) \left[v(b) - (u(w)-\theta)\right] - \psi dF(\theta).\end{aligned} \tag{A.5}$$

The first line of (A.5) captures the gain for the taxpayer (the fiscal cost reduction), the second line is the loss in insurance value for the always applicants (mechanically screened out). The third line is the insurance loss that marginal applicants experience and is the key difference to the marginal case. The Envelope theorem does not apply for a non-marginal change in θ^* and behavioral responses have a first order welfare effect. Note that for the limiting case of a marginal change $\theta^H \to \theta^L$ we have $\theta_H^A \to \theta_L^A$ and $\int_{\theta_L^A}^{\theta_H^A} p(\theta;\theta^L) \left[v(b) - (u(w)-\theta)\right] - \psi dF(\theta) \to p(\theta_L^A;\theta^L) \left[v(b) - (u(w)-\theta_L^A)\right] - \psi = 0$ by the definition of the marginal applicant θ_L^A.

Using the government budget constraint we can rewrite the fiscal effect again as the behavioral fiscal effect B_Δ plus the mechanical fiscal effect M_Δ:

$$\begin{aligned}\tau^L - \tau^H &= \underbrace{b \cdot \int_{\theta_L^A}^{\theta_H^A} p(\theta;\theta^L) dF(\theta)}_{\equiv B_\Delta} \\ &= \underbrace{+b \cdot \int_{\theta_H^A}^{\theta^R} \left[p(\theta;\theta^L) - p(\theta;\theta^H)\right] dF(\theta) - (b-z) \cdot \int_{\theta^R}^{\infty} \left[p(\theta;\theta^L) - p(\theta;\theta^H)\right] dF(\theta)}_{\equiv M_\Delta}.\end{aligned} \tag{A.6}$$

Moreover, we can write the welfare effect associated with the fiscal effect as $u(w-\tau^H) - u(w-\tau^L) = \lambda\left(\tau^L - \tau^H\right) = \lambda\left(B_\Delta + M_\Delta\right)$, where $\lambda = u'(w-\tau_\Delta)$ with τ_Δ such that this equality holds.[65] We can then rearrange (A.5) to $\Delta W \gtreqless 0 \Leftrightarrow$

$$1 + \frac{B_\Delta}{M_\Delta} \gtreqless \frac{L_{W\Delta} + L_{Z\Delta}}{u'(w-\tau_\Delta) M_\Delta} + \frac{L_{MA}}{u'(w-\tau_\Delta) M_\Delta} \tag{A.7}$$

where

$$L_{W\Delta} \equiv \int_{\theta_H^A}^{\theta^R} \left[p(\theta;\theta^L) - p(\theta;\theta^H)\right] \left[v(b) - (u(w)-\theta)\right] dF(\theta), \tag{A.8}$$

[65] Later on in the general model λ will denote the multiplier of the government budget constraint and therefore measures the social value of public funds.

$$L_{Z\Delta} \equiv \int_{\theta^R}^{\infty} \left[p(\theta;\theta^L) - p(\theta;\theta^H)\right] \left[v(b) - v(z)\right] dF(\theta) \tag{A.9}$$

and

$$L_{MA} \equiv \int_{\theta_L^A}^{\theta_H^A} p(\theta;\theta^L) \left[v(b) - (u(w) - \theta)\right] - \psi dF(\theta). \tag{A.10}$$

In summary, the fiscal multiplier is still key for evaluating welfare effects. Moreover, we show later on that our empirical method to estimate the multiplier is robust to non-marginal changes; and that the insurance value is the discrete analog of the marginal change with the additional term L_{MA}.

Welfare Effect DI Benefit Level. Starting from equation (2) we get

$$\begin{aligned} \frac{\partial W}{\partial b} &= -u'(w-\tau)\frac{\partial \tau}{\partial b} + \int_{\theta^A}^{\infty} p(\theta;\theta^*) v'(b) dF(\theta) \\ &\quad + \frac{\partial \theta^A}{\partial \theta^*} p(\theta^A) f(\theta^A) \underbrace{\left(u(w) - v(b) + \frac{\psi}{p(\theta^A)} - \theta^A \right)}_{=0 \text{ by definition of } \theta^A} \end{aligned} \tag{A.11}$$

where $\partial\tau/\partial b = -b\left(\partial\theta^A/\partial b\right)p(\theta^A)f(\theta^A)\big) + \int_{\theta^A}^{\infty} p(\theta;\theta^*)dF(\theta) = B(b) + M(b)$ is the change in taxes necessary to fund a DI system with a marginally higher DI benefit b. (6) then immediately follows from (A.11).

Optimal Policy Mix. So far, we have derived conditions for social optimality for each single DI policy parameter, holding the other policy parameter fixed. However, a natural question is how a DI reform should optimally combine these two policy parameters. More precisely: how strongly – and in which direction – should DI eligibility rules θ^* be changed per unit change of DI benefits b?

The slope of the gradient answers this question of the optimal policy mix. Figure A.3 illustrates the idea of the optimal policy mix. It depicts the current policy (θ_0^*, b_0). The dotted curve indicates the combinations of (θ^*, b) that generate the same level of social welfare. Consider the effect of a DI reform starting from the pre-reform DI policy is (θ_0^*, b_0). The vertical (horizontal) arrow shows how θ^* (resp. b) needs to be changed to increase welfare. In Figure A.3 the vertical arrow points up and the horizontal arrow points to the left, indicating that a welfare-enhancing DI reform implement lower benefits and stricter eligibility rules.[66] The length of the arrows correspond to the efficiency gains associated with the respective policy instrument. In Figure A.3 the horizontal arrow is short, while the vertical arrow is long, suggesting that the DI reform should strongly increase θ^* per unit reduction of b. The slope of the gradient – the arrow pointing to the northwest –

[66] Alternatively, if the horizontal arrow points to the right and the vertical arrow points up, the existing DI system is too restrictive in both dimensions and a welfare reform increasing DI benefits and implementing more lenient DI eligibility rules is welfare improving. Of course, all other permutations are possible.

yields the optimal policy mix.

Using the first order conditions (5) and (6), the gradient is given by

$$\nabla W = \begin{pmatrix} -\partial W/\partial b \\ \\ \partial W/\partial \theta^* \end{pmatrix} = \begin{pmatrix} \sigma \cdot M(b) \\ \\ \gamma \cdot M(\theta^*) \end{pmatrix} u'(w-\tau) \tag{A.12}$$

where σ and γ measure the gap between fiscal gains and insurance loss for changes in b and θ^*, respectively. Formally, we have $\sigma \equiv [1 + B(b)/M(b)] - [v'(c_d)] / [u'(w-\tau)M(b)]$ and $\gamma \equiv [1 + B(\theta^*)/M(\theta^*)] - [L_W + L_Z] / [u'(w-\tau)M(\theta^*)]$. Therefore, the optimal DI policy mix is given by the ratio

$$\left.\frac{\partial \theta^*}{\partial b}\right|_{opt} = \frac{\gamma}{\sigma} \cdot \frac{M(\theta^*)}{M(b)}. \tag{A.13}$$

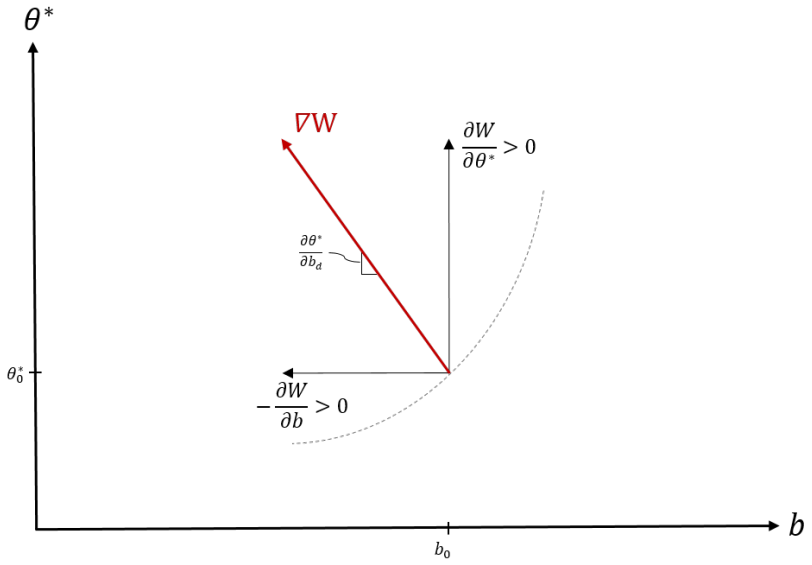
The sign of σ determines the direction in which benefits should be adjusted (if $\sigma \gtrless 0 \Leftrightarrow -\partial W/\partial b \gtrless 0 \Leftrightarrow \partial b \lessgtr 0$). Similarly, the sign of γ determines the direction of adjustment in θ^* (if $\gamma \gtrless 0 \Leftrightarrow \partial W/\partial \theta^* \gtrless 0 \Leftrightarrow \partial \theta^* \gtrless 0$). σ and γ determine the direction of welfare-enhancing adjustments in b and θ^*, while the ratio γ/σ determines the optimal DI policy mix, $(\partial \theta^*/\partial b)_{opt}$. b and θ^* have different units. Hence, for a meaningful interpretation of the optimal direction we normalize by the respective mechanical fiscal effects of a one unit change in b and θ^* respectively. This means that the optimal direction is expressed in terms of a mechanical 1 dollar change in fiscal costs: For a mechanical one-dollar reduction in fiscal costs due to lower DI benefits, DI eligibility rules should be adjusted such that fiscal costs are mechanically reduced by γ/σ dollars.

In the empirical implementation we focus on the relative comparison of the two instruments (reducing benefits or tightening eligibility rules). The implementation of the optimal policy mix requires estimating the insurance value. We lack the data on consumption, wealth and health to do this without strong assumptions. The relative comparison is more robust as we can use income losses and the estimated fiscal multipliers to rank the two policies as we show later on.

A.2 General Model

DI Applications and Labor Supply. We proceed in a similar way as the discussion of the static model in the main text. The setup mirrors the static model but extends it in two important dimensions. First, we extend the model to T periods, so agents need to make inter-temporal decisions. Second, we allow θ (and as well as other state variables) evolve stochastically over the agent's relevant time horizon. Let $X_{i,t} = \{\theta_{i,t}, A_{i,t}, \chi_{i,t}\}$

Figure A.3: Optimal Policy Mix – Gradient of Welfare Function



Notes: The figure illustrates the idea of the optimal policy mix. It shows the gradient in case DI eligibility should be stricter and benefits should be less generous. The dashed line is the indifference curve of the welfare function of the current benefit level and strictness of DI eligibility. The gradient of the welfare function is orthogonal to the indifference curve and points in the direction of greatest increase of the function.

denote the vector of state variables where $\theta_{i,t}$ denotes agent i's disability level in period t, $A_{i,t}$ denotes the asset level and $\chi_{i,t}$ is a vector of other state variables (which allows for heterogeneity across agents such as differences in wages etc.). The state vector $X_{i,t}$ summarizes all the information relevant for agent i's choices in period t. The laws of motion of assets in the disability, employment and welfare benefit state are

$$A_{i,t+1} = (1+r_t)A_{i,t} + b_{i,t}(X_{i,t}) - c^D_{i,t}(X_{i,t}) \tag{A.14}$$

$$A_{i,t+1} = (1+r_t)A_{i,t} + w_{i,t}(X_{i,t}) - \tau_{i,t}(X_{i,t}) - c^E_{i,t}(X_{i,t}) \tag{A.15}$$

$$A_{i,t+1} = (1+r_t)A_{i,t} + z_{i,t}(X_{i,t}) - c^Z_{i,t}(X_{i,t}). \tag{A.16}$$

$b_{i,t}(X_{i,t})$ denotes DI benefits of individual i in period t and can depend on the agent's state $X_{i,t}$. Analogously, $w_{i,t}(X_{i,t})$ denotes labor income, $\tau_{i,t}(X_{i,t})$ are taxes and $z_{i,t}(X_{i,t})$ denotes social welfare benefits. Agents make state contingent plans on how much to consume in each labor market state $\{c^D_{i,t}(X_{i,t}), c^W_{i,t}(X_{i,t}), c^Z_{i,t}(X_{i,t})\}$, whether they apply to DI benefits $\alpha_{i,t}(X_{i,t}) \in \{0,1\}$ and, if not, on DI whether they work or claim social welfare benefits $\omega_{i,t}(X_{i,t}) \in \{0,1\}$.

The within-period sequence of events and choices is identical to the static model in Section 2.1. At the beginning of the period, the shocks $\theta_{i,t}$ and $\chi_{i,t}$ are revealed. Having learned $X_{i,t}$, she decides whether to file a DI application and, if accepted, becomes a DI beneficiary for the rest of her life.[67] If her application is rejected, she either resumes work or claims social welfare, whatever yields higher utility.

[67] The assumption that DI is an absorbing state, is supported by the empirically observed negligibly low outflow rates, particularly among older workers.

Denote by $D_{i,t}, W_{i,t}$ and $Z_{i,t}$, respectively, the probability that, in period t, agent i is a DI benefit recipient, an employed worker, or a social welfare recipient. These probabilities are given by

$$D_{i,t} = 1 - \left[\prod_{k=0}^{t} \left(1 - \alpha_{i,k}(X_{i,k}) \cdot p\left(\theta_{i,k}, \theta_k^*\right)\right)\right] \tag{A.17}$$

$$W_{i,t} = \omega_{i,t}(X_{i,t}) \left[\prod_{k=0}^{t} \left(1 - \alpha_{i,k}(X_{i,k}) \cdot p\left(\theta_{i,k}, \theta_k^*\right)\right)\right] \tag{A.18}$$

$$Z_{i,t} = \left(1 - \omega_{i,t}(X_{i,t})\right) \left[\prod_{k=0}^{t} \left(1 - \alpha_{i,k}(X_{i,k}) \cdot p\left(\theta_{i,k}, \theta_k^*\right)\right)\right]. \tag{A.19}$$

The probability agent i transitions to DI in period k is $\alpha_{i,k}(X_{i,k}) \cdot p\left(\theta_{i,k}, \theta_k^*\right)$. Hence, the probability that an agent is not yet on DI in period t is $\left[\prod_{k=0}^{t} \left(1 - \alpha_{i,k}(X_{i,k}) \cdot p\left(\theta_{i,k}, \theta_k^*\right)\right)\right]$. From this pool, $\omega_{i,t}(X_{i,t})$ of the non DI individuals work and $1 - \omega_{i,t}(X_{i,t})$ are on social welfare benefits.[68] We assume that the first application bears a fix cost ψ and follow-up applications are costless. $\Lambda_{i,t} = \alpha_{i,t}(X_{i,t}) \prod_{k=0}^{t-1} \left(1 - \alpha_{i,k}(X_{i,k})\right) \in 0, 1$ indicates whether agent i applies for the first time in period t. The other state variables, disutility of work $\theta_{i,t}$ and $\chi_{i,t}$, follow stochastic processes that can, in principle, depend on agents' choices. The expectation operator $\mathbb{E}[\cdot]$ below captures the evolution of the state variables and encompasses aggregation across individuals and time.[69] The agent's problem is then given by

$$\begin{aligned} V_i(P) = \max E & \left[\sum_{t=0}^{T-1} \beta^t \left(v(c_{i,t}^D) \cdot D_{i,t} + v(c_{i,t}^Z) \cdot Z_{i,t} + \left(u(c_{i,t}^W) - \theta_{i,t}\right) \cdot W_{i,t} - \Lambda_{i,t} \cdot \psi\right)\right] \\ & + E\left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \left((1 + r_t) A_{i,t} + b_{i,t} - c_{i,t}^D - A_{i,t+1}\right) D_{i,t}\right] \\ & + E\left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^W \left((1 + r_t) A_{i,t} + w_{i,t} - \tau_{i,t} - c_{i,t}^W - A_{i,t+1}\right) W_{i,t}\right] \\ & + E\left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^Z \left((1 + r_t) A_{i,t} + z_{i,t} - c_{i,t}^Z - A_{i,t+1}\right) Z_{i,t}\right]. \end{aligned} \tag{A.20}$$

Welfare Effect of DI Reforms. We proceed in the same way as we did in the static framework. We first set up the utilitarian planner's problem and study the impact of more stringent DI eligibility rules. The social planner maximizes social welfare by choosing the strictness of DI eligibility θ_s^* and DI benefit function b_s in each period s. We denote this disability policy by $P = \{\theta_s^*, b_s\}_{s=0}^{T-1}$. The planner therefore solves

[68] We assume that social welfare, unlike DI, is not an absorbing state. This implies that an agent who has not yet entered DI is "at risk" of being employed or being on social welfare in period k.

[69] The operator $\mathbb{E}[Y]$ aggregates the variable Y over states of nature and across individuals, i.e. $\mathbb{E}[Y] = \int \int Y(X_{i,t}) dF(X_{i,t}) di$ where $F(\cdot)$ is the distribution of state variables $X(i,t)$. Notice that this is a flexible formulation: the only restriction we impose on this distribution of state variables is that it does not directly depend on the planner's policy instruments $P = \{\theta_t^*, b_t\}_{t=0}^{T-1}$. The evolution of $X(i,t)$, however, can depend on agent i's choices which themselves depend on the policy instruments P. We use the operator $E[Y_{i,t}] = \int Y(X_{i,t}) dF(X_{i,t})$ to denote the expectation w.r.t. state variables for a given individual.

$$\max_P W(P) = \int_i V_i(P)di + \lambda \left(G(P) - \bar{G}\right) \tag{A.21}$$

where

$$G(P) = \int_i E\left[\sum_{t=0}^{T-1} (1+r_t)^{-t} \left(W_{i,t} \cdot \tau_{i,t} - D_{i,t} \cdot b_{i,t} - Z_{i,t} \cdot z_{i,t}\right)\right] di \tag{A.22}$$

is the planners net revenue, $\bar{G}$ is an exogenous revenue constraint and λ denotes the Lagrange multiplier on the planner's budget constraint.

More Stringent DI Eligibility Rules. The following proposition characterizes the optimal DI policy $P = \{\theta_s^*, b_s\}_{s=0}^{T-1}$.

Proposition 1. *Assume the planner's budget constraint is differentiable in θ_s^*. Optimal strictness of DI eligibility rules in period s, θ_s^*, then fulfills*

$$1 + \frac{\mathbb{E}\left[B(\theta_s^*)\right]}{\mathbb{E}\left[M(\theta_s^*)\right]} = \frac{\mathbb{E}\left[L_W\right] + \mathbb{E}\left[L_Z\right]}{\lambda \mathbb{E}\left[M(\theta_s^*)\right]} \tag{A.23}$$

where

$$\mathbb{E}\left[M(\theta_s^*)\right] \equiv \mathbb{E}\left[\sum_{t=s}^{T-1} (1+r_t)^{-t} \left(M_{W_{i,t}}\left(b_{i,t} + \tau_{i,t}\right) + M_{Z_{i,t}}\left(b_{i,t} - z_{i,t}\right)\right)\right] \tag{A.24}$$

is the mechanical fiscal effect and $\mathbb{E}\left[B(\theta_s^*)\right] \equiv \partial G(P)/\partial \theta_s^* - \mathbb{E}\left[M(\theta_s^*)\right]$ is the behavioral fiscal effect. $M_{W_{i,t}}$ is the mechanical employment effect

$$M_{W_{i,t}} \equiv -\omega_{i,t} \left(\alpha_{i,s} \cdot \frac{\partial p(\theta_{i,s}, \theta_s^*)}{\partial \theta_s^*} \prod_{k=0, k\neq s}^{t} (1 - \alpha_{i,k} p_{i,k})\right) \tag{A.25}$$

and $M_{Z_{i,t}}$ is the mechanical benefit substitution effect

$$M_{Z_{i,t}} \equiv -\left(1 - \omega_{i,t}\right) \left(\alpha_{i,s} \cdot \frac{\partial p(\theta_{i,s}, \theta_s^*)}{\partial \theta_s^*} \prod_{k=0, k\neq s}^{t} (1 - \alpha_{i,k} p_{i,k})\right). \tag{A.26}$$

$\mathbb{E}\left[L_W\right]$ and $\mathbb{E}\left[L_Z\right]$ denote the insurance losses for individuals who return to work and substitute to welfare benefits respectively and are defined by

$$\mathbb{E}\left[L_W\right] \equiv \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{W_{i,t}}\left(v(c_{i,t}^D) - \left(u(c_{i,t}^W) - \theta_{i,t}\right)\right)\right)\right] \tag{A.27}$$

$$\mathbb{E}\left[L_Z\right] \equiv \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{Z_{i,t}}\left(v(c_{i,t}^D) - v(c_{i,t}^Z)\right)\right)\right]. \tag{A.28}$$

Proof. See below. □

Welfare Effect of Changing DI Benefits.

Proposition 2. *Assume the planner's budget constraint is differentiable in b_s for all periods s. The optimal DI benefit level in period s fulfills*

$$1 + \frac{\mathbb{E}\left[B(b_s)\right]}{\mathbb{E}\left[M(b_s)\right]} = \frac{\mathbb{E}\left[v'(c_s^D)\right]}{\lambda \cdot \mathbb{E}\left[M(b_s)\right]} \tag{A.29}$$

where $\mathbb{E}\left[M(b_s)\right] \equiv \mathbb{E}\left[(1+r_s)^{-s}\left(D_{i,s}\right)\right]$ is the mechanical fiscal effect of adjusting DI benefits, $\mathbb{E}\left[B(b_s)\right] \equiv -\partial G(P)/\partial b_s - \mathbb{E}\left[M(b_s)\right]$ denotes the behavioral fiscal effect and $\mathbb{E}\left[v'(c_s^D)\right] \equiv \mathbb{E}\left[\beta^s D_{i,s} v'(c^D)\right]$.

Proof. See below. □

Optimal Policy Mix: Gradient. Just like the static framework, we can also study the optimal policy mix. It turns out that the result is analogous to the static model. It directly derives from conditions (10) and (11) that characterize the social optimality of policy parameters θ_s^* and b_s.

From above we have

$$\frac{\partial W}{\partial \theta_s^*} = \gamma * \mathbb{E}\left[M(\theta_s^*)\right] * \lambda \tag{A.30}$$

where

$$\gamma \equiv 1 + \frac{\mathbb{E}\left[B(\theta_s^*)\right]}{\mathbb{E}\left[M(\theta_s^*)\right]} - \frac{\mathbb{E}\left[L_W\right] + \mathbb{E}\left[L_Z\right]}{\lambda \mathbb{E}\left[M(\theta_s^*)\right]} \tag{A.31}$$

and

$$-\frac{\partial W}{\partial b_s} = \sigma * \mathbb{E}\left[M(b_s)\right] * \lambda \tag{A.32}$$

where

$$\sigma \equiv 1 + \frac{\mathbb{E}\left[B(b_s)\right]}{\mathbb{E}\left[M(b_s)\right]} - \frac{\mathbb{E}\left[v'(c_s^D)\right]}{\lambda \cdot \mathbb{E}\left[M(b_s)\right]}. \tag{A.33}$$

The gradient is

$$\nabla W = \begin{pmatrix} -\partial W/\partial b_s \\ \\ \partial W/\partial \theta_s^* \end{pmatrix} = \begin{pmatrix} \sigma * \mathbb{E}\left[M(b_s)\right] \\ \\ \gamma * \mathbb{E}\left[M(\theta_s^*)\right] \end{pmatrix} \lambda. \tag{A.34}$$

The optimal change in the strictness of DI eligibility rules per unit change of DI benefits is then given by

$$\left.\frac{\partial \theta_s^*}{\partial b_s}\right|_{opt} = \frac{\gamma}{\sigma} \cdot \frac{\mathbb{E}\left[M(\theta_s^*)\right]}{\mathbb{E}\left[M(b_s)\right]}, \tag{A.35}$$

where γ and σ denote the difference between marginal social cost and marginal social benefit of changing θ_s^* and b_s, respectively (where the signs of γ and σ determine the direction of adjustment).

Comparing Static and Dynamic Models. Observe the similarity between the dynamic, general solution and the one of the static model. In the static model, we had

$$\frac{\partial W}{\partial \theta^*} \gtreqless 0 \iff 1 + \frac{B(\theta^*)}{M(\theta^*)} \gtreqless \frac{[v(b) - (u(w) - \tilde{\theta})]M_W + [v(b) - v(z)]\, M_Z}{u'(w-\tau)M(\theta^*)}$$

and in the dynamic model we have

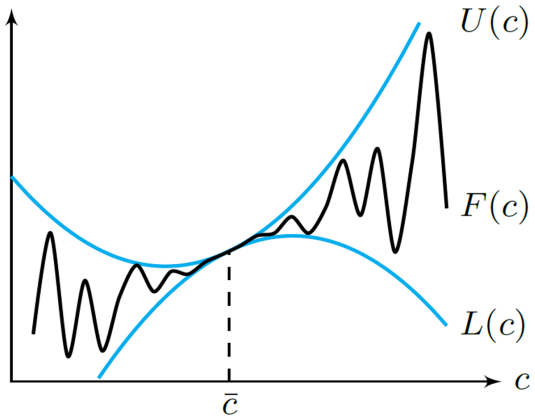
$$\frac{\partial W}{\partial \theta_s^*} \gtreqless 0 \iff 1 + \frac{\mathbb{E}\left[B(\theta_s^*)\right]}{\mathbb{E}\left[M(\theta_s^*)\right]} \gtreqless \frac{\mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(\left(v_i(c_{i,t}^D) - \left(u_i(c_{i,t}^W) - \theta_{i,t}\right)\right) M_{W_{i,t}} + \left(v_i(c_{i,t}^D) - v_i(c_{i,t}^Z)\right) M_{Z_{i,t}}\right)\right]}{\lambda \mathbb{E}\left[M(\theta_s^*)\right]}.$$

Proofs.

Proof. Proposition 1

The proof is a direct application of the Envelope Theorem. To derive conditions (A.23) and (A.29) we apply the differentiable sandwich lemma from Clausen and Strub (2020). Clausen and Strub (2020) establish that if a function $F(c)$ is sandwiched at some point $\bar{c}$ between two differentiable functions (upper and lower support functions $U(c)$ and $L(c)$), then this function F is differentiable at this point $\bar{c}$. Moreover, the derivative of the sandwiched function F equals the derivative of the upper and lower support functions at this point, i.e. $F'(\bar{c}) = U'(\bar{c}) = L'(\bar{c})$. Figure A.4 illustrates this idea nicely. The proof here therefore identifies differentiable upper and lower support functions of the welfare function $W(P)$.

Figure A.4: Illustration Differentiable Sandwich Lemma



Notes: This Figure illustrates the differentiable sandwich lemma of Clausen and Strub (2020), which is the key argument in the proof of Proposition 1 and 2.
Source:Clausen and Strub (2020).

Let $\overline{P}$ denote the optimal policy, i.e. the $P = \{\theta_s^*, b_s\}_{s=0}^{T-1}$ that maximizes welfare. By definition $W(\overline{P}) \geq W(P)\, \forall P$ and therefore the constant function $U(P) = W(\overline{P})$ is a natural upper support function. We have $U'(P) = 0$.

For the lower support function we use the idea of the "lazy" decision maker who does not take into account agents' behavioral responses to the policy change. Let $\bar{V}_i(P)$ denote the agent's indirect utility if she sticks to her behavior that is optimal for policy $\bar{P}$ even when the policy is changed to another P. That is, for all potential policies P the agent does not adjust her behavior. Therefore, $\bar{V}_i(P) \leq V_i(P)$. As a lower support function we then take $L(P) = \int \bar{V}_i(P)di + \lambda\left(G(P) - \bar{G}\right)$. The derivative of this lower support function with respect to θ_s^* is

$$\begin{aligned}\frac{\partial L(P)}{\partial \theta_s^*} &= \mathbb{E}\left[\sum_{t=s}^{T-1}\beta^t\left(\alpha_{i,s}\cdot\frac{\partial p(\theta_{i,s},\theta_s^*)}{\partial\theta_s^*}\prod_{k=0,k\neq s}^{t}(1-\alpha_{i,k}p_{i,k})\right)\left\{v(c_{i,t}^D) - v(c_{i,t}^Z)\cdot(1-\omega_{i,t})\ldots\right.\right.\\ &\left.\left.\ldots - \left(u(c_{i,t}^W) - \theta_{i,t}\right)\cdot\omega_{i,t}\right\}\right] + \lambda\frac{\partial G(P)}{\partial\theta_s^*}. \qquad \text{(A.36)}\end{aligned}$$

We can decompose the total fiscal effect $\partial G(P)/\partial\theta_s^*$ into the mechanical and behavioral fiscal effect. The mechanical effect is

$$\mathbb{E}\left[M(\theta_s^*)\right] = -\mathbb{E}\left[\sum_{t=s}^{T-1}(1+r_t)^{-t}\left(\alpha_{i,s}\cdot\frac{\partial p(\theta_{i,s},\theta_s^*)}{\partial\theta_s^*}\prod_{k=0,k\neq s}^{t}(1-\alpha_{i,k}p_{i,k})\right)\left(b_{i,t}+\tau_{i,t}\cdot\omega_{i,t}-z_{i,t}\cdot(1-\omega_{i,t})\right)\right]$$

and we define the behavioral fiscal effect as the residual $\mathbb{E}\left[B(\theta_s^*)\right] = \partial G(P)/\partial\theta_s^* - \mathbb{E}\left[M(\theta_s^*)\right]$.[70] The differentiable sandwich lemma then implies that $\partial W(P)/\partial\theta_s^* = \partial L(P)/\partial\theta_s^* = \partial U(P)/\partial\theta_s^* = 0$ at the optimal policy. It is then straightforward to rearrange (A.36) to (A.23). □

Proof. Proposition 2.

We apply the same logic to the optimal DI benefit policy as in the previous proof. We have

$$\frac{\partial L(P)}{\partial b_s} = \mathbb{E}\left[\beta^s\mu_{i,s}^D\right] + \lambda\frac{\partial G(P)}{\partial b_s}. \qquad \text{(A.37)}$$

The agent's first order condition implies $\mathbb{E}\left[\beta^s\mu_{i,s}^D D_{i,s}\right] = \mathbb{E}\left[\beta^s v'(c_{i,s}^D)D_{i,s}\right]$. Define the mechanical fiscal effect as $\mathbb{E}\left[M(b_s)\right] = \mathbb{E}\left[(1+r_s)^{-s}\left(D_{i,s}\right)\right]$ and the behavioral fiscal effect is again defined as the difference between total fiscal effect and mechanical fiscal effect $\mathbb{E}\left[B(b_s)\right] = -\partial G(P)/\partial b_s - \mathbb{E}\left[M(b_s)\right]$. We again have $\partial W(P)/\partial b_s = \partial L(P)/\partial b_s = \partial U(P)/\partial b_s = 0$. It is then straightforward to rearrange (A.37) to obtain (A.29). □

Welfare Effect Stricter Eligibility Rules: Non-marginal Change. Analogous to the discussion in the static model, consider a discrete change in eligibility rules in period s from θ_s^L to $\theta_s^H > \theta_s^L$. To resemble our empirical setup, assume that strictness of DI eligibility is high, $\theta^* = \theta^H$, until age s and lenient afterwards. This is denoted by policy $P^L = (\theta_0^H, \ldots, \theta_{s-1}^H, \theta_s^L, \theta_{s+1}^L, \theta_{T-1}^*; b_0, \ldots, b_{T-1})$. The reform we study empirically increased the age of relaxed screening from s to $s+1$. This corresponds to policy

[70] The behavioral fiscal effect contains the fiscal effect of all changes in behavior by agents. For instance, changes in DI application behavior.

$P^H = (\theta_0^H, \ldots, \theta_{s-1}^H, \theta_s^H, \theta_{s+1}^L, \theta_{T-1}^*; b_0, \ldots, b_{T-1})$. Let $a_{i,t}^H$ denote the application decision of individual i in period t if the policy is P^H and $a_{i,t}^L$ denote the application decision under policy P^L. The discrete welfare effect is

$$\begin{aligned} \Delta W &= W(P^H) - W(P^L) \\ &= \int_i V_i(P^H) - V_i(P^L) di + \lambda \left(G(P^H) - G(P^L) \right) \end{aligned} \tag{A.38}$$

assuming that λ is the same under both policies. We can again decompose the fiscal effect $G(P^H) - G(P^L)$ into the mechanical and behavioral fiscal effect. The mechanical fiscal effect is given by

$$\mathbb{E}\left[M_\Delta(\theta_s^*)\right] \equiv \mathbb{E}\left[\sum_{t=s}^{T-1} (1+r_t)^{-t} \left(M_{\Delta W_{i,t}} \left(b_{i,t} + \tau_{i,t}\right) + M_{\Delta Z_{i,t}} \left(b_{i,t} - z_{i,t}\right)\right)\right] \tag{A.39}$$

where $M_{\Delta W_{i,t}}$ is the mechanical employment effect

$$M_{\Delta W_{i,t}} \equiv \omega_{i,t} \left(\alpha_{i,s}^H \cdot \left[p_{i,s}^L - p_{i,s}^H\right] \prod_{k=0, k\neq s}^{t} \left(1 - \alpha_{i,k}^H p_{i,k}\right) \right) \tag{A.40}$$

and $M_{\Delta Z_{i,t}}$ is the mechanical benefit substitution effect

$$M_{\Delta Z_{i,t}} \equiv (1-\omega_{i,t}) \left(\alpha_{i,s}^H \cdot \left[p_{i,s}^L - p_{i,s}^H\right] \prod_{k=0, k\neq s}^{t} \left(1 - \alpha_{i,k}^H p_{i,k}\right) \right). \tag{A.41}$$

The mechanical fiscal effect is, as in the static model, driven by always applicants, $\alpha_{i,s}^H = 1$ (those who apply under the strict rules at age s), and the change in their award probability $\left[p_{i,s}^L - p_{i,s}^H\right]$. The share of always applicants at age s is given by $\pi^{AA} = \mathbb{E}\left[\alpha_{i,s}^H \cdot \prod_{k=0}^{s-1} \left(1 - \alpha_{i,k}^H p_{i,k}\right)\right]$.[71] We define the behavioral fiscal effect as the residual $\mathbb{E}\left[B_\Delta(\theta_s^*)\right] \equiv \left(G(P^H) - G(P^L)\right) - \mathbb{E}\left[M_\Delta(\theta_s^*)\right]$. The behavioral fiscal effect is driven by changes in the application behavior and potential other changes in behavior (which might affect the whole state distribution $F(X_{i,t})$). Writing out the behavioral fiscal effect is cumbersome because many margins can change. Empirically, we follow the same strategy by estimating the total fiscal effect and the mechanical fiscal effect and then calculate the behavioral fiscal effect as the residual.

Similarly, we can write the insurance loss as

$$\int_i V_i(P^H) - V_i(P^L) di = \mathbb{E}\left[L_{\Delta W}\right] + \mathbb{E}\left[L_{\Delta Z}\right] + \mathbb{E}\left[L_{MA}\right] \tag{A.42}$$

[71] The share of marginal applicants is $\pi^{MA} = \mathbb{E}\left[\left[\alpha_{i,s}^L - \alpha_{i,s}^H\right] \cdot \prod_{k=0}^{s-1} \left(1 - \alpha_{i,k}^H p_{i,k}\right)\right]$.

where

$$\mathbb{E}\left[L_{\Delta W}\right] \equiv \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{\Delta W_{i,t}}\left(v_i(c_{i,t}^D) - \left(u_i(c_{i,t}^W) - \theta_{i,t}\right)\right)\right)\right] \tag{A.43}$$

$$\mathbb{E}\left[L_{\Delta Z}\right] \equiv \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{\Delta Z_{i,t}}\left(v_i(c_{i,t}^D) - v_i(c_{i,t}^Z)\right)\right)\right] \tag{A.44}$$

and $\mathbb{E}\left[L_{MA}\right] \equiv \int_i V_i(P^H) - V_i(P^L)di - \mathbb{E}\left[L_{\Delta W}\right] - \mathbb{E}\left[L_{\Delta Z}\right] > 0$ is the utility loss associated with behavioral changes. The welfare effect of a discrete change is therefore $\Delta W \gtreqless 0 \Leftrightarrow$ if

$$1 + \frac{\mathbb{E}\left[B_\Delta(\theta_s^*)\right]}{\mathbb{E}\left[M_\Delta(\theta_s^*)\right]} \gtreqless \frac{\mathbb{E}\left[L_{\Delta W}\right] + \mathbb{E}\left[L_{\Delta Z}\right]}{\lambda \mathbb{E}\left[M_\Delta(\theta_s^*)\right]} + \frac{\mathbb{E}\left[L_{MA}\right]}{\lambda \mathbb{E}\left[M_\Delta(\theta_s^*)\right]}. \tag{A.45}$$

B Empirical Analysis: Additional Tables and Figures

B.1 Summary Statistics Samples

Table B.1: Summary Statistics, RSA Sample

	RSA 57	RSA 58	RSA 59
DI application ever (%)	17.23	14.83	11.84
DI application yearly (%)	4.66	3.96	3.54
w/ mental disorders	0.65	0.63	0.60
w/ musculoskeletal system	2.24	1.78	1.51
w/ other disorders (%)	1.78	1.56	1.44
Re-application yearly (%)	1.46	1.41	1.28
DI benefit receipt (%)	12.94	10.01	7.01
Employment (%)	75.54	77.69	81.60
Other benefit receipt (%)	7.51	7.85	7.92
Avg. annual earnings best 15 years (Euro)	41,148	42,193	43,007
	(10,743)	(10,933)	(10,956)
Insurance years by age 50	28.72	29.29	29.44
	(7.02)	(6.97)	(6.78)
Employment years by age 50	13.87	13.93	13.95
	(2.01)	(1.99)	(1.99)
Was ever on sick leave by age 50 (%)	33.27	32.02	31.15
Blue-collar (%)	57.61	56.51	56.28
No. Observations	1,557,723	887,252	809,342
No. Individuals	49,418	28,144	29,245

Notes: The table reports summary statistics for men between age 54 and age 62. The RSA-57 cohort comprises men born between December 1953 and November 1955, the RSA-58 cohort comprises men born between December 1955 and November 1956, and the RSA-59 cohort comprises men born between December 1956 and November 1957. Earnings are reported in 2018 Euro. Sample standard deviations for continuous variables are reported in parentheses.

B.2 Stricter DI Eligibility Rules: The 2013 Reform

Figure B.5 illustrates the step-wise increase in the RSA from age 57 to age 60 graphically. The RSA increased from age 57 to age 58 on January 1, 2013, followed by further increases to age 59 on January 1, 2015 and to age 60 on January 1, 2017. All changes were announced in November 2012.

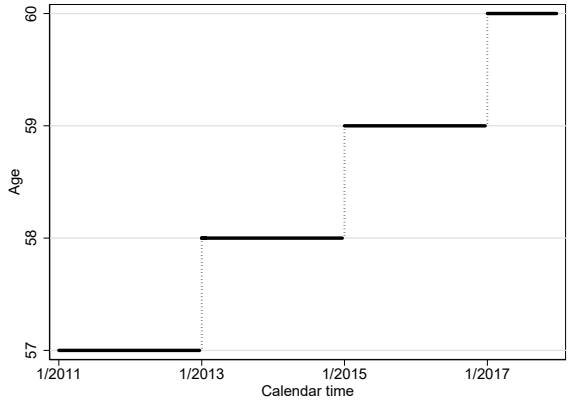
Figure B.6 provides descriptive evidence on the labor market effects of the RSA increases. We plot, by birth cohort, the percentage of males aged 54-61 receiving DI benefits, having ever applied for DI, working, and receiving other benefits. For each variable, trends across birth cohorts are remarkably similar until age 57 – the relaxed screening age for the RSA-57 cohort. At this age, the DI recipient rate rises sharply in the RSA-57 cohort. The percentage males having ever applied for DI also increases, suggesting that many are aware of the

Table B.2: Summary Statistics, Benefit-Generosity Sample

	Age Groups	
	57-60	30-56
DI application (%)	6.10	0.78
w/ mental disorders	0.47	0.11
w/ musculoskeletal system	2.02	0.16
w/ other disorders	3.61	0.52
DI inflow (%)	4.06	0.35
Employment outlfow (%)	1.55	0.06
Other benefit outflow (%)	2.51	0.29
Age (years)	57.35	42.39
	(6.98)	(3.37)
Insurance years	36.98	22.61
	(8.16)	(8.44)
Last annual earnings (Euro)	41,405	39,511
	(18,038)	(17,196)
Avg. annual earnings, best 15 years (Euro)	38,855	42,481
	(12,512)	(12,680)
Blue-collar (%)	49.25	50.68
No. Observations	1,453,448	15,968,003
No. Individuals	491,426	1,801,685

Notes: The table reports summary statistics for 30-56 year old men and 57-60 year old men using data for the years 2004 to 2017. Earnings are reported in 2018 Euro. Sample standard deviations for continuous variables are reported in parentheses.

Figure B.5: The 2013 Reform: Increase in the RSA

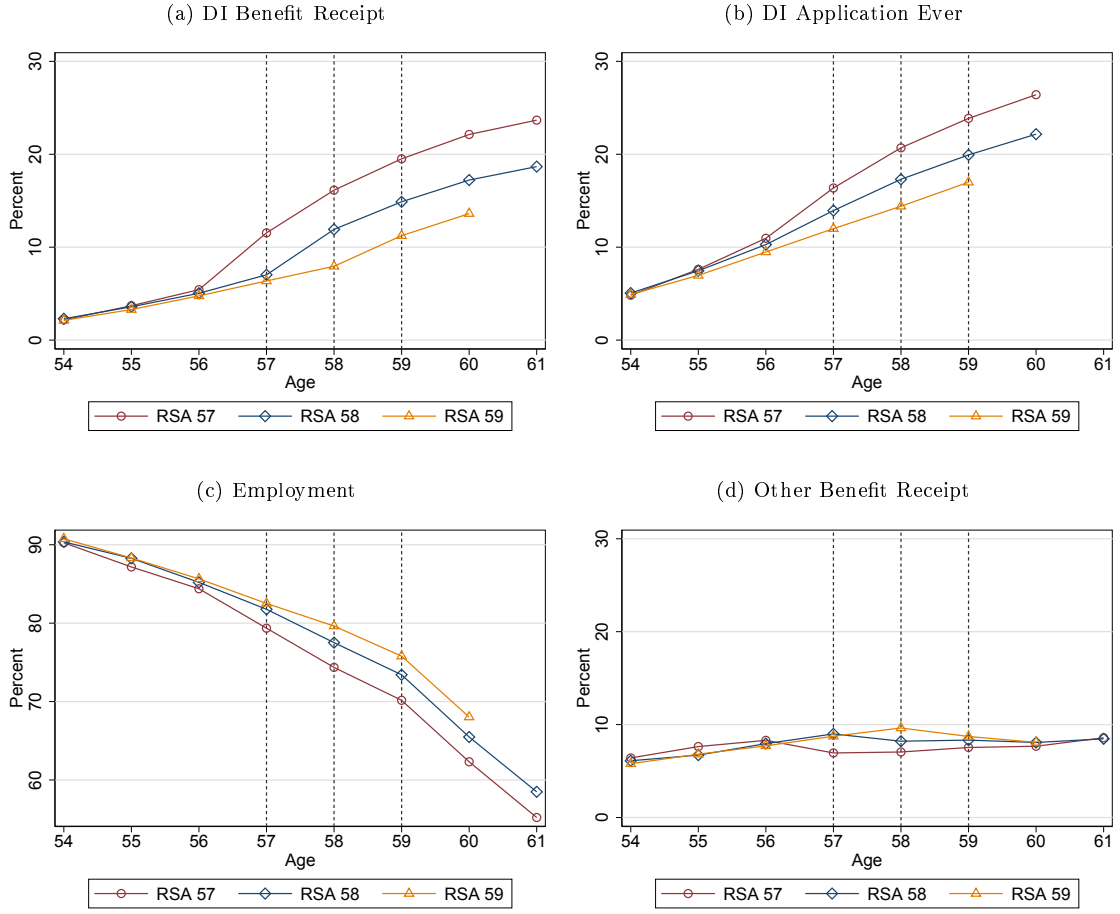


Notes: The figure displays the step-wise increase in the relaxed screening age (RSA) for DI benefits from age 57 to age 60, as mandated by the 2012 2nd Stability Act. Source: Austrian federal law (Bundesgesetzblatt) no. 35/2012.

RSA and time their DI application to this age. Conversely, the percentage males of the RSA-57 cohort who are employed or receive other benefits drops at age 57, pointing to the role of DI as a substitute for UI or SI.

We observe similar breaks in trends when the RSA-58 and RSA-59 cohorts reach their RSA, but interestingly cohorts with a higher RSA never catch up to cohorts with a lower RSA. For example, the DI recipiency rate rises sharply at age 58 for the RSA-58 cohort, but it never reaches the level of the RSA-57 cohort. A one year increase in the RSA changes labor market dynamics not only at the age where disability eligibility becomes

Figure B.6: Labor Market States and DI Applications Ever, by Age and RSA



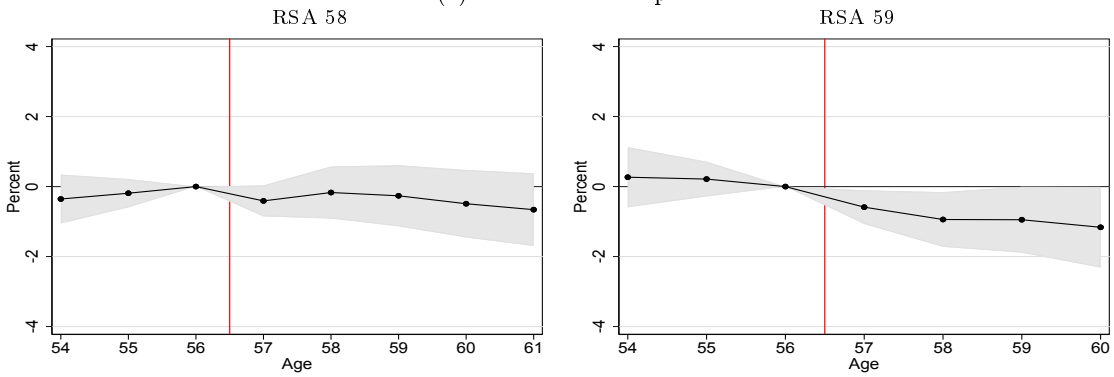
Notes: The figure shows trends in DI benefit receipt, DI application ever (measuring whether somebody has ever applied for DI), employment, and other benefit receipt by age for the different RSA cohorts.

stricter, but also at higher ages. Our empirical strategy is designed to separately identify the effect of the RSA at the ages where eligibility becomes stricter as well as at higher ages where eligibility is relaxed.

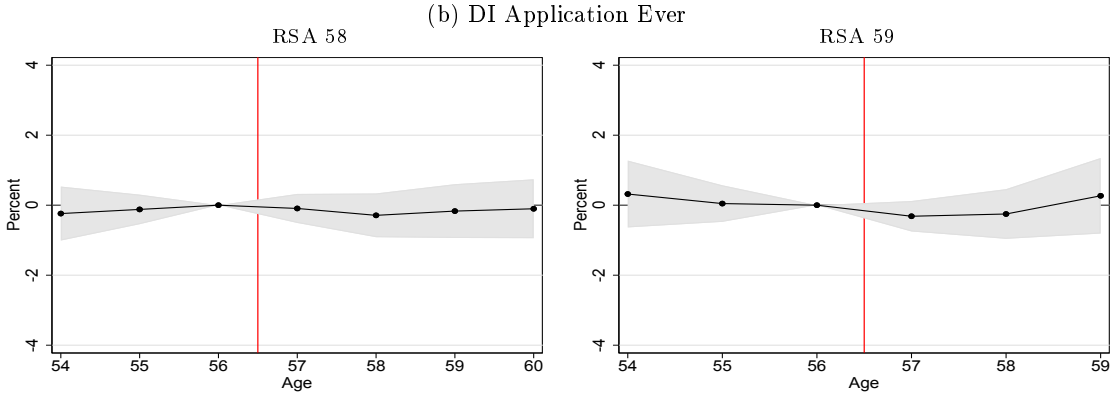
RSA Effect Among Non-eligible Males. In Figure B.7, we plot the estimated β_k-coefficients from equation (12) for men with too little work experience to be eligible for DI under relaxed eligibility rules. We find that DI benefit receipt, DI application ever, employment and other benefit receipt do not differ significantly across birth cohorts, even after age 56, providing strong support for the absence of differential trends across birth cohorts.

Figure B.7: Effects of RSA on DI Benefit Receipt and DI Applications by Age, Non-eligibles

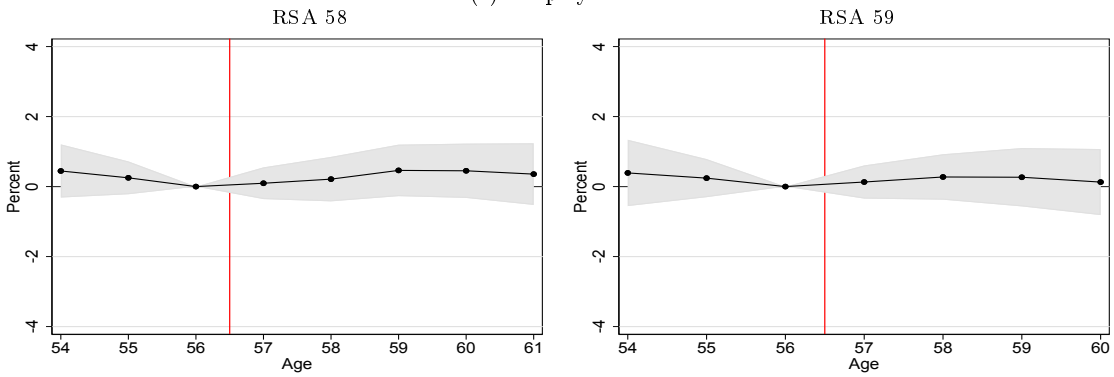
(a) DI Benefit Receipt



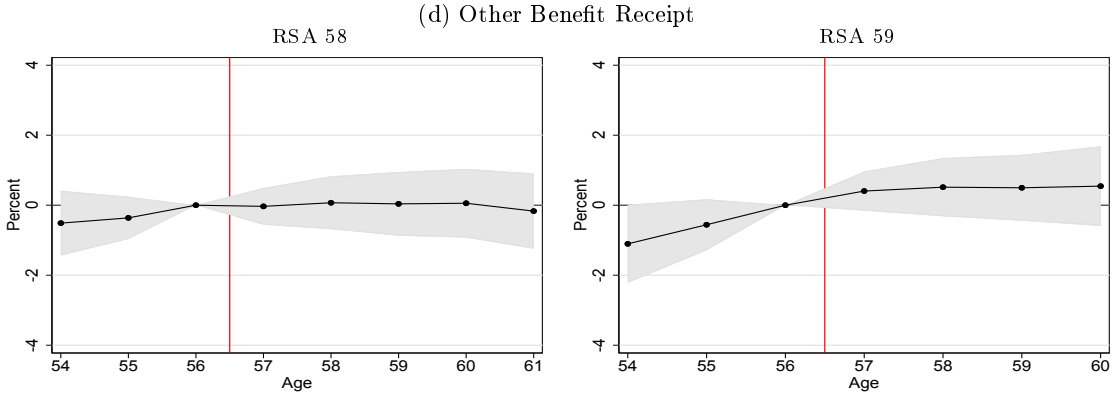
(b) DI Application Ever



(c) Employment



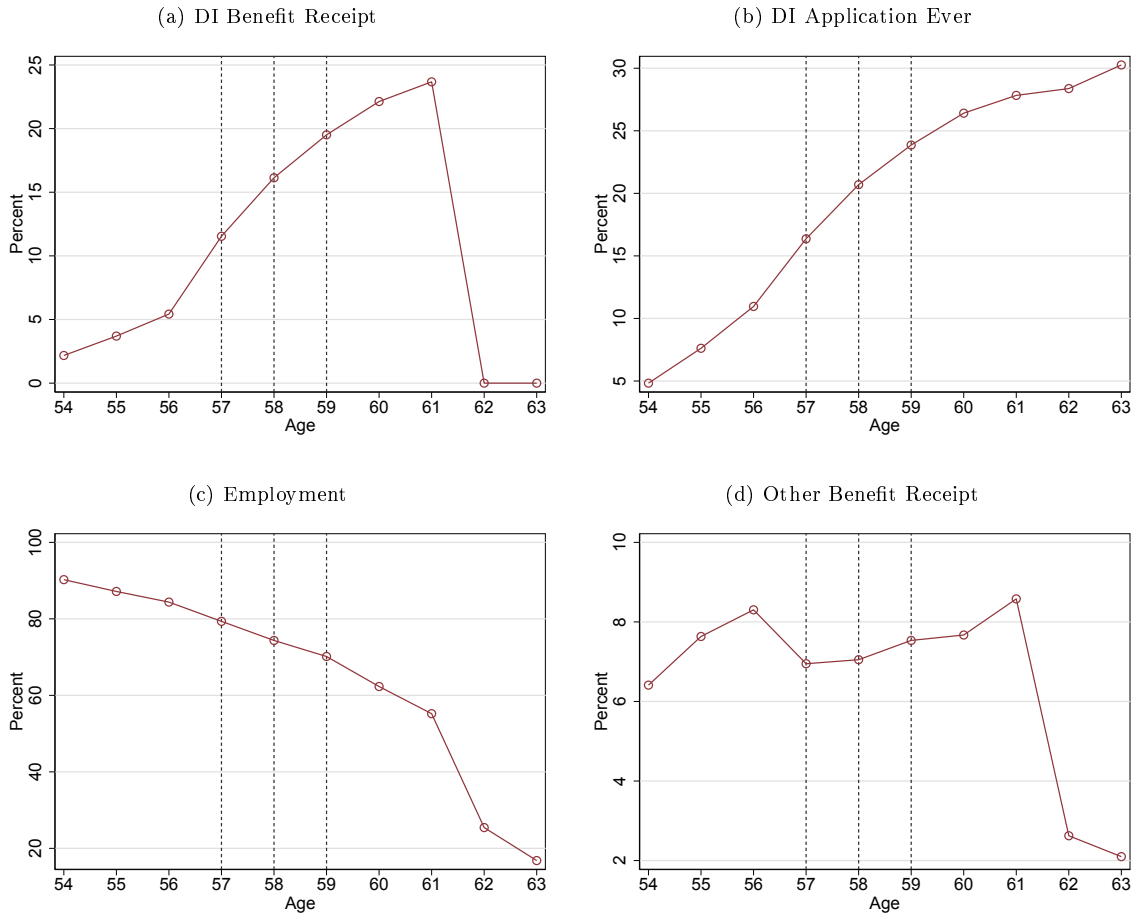
(d) Other Benefit Receipt



Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA 58 and RSA 59 increases using the sample of non-eligible men. The shaded area denotes the 95 percent confidence interval.

Tracking RSA Effects up to Age 62. The empirical results suggest that the RSA increases have long-term effect up to the last age, we can observe treated cohorts in the data (age 61 for the RSA-58 cohort and age 60 for the RSA-59 cohort). A natural question is whether these effects would continue even beyond the last age, we currently observe in the data? A simple way to shed light on this question is by looking at an older cohort, men born in 1954, whom we can track until age 63. We would expect that the effects of the RSA increases disappear at age 62. At this age most men in Austria retire, because they become eligible for retirement benefits. Indeed, if we plot age trends in labor market outcomes and DI application ever for men born in 1954 (Figure B.8), at age 62 we observe sharp drops in the precent of men receiving DI benefits, being employed, or receiving other benefits.

Figure B.8: Trends by Age for Eligible Men Born in 1954



Notes: The figure shows trends in DI benefit receipt, DI application ever (measuring whether somebody has ever applied for DI), employment, and other benefit receipt by age for men born 1954 who we can track until age 63.

We can estimate the effect of the RSA increases up to age 62 if we assume that the β_k-coefficient estimates in equation (12) are unchanged between the last age, we currently observe in the data, and age 62. This assumption is reasonable because, as Figure B.14 shows, the effects stabilize after age 58. Table B.3 reports the correpsonding average effects between age 57 and age 61, that is $\sum_{k=57}^{61} \beta_k/5$. For the RSA-58 cohort, we can

Table B.3: Average Effect Above RSA up to Age 62, Men

	Eligible				Non-eligible			
	2013		2015		2013		2015	
	Estimate	Mean	Estimate	Mean	Estimate	Mean	Estimate	Mean
A. Labor market effects (%)								
DI benefit receipt	-2.54***	18.56	-4.94***	17.3	-0.4	38.17	-0.96**	37.52
	(0.44)		(0.43)		(0.38)		(0.42)	
DI application ever	-1.17***	21.81	-2.86***	20.29	-0.15	38.61	0.05	37.89
	(0.36)		(0.36)		(0.34)		(0.42)	
Employment	1.85***	68.36	3.18***	71.59	0.32	14.34	0.19	14.73
	(0.39)		(0.43)		(0.3)		(0.35)	
Other benefit receipt	0.94***	7.55	2.20***	7.3	-0.01	19.8	0.5	20.08
	(0.25)		(0.30)		(0.38)		(0.42)	
B. Fiscal effects (euro)								
DI benefits	-884***	6756	-1793***	6245	-115	11012	-445***	10721
(A)	(161)		(159)		(120)		(123)	
Tax revenue	263***	11185	427***	11625	16	1582	-16	1608
(B)	(56)		(65)		(33)		(38)	
Other benefits	172***	1217	451***	1182	-5	2233	92	2277
(C)	(46)		(63)		(55)		(67)	
Total fiscal effect	-976***	-3213	-1769***	-4199	-135	11663	-338***	11389
(A-B+C)	(185)		(186)		(115)		(123)	
No. Observations	2,444,975		2,176,311		916,207		806,100	

Notes: The tabe reports the average effect of the RSA for the ages between above age 56 and below age 62. The estimates are constructed by taking the average of the β_k-coefficients from equation (12) for $k \geq 57$. The *eligible*-sample comprises men who, at age 56, have worked for at least 10 years in the last 15 years. The *non-eligible*-sample comprises men who, at age 56, have worked less than 10 years in the last 15 years. *Mean* denotes the mean above the RSA for the RSA-57 cohort. Fiscal effects are reported in 2018 Euro. Standard errors clustered at the year-month of birth level are reported in parentheses. Levels of significance: *1%, **5%, and ***1%.

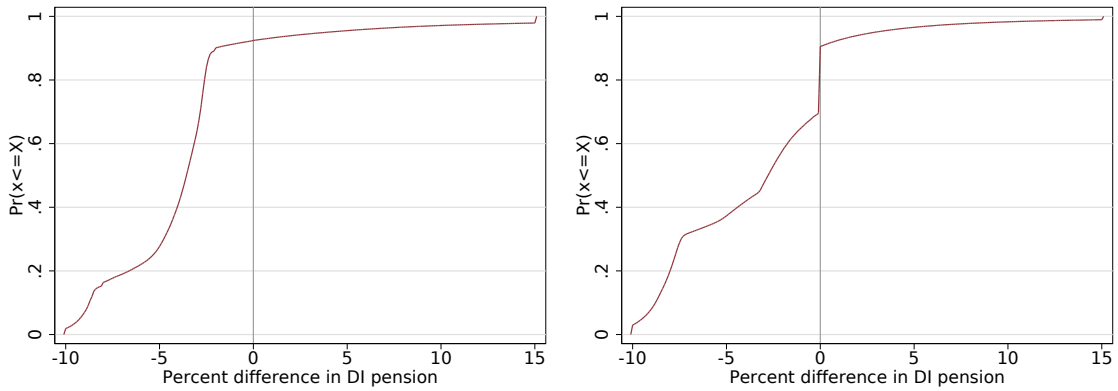
observe labor market outcomes up to age 61. Thus, the only estimate that changes compared to Table 1 is the estimate on DI application ever, which we observe only until age 60. We find that the application rate declines by 1.17 percentage points on average, which is almost identical to the estimate in Table 1 (-1.19 percentage points). We find equally small differences when comparing the estimates for the RSA-59 cohort.

B.3 Lower DI Benefits: The 2003 Reform

Figure B.9 illustrates the effect of the 2003 reform by showing the distribution of changes in potential DI benefits between 2004 and 2017.

Figure B.9: Cumulative Distribution Functions of %-change in DI Benefits

(a) Ages 57-60 (b) Ages 30-56

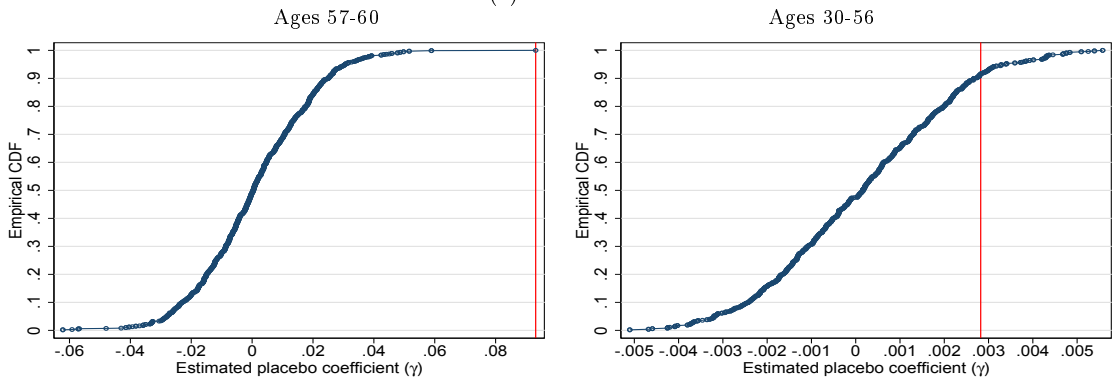


Notes: The figure shows the cumulative distribution in the percent change in DI benefits between 2004 and 2017 for men between ages 57-60 (panel a) and men between ages 30-56 (panel b).

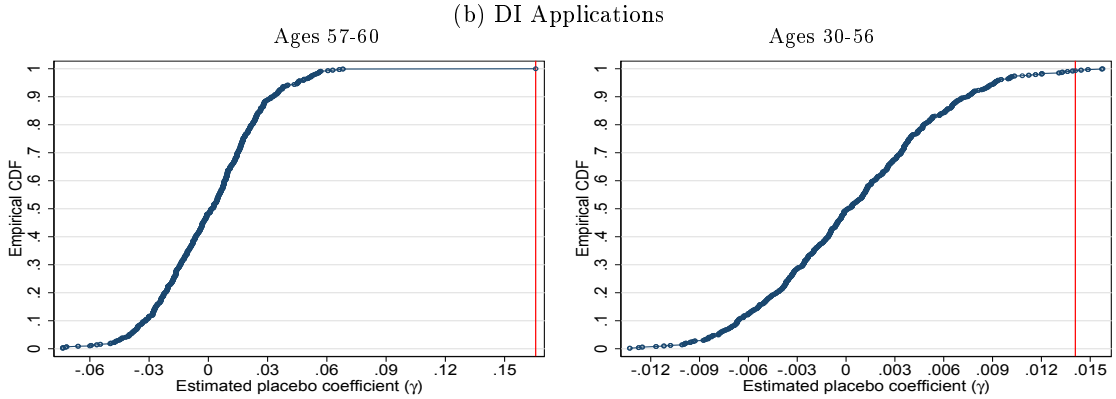
Placebo Estimates Benefit Generosity. To test the appropriateness of our strategy to identify the effect of DI benefits, we estimate 1,000 placebo regressions, in which we randomly assign individuals within each cell defined by year, insurance-year decile, and assessment decile potential benefits $b_r(Z_{it})$ from a different year. Appendix Figure B.10 plots the empirical cumulative distribution of the placebo estimates together with the true estimate for each outcome in Panel A of Table 2. The figures confirm that true increases in benefit generosity lead to unusually large increases in DI inflow, DI applications, and other benefit outflow, but have no effect on employment outflow.

Figure B.10: Distribution of Placebo Estimates

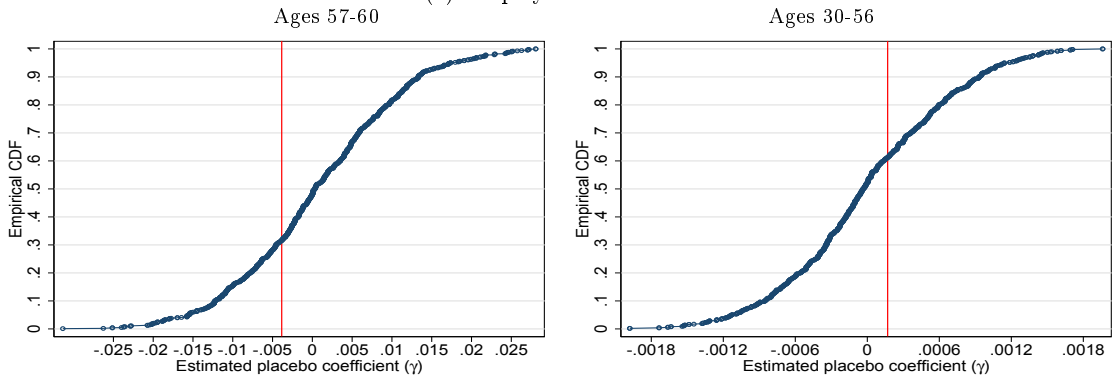
(a) DI Inflow



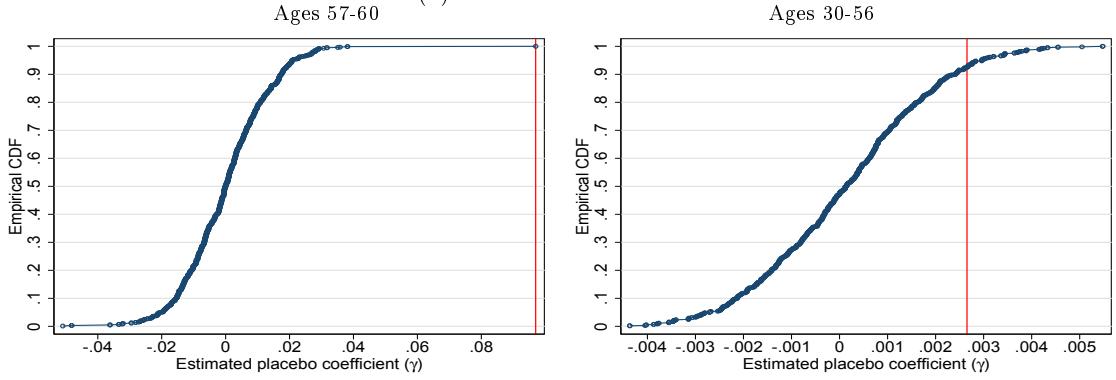
(b) DI Applications



(c) Employment Outflow



(d) Other Benefit Outflow

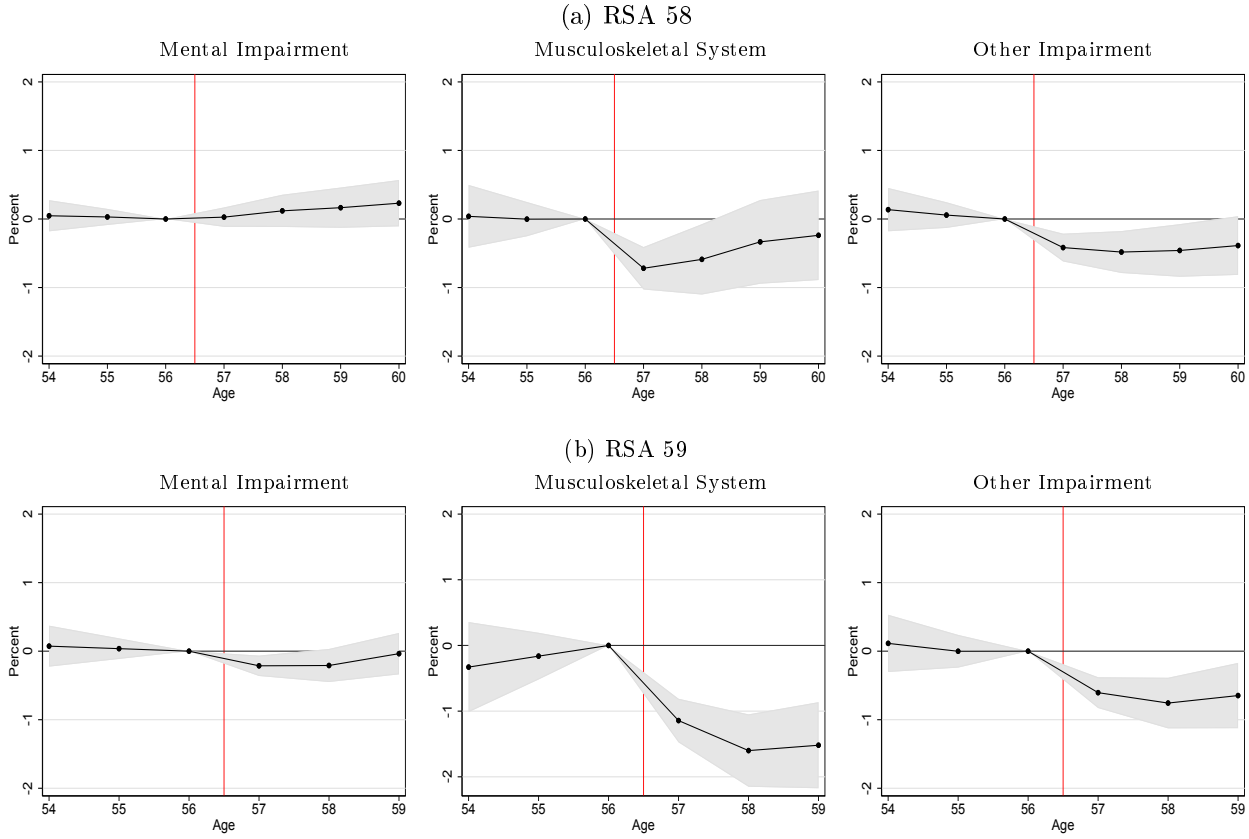


Notes: The figure plots the empirical distribution of placebo effects for labor market outcomes estimated. The CDF is constructed from 1,000 estimates of γ from estimation equation (14) when we randomly assign individuals within each call defined by year, insurance-year decile, and assessment-basis decile potential benefits b_{it} from a different year. The vertical line shows the treatment effect estimate responded in Table.

B.4 DI Application Effects by Medical Impairment

Our main results indicate that stricter DI eligibility rules and lower DI benefits reduce the propensity to apply for DI benefits. Here we examine the application effect on the type of impairments with which individuals apply, separately for each DI policy instrument. Figure B.11 plots the estimated β_k-coefficients from equation (12) for increases in the RSA to 58 and 59. As outcome variables we use whether individuals have ever applied with a mental impariment, a musculoskeletal impairment, or an other impairment. The shaded area denotes the 95 percent confidence interval. The estimates suggest that an increase in the RSA reduces mainly applications with musculoskeletal and other impairments, while applications with mental impairments do not change.

Figure B.11: RSA Effects by Application Impairment



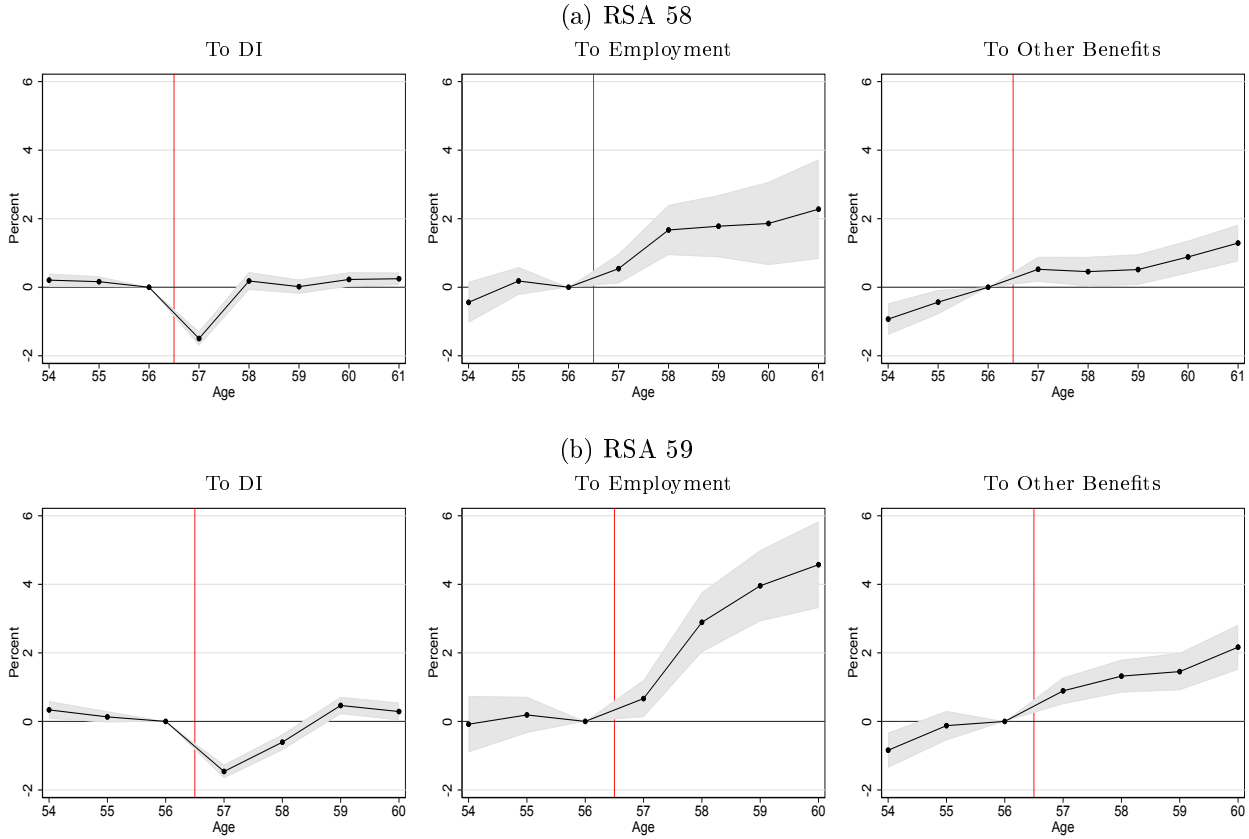
Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA 58 and RSA 59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

Table B.4 shows how changes in benefit generosity affect the type of impairment with which individuals apply for DI. Similar to stricter DI eligibility rules, we find no significant effect on the number of applications with a mental impairment, but signifcant increase in the number of applications with a musculoskeletal or any other impairment. The absence of an effect on applications with a mental impairment is interesting, because a mental illness is often considered a difficult-to-verify disorder (see, e.g., Autor and Duggan, 2006) and one would expect it to be responsive to changes in DI policy, but our findings do not provide support for this intuition.

B.5 The Effect of DI Reforms on Labor Market Transitions

Our estimates show that stricter DI eligibility rules increases employment and other benefit receipt among treated cohorts. The increases can result either from changes in the inflow into employment or other benefit receipt, or changes in the persistences in employment or other benefit receipt. To shed light on the importance of these two effects, Figures B.12 and B.13 plots the estimated β_k-coefficients from equation (12) using as outcome variable transitions from and persistence in employment and other benefit receipt.

Figure B.12: RSA Effects on Transitions from Employment



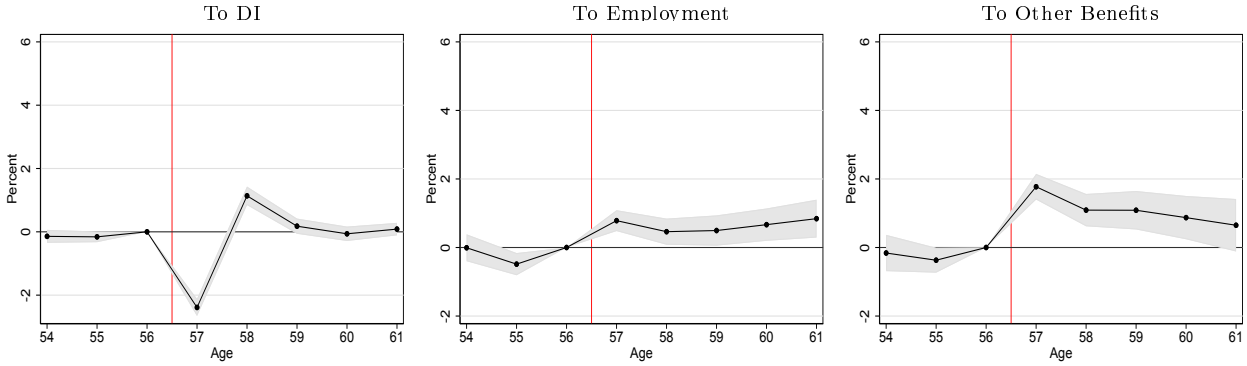
Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA-58 and RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

The first colum of figure B.12 shows a drop in transitions from employment to DI at the ages where DI eligibility rules become stricter but not at other ages. The middle column shows that stricter DI eligibility rules induce indivdiuals who already employed to stay employed longer. The last column shows an increase in ransitions from employment to other benefit receipt, but the magnitude of the effect is only about half as big as the increase in employment persistence. Similarly, the first colum of figure B.13 shows a drop in transitions from other benefits to DI at the ages where DI eligibility rules become stricter, but we also see an increase in transitions into DI as soon as a cohort reaches its RSA, suggesting that some individuals receive other benefits longer until they reach the new RSA. Consistent with this idea, the third column of B.13 shows a sharp

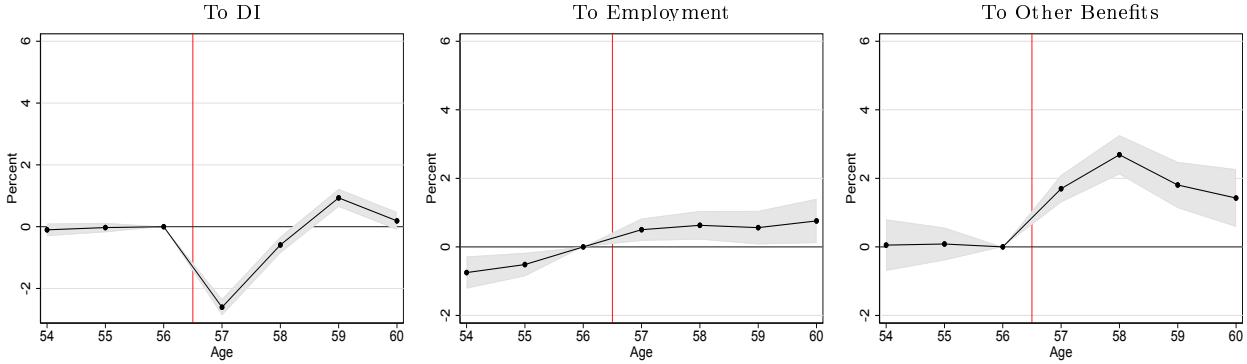
increase in persistence in other benefits at the ages where DI eligibility rules become stricter followed by a drop as soon as DI eligiblity rules are relaxed again. Transitions from other benefits to employment also increase after age 56, but the magnitude of the effect is smaller than the increase in persistence in other benefits (second column of Figure B.13)

Figure B.13: RSA Effects on Transitions from Other Benefits

(a) RSA 58



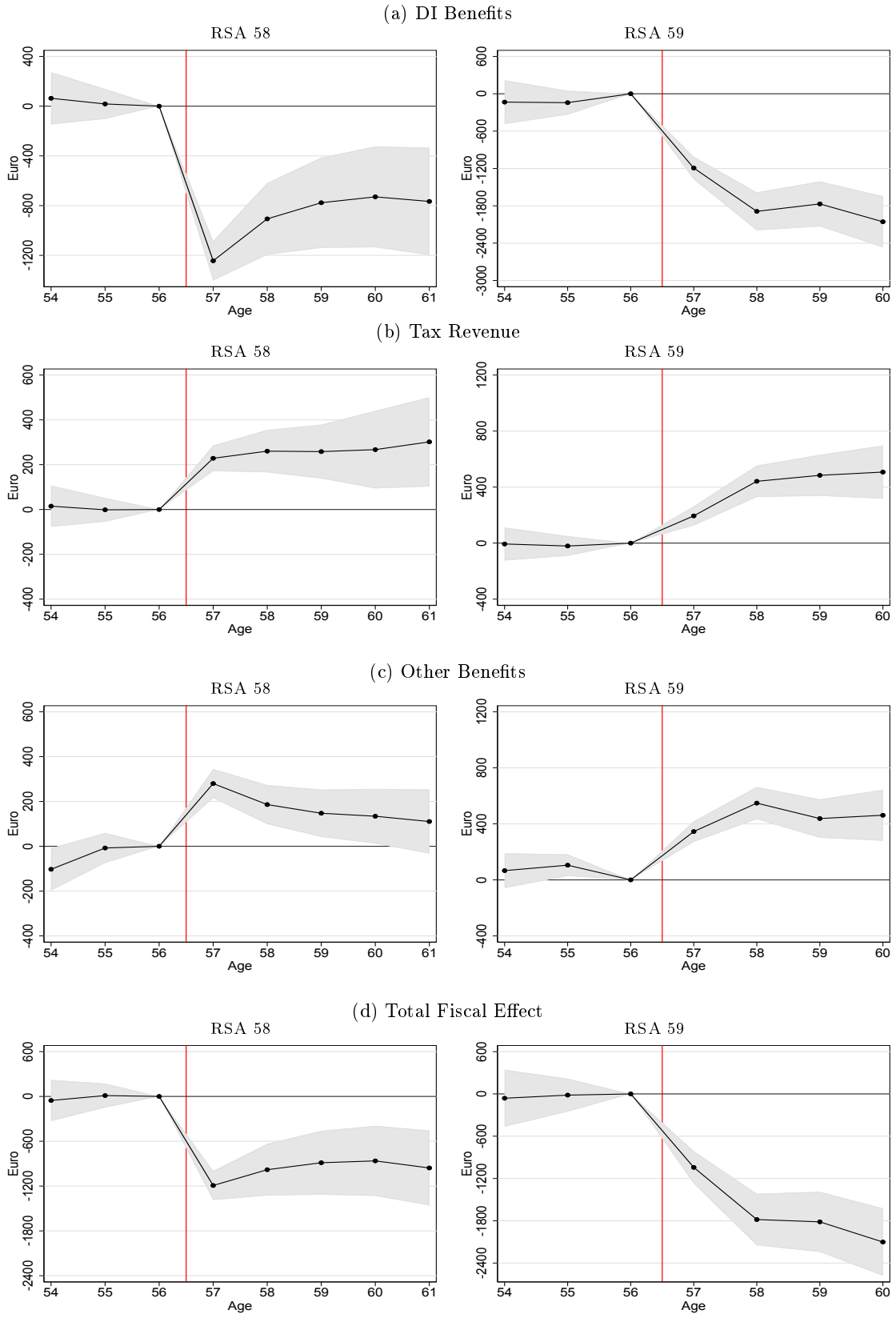
(b) RSA 59



Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA-58 and RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

B.6 Fiscal Effects

Figure B.14: Fiscal Effects of RSA by Age

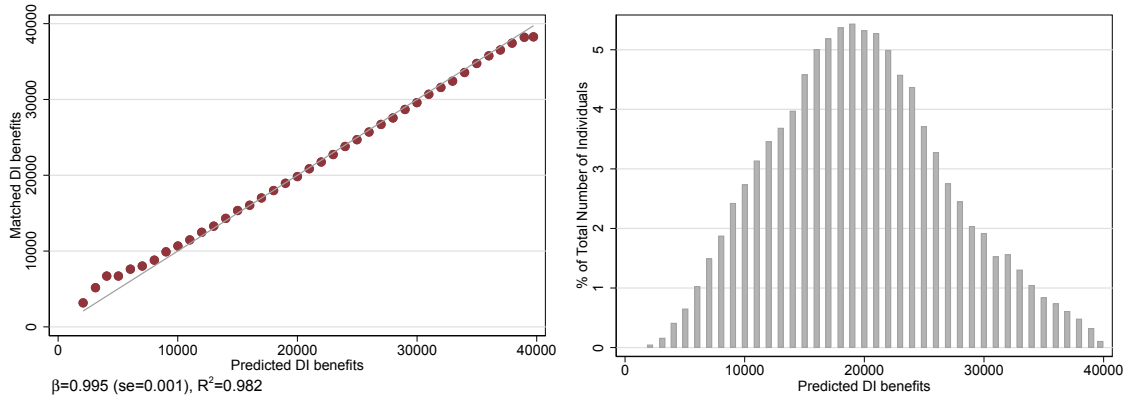


Notes: The figure shows the estimated β_k-coefficients from the econometric specification in (12) for the RSA-58 and RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

B.7 Further Evidence on the Effect of DI Benefit Generosity

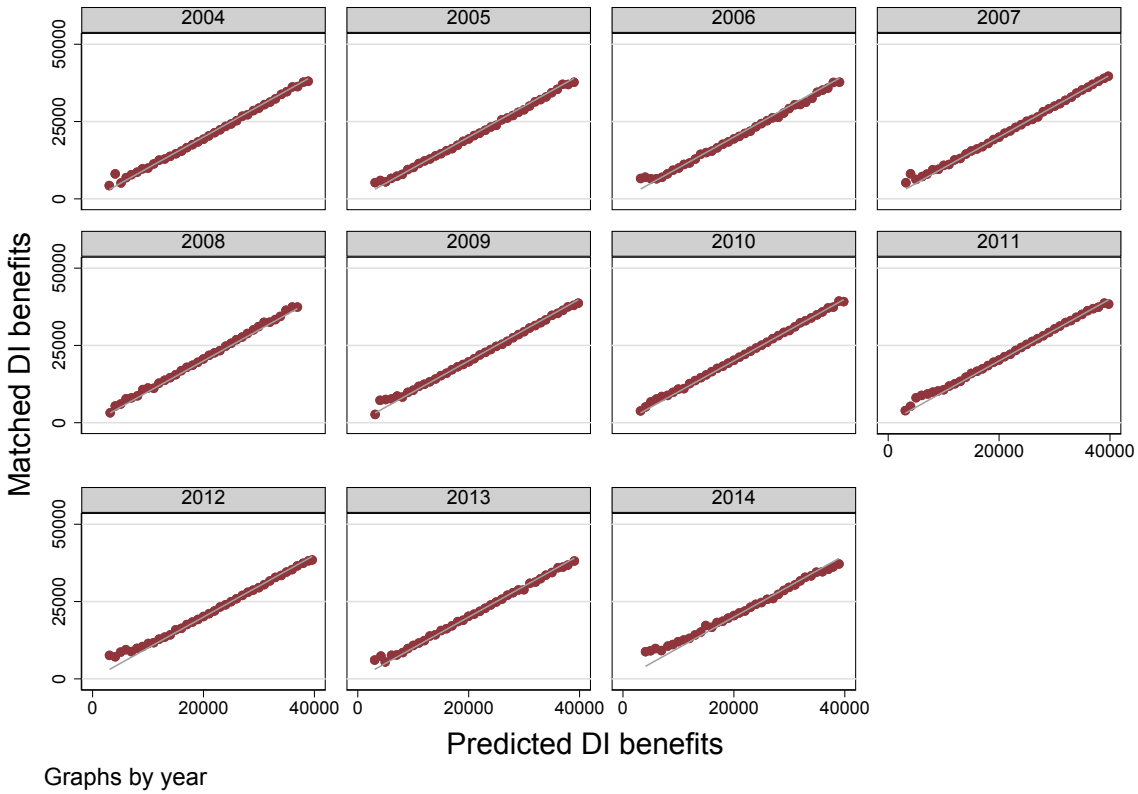
Figure B.15: Predicted and Matched DI benefits

(a) All Years Pooled (b) Sample Distribution



Notes: Panel (a) compares the predicted and matched DI benefits in 1,000 Euro bins. The β-coefficient and R^2 are from a linear regression of matched DI benefits on predicted DI benefits. Panel (b) shows the percent of individuals in each bin relative to the total number of individuals.

Figure B.16: Predicted and Matched DI Benefits, by Year



Notes: The figure compares the prediected and matched DI benefits in each year from 2004 to 2014, the last year in which matched DI benefits are available.

Table B.4: Application Effect of DI Benefit Generosity, by Health Impairment

	Ages 57-60		Ages 30-56	
	Estimate	Mean	Estimate	Mean
Mental impairments	0.007	0.42	<0.001	0.11
	(0.004)		(0.001)	
Musculoskeletal system	0.067***	1.77	0.002**	0.15
	(0.010)		(0.001)	
Other impairments	0.096***	3.17	0.012***	0.51
	(0.017)		(0.003)	
Observations	1,453,448		15,968,003	

Notes: The tabe reports estimates for γ from the econometric specification in (14). *Mean* denotes the mean in levels for the year 2004. Standard errors clustered at the year-month of birth level are reported in parentheses. Levels of significance: *1%, **5%, and ***1%.

C Characterizing Compliers

C.1 Framework for Complier Analysis

In this section, we describe the complier analysis for difference-in-differences settings, as outlined in De Chaisemartin and D'Haultfoeuille (2018); Jäger et al. (2019), to study the characteristics of marginal, always, and never applicants. (We follow the same steps to study the characterististcs of marginal, always, and never enrollees.) For the RSA-58 change, we focus on the ages 56 and 57 and compare the RSA-58 cohort to the RSA-57 cohort. The RSA-57 cohort faces relaxed DI eligibility standards at age 57, while eligibility standards for the RSA-58 cohort are strict at both ages. For the RSA-59 increase, we focus on the ages 55 to 58 and compare the RSA-59 cohort to the RSA-57 cohort. The RSA-57 cohort faces relaxed DI eligibility standards at age 57 and age 58, while eligibility standards for the RSA-59 cohort are strict at any age between 55 and 58.

We denote by a the age window. It can take two values: $a = A57$ is the age window above 56 and $a = B57$ is the age window below 57. We denote by c the cohort; $c = T$ is the RSA-57 cohort (the treatment cohort) and $c = C$ is the RSA-58 cohort (the control cohort). For the RSA-59 change, $c = C$ denotes the RSA-59 cohort. We have a binary instrument Z, equal to one if DI eligibility is relaxed and zero otherwise, that is $Z = 1$ for $(T, A57)$ and $Z = 0$ for $(T, B57), (C, A57)$, and $(C, B57)$. AP is an indicator whether an individual applies for DI benefits. Following the potential outcomes framework, AP can take two potential values: AP_0 is the potential value of AP for $Z = 0$ and AP_1 is the potential value of AP for $Z = 1$. We can now distinguish three groups of applicants: always applicants ($AP_0 = AP_1 = 1$), never applicants ($AP_0 = AP_1 = 0$), and marginal applicants who only apply when DI eligibility standards are relaxed ($AP_0 = 0$ and $AP_1 = 1$). We define the different groups of enrollees in the same way basend on an indicator DI, which is one if an individual is awarded DI benefits and zero otherwise. DI_0 and DI_1denote the potential values of DI for $Z = 0$ and $Z = 1$.

Estimating the expected value of a characteristic X for never applicants is straightforward. All individuals in $(T, A57)$ who do not apply for DI are never applicants if we assume $AP_1 - AP_0 \geq 0$, the standard monotonicity assumption in the instrumental variables literature.[72] We can estimate the conditional value of a never applicant characteristic $E(X|AP_1 = 0, T, A57)$ by the corresponding sample mean $\frac{1}{N^{na}_{T,A57}} \sum_{i \epsilon (T,A57)} X_i \cdot \mathbb{I}(AP_i = 0)$,where i is individual and $N^{na}_{T,A57}$ is the number of people in $(T, A57)$ who do not apply for DI and $\mathbb{I}(AP_i = 0)$ is equal to one if an individual has not applied for DI and zero otherwise. We use the same logic to estimate the expected value of a characteristic for a never enrollee.

Estimating the expected value of a characteristic for marginal and always applicants is more challenging and requires additional assumptions (Jäger et al., 2019). The key insight is that the expected value of a characteristic X for all applicants in $(T, A57)$ is a weighted average of the expected value for marginal and

[72] The monotonicity assumption rules out defying applicants who would apply when DI eligibility rules are strict but not when DI eligibility rules are relaxed.

always applicants, where the weights represent the share of marginal applicants and always applicants among all applicants. We can re-arrange the weighted average to get an expression for the expected value of a marginal applicant characteristic:

$$E(X|AP_0 = 0, AP_1 = 1, T, A57) = \frac{\pi^{ma} + \pi^{aa}}{\pi^{ma}} \cdot E(X|AP_1 = 1, T, A57) - \frac{\pi^{aa}}{\pi^{ma}} \cdot E(X|AP_0 = 1, T, A57) \quad \text{(C.46)}$$

where $\pi^{ma} = N^{ma}_{T,A57}/N_{T,A57}$ and $\pi^{aa} = N^{aa}_{T,A57}/N_{T,A57}$ are the shares of marginal and always applicants in $(T, A57)$.

We can estimate each term of the right-hand side of equation (C.46) empirically. We estimate the shares of each group of applicants with the following regression:

$$AP_{iac} = \alpha + \beta_a + \gamma_c + \delta\, Z_{ac} + \varepsilon_{iac}, \quad \text{(C.47)}$$

where β_a is a fixed effect for the age window $a = A57$ and γ_c is a fixed effect for the cohort $c = T$. If Z is independent from AP and application trends in the absence of relaxed DI eligibility are parallel across cohorts, then $\pi^{aa} = \alpha + \beta + \gamma$ is the share of always applicants, $\pi^{ma} = \delta$ is the share of marginal applicants, and $\pi^{na} = 1 - \pi^{aa} - \pi^{ma}$ is the share of never applicants (De Chaisemartin and D'Haultfoeuille (2018); Jäger et al. (2019)).[73] We estimate the share of always enrollees (π^{ae}), marginal enrollees (π^{me}) and never enrollees (π^{ne}) in the same way, but simply use DI as the dependent variable in equation (C.47). We estimate the conditional value of an applicant characteristic $E(X|AP_1 = 1, T, A57)$ by the corresponding sample mean $(1/N^a_{T,A57}) \cdot \sum_{i\epsilon(T,A57)} X_i \cdot \mathbb{I}(AP_i = 1)$,where $N^a_{T,A57}$ is the number of applicants in $(T, A57)$ and $\mathbb{I}(AP_i = 1)$ is one if an individual has applied for DI.

Calculating $E(X|AP_0 = 1, AP_1 = 1, T, A57)$ is more difficult, because we never get to see whether applicants in $(T, A57)$ would have applied if eligibility standards were strict, that is we never observe the potential outcome AP_0. But because of monotonicity we know that individuals who apply when eligibility standards are strict also apply when eligibility standards are relaxed, allowing us to write $E(X|AP_0 = 1, AP_1 = 1, T, A57) = E(X|AP_0 = 1, T, A57)$. If trends in X are parallel across cohorts under strict eligibility standards and Z is independent from AP and X, we can estimate $E(X|AP_0 = 1, T, A57)$ using the change in applications for cohort C, $E(X|AP_0 = 1, T, A57) = E(X|AP_0 = 1, T, B57) + E(X|AP_0 = 1, C, A57) - E(X|AP_0 = 1, C, B57)$. We can estimate each element on the right-hand side by the corresponding sample mean: $(1/N^a_{T,B57}) \cdot \sum_{i\epsilon(T,B57)} X_i \cdot \mathbb{I}(AP_i = 1) + (1/N^a_{C,A57}) \cdot \sum_{i\epsilon(C,A57)} X_i \cdot \mathbb{I}(AP_i = 1) - (1/N^a_{C,B57}) \cdot \sum_{i\epsilon(C,B57)} X_i \cdot \mathbb{I}(AP_i = 1)$.

[73] Formaly, the independence assumption is equal to $AP_0, AP_1 \perp Z \mid a, c$ and the parallel trend assumption is equal to $E(AP_0|A57, T)$-$E(AP_0|B57, T) = E(AP_0|A57, C) - E(AP_0|B57, C)$.

C.2 Complier Analysis for the RSA-59 Increase

Table C.5 shows the population shares and average characteristics of marginal applicants and enrollees, always applicants and enrollees, and never applicants and enrollees for the RSA-59 change. The differences between the various groups of applicants and enrollees mirror the patterns for the RSA-58 change (Table 3). Marginal applicants and enrollees, compared to always applicants and enrollees, are in better health (proxied by sick leave absence at age 56), are more likely to be employed at age 56, are more likely to work in blue-collar jobs and are more likely to apply with a musculoskeletal impairment. Overall, these differences provide evidence consistent with marginal applicants and enrollees having a higher work capacity than always applicants and enrollees (but a lower work capacity than never applicants and enrollees).

Table C.5: Characteristics of DI Applicants and DI Recipients, RSA 59

	Marginal (M)	Always (A)	Difference M-A	Never (N)	Difference M-N
A. Applicants					
Share in population (%)	0.033***	0.102***	-0.069***	0.866***	-0.833***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Sick Leave at age 56 (%)	-0.88	8.38***	-9.26***	0.92***	-1.80**
	(0.74)	(0.20)	(0.91)	(0.01)	(0.74)
Unemployed at age 56 (%)	10.65***	25.53***	-14.88***	4.26***	6.40***
	(1.57)	(0.43)	(1.96)	(0.03)	(1.57)
Employed at age 56 (%)	84.83***	62.80***	22.03***	87.15***	-2.31
	(1.94)	(0.52)	(2.40)	(0.06)	(1.94)
Avg. annual earnings (euro)	41,392***	41,054***	339	46,347***	-4,955***
	(365)	(109)	(465)	(20)	(367)
Blue-collar (%)	85.87***	80.34***	5.53***	53.89***	31.98***
	(1.63)	(0.48)	(2.07)	(0.09)	(1.63)
Musculoskeletal (%)	74.93***	36.37***	38.57***		
	(2.29)	(0.53)	(2.70)		
Mental (%)	4.55***	17.83***	-13.27***		
	(1.67)	(0.48)	(2.10)		
Other (%)	20.51***	45.80***	-25.29***		
	(2.15)	(0.57)	(2.65)		
B.Enrollees					
Share in population (%)	8.00***	3.82***	4.19***	88.18***	-80.17***
	(0.10)	(0.08)	(0.16)	(0.05)	(0.13)
Sick Leave at age 56 (%)	6.40***	9.90***	-3.50***	0.83***	5.58***
	(0.19)	(0.28)	(0.44)	(0.01)	(0.19)
Unemployed at age 56 (%)	23.20***	16.92***	6.28***	4.68***	18.52***
	(0.32)	(0.44)	(0.69)	(0.03)	(0.32)
Employed at age 56 (%)	70.05***	64.31***	5.74***	86.79***	-16.74***
	(0.41)	(0.69)	(1.03)	(0.05)	(0.41)
Avg. annual earnings (euro)	40,852***	42,207***	-1,355***	46,231***	-5,379***
	(94)	(149)	(229)	(19)	(94)
Blue-collar (%)	87.69***	72.73***	14.96***	54.24***	33.45***
	(0.44)	(0.66)	(1.02)	(0.08)	(0.44)
Musculoskeletal (%)	57.28***	24.91***	32.37***		
	(0.54)	(0.64)	(1.03)		
Mental (%)	7.43***	25.18***	-17.76***		
	(0.42)	(0.68)	(1.02)		
Other (%)	30.57***	55.02***	-24.45***		
	(0.54)	(0.77)	(1.19)		

Notes: The tabe reports the population shares and average characteristics of marginal applicants and enrollees, always applicants and enrollees, and never applicants and enrollees for the RSA-59 increase. Earnings are reported in 2018 Euro. Levels of significance: *1%, **5%, and ***1%.

D Fiscal Multipliers

D.1 Comparing Pre-57 Applicants to Always Applicants in the Treatment Group

Table D.6 presents summary statistics for the pre-57 applicants sample, which we use to estimate the mechanical effect of stricter eligibility rules. Around 50 percent of the pre-57 applicants are already on DI by age 57. Around 40 percent of the remaining individuals reapply at age 57. 60 to 77 percent of all pre-57 applicants are on DI benefits by age 60. In total we have around 6000 pre-57 applicants in the control group (RSA 57) and around 3000 individuals in each of the two treatment groups (RSA 58 and RSA 59). In terms of labor market characteristics before age 57 the treatment and control groups are very comparable.

Table D.6: Summary Statistics, Pre-57 Applicants

	RSA 57	RSA 58	RSA 59
On DI by age 57 (%)	50.89	49.36	47.63
Died by age 57 (%)	2.44	3.12	2.72
Apply at age 57 (%)	21.27	18.13	13.81
w/ mental disorders	2.69	2.85	2.46
w/ musculoskeletal system	11.49	8.78	6.55
w/ other disorders (%)	7.20	6.60	4.86
Apply ever after age 57 (%)	33.70	38.26	32.72
DI benefit receipt at age 60 (%)	77.36	69.67	65.09
Employment at age 60 (%)	7.39	10.53	13.33
Other benefit receipt at age 60 (%)	9.55	13.68	15.79
Avg. annual earnings best 15 years (Euro)	37,926	38,808	39,245
	(9,037)	(9,316)	(8,823)
Insurance years by age 50	29.11	29.59	29.94
	(7.07)	(6.97)	(6.49)
Employment years by age 50	13.60	13.67	13.69
	(1.99)	(1.94)	(1.93)
Was ever on sick leave by age 50 (%)	58.70	56.97	57.41
Blue-collar (%)	57.61	56.51	56.28
No. Observations	192,591	100,490	84,299
No. Individuals	6,282	3,304	3,128

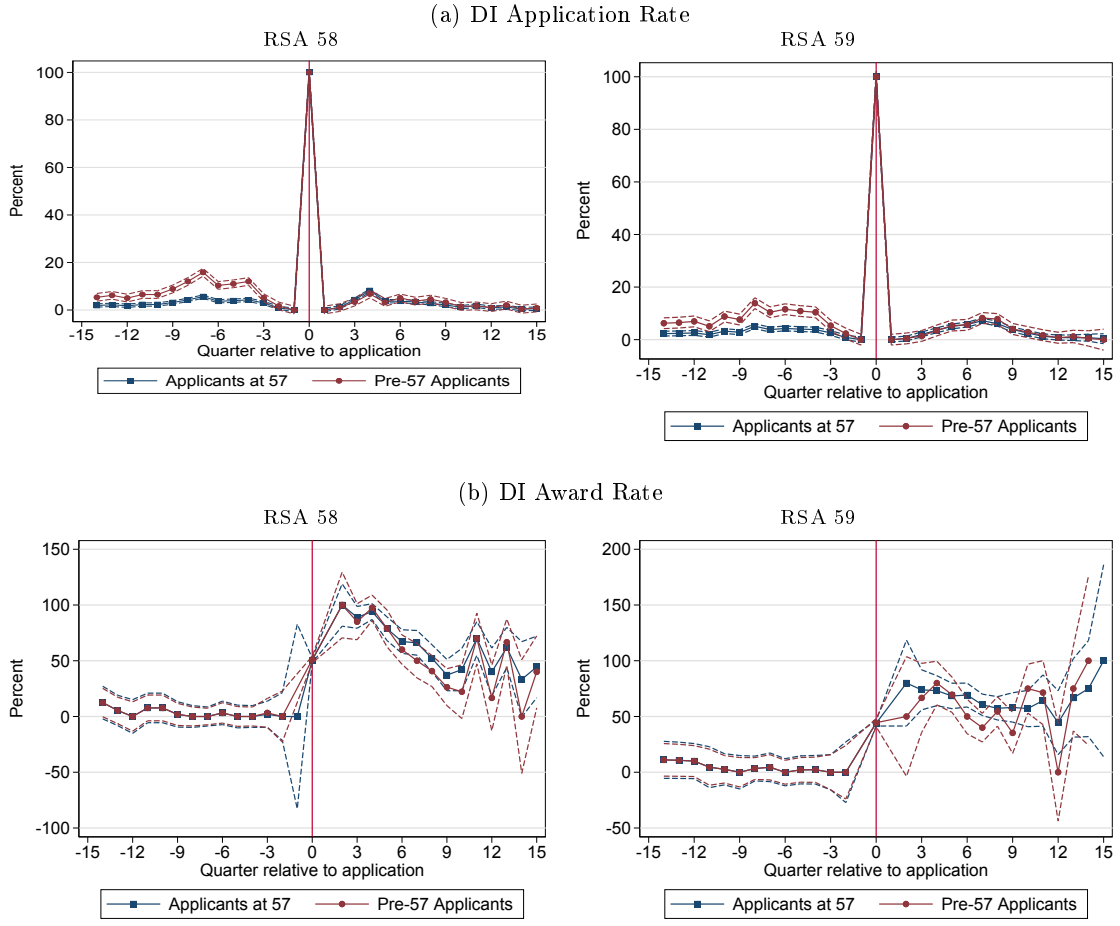
Notes: The table reports summary statistics for men between age 54 and age 62 who have applied for DI during ages 50-56 (pre-57 applicants). The RSA-57 cohort comprises men born between December 1953 and November 1955, the RSA-58 cohort comprises men born between December 1955 and November 1956, and the RSA-59 cohort comprises men born between December 1956 and November 1957. Earnings are reported in 2018 Euro. Sample standard deviations for continuous variables are reported in parentheses.

The DI applicants at age 57 in the whole population consist of (i) DI applicants who file an application for the first time and (ii) DI re-applicants whose pre-57 application got rejected (= those in the subpopulation of pre-57 applicants whose previous DI application was rejected and who re-apply at age 57). Hence, our comparison of pre-57 applicants who re-apply at age 57 to applicants at age 57 in the whole population boils down to comparing group (ii) to the sum of groups (i) and (ii). The share of pre-57 applicants among all applicants at age 57 in the RSA-58 sample is 35 percent, the share of group of pre-57 applicants in the RSA-59 sample is 37

percent. Hence, around one third of all applicants at age 57 already filed an application before age 57 and two thirds are first-time applicants.

Figure 4, Panel (a) in the main text shows that these two groups have very similar DI benefit receipt rates. Figure D.17 decomposes DI benefit receipt further into application rates and award rates and shows that the two groups are almost identical in these dimensions as well. The very similar award rates indicate that the pre-57 applicants seem to be representative for applicants at age 57.

Figure D.17: Applicants at 57 versus Pre-57 Applicants, Treated Cohorts Only

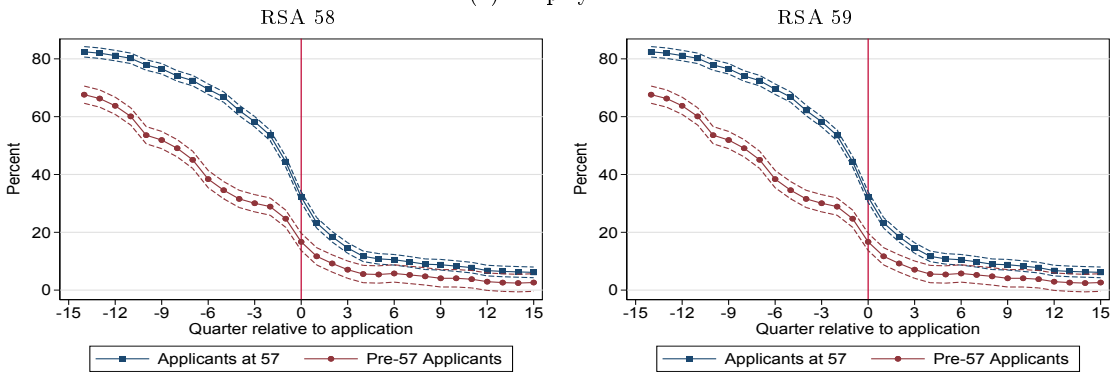


Notes: The figure compares trends in DI application rates (panel a) and the DI award rates (panel b) for applicants at age 57 (always applicants) and pre-57 applicants in the treatment groups. Always applicants are individuals who apply for for DI at age 57. Pre-57 applicants comprise individuals who applied for DI between age 50 and age 56 and re-apply for DI at age 57. The comparison shows that the two groups are very similar in outcomes after their application at age 57.

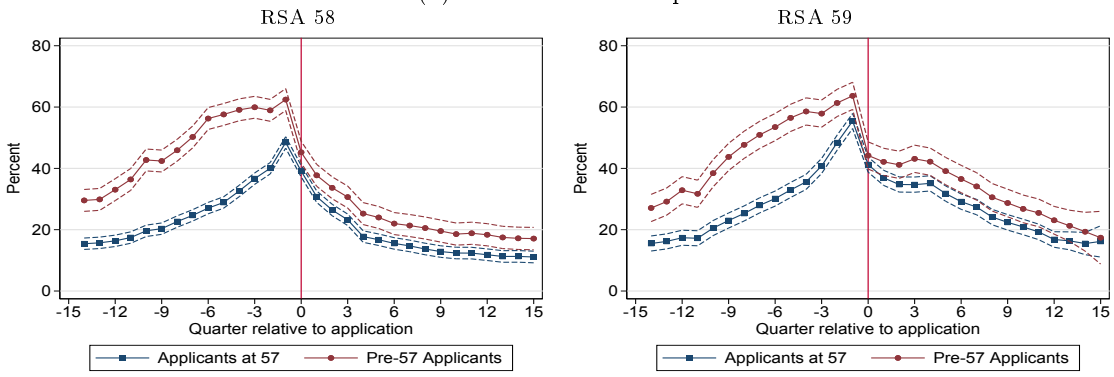
Figure D.18 compares the two groups with respect to employment, other benefit receipt, social security contributions and other benefit payments. The two groups show a similar pattern after their application at 57 but they are not identical. Applicants at age 57 have slightly higher employment and a lower chance to receive other welfare benefits than pre-57 applicants. This arises because first time applicants at age 57 have a higher employment probability than pre-57 applicants. The differences are, however, not large.

Figure D.18: Pre-57 Applicants versus Age-57 Applicants

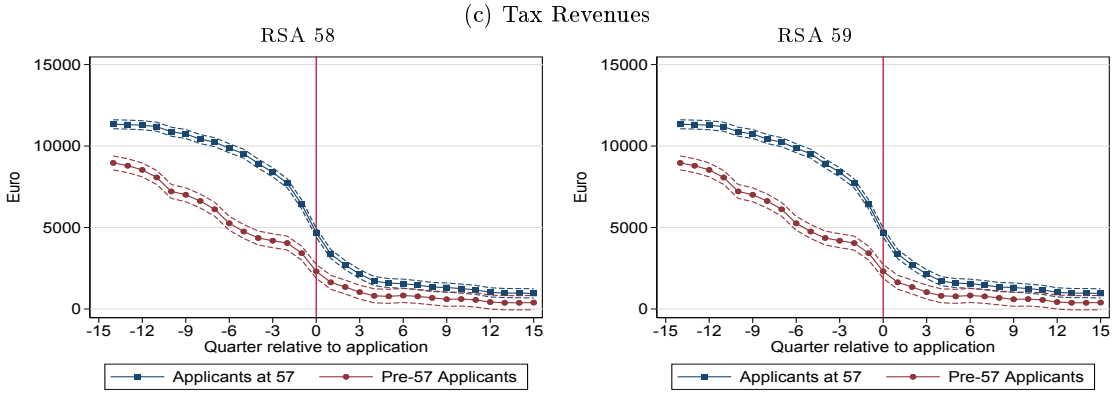
(a) Employment



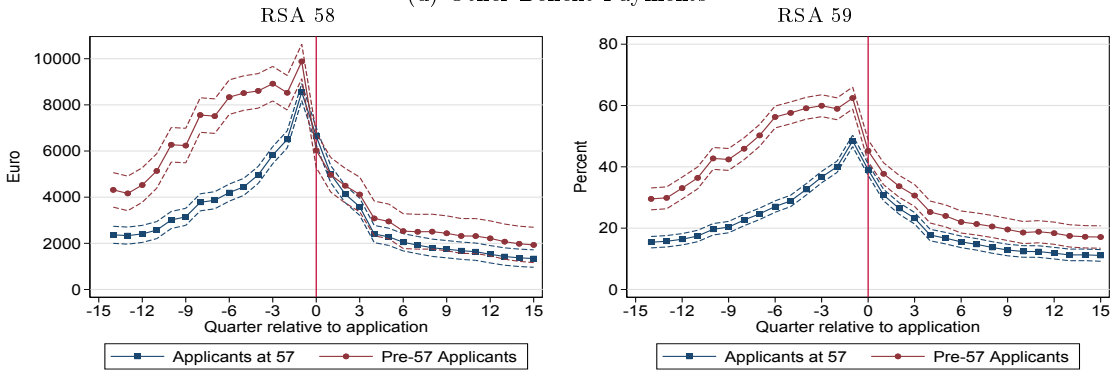
(b) Other Benefit Receipt



(c) Tax Revenues



(d) Other Benefit Payments



D.2 Insurance Loss of Stricter DI Eligibility Rules

Implementing the insurance value (the rhs of (10)) is associated with several challenges. In contrast to the sufficient statistics literature on UI, the utility loss is expressed in differences in utility levels rather than in marginal utilities. Moreover, the insurance value also depends on the abstract quantity θ. We tackle this challenges by deriving bounds of the insurance value that do not depend on the unobserved disability level θ. Furthermore, we assume utility is state-independent and CRRA, i.e. $v(c) = u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$, and that we have hand-to-mouth consumers.[74] In the following we discuss the derivation of the bounds and the implications of our assumptions. For this define the insurance value as

$$\Delta V \equiv \frac{1}{\lambda \mathbb{E}\left[M(\theta_s^*)\right]} \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{W_{i,t}} \left(v(c_{i,t}^D) - \left(u(c_{i,t}^W) - \theta_{i,t}\right)\right) + M_{Z_{i,t}} \left(v(c_{i,t}^D) - v(c_{i,t}^Z)\right)\right)\right] \tag{D.48}$$

In the following we derive lower and upper bounds on this insurance value ΔV.

Upper Bound Insurance Value. The social welfare benefits act as safety net. An agent cannot do worse than being on social welfare benefits in all periods. The insurance loss can therefore not be larger than

$$\Delta V \leq \frac{1}{\lambda \mathbb{E}\left[M(\theta_s^*)\right]} \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{W_{i,t}} + M_{Z_{i,t}}\right) \left(v(c_{i,t}^D) - v(c_{i,t}^Z)\right)\right]. \tag{D.49}$$

This bound assumes that individuals who are screened out are all on social welfare benefits. Individuals who decide to work at some points can only do better than being on social welfare benefits in all periods and hence experience a lower insurance loss than assumed by this bound.

Lower Bound Insurance Value. Since $M_{W_{i,t}} \left(v(c_{i,t}^D) - \left(u(c_{i,t}^W) - \theta_{i,t}\right)\right) \geq 0$ we have

$$\Delta V \geq \frac{1}{\lambda \mathbb{E}\left[M(\theta_s^*)\right]} \mathbb{E}\left[\sum_{t=s}^{T-1} \beta^t \left(M_{Z_{i,t}} \left(v(c_{i,t}^D) - v(c_{i,t}^Z)\right)\right)\right]. \tag{D.50}$$

This lower bound simply assumes that individuals who are screened out and then return to work have no loss in insurance value, i.e. they are indifferent between working and receiving DI benefits.

Implementation. To implement (D.49) we make four assumptions. First, we measure the insurance loss relative to an increase in resources during an employment spell ($\lambda = \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t u'(c_{i,t}^W)\right]$ where $c_{i,t}^W$ is the consumption level of working individuals). This is the standard to measure the insurance value in the UI literature. Second, we assume utility is state-independent and CRRA, i.e. $v(c) = u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$. Third, we assume individuals are hand-to-mouth and set consumption equal to current income ($c_{i,t}^D = b_{i,t}, c_{i,t}^Z = z_{i,t}, c_{i,t}^W =$

[74] We only observe transfers and incomes in our data and cannot measure consumption.

$w_{i,t} - \tau_{i,t}$). We assume this because we cannot observe consumption in our data. This assumption provides an upper bound on the insurance value. If individuals can self-insure through savings the insurance loss is smaller than if they were hand-to-mouth and simply consumed their income. Hence, in our implementation we tend to overestimate the insurance loss. Fourth, we assume no discounting $\beta = (1 + r) = 1$. All effects are within a 5 years horizon and hence discounting does not play a major role.

With these assumptions we have

$$\Delta V \leq \frac{1}{\mathbb{E}\left[\sum_{t=0}^{T-1} u'(w_{i,t} - \tau_{i,t})\right] \mathbb{E}\left[M(\theta_s^*)\right]} \mathbb{E}\left[\sum_{t=s}^{T-1} \left(M_{W_{i,t}} + M_{Z_{i,t}}\right) \frac{1}{1-\gamma} \left((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}\right)\right] \tag{D.51}$$

and

$$\Delta V \geq \frac{1}{\mathbb{E}\left[\sum_{t=0}^{T-1} u'(w_{i,t} - \tau_{i,t})\right] \mathbb{E}\left[M(\theta_s^*)\right]} \mathbb{E}\left[\sum_{t=s}^{T-1} M_{Z_{i,t}} \frac{1}{1-\gamma} \left((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}\right)\right]. \tag{D.52}$$

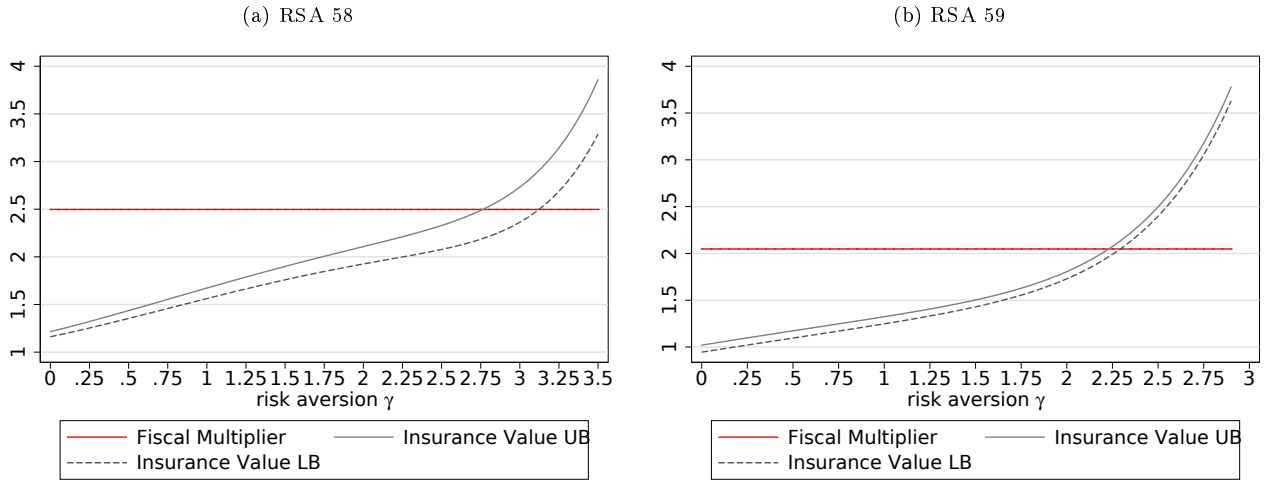
We calculate $\mathbb{E}\left[\sum_{t=0}^{T-1} u'(w_{i,t} - \tau_{i,t})\right]$ from the data for a given value of risk aversion. We then estimate the mechanical fiscal effect $\mathbb{E}\left[M(\theta_s^*)\right]$ in Section 6. $\mathbb{E}\left[\sum_{t=s}^{T-1} \left(M_{W_{i,t}} + M_{Z_{i,t}}\right) \frac{1}{1-\gamma} \left((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}\right)\right]$ in (D.51) only depends on the mechanical effect. We therefore use the same pre-57 applicants strategy as in the main text to estimate the mechanical fiscal effect. Here we just apply this strategy to a different outcome. For each individual we create a variable $q_{i,t}$ which is equal to the DI benefits $b_{i,t}$ if this individual is on DI benefits and equal to the individuals (hypothetical) social welfare benefits $z_{i,t}$ if this individual is not on DI benefits. This ensures that an individual who returns to work experiences a utility loss as if she was on social welfare benefits. We then calculate for a given risk aversion γ the utility $v_{i,t} = \frac{1}{1-\gamma}(q_{i,t})^{1-\gamma}$ and run our DiD strategy on this outcome variable $v_{i,t}$. Analogously to the mechanical fiscal effect this identifies the mechanical utility loss $\mathbb{E}\left[\sum_{t=s}^{T-1} \left(M_{W_{i,t}} + M_{Z_{i,t}}\right) \frac{1}{1-\gamma} \left((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}\right)\right]$.

To estimate $\mathbb{E}\left[\sum_{t=s}^{T-1} M_{Z_{i,t}} \frac{1}{1-\gamma} \left((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}\right)\right]$ in (D.52) we create a variable $l_{i,t}$ which is equal to the actual DI benefit $b_{i,t}$ if this individual is on DI, and equal to the hypothetical DI benefit $b_{i,t}$ if the individual is employed (= the DI benefit the individual would receive in case of a DI reward); and $l_{i,t}$ is equal to the individual social welfare benefits $z_{i,t}$ if this individual is on social welfare benefits. This ensures that an individual experiences no utility loss if she returns to work after being screened out (i.e. there is only a insurance loss if $M_{Z_{i,t}} = 1$). We then calculate for a given risk aversion γ the utility $u_{i,t} = \frac{1}{1-\gamma}(l_{i,t})^{1-\gamma}$ and run our DiD strategy on this outcome variable $u_{i,t}$.

Using this approach we estimate the upper and lower bound of the insurance loss for different values of risk aversion and plot the two bounds in Figure D.19. We find that shifting the RSA by one year is welfare-improving if risk aversion $\gamma < 2.8$ and it is welfare-reducing if $\gamma > 3.1$. Increasing the RSA by two years is

welfare-improving if risk aversion $\gamma < 2.2$. Estimates from the literature suggest that the coefficient of relative risk aversion is below 2, Chetty (2006b) finds an upper bound of $\gamma \leq 1.78$. Hence, our implementation implies that the increase in the RSA was welfare-improving for reasonable values of risk aversion.

Figure D.19: Stricter Eligibility Rules



Notes: Figure plots the LHS and the upper and lower bounds of the RHS of inequality (10) for the one year increase in the RSA from 57 to 58 in panel (a) and two year increase in RSA in panel (b) against different levels of risk aversion. If risk aversion is lower than the point where the solid gray line crosses the red line, then the reform is welfare improving. If risk aversion is higher than the point where the dashed gray line crosses the red line, then increasing the RSA is welfare reducing. For levels of risk aversion between these two points our sufficient statistics condition do not allow for a welfare statement.

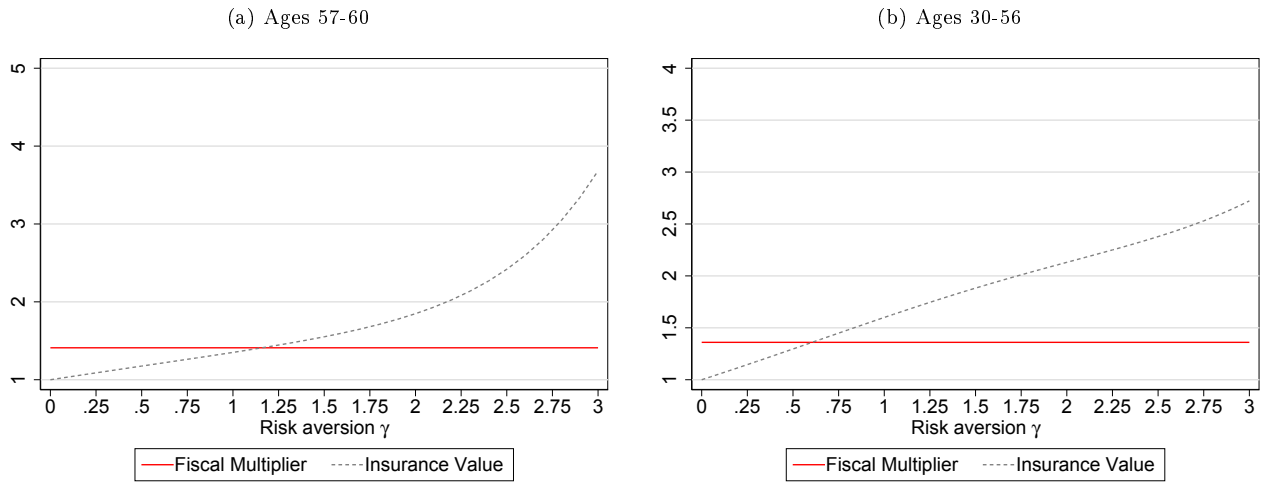
D.3 Insurance Loss of Lower DI Benefits

The effect we estimate empirically is a benefit reduction from age s to $T-1$. For the welfare effect this simply implies that we need to sum up the welfare effects of changing benefits in each period. To implement the welfare effects we impose the same four assumptions as in the above implementation for stricter eligibility rules. This yields for the insurance value

$$\frac{\mathbb{E}\left[\sum_{t=s}^{T-1}\left(b_{i,s}\right)^{-\gamma}\right]}{\mathbb{E}\left[\sum_{t=0}^{T-1}(w_{i,t})^{-\gamma}\right]}. \tag{D.53}$$

We can directly calculate this for different values of risk aversion based on the pre-reform benefit levels. Figure D.20 plots the fiscal multiplier and the insurance value for different values of risk aversion. We find that for risk aversion around $\gamma = 1.1$ the benefit levels are optimal for the age group 57-60. Younger individuals have lower multipliers with similar insurance values and hence a lower critical risk aversion level of around $\gamma = 0.6$. Hence, benefit generosity is optimal for reasonable values of risk aversion.

Figure D.20: Welfare Effects Benefit Generosity, Men

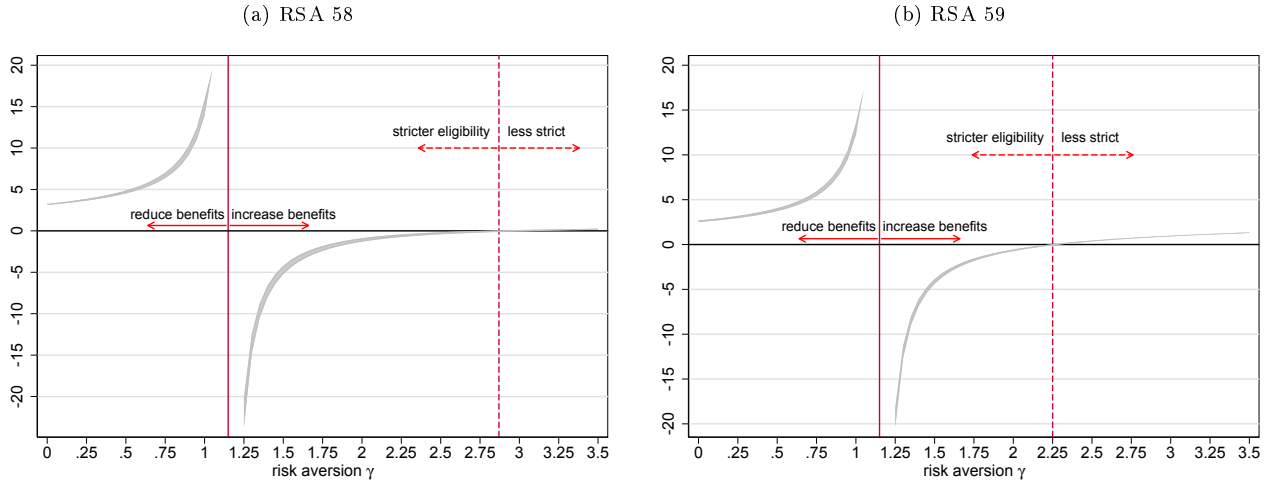


Notes: Figure plots the LHS and RHS of inequality (11) for men aged 57-60 in panel (a) and 30-56 in panel (b) against different levels of risk aversion. If risk aversion is higher than the point where the gray line crosses the red line, then it is welfare improving to increase benefit generosity. If risk aversion is lower than this point, it is welfare improving to reduce benefit generosity.

D.4 Optimal Policy Mix

For the gradient we can use the implementation from above to express γ and σ from equation (A.35) as a function of risk aversion. For γ we use the upper and lower bounds and therefore get a range of optimal directions for a given level of risk aversion. The optimal direction $\frac{\gamma}{\sigma}$ measures the direction in units of mechanical cost reductions. Intuitively, the gradient says that for a one dollar mechanical reduction in fiscal costs due to lower benefits, eligibility rules should be stricter such that $\frac{\gamma}{\sigma}$ dollars are saved mechanically. Figure D.21 plots the gradient for different values of risk aversion. Panel (a) plots the optimal combination of changing benefit generosity and eligibility rules at age 57. For risk aversion below 1.15 benefits should be reduced and eligibility rules should be stricter. For instance with risk aversion of 0.5 the optimal combination reduces spending through stricter eligibility rules by 4 dollars for a one dollar reduction in spending due to lower benefits. Hence, at this level of risk aversion DI eligibility is optimally tightened more than benefits. For levels of risk aversion above 2.8, benefits should be increased and eligibility rules should be less strict. In this region it is more effective to increase benefits than making eligibility less strict (optimal direction of around $\frac{d\theta^*}{db} \approx$0.1-0.2). For risk aversion between 1.15 and 2.8 eligibility criteria should be tightened but benefits more generous. Panel (b) plots the optimal direction for the adjustment of eligibility between 57 and 59. This gradient looks qualitatively similar to the gradient in Panel (a).

Figure D.21: Optimal Policy Mix: Gradient



Notes: Figure plots the optimal combination of changing benefit generosity and eligibility, i.e. the direction $\frac{\gamma}{\sigma}$ from equation (A.35). The red lines indicate the critical risk aversion values for optimal DI benefits and eligibility rules.

D.5 Relation between Insurance Loss and Income Loss

Stricter DI Eligibility Rules. Analogously to the discussion in Appendix A.2 consider a discrete change in eligibility rules in period s from θ_s^L to $\theta_s^H > \theta_s^L$. We can write the insurance loss of stricter eligibility rules as

$$\begin{aligned}
\Delta V_{\theta^*} &\equiv \int_i V_i(\theta_s^L) - V_i(\theta_s^H) di && \text{(D.54)} \\
&\leq \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \left(\Delta D_{i,t}^W \left(v_i(c_{i,t}^D) - \left\{u(c_{i,t}^W) - \theta_{i,t}\right\}\right) + \Delta D_{i,t}^Z \left(v_i(c_{i,t}^D) - v_i(c_{i,t}^Z)\right)\right)\right] && \text{(D.55)} \\
&\quad -\mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \Delta \Lambda_{i,t} \cdot \psi\right].
\end{aligned}$$

Expression (D.55) captures the insurance loss due to changes in the labor market status. $\Delta D_{i,t}^W$ and $\Delta D_{i,t}^Z$ denote the change in labor market status of individual i in period t from disability to employment and from disability to other welfare benefits, respectively. This change in disability status can arise because individual i is rejected (mechanical effect) or no longer applies (behavioral effect). $\Delta \Lambda_{i,t} \cdot \psi$ denotes the reduction in application costs of individual i due to changing her application decision in period t. Other behavioral changes, such as adaption of consumption and savings decision in anticipation of the stricter eligibility rules, would reduce the insurance loss. Therefore, (D.55) is an inequality. Note that for a marginal change in θ_s^* behavioral changes would not have a first order welfare effect due to the envelope theorem. Then $\Delta D_{i,t}^W$ and $\Delta D_{i,t}^Z$ would only account for the mechanical changes. In this sense, our relative comparison of the insurance losses is robust to non-marginal changes as $\Delta D_{i,t}^W$ and $\Delta D_{i,t}^Z$ also capture changes in DI levels due to changes in behavior.

We now further bound the insurance loss ΔV_{θ^*} and relate it to the income loss. We have

$$\begin{aligned}
\Delta V_{\theta^*} &\leq \mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\left(\Delta D_{i,t}^W\left(v_i(c_{i,t}^D)-\left\{u(c_{i,t}^W)-\theta_{i,t}\right\}\right)+\Delta D_{i,t}^Z\left(v_i(c_{i,t}^D)-v_i(c_{i,t}^Z)\right)\right)\right] && \text{(D.56)}\\
&\leq \mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\Delta D_{i,t}\left(v_i(c_{i,t}^D)-v_i(c_{i,t}^Z)\right)\right] && \text{(D.57)}\\
&\approx \mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\Delta D_{i,t}v_i'(c_{i,t}^D)\left(c_{i,t}^D-c_{i,t}^Z\right)\right] && \text{(D.58)}\\
&\leq \mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\Delta D_{i,t}v_i'(c_{i,t}^D)\left(b_{i,t}-z_{i,t}\right)\right] && \text{(D.59)}
\end{aligned}$$

where (D.56) simply drops the reduction in application costs, $\Delta\Lambda_{i,t}\cdot\psi\geq 0$, in (D.55). (D.57) uses that other welfare benefits act as a safety net, i.e. individuals who choose to work cannot be worse of than on other welfare benefits, and we define $\Delta D_{i,t}\equiv\Delta D_{i,t}^W+\Delta D_{i,t}^Z$. (D.58) follows from a first order Taylor approximation and (D.59) uses that the consumption drop cannot be larger than the income drop (savings dampen the income loss).

Therefore, we have an upper bound on the insurance loss given by

$$\frac{\Delta V_{\theta^*}}{\lambda\mathbb{E}\left[M_\Delta(\theta_s^*)\right]}\leq\mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t\underbrace{\frac{v_i'(c_{i,t}^D)}{\lambda}}_{\text{"welfare weight"}}\underbrace{\frac{\Delta D_{i,t}\left(b_{i,t}-z_{i,t}\right)}{\mathbb{E}\left[M_\Delta(\theta_s^*)\right]}}_{\text{bound normalized income loss}}\right]. \qquad \text{(D.60)}$$

In Figure 6 we estimate the bounds of the normalized income loss by income quintile. The income quintiles can be thought of as an approximation of the welfare weight in the above inequality.

Lower DI Benefits . Similarly to the above discussion consider a discrete change in DI benefits $b_{i,t}^H$ to $b_{i,t}^L<b_{i,t}^H$. We have

$$\begin{aligned}
\Delta V_b &\equiv \int_i V_i(b^H)-V_i(b^L)di && \text{(D.61)}\\
&\geq \mathbb{E}\left[\sum_{t=0}^{T-1}\beta^t D_{i,t}v_i'(c_{i,t}^D)\left(b_{i,t}^H-b_{i,t}^L\right)\right]. && \text{(D.62)}
\end{aligned}$$

Inequality (D.62) holds because changes in behavior lead to additional utility losses for non-marginal changes in DI benefits. As a lower bound for the insurance loss of reducing benefits we have

$$\frac{\Delta V_b}{\lambda \mathbb{E}\left[M_\Delta(b)\right]} \geq \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \underbrace{\frac{v_i'(c_{i,t}^D)}{\lambda}}_{\text{"welfare weight"}} \underbrace{\frac{D_{i,t}\left(b_{i,t}^H - b_{i,t}^L\right)}{\mathbb{E}\left[M_\Delta(b)\right]}}_{\text{"bound normalized income loss}}\right]. \tag{D.63}$$

Comparing Insurance Loss of DI Eligibility Rules and DI Benefits. From (D.60) and (D.63) follows that the insurance loss of stricter eligibility rules is smaller than the insurance loss of reduced benefits if

$$\mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \frac{v_i'(c_{i,t}^D)}{\lambda} \frac{\Delta D_{i,t}\left(b_{i,t} - z_{i,t}\right)}{\mathbb{E}\left[M_\Delta(\theta_s^*)\right]}\right] \leq \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \frac{v_i'(c_{i,t}^D)}{\lambda} \frac{D_{i,t}\left(b_{i,t}^H - b_{i,t}^L\right)}{\mathbb{E}\left[M_\Delta(b)\right]}\right] \tag{D.64}$$

holds. Figure 6 implements this inequality empirically by estimating the bounds of the income losses by income quintiles and we discuss the implications in the main text.

Income Losses of DI Eligibility Rules and DI Benefits. As an additional interesting exercise we also look at the (unbounded) income losses of the two policy instruments taking labor earnings into account for stricter eligibility and behavioral changes (like lower DI take-up) for the reduction in DI benefits. Figure D.22 displays the results of this exercise by income sources for the RSA-58 and RSA-59 increases. Panel (a) shows that mostly individuals in the lower three income quintiles are affected by stricter eligibility rules. For the RSA-58 increase, the lowest three quintiles loose on average around 1000 Euros in DI benefits between age 57 and 61. The DI benefit loss is similarly distributed for the two year increase to RSA 59 with an approximately twice as large magnitude. This loss in DI benefits is to some degree offset with higher labor income (Panel (b)) and more income from other benefits (Panel (c)). Panel (d) presents the net effect on total income. The higher income quintiles can offset their DI income loss to a large degree with higher labor income. Only the lowest quintile experiences a substantial loss. We do the same exercise for lower DI benefits in D.23. We find that for the income losses associated with lower DI benefits are more pronounced at the lower quintiles. These losses are only weakly compensated by other transfers, while labor earnings are not affected. There are also some losses due to lower DI income in the second and third income quintile though these losses are substantially smaller, while other income sources are largely unaffected by the the DI benefit cut.

In Figure D.24 we present the normalized income losses in each income quintile for the RSA increases and benefit generosity from the unbounded income losses from Figures D.22 and D.23.[75] Figure D.24 shows that for the (unbounded) income losses the conclusion is even clearer as for the bounded income losses. The normalized income loss of benefit generosity is for all quintiles larger than that of stricter eligibility rules (the red line is above the blue line, except for the third quintile where the losses are very similar).

If the objective function of the planner would not take disutility of work into account, i.e. $\theta_{i,t}$ would not

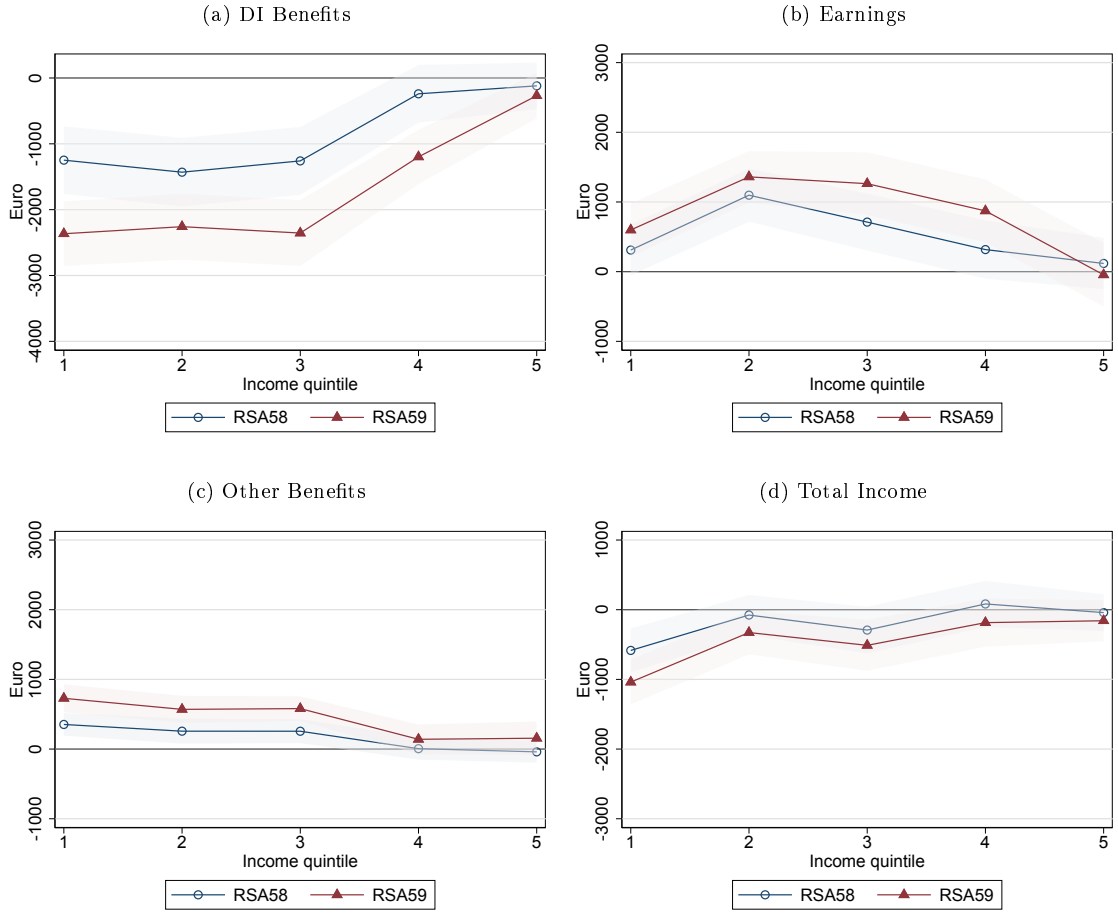
[75] The normalized income loss is simply the the total income loss by quintile in Panel (d) in Figures D.22 and D.23 divided by the respective mechanical fiscal effect from Table 4.

show up in (D.56), we would have the following relationship between the insurance loss and the normalized income loss

$$\frac{\Delta V_{\theta^*}}{\lambda \mathbb{E}\left[M_\Delta(\theta_s^*)\right]} \leq \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t \underbrace{\frac{v_i'(c_{i,t}^D)}{\lambda}}_{\text{"welfare weight"}} \underbrace{\frac{\left(\Delta D_{i,t}^W \left(b_{i,t} - (w_{i,t} - \tau_{i,t})\right) + \Delta D_{i,t}^Z \left(b_{i,t} - z_{i,t}\right)\right)}{\mathbb{E}\left[M_\Delta(\theta_s^*)\right]}}_{\text{normalized income loss}}\right].$$

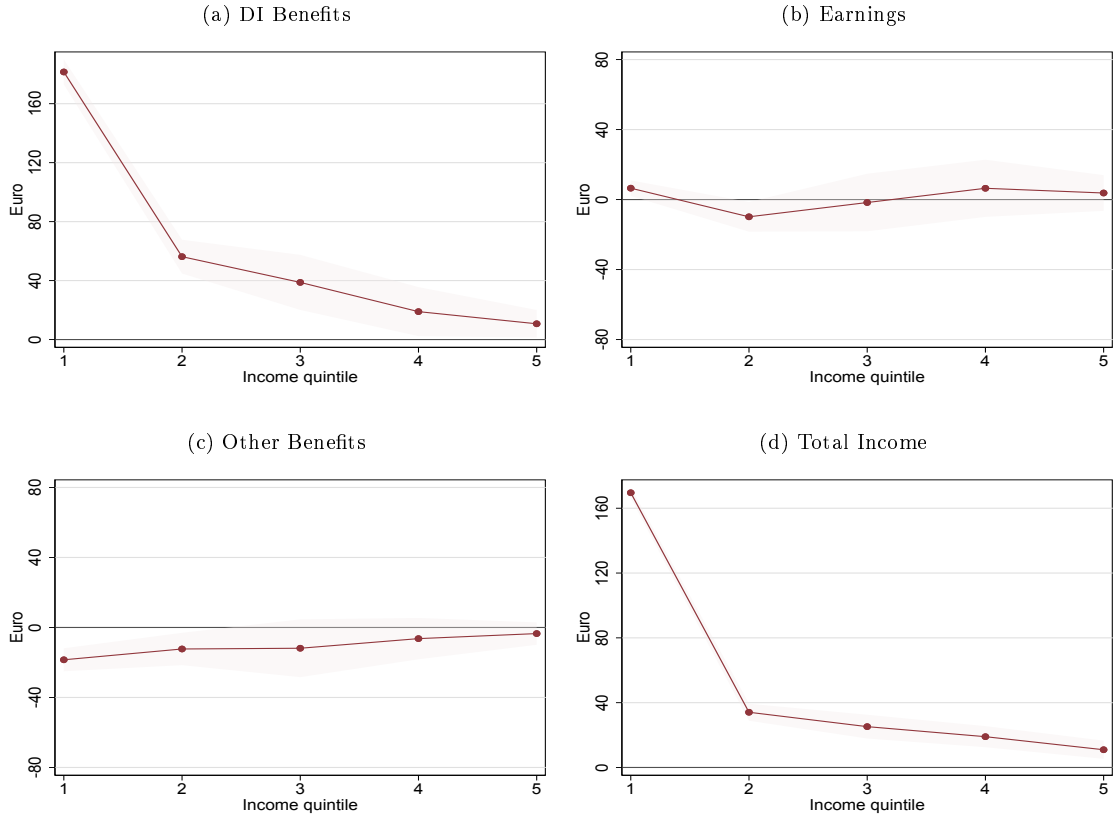
This bound corresponds to Figure D.24 where we estimate the normalized income loss (taking actual labor earnings into account instead of potential benefits as in (D.60)) by income quintile. Hence, a policy maker who is mainly concerned about income replacement therefore finds stricter eligibility rules less costly than reducing DI benefits (irrespective of distributional preferences). Such a policy maker clearly prefers tightening eligibility rules over reducing benefit generosity as it has lower costs (lower income losses across the income distribution) while it creates a higher fiscal cost reduction (higher multiplier).

Figure D.22: Effect of Stricter DI Eligibility Rules, by Income Bins



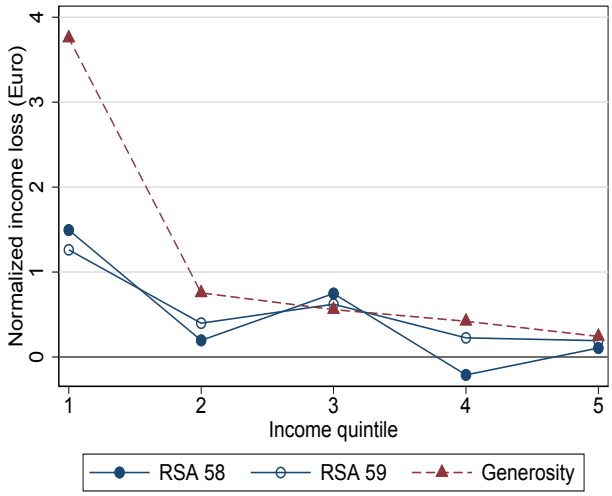
Notes: The Figure plots the estimated income losses in each income quintile for the RSA increases by income source. In contrast to Figure 6 we do not implement the bounds on the income loss here. That is, we do not replace labor earnings with potential welfare benefits for individuals resuming to work. Higher labor earnings offset a large part of the DI income loss for the upper four quintiles.

Figure D.23: Effect of benefit generosity by income bins



Notes: The Figure plots the estimated income losses in each income quintile for reduced benefit generosity by income source. In contrast to Figure 6 we do not implement the bounds on the income loss here. That means we also consider changes in behavior, like reduced DI take-up, that trigger changes in income.

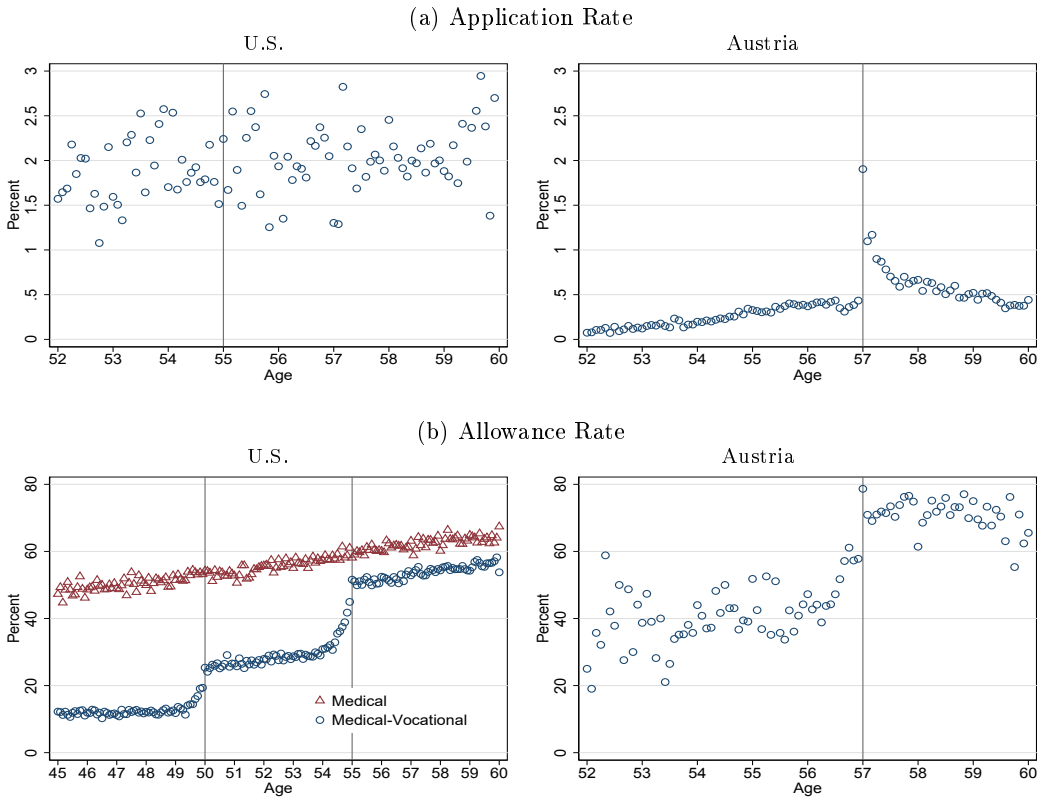
Figure D.24: Normalized Income Loss Eligibility Rules vs. Benefit Generosity



Notes: The figure plots the normalized income losses in each income quintile for the RSA increases and a reduction in benefit generosity. Income quintiles are measured at age 55. The normalized income loss is simply the total income loss by quintile in Panel (d) in Figures D.22 and D.23 divided by the respective mechanical fiscal effect from Table 4. The normalized income loss measures the quintile's income reduction for a 1 Euro mechanical reduction in fiscal spending.

D.6 Austria versus U.S.: DI Applications and DI Inflow by Age

Figure D.25: Application and Award Rate by Age: U.S. vs. Austria



Notes: Panel (a) contrasts the U.S. application behavior to the Austrian application behavior around the age cutoffs of relaxed eligibility rules. In Austria there is large spike of DI applications exactly at the RSA. In the U.S. there is no strategic bunching of DI applications at the age cutoffs. Panel (b) shows that in both countries there is a discontinuous jump in award rates at the age cutoffs.
Source: The U.S. Figures are based on Strand (2016).