DESIGNING DISABILITY INSURANCE REFORMS:
TIGHTENING ELIGIBILITY RULES OR REDUCING BENEFITS?*

Andreas Haller, Norwegian School of Economics and IZA
Stefan Staubli, University of Calgary and NBER
Josef Zweimüller, University of Zurich and CEPR

October 2020

Abstract

We study the welfare effects of disability insurance (DI) and derive social-optimality conditions for the two main DI policy parameters: (i) eligibility rules and (ii) benefit levels. 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. DI reforms are associated with income losses (earnings + transfers) of the affected individuals for both policy parameters, but these losses are lower when DI eligibility rules are tightened, particularly at the lower end of the income distribution. This suggests that the welfare cost of rolling back the Austrian DI program is lower through tightening eligibility rules than through lowering benefits.

Keywords: Disability insurance, screening, benefits, policy reform
JEL Codes: H53; H55; J14; J21; J65.

*Address: Andreas Haller, Norwegian School of Economics, 5045 Bergen, Norway; email andreas.haller@nhh.no. Stefan Staubli, University of Calgary, Calgary, AB T2N 1N4, Canada; email sstaubli@ucalgary.ca. Staubli is also affiliated with CEPR and IZA. Josef Zweimüller, University of Zurich, CH-8000 Zurich, Switzerland; email josef.zweimueller@econ.uzh.ch. Zweimüller is also affiliated with CESifo and IZA. 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 improving health, higher material living standards, and less physically demanding working conditions. The increasing financial burden of DI programs 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 new light on the welfare consequences of DI reforms. The paper contributes to the literature both on the theoretical and the empirical front. On the theoretical side, we provide a novel and rather general framework to explore the incentive-insurance trade-off associated with the two main DI policy parameters: DI eligibility rules and DI benefits. While optimal DI benefits have been studied in previous work, our analysis of optimal DI eligibility rules is new. Hence, our framework highlights the sufficient statistics needed to rank the two DI policy instruments with respect to their incentive costs and their insurance values. On the empirical side, we study the fiscal effects and the distributional implications of two Austrian DI reforms, exploiting exceptionally rich administrative register data. A novel contribution on the empirical side is our analysis of tighter DI eligibility rules, where existing empirical evidence is particularly scarce. Implementing our empirical results in the context of our theoretical framework suggests a clear ranking for Austrian DI policies: stricter DI eligibility rules dominate lower DI benefits as a policy tool for rolling back the Austrian DI program.

Our analysis comes in three steps. In the first step, 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. To assess the insurance losses associated with restrictive DI reforms, we derive, separately for each policy instrument, sufficient conditions for upper and lower bounds of the associated insurance losses. We show that these bounds can be expressed in terms of the income losses (earnings plus transfers) associated with each policy change.

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 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.

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 estimate the fiscal effect of always applicants (who still apply under stricter DI eligibility rules and drive the
mechanical effect). We cannot directly infer in the data who is an always applicant. 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. The relative size of fiscal multipliers suggests that stricter DI eligibility rules are more effective than lower DI benefits in reducing program expenditures. To estimate the insurance losses, we look at income losses associated with each policy instrument. We find that the income losses resulting from 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 – which provides us with lower and upper bounds for the insurance losses in terms of income losses – this implies lower insurance losses of stricter DI eligibility rules compared to lower DI benefits. Hence, stricter DI eligibility rules dominate lower DI benefits as a policy tool for rolling back the Austrian DI program.

The sufficient-statistics approach has several obvious and well-known limitations. A first limitation is that the welfare implications are drawn from reduced-form estimates, which apply only locally to the particular context. We think that our analysis of Austrian DI reforms is nevertheless of 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 younger applicants. Moreover, we can apply our framework to other contexts in a rather straightforward way. In light of our framework estimates from U.S. studies suggest a much lower fiscal multiplier of 1 for stricter DI eligibility rules compared to Austria with 2-2.5. For reduced benefits we find a similar fiscal multiplier of 1.4-1.5 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).\footnote{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 reduced-form methods.}

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, 2020). Our sufficient-statistics approach complements the literature which studies the incentive-insurance trade-off of DI programs using structural models (Benitez-Silva et al., 2004; Bound et al., 2010; Autor et al., 2019). Most closely related is the U.S. study by Low and Pistaferri (2015). They assess optimal DI benefits and DI eligibility criteria, and conclude that eligibility criteria are too strict and benefits too low in the U.S. We reach a similar conclusion using a very different approach. An interesting difference between our approach and their structural approach is that we do not need to estimate type-I and type-II errors (false rejections and false acceptances) in the screening process. 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 research question. A more structural app-
proach, as in Low and Pistaferri (2015; 2019), is able to directly estimate the type-I and type-II errors at the cost of more structural assumptions.\footnote{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.}

Our paper also relates to a theoretical literature on optimal disability insurance. We generalize and extend the model of Diamond and Sheshinski (1995) by expressing the social optimality conditions of DI eligibility rules and DI benefit levels as a set of sufficient statistics, which we can estimate empirically using program evaluation methods. Our approach is a local evaluation of existing DI systems and does not analyze the optimal system from a mechanism design perspective like Golosov and Tsyvinski (2006). 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 important as the discussion about DI reforms focuses on whether individuals are truly eligible for DI benefits.

Another strand of the DI literature uses reduced-form methods to estimate the labor market impact of DI programs without considering welfare effects. We add to this literature by providing novel evidence on the causal effect of DI eligibility rules. In contrast to previous work, which primarily relies on aggregate (state-level) evidence, our empirical analysis uses register data from the universe of social security records and the universe of DI applications. This combination allows to shed new light on the impact of DI reforms on DI applications, DI inflow as well as earnings and enrollment in other social insurance programs.\footnote{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), by exploiting 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

5
who find that relaxed eligibility rules and higher DI replacement rates explain the stark growth of DI rolls in the U.S. and lead to a lower unemployment rate. Parsons (1991) and Gruber and Kubik (1997) exploit variation in DI rejection rates across U.S. states over time and find that higher rejection rates reduce DI applications and increase labor force participation. We find similar evidence for self-screening in response to stricter eligibility rules, i.e. a decline in applications, and show that stricter eligibility rules target healthier individuals.\(^4\) Another set of studies explore the effects of DI benefit levels on 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. Section 6 estimates the fiscal multipliers of these two policy instruments and discusses how our estimates can be used for welfare evaluation. Section 7 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.\footnote{By increasing the “strictness of DI eligibility rules” we mean any policy making it more difficult that a DI application for a given disability gets accepted. This is what Low and Pistaferri (2015) and Diamond and Sheshinski (1995) call, respectively, “strictness of screening” and “disability standard”. The formal definition of strictness is discussed in section 2.1.} Section 2.1 starts with the static framework of Diamond and Sheshinski (1995) and Section 2.2 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 \( \tau \) (which finances the DI program) and enjoys utility \( u(w - \tau) \). There are no savings or any other choices in the first period.\footnote{The setup follows Chetty (2006a) who reconsiders Baily’s (1978) formula of optimal unemployment insurance (UI). The stylized two-period framework simplifies the formula without changing the substance of the argument.} In the second period, the agent suffers a disability shock \( \theta \), modeled as a random draw from a continuous distribution \( F(\theta) \). If \( \theta \) is small (=the disability is not very severe), the agent continues working and enjoys second-period utility \( u(w) - \theta \). If \( \theta \) is sufficiently large (= the disability is severe), the agent applies for DI benefits. A DI application causes disutility \( \psi \), capturing the extensive medical checks, the bureaucratic hassle, etc. associated with the DI assessment process. With probability \( p(\theta) \) the application is accepted, where \( p'(\theta) > 0 \). 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 welfare benefits \( z < b \) and gets second-period utility \( v(z) - \psi \). (No disutility or uncertainty are associated with claiming welfare benefits.)

**DI Applications and Labor Supply.** An agent prefers working over claiming welfare benefits if her disability is \( \theta < \theta^R \equiv u(w) - v(z) > 0 \), i.e. the utility of claiming welfare benefits falls short of the utility of working. Hence, \( \theta^R \) denotes the “marginal welfare benefit claimant”. The “marginal applicant,” the agent who is indifferent between filing
a DI application and remaining employed, has disability

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

Agents with disability $\theta \geq \theta^A$ apply for DI, while agents with disability $\theta < \theta^A$ remain employed. Panel (a) of Figure 1 characterizes the outcome of an agent’s DI application choice. It draws the probability of a DI award $p(\theta)$ against $\theta$ and indicates the disability cutoff-levels $\theta^A$ and $\theta^R$. 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 welfare benefits.\(^7\)

**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 $\theta^*$ needs further discussion. The inherent problem of the DI assessment process is that the true disability $\theta$ 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, $\theta$ is the applicant’s true disability and $e(\theta)$ is the noise. 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 $\theta^*$, 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 $p(\theta; \theta^*)$.\(^8\) In what follows, we consider the case where the government can change $\theta^*$ but takes the signal as given. This is the context of our empirical analysis, which exploits variation across

---

\(^7\)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 the Supplementary Material S.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 in 2.2.

\(^8\)We assume that the DI assessment process is informative, so that the award probability is increasing in the severity of the disability, or $\partial p(\theta; \theta^*)/\partial \theta \geq 0$. This implies that in an applicant pool with more severely disabled a smaller fraction of DI assessments falls short of the cutoff $\theta^*$. 

8
cohorts in the “relaxed screening age” (RSA); the age 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 $\theta^H$, while the control cohort is subject to the lenient standard $\theta^L$. 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 of an increase in $\theta^*$ on the outcomes of interest.

**Welfare Effects of DI Reforms.** We follow the literature assuming society’s objective can be represented by a utilitarian social welfare function. Abstracting from discounting, the social welfare function for a population of mass unity is given by

$$W (\theta^*, b) = u(w - \tau) + \int_{\theta^L}^{\theta^A} (u(w) - \theta) dF(\theta) + \int_{\theta^L}^{\theta^R} (1 - p(\theta; \theta^*)) (u(w) - \theta) dF(\theta) + \int_{\theta^L}^{\theta^A} p(\theta; \theta^*) v(b) dF(\theta) + \int_{\theta^L}^{\theta^R} (1 - p(\theta; \theta^*)) v(z) dF(\theta) - \int_{\theta^L}^{\theta^R} \psi dF(\theta). \quad (2)$$

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); the working healthy (second term), the rejected DI applicants resuming work (third term); the DI recipients (fourth term); and the welfare benefit recipients (fifth term). The last term captures 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 welfare benefit payments have to be covered by the taxes raised in the first period,

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

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 $\theta^*$.** The utilitarian government sets DI eligibility rules $\theta^*$ 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 $\theta^*$ is

$$\frac{\partial W}{\partial \theta^*} = u'(w - \tau) [B(\theta^*) + M(\theta^*)] - \left[ (v(b) - (u(w) - \bar{o})] M_w + [v(b) - v(z)] M_Z \right]$$

(4)

Condition (4) highlights the two opposing effects of stricter DI eligibility rules $\theta^*$ on social welfare. On the one hand, a higher $\theta^*$ raises social welfare because it saves taxpayers money (fiscal cost reduction). On the other hand, a higher $\theta^*$ 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 being 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 is accepted, $Pr(DI) = Pr(Apply) \ast Pr(Accept|Apply)$. In the above notation, the application probability is $Pr(Apply) = 1 - F(\theta^A)$, while the acceptance probability is $Pr(Accept|Apply) = \left[ \int_{0}^{\infty} p(\theta; \theta^*) dF(\theta) \right] / [1 - F(\theta^A)]$. The derivative of the application probability with respect to $\theta^*$ 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 $B(\theta^*)$. The derivative of the acceptance probability with respect to $\theta^*$ equals $-\int_{0}^{\infty} (\partial p(\theta; \theta^*) / \partial \theta^*) dF(\theta)$, which is the sum of the gray and the blue area in Panel (b) of Figure 1. The gray area captures
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 $\theta$ 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 $\theta$. 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 $\theta^A$ and individuals with $\theta \geq \theta^A$ apply to DI. The marginal welfare benefits type is denoted by $\theta^R$ 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).

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 welfare benefits$M_Z \equiv -\int_{\theta^A}^{\infty} (\partial p(\theta; \theta^*)/\partial \theta^*)dF(\theta)$. 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 $\theta^*$ does not only save money to taxpayers, it also reduces the insurance value of the DI program. 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 welfare benefits, $M_Z$. The average utility loss of rejected applicants who go back to work is $v(b) - (u(w) - \tilde{\theta}) > 0$, where $\tilde{\theta} \equiv \int_{\theta^A}^{\theta^R} (\partial p(\theta; \theta^*)/\partial \theta^*)\theta dF(\theta)/ \int_{\theta^A}^{\theta^R} (\partial p(\theta; \theta^*)/\partial \theta^*)dF(\theta)$ is their average disability level. The utility loss of rejected applicants who claim welfare benefits is $v(b) - v(z) > 0$. Note the reduction in the insurance value depends only on the mechanical effect but not...
on the behavioral effect; a direct implication of the Envelope theorem.

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

\[
\frac{\partial W}{\partial \theta^*} \geq 0 \iff 1 + \frac{B(\theta^*)}{M(\theta^*)} \geq \frac{L_W + L_Z}{u'(w - \tau)M(\theta^*)},
\]

where \( L_W \equiv [v(b) - (u(w) - \bar{\theta})]M_W > 0 \) 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 welfare benefits \( (L_Z) \). The two sides of the inequality have an intuitive interpretation. The left-hand-side is the fiscal multiplier, \( 1 + B(\theta^*)/M(\theta^*) \). It 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 measures the 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 \( \theta^* \).** The welfare implications of stricter DI eligibility rules, as summarized in condition (5), hold true for a marginal increase in \( \theta^* \). However, our empirical application exploits an RSA increase from age \( R \) to some higher age \( R + \Delta \), implying a discrete increase in the DI eligibility standard from an initially lenient standard \( \theta^L \) to a strict standard \( \theta^H \) during the age window \( [R, R + \Delta] \). Appendix A.1 shows that, for a discrete increase in \( \theta^* \), we need to rewrite condition (5) as

\[
\Delta W \geq 0 \iff 1 + \frac{B_\Delta(\theta^*)}{M_\Delta(\theta^*)} \geq \frac{L_{W\Delta} + L_{Z\Delta}}{u'(w - \tau_\Delta)M_\Delta} + \frac{L_{M\Delta}}{u'(w - \tau_\Delta)M_\Delta},
\]

where the subscript \( \Delta \) highlights that the corresponding effect has been generated by a discrete change in \( \theta^* \). The main takeaway is that the fiscal multiplier, \( 1 + B_\Delta/M_\Delta \), still captures the fiscal cost reduction. Note that the mechanical effect \( M_\Delta \) only depends on the always applicants (those who still apply under the stricter rules) as illustrated in
Figure A.1 In the empirical analysis we implement this non-marginal fiscal multiplier. However, the insurance losses of marginal applicants (those abstaining from a DI application under the now stricter rules) can no longer be ignored. These 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

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

(7)

A lower $b$ reduces the financial burden for taxpayers, because fewer agents apply for DI benefits. The associated welfare-gain is captured in condition (7) by the fiscal multiplier, $1 + B(b)/M(b)$. But a lower $b$ also creates an insurance loss for disabled workers, because it reduces their consumption possibilities when hit by a disability shock. The associated welfare-loss is captured in condition (7) by the ratio of the marginal utility of a DI benefit recipient relative to the marginal utility of a taxpayer.$^9$ A reduction in DI benefits is therefore welfare-improving if the fiscal gain exceeds the insurance loss, and vice versa.

### 2.2 The General Model

The static model highlights the basic trade-offs but misses two important features for evaluating DI reforms: heterogeneity across individuals and intertemporal choices. In the static model, agents differ only in $\theta$ and all actions happen within one period. In what follows, we allow for multiple sources of heterogeneity, such as wage heterogeneity, and we extend the model to multiple periods. In our empirical setting, where we exploit an RSA increase from $R$ to some higher age $R + \Delta$, the question “when

$^9$As Appendix A.1 shows, the ratio of behavioral over mechanical fiscal effect is equal to $\eta$, the elasticity of the DI inflow with respect to the DI benefit level. Hence, condition (7) can also be written as $1 + \eta = v'(b)/u'(w - \tau)$, which corresponds to the Bailey-Chetty formula for the optimal UI benefit level with $\eta$ being the elasticity of the unemployment duration with respect to the UI benefit level.
should I apply?” becomes crucial. Having a dynamic framework allows us to capture the intertemporal nature of the DI application choice.

**Agents’ Choices and Social Welfare.** Assume that an 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 wage/productivity shocks) 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 welfare benefits. 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 \( t \). 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.

The utilitarian government chooses DI policy parameters in each period \( P = (\theta^*_0, \ldots, \theta^*_{T-1}; b_0, \ldots, b_{T-1}) \) to maximize welfare \( W(P) \), defined as the aggregated indirect lifetime utility of all agents, subject to a budget constraint. The model environment allows for rich dynamics and heterogeneity across agents. We formally discuss the model environment in Appendix A.2.

**Stricter DI Eligibility Rules.** We now explore the welfare effects of a marginal change in the strictness of DI eligibility rules \( \theta^*_s \), while leaving all other elements of the DI policy vector \( P = (\theta^*_0, \ldots, \theta^*_{T-1}; b_0, \ldots, b_{T-1}) \) unchanged. 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 \( \theta^*_t \) 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 \( \theta^*_s \) but unchanged values of \( \theta^*_{t \neq s} \). In Appendix A.2, we show that \( \partial W(P) / \partial \theta^*_s \gtrsim 0 \) is equivalent to

\[
1 + \frac{\mathbb{E}[B(\theta^*_s)]}{\mathbb{E}[M(\theta^*_s)]} \gtrsim \frac{\mathbb{E}[L_W] + \mathbb{E}[L_Z]}{\lambda \cdot \mathbb{E}[M(\theta^*_s)]},
\]

(8)
The left-hand-side is the fiscal multiplier of increasing $\theta^*_s$, where $\mathbb{E}[M(\theta^*_s)]$ is the mechanical fiscal effect and $\mathbb{E}[B(\theta^*_s)]$ is the behavioral fiscal effect. On the right-hand side, $\mathbb{E}[L_W] + \mathbb{E}[L_Z]$ are the dynamic utility losses arising from fewer agents being admitted to the DI program in period $s$, and $\lambda$ denotes the Lagrange multiplier on the government’s budget constraint. The Lagrange multiplier measures the value to society of relaxing the government budget constraint and normalizing by it yields the money-metric of the utility losses. In Appendix A.2, we make explicit how $\mathbb{E}[B(\theta^*_s)]$, $\mathbb{E}[M(\theta^*_s)]$, $\mathbb{E}[L_W]$ and $\mathbb{E}[L_Z]$ are determined.

Notice the similarity of the social optimality condition (8) to the social optimality condition (5) of the simple static framework. A key difference between the static and the general model is that an increase in $\theta^*_s$ – stricter DI eligibility rules at some age $s$ – does not only affect the DI inflow at age $s$, but also at other ages. The behavioral fiscal effect of an increase in $\theta^*_s$, $\mathbb{E}[B(\theta^*_s)]$, 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, $\mathbb{E}[M(\theta^*_s)]$, 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 $\mathbb{E}[M(\theta^*_s)]$ spreads out over the age window $[s,T-1]$. In Section 6, we 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 $\theta^*$.** The RSA increase, we exploit empirically to identify the impact of stricter DI eligibility rules, translates to a discrete (rather than a marginal) increase in $\theta^*$. In Appendix A.2, we show that the logic from the static model for a discrete increase in $\theta^*$ also applies in the general model. The left-hand side of the social optimality condition remains a fiscal multiplier, where the mechanical fiscal effect measures the fiscal effect of always applicants (those who still apply under strict rules). We estimate the fiscal multiplier in Section 6. The right-hand side includes an additional term, which measures the insurance loss of marginal applicants. This insurance loss can be ignored for a marginal increase in $\theta^*$ because of the envelope theorem.
Lower DI Benefits. Alternatively, a 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 Appendix A.2, we show that that condition
\[-\frac{\partial W(P)}{\partial (b_s)} \geq 0\]
is equivalent to
\[1 + \frac{\mathbb{E}[B(b_s)]}{\mathbb{E}[M(b_s)]} \geq \lambda \cdot \mathbb{E}[\nu'(c^D_s)]\]
where \( \mathbb{E}[B(b_s)] \) and \( \mathbb{E}[M(b_s)] \) are the behavioral and mechanical fiscal effects of a marginal reduction of \( b_s \). Again, this optimality condition looks very similar to the one in the static model.

3 Institutional Background and Data

3.1 Institutional Background and Policy Variation

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.\(^\text{10}\) 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. Once benefits are awarded, DI beneficiaries receive monthly payments until their return to work, medical recovery or death.

\(^{10}\)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.
DI benefits can be granted for a temporary period, but less than 4 percent of claimants ever leave the DI rolls.

**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 than 50% of the earnings capacity of a healthy person in any reasonable occupation the individual could be expected to carry out. 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.

The RSA was 57 until the end of 2012, but it was increased in three one-year steps to age 60 as part of the 2. Stability Act (2. Stabilitätsgesetz).\(^{11}\) We exploit the variation in the RSA to identify the labor market effects of stricter DI eligibility rules (Section 4). The 2. Stability Act was announced in April 2012 and had two objectives: reduce expenditures in the public pension systems and foster employment among older workers. The only change to the DI program was the stepwise increase in the RSA 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. 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.

**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

\(^{11}\)Staubli (2011) studies the labor market effects of an earlier increase in the RSA for men from 55 to 57 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.
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.

In January 2004, the Austrian government implemented several changes to the calculation of DI benefits as part of a larger reform (Pensionsreform 2003). We use this exogenous variation to identify the labor market effects of changes in benefit levels (Section 5). The changes were phased in between 2004 and 2017 and lowered potential benefit levels for most individuals, by reducing the pension coefficient and increasing the penalty for claiming benefits before the normal retirement age. The reform also gradually increased the length of the assessment basis from 16 years to 40 years. 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. The large 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 5 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.\footnote{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 Figure T.10 in the Supplementary Material illustrates the effect of the reform by showing the distribution of changes in potential DI benefits between 2004 and 2017.}

12
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 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.\(^{13}\) 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 conditions.

\(^{13}\)Staubli and Zweimüller (2013) show that this increase had sizeable employment and unemployment effects. Men in our sample were not affected by this increase; their eligibility age for an old age pension was always age 62.
eligibility criteria above the RSA.

To study the effect of stricter DI eligibility, we focus on 54-61 year old men who are born between 1954 and 1957. We split the sample into men with more and less than 10 employment years in the past 15 years (measured at age 56). Only men with more than 10 employment years in the past 15 years are considered eligible for relaxed DI eligibility.\footnote{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.} We will use the sample of eligible men for our main effects and the sample of ineligible men for placebo tests. Individuals are observed on the 1\textsuperscript{st} of March, June, September, and December in each year.

To study the effect of lower DI benefits, we focus on 30-60 year old men during the time period 2004-2017. 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.\footnote{Tables T.4 and T.5 in the Supplementary Material show summary statistics for the RSA and benefit generosity samples.}

4 The Effect of Tighter DI Eligibility Rules

4.1 Estimation Strategy

We exploit the policy-induced variation in the RSA across birth cohorts in a difference-in-differences design. The RSA is 57 for men who turn 57 before December 2012 (those born before December 1955). We label this cohort the RSA-57 cohort.\footnote{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.} Conversely, the RSA is 58 for men who turn 57 between December 2012 and November 2017.
ber 2013 (those born after November 1955 and before December 1956). We label this cohort the RSA-58 cohort. Finally, the RSA is 59 for men who turn 57 after November 2013 (those born after November 1956). We label this cohort the RSA-59 cohort.\footnote{Figure U.11 in the Supplementary Material illustrates the step-wise increase graphically and Figure U.12 provides descriptive evidence on the labor market effects of the RSA increases.}

The RSA increases imply that the strictness of DI eligibility rules varies at certain ages across birth cohorts. The RSA-57 birth cohort – our control cohort – is eligible to the more lenient DI eligibility rules already at age 57, while the RSA-58 and RSA-59 birth cohorts – our treated cohorts – are eligible only at age 58 and age 59, respectively. Thus, we can identify the effect of stricter DI eligibility rules by comparing the age profiles of treated and control birth cohorts. This comparison can be implemented by estimating regressions of the following type:

\[
y_{ict} = \alpha + \theta_a + \pi_c + \lambda_t + \sum_{k=54,56}^{61} \beta_k \cdot T_{ic} \cdot I[\text{age} = k] + X'_{ict} \delta + \epsilon_{ict},
\]

where \(i\) denotes individual, \(c\) denotes the year and month of birth, and \(t\) denotes the year and quarter of calendar time; \(y_{ict}\) is the outcome variable of interest (such as an indicator for receiving DI benefits), \(\theta_a\) are age-in-years fixed effects to control for age-specific levels in the outcome variable, \(\pi_c\) are year-month of birth fixed effects to capture time-constant differences across birth cohorts, \(\lambda_t\) are year fixed effects to capture common time shocks and quarter fixed effects to capture seasonal effects, and \(X_{ict}\) represent individual or region specific characteristics to control for any observable differences that might confound the analysis.\footnote{The fixed effects are identified, because we have quarterly reference dates, so that we have variation in the age-in-years (\(\theta_a\)) conditional on year-month of birth (\(\pi_c\)) and calendar year (\(\lambda_t\)). For example, men born in July 1955 are age 56 in March and June 2012 and age 57 in September and December 2012.} We cluster standard errors at year-month of birth and state of residence.

The key variables of interest are the indicators \(T_{ic} \cdot I[\text{age} = k]\), which are equal to one if an individual belongs to a treated cohort \((T_{ic} = 1)\) and the age is equal to \(k\), where \(k\) runs from 54 to 61 using \(k = 56\) as the reference age. Each \(\beta_k\)-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 $\beta_k$-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 cohort. This way we can directly compare the effects of a one-year and a two-year RSA-increase. The identification assumption is that, absent the increase in the RSA, the change in $y_{ict}$ across age would have been comparable between treated and control birth cohorts. A potential concern is that age-specific trends in the outcome variable could change across birth cohorts for reasons unrelated to the RSA increase. The estimated $\beta_k$-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 (10) for men with less than 10 employment years in the past 15 years who are not eligible for the lenient DI eligibility rules. They should not respond to the changes in the RSA.

4.2 Empirical Results

For brevity, we focus on the one year increase of the RSA (the RSA-58 cohort) in the main text. Appendix B.1 shows the same set of results for the RSA-59 cohort. Figure 2 shows the estimated $\beta_k$-coefficients from equation (10) for four key outcomes: DI benefit receipt, DI application ever, employment, and other benefits.\textsuperscript{19} 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. Panel (a) shows that fewer men receive DI benefits between age 57 and age 61. DI recipiency rates drop by about 4 percentage points at age 57 and remain

\textsuperscript{19} DI benefit receipt is an indicator for whether an individual is receiving DI benefits, DI application ever is an indicator for whether an individual has ever applied for 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.
lower after age 57, even though DI eligibility rules have become more lenient. If applying for DI imposes costs, we would expect that fewer people apply when eligibility criteria are strict. Indeed, panel (b) shows that DI application rates drop at all ages above 56.\textsuperscript{20} Panels (c) and (d) show that stricter DI eligibility rules increase employment and other benefit receipt above age 56.\textsuperscript{21} The expansion in employment persists until the last age we can observe in the data, while the rise in other benefit receipt is temporary.

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

(a) DI Benefit Receipt  
(b) DI Application Ever

(c) Employment  
(d) Other Benefit Receipt

Notes: The figure shows the estimated $\beta_k$ coefficients from the econometric specification in (10) for the RSA-58 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

It is interesting to look at the timing and dynamics of the estimates in Figure 2. First, the DI application rate falls at age 57, implying that individuals are aware of the RSA

\textsuperscript{20}Supplementary Material U.2 provides more evidence on DI applications.  
\textsuperscript{21}Supplementary Material Figure U.15 shows that the expansion in employment is primarily driven by individuals who are already employed and who stay employed longer. Similarly, the increase in other benefit receipt is primarily driven by individuals who are already receiving other benefits (Figure U.16).
and adjust their behavior. If the estimated effects were purely mechanical, applications at age 57 should not react. Second, 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.

We can estimate the average effect of tighter DI eligibility rules between age 57 and age 61 by taking the average of the $\beta_k$-coefficients over these ages. Table 1 reports the average effects for the outcomes from Figure 2 as well as the corresponding fiscal effects, 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 fiscal effect – 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. Table 1 shows that tighter DI eligibility rules lessen spending on DI benefits (884 Euro per individual and year) and raise tax revenues from increased work activity (263 Euro), but they also raise spending on other benefits because of benefit substitution (172 Euro). Overall, total fiscal costs at each age above 56 declines by 976 Euro per individual and year.

The results for the RSA-59 cohort are qualitatively similar but are about twice as large compared to the RSA-58 cohort (Appendix Figure B.3 and Appendix Table B.1). Interestingly, for the RSA-59 we also find no evidence for anticipation effects even though this cohort learned about the reform at age 55.

As a placebo test, we show in Supplementary Material U.5 the estimates for men with too little work experience to be eligible for the lenient DI eligibility rules. For these “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.

---

22 Supplementary Material U.4 contains the Figures of the estimated $\beta_k$-coefficients for these fiscal outcomes.
Table 1: Average Effect of Stricter DI Eligibility Rules, RSA 58

<table>
<thead>
<tr>
<th>Labor market effects (%)</th>
<th>Fiscal effects (Euro)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
</tr>
<tr>
<td>DI benefit receipt</td>
<td>-2.54***</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
</tr>
<tr>
<td>DI application ever</td>
<td>-1.19***</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
</tr>
<tr>
<td>Employment</td>
<td>1.85***</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
</tr>
<tr>
<td>Other benefit receipt</td>
<td>0.94***</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
</tr>
</tbody>
</table>

No. Observations 2,444,975

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 $\beta_k$-coefficients from equation (10) for $k \geq 57$. 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%.

This provides strong support that our main estimates are not confounded by differential trends across birth cohorts.

5 Impact of Benefit Generosity

The ideal experiment to analyze the impact of a change in 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 Estimation Strategy

We exploit the variation in DI benefit levels stemming from the 2003 pension reform described in Section 3 to estimate the causal impact of changes in benefit levels on labor market and fiscal outcomes. We are interested in estimating the following regression:

\[ y_{it} = \alpha + X_{it}'\beta + \gamma b_{it}(Z_{it}) + \lambda_t + \epsilon_{it}, \quad (11) \]

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, \( b_{it}(Z_{it}) \) 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), \( \lambda_t \) are year fixed effects, and \( \epsilon_{it} \) are any unobserved factors affecting the outcome such as taste for work. The parameter of interest is \( \gamma \), which measures the average effect of a change in benefit levels on the outcome variable.

Regression (11) has an endogeneity problem as labor market characteristics \( Z_{it} \) affect benefits and outcomes simultaneously. This problem can be solved by exploiting the 2003 reform, because it creates variation in \( b \) that is independent from \( Z_{it} \). 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 (11). Because of the phased-in nature of the 2003 policy reform, we have 14 different hypothetical benefits for each year from 2004 to 2017:

\[ y_{it} = \alpha + X_{it}'\beta + \gamma b_{it}(Z_{it}) + \sum_{r=2005}^{2017} \delta_r b_r(Z_{it}) + \lambda_t + \epsilon_{it}, \quad (12) \]

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 \( \gamma \) identifies the causal effect of DI benefits.\(^{23}\) We cluster standard errors at the year-month of birth.

\(^{23}\)Figure V.20 in the Supplementary Material assesses the quality of our prediction of hypothetical DI benefits under different policy regimes. Actual benefits track our predicted benefits very closely.
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.

5.2 Empirical Results

Table 2 summarizes our main results providing estimates of equation (12) for labor market and 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 which correspond to a 0.64 percent increase in the application level. DI inflow increases by 0.093 percentage points. 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. 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 annual spending on DI benefits by 36.95 Euro for a 57-60 year old individual, lowers tax revenue by 2.37 Euro, but also lessens spending on other benefits by 20.62 Euro because of benefit substitution. 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. We report the estimates for the age group 30-56 in Appendix Table C.2. The effects in this younger age group are much smaller.

Figure V.22 in the Supplementary Material presents the results of our placebo regressions where we randomly assign individuals potential benefits from a different year. The figure confirms that true increases in benefit generosity lead to unusually large in-
### Table 2: Average Effect of Benefit Generosity for 57-60 old Individuals

<table>
<thead>
<tr>
<th></th>
<th>Labor market effects (%)</th>
<th>Fiscal effects (Euro)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Mean</td>
</tr>
<tr>
<td>DI application ever</td>
<td>0.171***</td>
<td>26.71</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>DI inflow</td>
<td>0.093***</td>
<td>18.68</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Employment outflow</td>
<td>-0.004</td>
<td>71.43</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td></td>
</tr>
<tr>
<td>Other benefit outflow</td>
<td>0.097***</td>
<td>9.89</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,453,448</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table reports estimates for γ from the econometric specification in (12). 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%.

Increases in DI inflow, DI applications, and other benefit outflow (but have no effect on employment outflow), while 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 reforms. 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 one dollar from DI recipients yields an insurance loss of less than two dollars, or put differently, if an additional dollar in the hands of DI recipients has a lower social value than two dollars in the hands of the government. Fiscal multipliers also allow us to compare the effectiveness of alternative DI policy instruments. For instance, if the fiscal multiplier of stricter eligibility rules is substantially larger than the fiscal multiplier of reducing benefits, then a DI reform affecting the same group of agents should implement stricter eligibility rules. In what follows, we first estimate the fiscal multiplier of the Austrian

---

24: “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 (8) and (9) which are normalized by mechanically saved dollars.
DI reforms, separately for stricter eligibility rules and lower benefits. We then compare the effectiveness of the two DI policy instruments, taking account of who is mainly affected by them.

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 $\mathbb{E}[\Delta G(\theta^*_R)]$. In Section 4, we have explored the effects of increasing the RSA and came up with an estimate of total fiscal cost savings. The next step is 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, because the two quantities cannot be directly estimated from the data. As discussed in Section 2 and Appendix A, the mechanical fiscal effect of a change in the strictness of eligibility rules is driven by always applicants, individuals who apply under both strict and lenient eligibility rules. The problem is that we cannot directly observe who is an always applicant and who is a marginal applicant.

To make progress, we proceed in two steps. In a first step, 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). In a second step, we propose an empirical strategy estimating the mechanical fiscal effect based on a group that is very similar to always applicants in the whole population.

**Always- versus Marginal Applicants: A Complier Analysis.** In the complier analysis, 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 characteristics uncover how marginal applicants (who apply only under lenient rules) differ from always applicants (who apply under strict and lenient rules), and also from those of never applicants. Appendix D.1 provides the details of our complier analysis. We estimate a share of always applicants of $\pi^{AA} = 0.070$. The shares of marginal and never
applicants are $\pi^{MA} = 0.014$ and $\pi^{NA} = 0.916$. Moreover, the complier analysis shows that marginal and always applicants are different on a number of characteristics. For instance, marginal applicants are less likely to be on sick leave at age 56 than always applicants. This suggests that healthier individuals react to stricter eligibility criteria by no longer applying, because being on sick leave is a good proxy for underlying health problems.

**Representative Group for Always Applicants: Pre-57 Applicants.** Applicants at age 57 in the treatment group (the RSA-58 cohort) are always applicants, because they face strict eligibility rules and apply. The problem is that applicants at age 57 in the control group (the RSA-57 cohort) are a mix of always and marginal applicants, and we cannot directly separate the two types of applicants. If we knew who is an always applicant in the control group, we could restrict our sample to always applicants and estimate the mechanical effect directly in this subsample. We therefore propose a strategy, which identifies a group that is similar to always applicants and then run our difference-in-difference strategy for this subgroup. We argue and provide evidence that the subpopulation of previously rejected DI applicants, who filed a DI application between age 50 and age 57, is representative for always applicants. In what follows, we refer to this subpopulation as *pre-57 applicants.*

The question is whether pre-57 applicants are indeed representative for always applicants. 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). If rejection decisions were random in this group and individual characteristics, like health, were time-invariant, then pre-57 applicants would be representative for always applicants at age 57. However, rejection decisions are likely not random. For example, healthier individuals are more likely to be rejected, so that pre-57 applicants who reapply at age 57 tend to be healthier than always applicants. The mechanical effect from pre-57 applicants is then overestimated (and the fiscal multiplier underestimated), because we expect a larger change in award probabilities of healthier individuals. In

---

25Supplementary Material Table W.8 shows summary statistics for the pre-57 applicant sample.
this regard, the strategy is conservative. However, this logic only applies if individual characteristics do not change over time. If individual characteristics, like health and skills, also dynamically evolve, it is not clear how pre-57 applicants differ from always applicants. Health could improve or worsen and skills could deteriorate or improve, leading to complex selection effects. Hence, how representative pre-57 applicants are for always applicants becomes an empirical question.

We provide two pieces of evidence that pre-57 applicants are representative for always applicants. As a first piece of evidence, we look at the (re-)application behavior of pre-57 applicants in the treated and control groups. 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). If pre-57 applicants are representative for always applicants, then stricter DI eligibility rules should not change their application behavior. Figure 3 plots the coefficients from our difference-in-differences strategy among the subpopulation of pre-57 applicants. The application behavior at age 57 is not different between the treatment and the control group (Panel a), but DI inflow drops at age 57 (Panel b) due to the stricter eligibility criteria. After age 57, we see significant increases in DI applications and DI inflow among the treated cohort. This pattern aligns with the theoretical considerations on the persistence of the mechanical effect, discussed in Section 2.2: If many applicants are screened out today, more applicants will reapply tomorrow.

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

(a) DI Application Yearly

(b) DI Inflow

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA-58 increase 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.
While the re-application behavior of pre-57 applicants mirrors what one would expect for always applicants, characteristics of pre-57 applicants and always applicants could still be quite different. As a second piece of evidence, we focus on the treated cohort (RSA=58) and compare key characteristics of pre-57 applicants, who re-apply at age 57, with all age-57 applicants. Age-57 applicants are always applicants by definition, since screening is strict at age 57. This comparison therefore reveals whether always applicants among pre-57 applicants are comparable with all always applicants in the treated cohort. Figure 4 shows DI benefit receipt and net fiscal expenditures for all age-57 applicants (blue squares) and for 57 re-applicants in the subgroup of pre-57 applicants (red triangle) during 15 quarters before and after the application at age 57. After date 0, DI benefit receipt and net fiscal expenditures of the two groups are very similar, and 3 quarters after the application they 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 re-applicants, by construction, already filed a DI application before date 0, while age-57 applicants did not necessarily.\textsuperscript{26}

Figure 4: Comparison of Applicants at 57 and Pre-57 Applicants in Treatment Group, RSA 58

(a) DI Benefit Receipt

(b) Net Fiscal Expenditures

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.

\textsuperscript{26}Appendix D.2 performs the same tests for the the RSA-59 cohort. We find that also for the two year increase in the RSA to age 59, pre-57 applicants are representative for always applicants.
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. In Appendix Figure D.6 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 3 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 $E\left[\Delta G(\theta^*_R)|\text{pre-57}\right] = 5,585$ Euro.\(^{27}\) If our identifying assumption is satisfied, we have $E\left[\Delta G(\theta^*_R)|\text{pre-57}\right] = E\left[M(\theta^*_R)|\text{pre-57}\right] / \pi^{AA}$, 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.\(^{28}\) The mechanical fiscal effect in the whole population can then be calculated as $E\left[M(\theta^*_R)\right] = 5,585$ Euro $\times 0.070 = 391$ Euro where $\pi^{AA} = 0.070$

\(^{27}\) Notice that $E[\Delta G|\text{pre-57}]$ 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 $E[\Delta G|\text{pre-57}] \times Pr(\text{re-apply at 57}) = 1,167$ 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 $E[\Delta G|\text{pre-57}] = 5,585$ Euro.

\(^{28}\) Notice that, by definition, the mechanical fiscal effect $E\left[M(\theta^*_R)\right]$ comprises the total (mechanical) cost savings per individual in the whole population, i.e. unconditional on a DI application, while $E[\Delta G|\text{pre-57}]$ conditions on an application at age 57.
is the share of always applicants from Table D.3. The behavioral fiscal effect is calculated as the difference between the fiscal cost effect from Table 1 and the mechanical fiscal effect, $\mathbb{E} \left[ B(\theta^*_R) \right] = 976 - 391 = 585$ Euro. This decomposition implies a fiscal multiplier of 2.50. For the RSA increase from 57 to 59 we follow the same steps and find a multiplier of 2.05 as displayed in Table 3.

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 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
Table 3: Fiscal Multiplier for Eligibility Rules and Benefit Generosity

<table>
<thead>
<tr>
<th></th>
<th>Eligibility Rules</th>
<th>Benefit Generosity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RSA 58 RSA 59</td>
<td>Ages 57-60 Ages 30-56</td>
</tr>
<tr>
<td>Total fiscal effect</td>
<td>976 1,686</td>
<td>63.85 4.42</td>
</tr>
<tr>
<td>Mechanical fiscal effect (M)</td>
<td>391 823</td>
<td>45.16 3.24</td>
</tr>
<tr>
<td>Behavioral fiscal effect (B)</td>
<td>585 863</td>
<td>18.69 1.18</td>
</tr>
<tr>
<td>Fiscal multiplier (1+B/M)</td>
<td>2.50 2.05</td>
<td>1.41 1.36</td>
</tr>
</tbody>
</table>

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 Tables 1 and B.1. 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 Tables 2 and C.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 Tables 2 and C.2 with 0.01. The total fiscal effect is the sum of the mechanical and behavioral fiscal effects.

the insurance value. It is not straightforward to estimate the insurance losses of stricter eligibility rules, the right-hand-side of the social optimality condition (8). Nevertheless, we discuss in Supplementary Material W.4 how we can bound the insurance losses 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).

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 straightforward. In section 5, we have directly estimated the behavioral fiscal effect $\mathbb{E}[B(b_s)]$ 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, 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). The total fiscal effect, the sum of behavioral and mechanical fiscal effects, is then 63.85 Euro for a benefit cut for the age group 57-60 (and 4.42 Euro for the age group 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 3 summarizes our results.\textsuperscript{29}

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 social welfare. The above analysis has shown that the fiscal multiplier of stricter eligibility rules is significantly larger than the fiscal multiplier of reducing benefits. A mechanical one-dollar reduction of DI expenditures reduces overall expenditures 1.8 times (= 2.50/1.41) more when the reduction is due to stricter eligibility rules (RSA increase to age 58) rather than lower 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.\textsuperscript{30}

Clearly, the insurance losses of the two policy instruments are not identical. Comparing the right-hand-sides of the social optimality conditions (8) and (9) shows that stricter eligibility rules affect only the additionally screened-out DI applicants, while lower benefits affect all DI recipients alike. For the relative comparison of the insurance losses, we proceed in two steps. We first use our theoretical framework from section 2.2 to derive a condition that compares the insurance loss of tighter eligibility rules with the loss of lower benefits. We then 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.4, we show that a sufficient condition for the insurance loss of stricter

\textsuperscript{29}To assess the overall welfare consequences of a DI benefit cut, the fiscal multiplier needs to be compared to the associated insurance loss. We lack consumption data to follow the standard procedures to implement the insurance loss of lower benefits. We discuss bounds on the insurance value in Supplementary Material W.5.

\textsuperscript{30}We focus here on the relative performance of the two policy instruments if a policy maker wants to cut DI program costs.
DI eligibility rules being smaller than the insurance loss of lower DI benefits is

\[
\mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \cdot \frac{v'(c_{ij}^D)}{\lambda} \cdot \frac{\Delta D_{i|t} (b_{ij} - z_{ij})}{\mathbb{E}[M_N(\theta^*_i)]} \right] \leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \cdot \frac{v'(c_{ij}^D)}{\lambda} \cdot \frac{D_{i|t} (b_{ij}^H - b_{ij}^L)}{\mathbb{E}[M_N(b)]} \right]. 
\]

(13)

The left-hand-side of condition (13) is an upper bound of the insurance loss associated with tighter eligibility rules, which comes from additionally rejected DI applicants (indicated by \( \Delta D_{i|t} = 1 \)).\(^{31}\) The right-hand-side of the condition is 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 (13) splits the insurance loss into a (bounded) income loss and a welfare weight, \( v'(c_{ij}^D)/\lambda \). 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). 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 this case, we would conclude that tightening DI eligibility rules generates a smaller insurance loss than cutting DI benefits.

The second (empirical) step is to implement condition (13). For the left-hand-side of (13), we estimate the bounds on income losses for each income quintile using the empirical model (10). An individuals’ income is the sum of labor earnings net of payroll taxes, DI benefits, UI benefits, and SI benefits. We fix the income quintile of an individual at age 55 and convert the estimated treatment-coefficients into an average yearly

\(^{31}\)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. By a revealed preference argument assuming all rejected applicants receive social welfare provides an upper bound on the insurance loss. Note this bound also holds for an non-marginal policy change.
income loss during ages 57-61. For rejected DI applicants resuming work, we implement the bound on income loss \((b_{it} - z_{it})\) by replacing their labor earnings with their potential welfare benefits.\(^{32}\) For the right-hand-side of (13) we can directly calculate the reduction in DI benefits \((b^H_{it} - b^L_{it})\) of DI recipients. Following condition (13), we normalize the absolute income losses by the mechanical fiscal effect of the respective DI policy instrument from Table 3. 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.\(^{33}\)

Figure 6 shows the normalized (bounded) income losses by income quintiles. In the lowest income quintile, stricter DI eligibility rules (blue lines) generate normalized income losses about 1.50 Euros lower than the income losses generated by lower DI benefits (red line). The income losses are almost identical in the second, fourth, and fifth quintile. Only in the third quintile, stricter eligibility rules generate a larger income loss than lower benefits but the difference is small, about 0.20-0.40 Euros. Roughly speaking, condition (13) 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. 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.\(^{34}\)

In sum, our analysis of the Austrian DI reforms lead to the conclusion that fiscal multipliers of stricter eligibility rules are substantially larger than fiscal multipliers of lowering benefits. Moreover, the overall insurance loss of tighter DI eligibility rules is smaller than the average insurance loss of lower DI benefits. In the Austrian context, our estimates therefore suggest that the welfare cost of tightening DI eligibility is smaller

\(^{32}\)We construct the potential welfare benefits as a weighted mix of UI and SI benefits where the weights are the relative share of UI and SI benefit recipients. This ways leads to an average replacement rate of 59 percent of previous earnings.

\(^{33}\)Supplementary Material Figures W.25 and W.26 show the income losses by income source.

\(^{34}\)In Supplementary Material W.3, we show that a policy maker who is only concerned about income replacement finds stricter eligibility rules less costly than reducing DI benefits, irrespective of distributional preferences.
Comparison to U.S. DI system. The U.S. DI eligibility criteria are also subject to vocational factors similar to the RSA in Austria. Based on estimates from the literature, we can also apply our framework to the U.S. In Appendix D.5, we discuss how previous estimates can shed new light on welfare effects through the lens of our sufficient statistics framework. We find that the evidence from Chen and van der Klaauw (2008) implies 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. Hence, 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. U.S. estimates of the DI take-up elasticity suggest a fiscal multiplier of 1.4 for DI benefit levels which is very similar to the multiplier we estimate for Austria. The DI replacement rates in the U.S. are lower than in Austria and hence the insurance value in the U.S. might be higher than in Austria. Our analysis in Supplementary Material W.5 cannot reject that DI benefits in Austria are optimal for reasonable values of risk aversion. This implies that benefits in the U.S. might be too low.

35This is almost certainly the case to the extent that DI applicants are, on average, more deserving than the whole population.
7 Conclusion

In this paper, we pursue the sufficient-statistics approach of policy evaluation to shed new light on the welfare consequences of DI reforms. As a novelty on the theoretical side, we develop a general framework to explore the incentive-insurance trade-off associated with changes in DI eligibility rules versus DI benefits. The main innovation of our analysis relates to both the theoretical analysis of optimal DI eligibility rules and the empirical analysis of the causal effect on tighter DI eligibility rules on the (behavioral and mechanical) fiscal expenditures of the DI program. This is progress over the existing literature which has – both theoretically and empirically – mainly focused on the effects of changes in DI benefits. Complementing our analysis of tighter DI eligibility rules with a corresponding exercise on the effects of lower DI benefits, allows us to rank the two important DI policy instruments both with respect to their relative fiscal cost savings and their relative welfare losses. Our theoretical framework highlights the sufficient statistics needed for such a ranking and our empirical analysis studies the fiscal effects and the distributional implications (with respect to both DI policy instruments) exploiting two Austrian DI reforms and exceptionally rich administrative register data.

We find that stricter DI eligibility rules are more effective in reducing the fiscal burden of the DI system. Furthermore, 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. Hence, our empirical results combined with our theoretical framework imply a clear ranking for Austrian DI policies: stricter DI eligibility rules dominate lower DI benefits.

It is worth emphasizing that our framework is rather general and can be applied to study the welfare effects of DI reforms in other countries, including countries with different institutional rules. For instance, we can readily apply our framework to the U.S. DI system (where only non-employed individuals can apply for DI): based on previous estimates on the impact of DI eligibility rules and DI benefits, our framework suggests that, in the U.S., DI benefits are too low and DI eligibility rules are too tight. Hence our sufficient-statistics framework comes to a similar conclusion than previous
work using the structural approach (Low and Pistaferri, 2015).

Finally, it is important to keep in mind that the two DI policy instruments studied here are not the only policy instruments that affect costs and benefits of the DI system. Other relevant instruments concern alternative social insurance programs, such as UI, social welfare and (early) retirement programs. To the extent that benefit substitution is quantitatively important (as in the Austrian context), changing the generosity of those programs may have an important impact on the DI system. Moreover, a DI reform could target the precision of the imperfectly functioning disability-assessment system, in order to minimize false acceptances and false rejections (type-I and type-II errors). Unlike changing DI eligibility criteria, improving the precision of DI screening requires resources, such as more extensive medical checks, better equipment, additional monitoring of DI applicants, and the like. The welfare calculations of such a policy need to take into account society’s willingness to pay for improved DI screening. Studying this trade-off is an interesting direction for future research.
References


44


Online Appendix

A Theoretical Framework

A.1 Static Model

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

$$\frac{\partial W}{\partial \theta^*} = -u'(w - \tau) \frac{\partial \tau}{\partial \theta^*} + \int_{\theta^A}^{\infty} \frac{\partial p(\theta; \theta^*)}{\partial \theta^*} v(b) dF(\theta) - \int_{\theta^R}^{\theta^A} \frac{\partial p(\theta; \theta^*)}{\partial \theta^*} (u(w) - \theta) dF(\theta)$$

(A.1)

$$= 0 \text{ by definition of } \theta^A \text{ in eq. (1)}$$

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

Non-marginal Change of Strictness of DI Eligibility. Condition (4) in the main text holds for a marginal change in $\theta^*$. Here we consider the welfare effect of a discrete change. Suppose strictness of eligibility rules increase from $\theta^L$ to $\theta^H > \theta^L$. This implies that the award probability falls $p(\theta; \theta^H) < p(\theta; \theta^L)$ and fewer individuals apply $\theta^A_H > \theta^A_L$ ($\theta^A_L$ denotes the marginal applicant with the lenient criteria $\theta^L$ and $\theta^A_H$ denotes the marginal applicant with the high standard $\theta^H$). Note $\theta^R$ is still independent of the
eligibility rules.

Figure A.1 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 \( \theta < \theta_A^H \) no longer apply under the stricter rules. Individuals with \( \theta \in [\theta_A^L, \theta_A^H] \) are therefore “marginal applicants” as they only apply under the lenient rules. The share of these marginal applicants is \( \pi^M = F(\theta^H_A) - F(\theta^L_A) \). The behavioral effect is the area under the old award curve of these marginal applicants. Individuals with a disability level above \( \theta_A^H \) continue to apply. These are always applicants and their share is \( \pi^A = 1 - F(\theta^H_A) \). 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 (\( \theta \) to the left of \( \theta_R^L \)) and some substitute to welfare benefits (\( \theta \) to the right of \( \theta_R^L \)). Hence, we have the same effects as in the marginal case, but these effects are slightly differently defined.

Figure A.1: Illustration Non-Marginal Change

Note: This figure illustrates the effects of a non-marginal change in strictness of DI eligibility rules. The mechanical effect is driven by always applicants, while the behavioral fiscal effect is driven by marginal applicants.

Let \( W_H \) and \( W_L \) denote welfare in the two regimes. Welfare in the two regimes
The welfare effect of this discrete change $\Delta W \equiv W_H - W_L$ is given by

$$
\Delta W = u(w - \tau_H) - u(w - \tau_L) - \int_{\theta^A_H}^{\theta^A_L} \left[ p(\theta; \theta^L) - p(\theta; \theta^H) \right] dF(\theta) - \int_{\theta^R_H}^{\theta^R_L} p(\theta; \theta^L) \left[ v(b) - (u(w) - \theta) \right] - \psi dF(\theta).
$$

The first line of (A.4) captures the gain for the taxpayer (the fiscal cost reduction), the second line measure 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 $\theta^*$ 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 $\theta^A_H \to \theta^A_L$ and $\int_{\theta^A_H}^{\theta^A_L} p(\theta; \theta^L) \left[ v(b) - (u(w) - \theta) \right] - \psi dF(\theta) \to p(\theta^A_L; \theta^L) \left[ v(b) - (u(w) - \theta^A_L) \right] - \psi = 0$ by the definition of the marginal applicant $\theta^A_L$.

Using the government budget constraint we can rewrite the fiscal effect again as the behavioral fiscal effect $B_\Delta$ plus the mechanical fiscal effect $M_\Delta$:

$$
\tau^L - \tau^H = b \cdot \int_{\theta^A_H}^{\theta^A_L} p(\theta; \theta^L) dF(\theta)
$$

$$
\equiv B_\Delta
$$

$$
= +b \cdot \int_{\theta^A_H}^{\theta^A_L} \left[ p(\theta; \theta^L) - p(\theta; \theta^H) \right] dF(\theta) - (b - z) \cdot \int_{\theta^R_H}^{\theta^R_L} \left[ p(\theta; \theta^L) - p(\theta; \theta^H) \right] dF(\theta).
$$

$$
\equiv M_\Delta
$$
Moreover, we can write the welfare effect associated with the fiscal effect as \( u(w - \tau^H) - u(w - \tau^L) = \lambda (\tau^L - \tau^H) = \lambda (B_\Delta + M_\Delta) \), where \( \lambda = u'(w - \tau_\Delta) \) with \( \tau_\Delta \) such that this equality holds. We can then rearrange (A.4) to \( \Delta W \geq 0 \Leftrightarrow \\
\begin{align*}
1 + \frac{B_\Delta}{M_\Delta} & \geq \frac{L_{W\Delta} + L_{Z\Delta}}{u'(w - \tau_\Delta)M_\Delta} + \frac{L_{M\Delta}}{u'(w - \tau_\Delta)M_\Delta} \\
\end{align*}
(A.6)
\]
where \( L_{W\Delta} \equiv \int_{\theta^*_L}^{\theta^*_R} [p(\theta; \theta^L) - p(\theta; \theta^H)] [v(b) - (u(w) - \theta)] dF(\theta) \), \( L_{Z\Delta} \equiv \int_{\theta^*_L}^{\theta^*_R} [p(\theta; \theta^L) - p(\theta; \theta^H)] [v(b) - v(z)] dF(\theta) \) and \( L_{M\Delta} \equiv \int_{\theta^*_L}^{\theta^*_R} p(\theta; \theta^L) [v(b) - (u(w) - \theta)] - \psi dF(\theta) \).

**Welfare Effect DI Benefit Level.** Starting from equation (2) we get
\[
\frac{\partial W}{\partial b} = -u'(w - \tau) \frac{\partial \tau}{\partial b} + \int_{\theta^*_L}^{\infty} p(\theta; \theta^*) v'(b) dF(\theta) \\
+ \frac{\partial \theta^A}{\partial \theta^*} p(\theta^A) f(\theta^A) \left\{ u(w) - v(b) + \frac{\psi}{p(\theta^A)} - \theta^A \right\} \\
= 0 \text{ by definition of } \theta^A
\]
(A.7)
where \( \frac{\partial \tau}{\partial b} = -b (\partial \theta^A / \partial b) p(\theta^A) f(\theta^A) + \int_{\theta^*_L}^{\infty} p(\theta; \theta^*) dF(\theta) = B(b) + M(b) \) is the change in taxes necessary to fund a DI system with a marginally higher DI benefit \( b \). (7) then immediately follows from (A.7). The behavioral fiscal effect is \( B(b) \equiv - (\partial \theta^A / \partial b) p(\theta^A) f(\theta^A) \cdot b \) and the mechanical fiscal effects is \( M(b) \equiv \int_{\theta^*_L}^{\infty} p(\theta) dF(\theta) \). The ratio of behavioral over mechanical fiscal effect corresponds to the DI inflow elasticity, i.e. \( \xi = (\partial DI / \partial b)(b/\Delta DI) = - (\partial \theta^A / \partial b) p(\theta^A) f(\theta^A) b / \int_{\theta^*_L}^{\infty} p(\theta) dF(\theta) = B(b)/M(b) \). This yields an interesting analogy of the optimal DI formula to the famous Baily (1978) formula for optimal unemployment insurance (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, \( \eta \) 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

---

36 Later on in the general model \( \lambda \) will denote the multiplier of the government budget constraint and therefore measures the social value of public funds.
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.

A.2 General Model

**Setup.** 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 \( \theta \) (and as well as other state variables) evolve stochastically over the agent’s relevant time horizon. Let \( X_{i,t} = \{ \theta_{i,t}, A_{i,t}, X_{i,t} \} \) 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 \( X_{i,t} \) 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

\[
A_{i,t+1} = (1 + r_t)A_{i,t} + b_{i,t}(X_{i,t}) - c_{D,i,t}^D(X_{i,t}) \quad (A.8)
\]

\[
A_{i,t+1} = (1 + r_t)A_{i,t} + w_{i,t}(X_{i,t}) - \tau_{i,t}(X_{i,t}) - c_{E,i,t}^E(X_{i,t}) \quad (A.9)
\]

\[
A_{i,t+1} = (1 + r_t)A_{i,t} + z_{i,t}(X_{i,t}) - c_{Z,i,t}^Z(X_{i,t}). \quad (A.10)
\]

\( b_{i,t}(X_{i,t}) \) 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 \( \{ c_{D,i,t}^D(X_{i,t}), c_{W,i,t}^W(X_{i,t}), c_{Z,i,t}^Z(X_{i,t}) \} \), whether they apply to DI benefits \( a_{i,t}(X_{i,t}) \in \{ 0, 1 \} \) 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 \( X_{i,t} \) 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.\(^{37}\) If her application is rejected, she either resumes

\(^{37}\)The assumption that DI is an absorbing state, is supported by the empirically observed negligibly low outflow rates, particularly among older workers.
work or claims social welfare benefits, whatever yields higher utility. Note that the general model also admits the possibility that $\theta^A \geq \theta^R$. This might occur for agents with low wage realization and low DI acceptance probabilities.

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

$$D_{i,t} = 1 - \left[ \prod_{k=0}^{t} (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k}, \theta^*_k)) \right]$$  \hspace{1cm} (A.11)$$

$$W_{i,t} = \omega_{i,t}(X_{i,t}) \left[ \prod_{k=0}^{t} (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k}, \theta^*_k)) \right]$$  \hspace{1cm} (A.12)$$

$$Z_{i,t} = \left( 1 - \omega_{i,t}(X_{i,t}) \right) \left[ \prod_{k=0}^{t} (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k}, \theta^*_k)) \right] .$$  \hspace{1cm} (A.13)$$

The probability agent $i$ transitions to DI in period $k$ is $\alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k}, \theta^*_k)$. Hence, the probability that an agent is not yet on DI in period $t$ is $\left[ \prod_{k=0}^{t} (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k}, \theta^*_k)) \right]$. 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.\(^{38}\) We assume that the first application bears a fix cost $\psi$ and follow-up applications are costless. $\Lambda_{i,t} = \alpha_{i,t}(X_{i,t}) \prod_{k=0}^{t-1} (1 - \alpha_{i,k}(X_{i,k})) \in [0,1]$ indicates whether agent $i$ applies for the first time in period $t$. The other state variables, disutility of work $\theta_{i,t}$ and $X_{i,t}$, 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.\(^{39}\) The agent’s problem is then given by

\(^{38}\)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 $k$.

\(^{39}\)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 $\mathbb{E}[Y_{i,t}] = \int Y(X_{i,t}) dF(X_{i,t})$ to denote the expectation w.r.t. state variables for a given individual.
\[ V_i(P) = \max E \left[ \sum_{t=0}^{T-1} \beta^t \left( v(c_{ij}^D) \cdot D_{ij} + v(c_{ij}^Z) \cdot Z_{ij} + \left( u(c_{ij}^W) - \theta_{ij} \right) \cdot W_{ij} - \Lambda_{ij} \cdot \psi \right) \right] + E \left[ \sum_{t=0}^{T-1} \beta^t \mu_i^D \left( (1 + r_t)A_{ij} + b_{ij} - c_{ij}^D - A_{ij+1} \right) D_{ij} \right] + E \left[ \sum_{t=0}^{T-1} \beta^t \mu_i^W \left( (1 + r_t)A_{ij} + w_{ij} - \tau_{ij} - c_{ij}^W - A_{ij+1} \right) W_{ij} \right] + E \left[ \sum_{t=0}^{T-1} \beta^t \mu_i^Z \left( (1 + r_t)A_{ij} + z_{ij} - c_{ij}^Z - A_{ij+1} \right) Z_{ij} \right]. \] (A.14)

The social planner maximizes social welfare by choosing the strictness of DI eligibility \( \theta_{ij}^* \) and DI benefit function \( b_{ij} \) in each period \( s \). We denote this disability policy by \( P = \{ \theta_{ij}^*, b_{ij} \}_{s=0}^{T-1} \). The planner therefore solves

\[ \max_P W(P) = \int V_i(P) dE + \lambda (G(P) - \bar{G}) \] (A.15)

where

\[ G(P) = \int E \left[ \sum_{t=0}^{T-1} (1 + r_t)^{-t} \left( W_{ij} \cdot \tau_{ij} - D_{ij} \cdot b_{ij} - Z_{ij} \cdot z_{ij} \right) \right] dE \] (A.16)

is the planner’s net revenue, \( \bar{G} \) is an exogenous revenue constraint and \( \lambda \) denotes the Lagrange multiplier on the planner’s budget constraint.

**Welfare Effects of DI Eligibility Rules and DI Benefits.** The following propositions characterize the optimal DI policy \( P = \{ \theta_{ij}^*, b_{ij} \}_{s=0}^{T-1} \).

**Proposition 1.** Assume the planner’s budget constraint is differentiable in \( \theta_{ij}^* \). Optimal strictness of DI eligibility rules in period \( s, \theta_{ij}^* \), then fulfills

\[ 1 + \frac{\mathbb{E} \left[ B(\theta_{ij}^*) \right]}{\mathbb{E} \left[ M(\theta_{ij}^*) \right]} = \frac{\mathbb{E} \left[ L_W \right] + \mathbb{E} \left[ L_Z \right]}{\lambda \mathbb{E} \left[ M(\theta_{ij}^*) \right]} \] (A.17)

where

\[ \mathbb{E} \left[ M(\theta_{ij}^*) \right] = \mathbb{E} \left[ \sum_{t=s}^{T-1} (1 + r_t)^{-t} \left( M_{W_{ij}}(b_{ij} + \tau_{ij}) + M_{Z_{ij}}(b_{ij} - z_{ij}) \right) \right] \] (A.18)
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. \( M_{Wi,s} \) is the mechanical employment effect

\[
M_{Wi,s} \equiv -\omega_{i,t} \left( \alpha_{i,s} \cdot \frac{\partial p(\theta_{is}, \theta_{s}^*)}{\partial \theta_{s}^*} \prod_{k=0, k \neq s}^{T} (1 - \alpha_{i,k} p_{i,k}) \right)
\]

(A.19)

and \( M_{Zi,s} \) is the mechanical benefit substitution effect

\[
M_{Zi,s} \equiv -\left( 1 - \omega_{i,t} \right) \left( \alpha_{i,s} \cdot \frac{\partial p(\theta_{is}, \theta_{s}^*)}{\partial \theta_{s}^*} \prod_{k=0, k \neq s}^{T} (1 - \alpha_{i,k} p_{i,k}) \right).
\]

(A.20)

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

\[
\mathbb{E}[L_W] \equiv \mathbb{E} \left[ \sum_{t=s}^{T-1} \beta^t \left( M_{Wi,s} \left( v(c_{i,t}^P) - \left( u(c_{i,t}^W) - \theta_{i,t} \right) \right) \right) \right]
\]

(A.21)

\[
\mathbb{E}[L_Z] \equiv \mathbb{E} \left[ \sum_{t=s}^{T-1} \beta^t \left( M_{Zi,s} \left( v(c_{i,t}^P) - v(c_{i,t}^Z) \right) \right) \right].
\]

(A.22)

**Proof.** See below.

**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

\[
1 + \frac{\mathbb{E}[B(b_s)]}{\mathbb{E}[M(b_s)]} = \frac{\mathbb{E}[v'(c_s^P)]}{\lambda \cdot \mathbb{E}[M(b_s)]}
\]

(A.23)

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

**Proof.** See below.

**Proofs.**
Proof. Proposition 1. The proof is a direct application of the Envelope Theorem. To derive conditions (A.17) and (A.23) 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.2 illustrates this idea nicely. The proof here therefore identifies differentiable upper and lower support functions of the welfare function $W(P)$.

Figure A.2: 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 $\bar{P}$ denote the optimal policy, i.e. the $P = \{\theta_s^*, b_s\}_{s=0}^{T-1}$ that maximizes welfare. By definition $W(\bar{P}) \geq W(P) \forall P$ and therefore the constant function $U(P) = W(\bar{P})$ 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 $L(P) = \int \bar{V}_i(P)di + \lambda (G(P) - \bar{G})$. The derivative of this lower
support function with respect to \( \theta_s^* \) is

\[
\frac{\partial L(P)}{\partial \theta_s^*} = \mathbb{E} \left[ \sum_{t=s}^{T-1} (1 + r_t)^{-t} \left( \alpha_{i,t} \cdot \frac{\partial p(\theta_{i,s}, \theta_s^*)}{\partial \theta_s^*} \right) \prod_{k=0,k\neq s}^{t} (1 - \alpha_{i,k} p_{i,k}) \left\{ v(c_{i,t}^P) - v(c_{i,t}^Z) (1 - \omega_{i,t}) \right\} \ldots \right.
\]

\[
- \left( u(c_{i,t}^W) - \theta_s^* \right) \omega_{i,t} \right] + \lambda \frac{\partial G(P)}{\partial \theta_s^*}. \tag{A.24}
\]

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

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

and we define the behavioral fiscal effect as the residual \( \mathbb{E} \left[ B(\theta_s^*) \right] = \frac{\partial G(P)}{\partial \theta_s^*} - \mathbb{E} \left[ M(\theta_s^*) \right]. \tag{A.40} \)

The differentiable sandwich lemma then implies that \( \frac{\partial W(P)}{\partial \theta_s^*} = \frac{\partial L(P)}{\partial \theta_s^*} = \frac{\partial U(P)}{\partial \theta_s^*} = 0 \) at the optimal policy. It is then straightforward to rearrange (A.24) to (A.17).

**Proof.** Proposition 2. We apply the same logic to the optimal DI benefit policy as in the previous proof. We have

\[
\frac{\partial L(P)}{\partial b_s} = \mathbb{E} \left[ \beta^s \mu_{i,s}^P \right] + \lambda \frac{\partial G(P)}{\partial b_s}. \tag{A.25}
\]

The agent’s first order condition implies \( \mathbb{E} \left[ \beta^s \mu_{i,s}^P D_{i,s} \right] = \mathbb{E} \left[ \beta^s v'(c_{i,s}^P) D_{i,s} \right]. \) Define the mechanical fiscal effect as \( \mathbb{E} \left[ M(b_s) \right] = \mathbb{E} \left[ (1 + r_s)^{-s} (D_{i,s}) \right] \) and the behavioral fiscal effect is again defined as the difference between total fiscal effect and mechanical fiscal effect \( \mathbb{E} \left[ B(b_s) \right] = -\frac{\partial G(P)}{\partial b_s} - \mathbb{E} \left[ M(b_s) \right]. \) We again have \( \frac{\partial W(P)}{\partial b_s} = \frac{\partial L(P)}{\partial b_s} = \frac{\partial U(P)}{\partial b_s} = 0. \) It is then straightforward to rearrange (A.25) to obtain (A.23).
by policy \( P^L = (\theta^H_0, \ldots, \theta^H_s, \theta^L_{s+1}, \theta^*_{T-1}; b_0, \ldots, b_{T-1}) \). The reform we study empirically increased the age of relaxed screening from \( s \) to \( s + 1 \). This corresponds to policy \( P^H = (\theta^H_0, \ldots, \theta^H_s, \theta^L_{s+1}, \theta^*_{T-1}; b_0, \ldots, b_{T-1}) \). Let \( a^H_{i,t} \) denote the application decision of individual \( i \) in period \( t \) if the policy is \( P^H \) and \( a^L_{i,t} \) denote the application decision under policy \( P^L \). The discrete welfare effect is

\[
\Delta W = W(P^H) - W(P^L)
\]

\[
= \int V_i(P^H) - V_i(P^L) di + \lambda \left( G(P^H) - G(P^L) \right)
\]

assuming that \( \lambda \) 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

\[
\mathbb{E} \left[ M_A(\theta^*_s) \right] \equiv \mathbb{E} \left[ \sum_{t=s}^{T-1} (1 + r_t)^{-t} \left( M_{\Delta W_{i,t}} (b_{i,t} + \tau_{i,t}) + M_{\Delta Z_{i,t}} (b_{i,t} - z_{i,t}) \right) \right]
\]

(A.27)

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

\[
M_{\Delta W_{i,t}} \equiv \omega_{i,t} \left( \alpha^H_{i,s} \cdot \left[ p^L_{i,s} - p^H_{i,s} \right] \prod_{k=0, k \neq s}^t \left( 1 - \alpha^H_{i,k} p_{i,k} \right) \right)
\]

(A.28)

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

\[
M_{\Delta Z_{i,t}} \equiv (1 - \omega_{i,t}) \left( \alpha^H_{i,s} \cdot \left[ p^L_{i,s} - p^H_{i,s} \right] \prod_{k=0, k \neq s}^t \left( 1 - \alpha^H_{i,k} p_{i,k} \right) \right)
\]

(A.29)

The mechanical fiscal effect is, as in the static model, driven by always applicants, \( \alpha^H_{i,s} = 1 \) (those who apply under the strict rules at age \( s \)), and the change in their award probability \( \left[ p^L_{i,s} - p^H_{i,s} \right] \). The share of always applicants at age \( s \) is given by

\[
\pi^{AA} = \mathbb{E} \left[ \alpha^H_{i,s} \cdot \prod_{k=0}^{s-1} \left( 1 - \alpha^H_{i,k} p_{i,k} \right) \right].
\]

(A.28)

We define the behavioral fiscal effect as the resid-

\[\text{[41]}\text{The share of marginal applicants is } \pi^{MA} = \mathbb{E} \left[ \left( \alpha^L_{i,s} - \alpha^H_{i,s} \right) \cdot \prod_{k=0}^{s-1} \left( 1 - \alpha^H_{i,k} p_{i,k} \right) \right].\]
ual $\mathbb{E}[B_\Delta(\theta^*_s)] \equiv (G(P^H) - G(P^L)) - \mathbb{E}[M_\Delta(\theta^*_s)]$. 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

$$\int_i V_i(P^H) - V_i(P^L) di = \mathbb{E}[L_{AW}] + \mathbb{E}[L_{AZ}] + \mathbb{E}[L_{MA}]$$

(A.30)

where

$$\mathbb{E}[L_{AW}] \equiv \mathbb{E} \left[ \sum_{t=s}^{T-1} \beta^t \left( M_{AW_{i,t}} \left( v_i(c^D_{i,t}) - \left( u_i(c^W_{i,t}) - \theta_{i,t} \right) \right) \right) \right]$$

(A.31)

$$\mathbb{E}[L_{AZ}] \equiv \mathbb{E} \left[ \sum_{t=s}^{T-1} \beta^t \left( M_{AZ_{i,t}} \left( v_i(c^D_{i,t}) - v_i(c^Z_{i,t}) \right) \right) \right]$$

(A.32)

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

$$1 + \frac{\mathbb{E}[B_\Delta(\theta^*_s)]}{\mathbb{E}[M_\Delta(\theta^*_s)]} \geqslant \frac{\mathbb{E}[L_{AW}] + \mathbb{E}[L_{AZ}]}{\lambda \mathbb{E}[M_\Delta(\theta^*_s)]} + \frac{\mathbb{E}[L_{MA}]}{\lambda \mathbb{E}[M_\Delta(\theta^*_s)]}.$$  

(A.33)
B The Effect of Tighter DI Eligibility Rules

B.1 Empirical Results RSA-59

Figure B.3: Effects of RSA 59 on Labor Market States and DI Application by Age

(a) DI Benefit Receipt  (b) DI Application Ever

(c) Employment  (d) Other Benefit Receipt

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.
Table B.1: Average Effect of Stricter DI Eligibility Rules RSA-59

<table>
<thead>
<tr>
<th>Labor market effects (%)</th>
<th>Fiscal effects (Euro)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
</tr>
<tr>
<td>DI benefit receipt</td>
<td>-4.82***</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
</tr>
<tr>
<td>DI application ever</td>
<td>-2.67***</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
</tr>
<tr>
<td>Employment</td>
<td>3.01***</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
</tr>
<tr>
<td>Other benefit receipt</td>
<td>2.26***</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
</tr>
</tbody>
</table>

No. Observations: 2,176,311

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 $\beta_k$-coefficients from equation (10) for $k \geq 57$. 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%.

C Impact of Benefit Generosity

Table C.2: Average Effect of Benefit Generosity for 30-56 old Individuals

<table>
<thead>
<tr>
<th>Labor market effects (%)</th>
<th>Fiscal effects (Euro)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
</tr>
<tr>
<td>DI application ever</td>
<td>0.014***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>DI inflow</td>
<td>0.003***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Employment outflow</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Other benefit outflow</td>
<td>0.003***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

Observations: 15,968,003

Notes: The table reports estimates for $y$ from the econometric specification in (12). 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%.
D Estimating the Fiscal Multiplier of DI Reforms

D.1 Complier Analysis Stricter Eligibility Rules

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 characteristics of marginal, always, and never enrollees. We focus here on the complier characteristics of the RSA-58 increase. Supplementary Material W.1 provides the formal framework and the results for the RSA-59 increase.

Table D.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 D.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 probability of a DI award for applicants who are in better health but have poor labor market prospects.

The analogous results for the RSA-59 change resemble qualitatively the results for the RSA-58 change. This is shown in Supplementary Material Table W.7.

\[42\]
Table D.3: Applicant and enrollee characteristics, RSA 58

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

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 Supplementary Material W.1. Earnings are reported in 2018 Euro. Avg. annual earnings measures the average earnings of the best 15 years. Levels of significance: *1%, **5%, and ***1%.
D.2 Pre-57 Applicants: Representative for Always Applicants?

Figure D.4: Effect of RSA on DI Application Yearly and DI Inflow by Age for Pre-57 Applicants RSA-59

(a) DI Application Yearly

(b) DI Inflow

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA-59 increase 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.

Figure D.5: Comparison of Applicants at 57 and Pre-57 Applicants in Treatment Group RSA-59

(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.
D.3 Mechanical Effect RSA-59

Figure D.6: Mechanical Effect of RSA-59 on Labor Market States and Net Fiscal Revenue by Age

(a) DI Benefit Receipt

(b) Other Benefit Receipt

(c) Employment

(d) Net Fiscal Effect

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA-59 increase 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.

D.4 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 $\theta^L_s$ to $\theta^H_s > \theta^L_s$. We can write the insurance loss of stricter eligibility rules as
\[ \Delta V_{\theta^*} = \int_i V_i(\theta^L_s) - V_i(\theta^H_s) di \]  
\[ \leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \left( \Delta D_i^W \left( v_i(c^D_{i,t}) - \{u(c^W_{i,t}) - \theta_{i,t}\} \right) + \Delta D_i^Z \left( v_i(c^D_{i,t}) - v_i(c^Z_{i,t}) \right) \right) \right] \] (D.35)
\[ - \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \Delta \Lambda_{i,t} \cdot \psi \right] . \] (D.36)

Expression (D.35) captures the insurance loss due to changes in the labor market status. \( \Delta D_i^W \) and \( \Delta D_i^Z \) 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.35) is an inequality. Note that for a marginal change in \( \theta^* \) behavioral changes would not have a first order welfare effect due to the envelope theorem. Then \( \Delta D_i^W \) and \( \Delta D_i^Z \) would only account for the mechanical changes. In this sense, our relative comparison of the insurance losses is robust to non-marginal changes as \( \Delta D_i^W \) and \( \Delta D_i^Z \) also capture changes in DI levels due to changes in behavior.

We now further bound the insurance loss \( \Delta V_{\theta^*} \) and relate it to the income loss. We have

\[
\Delta V_{\theta^*} \leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \left( \Delta D_i^W \left( v_i(c^D_{i,t}) - \{u(c^W_{i,t}) - \theta_{i,t}\} \right) + \Delta D_i^Z \left( v_i(c^D_{i,t}) - v_i(c^Z_{i,t}) \right) \right) \right] \] (D.37)
\[
\leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \Delta D_i^W \left( v_i(c^D_{i,t}) - v_i(c^Z_{i,t}) \right) \right] \] (D.38)
\[
\approx \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \Delta D_i^W \left( v'_i(c^D_{i,t}) (c^D_{i,t} - c^Z_{i,t}) \right) \right] \] (D.39)
\[
\leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \Delta D_i^W \left( v'_i(c^D_{i,t}) (b_{i,t} - z_{i,t}) \right) \right] \] (D.40)
where (D.37) simply drops the reduction in application costs, $\Delta \Lambda_i \cdot \psi \geq 0$, in (D.35). (D.38) uses that other welfare benefits act as a safety net, i.e. individuals who choose to work cannot be worse off than on other welfare benefits, and we define $\Delta D_{it} \equiv \Delta D^W_{it} + \Delta D^Z_{it}$. (D.39) follows from a first order Taylor approximation and (D.40) 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

$$
\frac{\Delta V_b}{\lambda \mathbb{E}[M_\lambda(\theta^*_i)]} \leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \left( \frac{\Delta D_{it} (b_{it} - z_{it})}{\lambda} \right) \right].
$$

(D.41)

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 $b_{it}^H$ to $b_{it}^L < b_{it}^H$. We have

$$
\Delta V_b \equiv \int \Delta V_i(b^H) - V_i(b^L) di \geq \mathbb{E} \left[ \sum_{t=0}^{T-1} D_{it} v_i'(c_{it}^D) (b_{it}^H - b_{it}^L) \right].
$$

(D.43)

Inequality (D.43) 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

$$
\frac{\Delta V_b}{\lambda \mathbb{E}[M_\lambda(b)]} \geq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \left( \frac{v_i'(c_{it}^D)}{\lambda} \frac{D_{it} (b_{it}^H - b_{it}^L)}{\mathbb{E}[M_\lambda(b)]} \right) \right].
$$

(D.44)
Comparing Insurance Loss of DI Eligibility Rules and DI Benefits. From (D.41) and (D.44) follows that the insurance loss of stricter eligibility rules is smaller than the insurance loss of reduced benefits if

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

(D.45)

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.

D.5 Austria versus US

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 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.. Figure D.7 contrasts the U.S. application behavior to the Austrian application behavior and reveals that the Austrian rules are well known as there is a 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 elastic-
ity $\xi$ 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.\footnote{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”.}

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.
Figure D.7: Application and Award Rate by Age: US vs. Austria

(a) Application Rate

US

Austria

(b) Allowance Rate

US

Austria

Notes: Panel (a) contrasts the US 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 US, 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 US Figures are based on Strand (2016).
S Theoretical Framework

S.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 S.8 – 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, $\theta^R$ 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

$$\psi < p(\theta^R)[v(b) - v(z)].$$

(S.46)

The condition is intuitive: if the DI program is generous (low $\psi$, 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 (S.46) is satisfied.

Condition (S.46) 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 (S.46) 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 S.8: DI Application Model: Decision Tree

![Decision Tree](image)

Note: The figure shows the decision tree of the second period in the static model. If the disability level $\theta$ 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 $v(z) - \psi$).

**Optimal Policy Mix.** In the main text 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 $\theta^*$ be changed per unit change of DI benefits $b$?

The slope of the gradient answers this question of the optimal policy mix. Figure S.9 illustrates the idea of the optimal policy mix. It depicts the current policy $(\theta^*_0, b_0)$. The dotted curve indicates the combinations of $(\theta^*, b)$ that generate the same level of social welfare. Consider the effect of a DI reform starting from the pre-reform DI policy is $(\theta^*_0, b_0)$. The vertical (horizontal) arrow shows how $\theta^*$ (resp. $b$) needs to be changed to increase welfare. In Figure S.9 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.\(^{44}\) The length of the arrows correspond to

\(^{44}\)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.
the efficiency gains associated with the respective policy instrument. In Figure S.9 the horizontal arrow is short, while the vertical arrow is long, suggesting that the DI reform should strongly increase $\theta^*$ per unit reduction of $b$. The slope of the gradient – the arrow pointing to the northwest – yields the optimal policy mix.

Using the first order conditions (5) and (7), the gradient is given by

$$
\nabla W = \begin{pmatrix}
-\partial W/\partial b \\
\partial W/\partial \theta^*
\end{pmatrix} = \begin{pmatrix}
\sigma \cdot M(b) \\
\gamma \cdot M(\theta^*)
\end{pmatrix} u'(w - \tau) \quad \text{(S.47)}
$$

where $\sigma$ and $\gamma$ measure the gap between fiscal gains and insurance loss for changes in $b$ and $\theta^*$, 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

$$
\left. \frac{\partial \theta^*}{\partial b} \right|_{opt} = \frac{\gamma \cdot M(\theta^*)}{\sigma \cdot M(b)}. \quad \text{(S.48)}
$$

The sign of $\sigma$ determines the direction in which benefits should be adjusted (if $\sigma \gtrless 0 \iff -\partial W/\partial b \gtrless 0 \iff \partial b \lessgtr 0$). Similarly, the sign of $\gamma$ determines the direction of adjustment in $\theta^*$ (if $\gamma \gtrless 0 \iff \partial W/\partial \theta^* \gtrless 0 \iff \partial \theta^* \lessgtr 0$). $\sigma$ and $\gamma$ determine the direction of welfare-enhancing adjustments in $b$ and $\theta^*$, while the ratio $\gamma/\sigma$ determines the optimal DI policy mix, $(\partial \theta^* / \partial b)_{opt}$. $b$ and $\theta^*$ 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 $\theta^*$ 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 $\gamma/\sigma$ 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 multi-
pliers to rank the two policies as we show later on.

Figure S.9: 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.

S.2 General Model

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 (8) and (9) that characterize the social optimality of policy parameters $\theta_s^*$ and $b_s$. From above we have

$$\frac{\partial W}{\partial \theta_s^*} = \gamma^* \mathbb{E}[M(\theta_s^*)] \ast \lambda$$  \hspace{1cm} (S.49)

where

$$\gamma \equiv 1 + \frac{\mathbb{E}[B(\theta_s^*)]}{\mathbb{E}[M(\theta_s^*)]} - \frac{\mathbb{E}[L_W] + \mathbb{E}[L_Z]}{\lambda \mathbb{E}[M(\theta_s^*)]}$$  \hspace{1cm} (S.50)

and

$$-\frac{\partial W}{\partial b_s} = \sigma^* \mathbb{E}[M(b_s)] \ast \lambda$$  \hspace{1cm} (S.51)

where

$$\sigma \equiv 1 + \frac{\mathbb{E}[B(b_s)]}{\mathbb{E}[M(b_s)]} - \frac{\mathbb{E}[v'(c_s^D)]}{\lambda \mathbb{E}[M(b_s)]}.$$  \hspace{1cm} (S.52)

The gradient is

$$\nabla W = \begin{pmatrix} -\frac{\partial W}{\partial b_s} \\ \frac{\partial W}{\partial \theta_s^*} \end{pmatrix} = \begin{pmatrix} \sigma^* \mathbb{E}[M(b_s)] \\ \gamma^* \mathbb{E}[M(\theta_s^*)] \end{pmatrix} \lambda.$$  \hspace{1cm} (S.53)
The optimal change in the strictness of DI eligibility rules per unit change of DI benefits is then given by
\[
\frac{\partial \theta^*_s}{\partial b_s}_{opt} = \frac{\gamma}{\sigma} \cdot \frac{E[M(\theta^*_s)]}{E[M(b_s)]},
\]
(S.54)
where \( \gamma \) and \( \sigma \) denote the difference between marginal social cost and marginal social benefit of changing \( \theta^*_s \) and \( b_s \). The signs of \( \gamma \) and \( \sigma \) determine the direction of adjustment.

T Institutional Background and Data

T.1 Lower DI Benefits: The 2003 Reform

Figure T.10 illustrates the effect of the 2003 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 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.

Figure T.10: 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).
## T.2 Summary Statistics Samples

Table T.4: Summary Statistics, RSA Sample

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


Table T.5: Summary Statistics, Benefit-Generosity Sample

<table>
<thead>
<tr>
<th></th>
<th>Ages 57-60</th>
<th>Ages 30-56</th>
</tr>
</thead>
<tbody>
<tr>
<td>DI application (%)</td>
<td>6.10</td>
<td>0.78</td>
</tr>
<tr>
<td>w/ mental disorders</td>
<td>0.47</td>
<td>0.11</td>
</tr>
<tr>
<td>w/ musculoskeletal system</td>
<td>2.02</td>
<td>0.16</td>
</tr>
<tr>
<td>w/ other disorders</td>
<td>3.61</td>
<td>0.52</td>
</tr>
<tr>
<td>DI inflow (%)</td>
<td>4.06</td>
<td>0.35</td>
</tr>
<tr>
<td>Employment outflow (%)</td>
<td>1.55</td>
<td>0.06</td>
</tr>
<tr>
<td>Other benefit outflow (%)</td>
<td>2.51</td>
<td>0.29</td>
</tr>
<tr>
<td>Age (years)</td>
<td>57.35</td>
<td>42.39</td>
</tr>
<tr>
<td>Insurance years</td>
<td>36.98</td>
<td>22.61</td>
</tr>
<tr>
<td>Last annual earnings (Euro)</td>
<td>41,405 39,511</td>
<td></td>
</tr>
<tr>
<td>Avg. annual earnings, best 15 years (Euro)</td>
<td>38,855 42,481</td>
<td></td>
</tr>
<tr>
<td>Blue-collar (%)</td>
<td>49.25</td>
<td>50.68</td>
</tr>
<tr>
<td>No. Observations</td>
<td>1,453,448</td>
<td>15,968,003</td>
</tr>
<tr>
<td>No. Individuals</td>
<td>491,426</td>
<td>1,801,685</td>
</tr>
</tbody>
</table>

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 in 2018 Euro.
U The Effect of Tighter DI Eligibility Rules

Figure U.11 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 U.11: 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.

U.1 Descriptive Figures

Figure U.12 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 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.
Figure U.12: Labor Market States and DI Applications Ever, by Age and RSA

(a) DI Benefit Receipt  
(b) DI Application Ever

(c) Employment  
(d) Other Benefit Receipt

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.

U.2 DI Application Effects

Figure U.13 shows the results for DI application yearly, an indicator for whether an individual has applied for DI benefits at a particular age, Total Applications, the cumulative sum of applications for each individual, and Re-applications, an indicator whether an individual who files a DI application applied before. For the RSA-58 increase in Panel (a) yearly applications drop at age 57 and are higher at age 58. The drop at age 57 is not driven by reapplications implying that only first-time applicants react to the stricter rules. At age 60 the total number of applications is not different under the stricter rules. For the RSA-59 increase in Panel (b) we see that the yearly application are also lower at age 57. However, at age 58 they are on the same level as under lenient rules. With lenient rules at age 57 there are few applications at age 58 as most individuals are awarded DI and no longer apply. This is why the yearly application effect is small at that age. We cannot observe applications of the RSA-59 group above age 59 and total applications did not catch up at that age.
Figure U.13: RSA Effects on Applications

(a) RSA 58

DI application yearly

Total Applications

Re-applications

(b) RSA 59

DI application yearly

Total Applications

Re-applications

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA 58 and RSA 59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval. DI application yearly is an indicator for whether an individual has applied for DI benefits at a particular age. Total applications measure the cumulative sum of applications for each individual at a specific age. Re-applications is an indicator whether an individual who files a DI application applied before.

Figure U.14: RSA Effects by Application Impairment

(a) RSA 58

Mental Impairment

Musculoskeletal System

Other Impairment

(b) RSA 59

Mental Impairment

Musculoskeletal System

Other Impairment

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA 58 and RSA 59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.
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. Figure U.14 plots the estimated $\beta_k$-coefficients from equation (10) for increases in the RSA to 58 and 59. As outcome variables we use whether individuals have ever applied with a mental impairment, 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.

### U.3 Labor Market Transitions

Our estimates show that stricter DI eligibility rules increases employment and other benefit receipt. The increases can result either from changes in the inflow into employment or other benefit receipt, or changes in the persistence in employment or other benefit receipt. To shed light on the importance of these two effects, Figures U.15 and U.16 plots the estimated $\beta_k$-coefficients from equation (10) using as outcome variable transitions from and persistence in employment and other benefit receipt.

The first column of figure U.15 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 individuals who are already employed to stay employed longer. The last column shows an increase in transitions 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 column of figure U.16 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 U.16 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 eligibility 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.

Figure U.15: RSA Effects on Transitions from Employment

To DI

(a) RSA 58

To Employment

To Other Benefits

(b) RSA 59

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA-58 and RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

Figure U.16: RSA Effects on Transitions from Other Benefits

To DI

(a) RSA 58

To Employment

To Other Benefits

(b) RSA 59

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA-58 and RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.
U.4 Fiscal Effects of Stricter Eligibility

Figure U.17: Fiscal Effects of RSA 58 by Age

(a) DI Benefits

(b) Tax Revenue

(c) Other Benefits

(d) Total Fiscal Effect

Notes: The figure shows the estimated \( \beta_k \)-coefficients from the econometric specification in (10) for the RSA-58 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.

Figure U.18: Fiscal Effects of RSA 59 by Age

(a) DI Benefits

(b) Tax Revenue

(c) Other Benefits

(d) Total Fiscal Effect

Notes: The figure shows the estimated \( \beta_k \)-coefficients from the econometric specification in (10) for the RSA-59 increases using the sample of eligible men. The shaded area denotes the 95 percent confidence interval.
U.5 Effects Non-Eligibles (Placebo Regressions)

In Figure U.19, we plot the estimated $\beta_k$-coefficients from equation (10) for men with too little work experience to be eligible for DI under relaxed eligibility rules.

Figure U.19: Effects of RSA by Age, Non-eligibles

(a) DI Benefit Receipt

(b) DI Application Ever

(c) Employment

(d) Other Benefit Receipt

Notes: The figure shows the estimated $\beta_k$-coefficients from the econometric specification in (10) for the RSA 58 and RSA 59 increases using the sample of non-eligible men. The shaded area denotes the 95 percent confidence interval.
V Impact of Benefit Generosity

V.1 Predicted and Matched 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. Figure V.20 plots mean matched DI benefits against mean predicted DI benefits. Actual benefits track our predicted benefits very closely. Figure V.21 plots mean matched DI benefits against mean predicted DI benefits for each year separately.

Figure V.20: Predicted and Matched DI benefits
(a) All Years Pooled
(b) Sample Distribution

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

Figure V.21: Predicted and Matched DI Benefits, by Year

Notes: The figure compares the predicted and matched DI benefits in each year from 2004 to 2014, the last year in which matched DI benefits are available.
V.2 DI Application Effects by Impairment Type

Table V.6 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 significant 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.

Table V.6: Application Effect of DI Benefit Generosity, by Health Impairment

<table>
<thead>
<tr>
<th></th>
<th>Ages 57-60</th>
<th>Ages 30-56</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Mean</td>
</tr>
<tr>
<td>Mental impairments</td>
<td>0.007</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Musculoskeletal system</td>
<td>0.067***</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Other impairments</td>
<td>0.096***</td>
<td>3.17</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,453,448</td>
<td>15,968,003</td>
</tr>
</tbody>
</table>

Notes: The table reports estimates for $\gamma$ from the econometric specification in (12). 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%.

V.3 Placebo Regressions 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. Figure V.22 plots the empirical cumulative distribution of the placebo estimates together with the true estimate for each labor market outcome reported in Tables 2 in the main text and C.2 in the Appendix. 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 V.22: Distribution of Placebo Estimates

(a) DI Inflow

Empirical CDF

(b) DI Applications

Empirical CDF

(c) Employment Outflow

Empirical CDF

(d) Other Benefit Outflow

Empirical CDF

Notes: The figure plots the empirical distribution of placebo effects for labor market outcomes estimated. The CDF is constructed from 1,000 estimates of $\gamma$ from estimation equation (12) 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 reported in Table 2 in the main text and Table C.2 in the Appendix.


W Estimating the Fiscal Multiplier of DI Reforms

W.1 Complier Analysis Stricter Eligibility Rules

Framework. In this section, we describe the complier analysis for difference-indifferences 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 characteristics 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 between age 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 using an indicator $DI$, which is one if an individual is awarded DI benefits and zero otherwise. $DI_0$ and $DI_1$ denote 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 appli-
cants if we assume $AP_1 - AP_0 \geq 0$, the standard monotonicity assumption in the instrumental variables literature.\footnote{The monotonicity assumption rules out defying applicants who would apply when DI eligibility rules are strict but not when DI eligibility rules are relaxed.} We can estimate the conditional value of a never applicant characteristic $E(X|AP_0 = 0, T, A57)$ by the corresponding sample mean

$$\frac{1}{N_{T,A57}} \sum_{i \in (T,A57)} X_i \cdot \mathbb{I}(AP_i = 0),$$

where $i$ is individual and $N_{T,A57}^{na}$ 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 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:

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

(W.55)

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

We can estimate each term of the right-hand side of equation (W.55) empirically. We estimate the shares of each group of applicants with the following regression:

$$AP_{iac} = \alpha + \beta_a + \gamma_c + \delta Z_{ac} + \varepsilon_{iac},$$

(W.56)

where $\beta_a$ is a fixed effect for the age window $a = A57$ and $\gamma_c$ 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)).

We estimate the share of always enrollees ($\pi^{ae}$), marginal enrollees ($\pi^{me}$) and never enrollees ($\pi^{ne}$) in the same way, but simply use $DI$ as the dependent variable in equation (W.56). We estimate the conditional value of an applicant characteristic $E(X|A_{P1} = 1,T,A57)$ by the corresponding sample mean $(1/N_{T,A57}^{a}) \cdot \sum_{i \in (T,A57)} X_{i} \cdot I(A_{P1} = 1)$, where $N_{T,A57}^{a}$ is the number of applicants in $(T,A57)$ and $I(A_{P1} = 1)$ is one if an individual has applied for DI.

Calculating $E(X|A_{P0} = 1,A_{P1} = 1,T,A57)$ 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 $A_{P0}$. 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 $E(X|A_{P0} = 1,A_{P1} = 1,T,A57) = E(X|A_{P0} = 1,T,A57)$. If trends in $X$ are parallel across cohorts and $Z$ is independent from $A_{P}$ and $X$, we can estimate $E(X|A_{P0} = 1,T,A57)$ using the change in applications for cohort $C$, $E(X|A_{P0} = 1,T,A57) = E(X|A_{P0} = 1,T,B57) + E(X|A_{P0} = 1,C,A57) - E(X|A_{P0} = 1,C,B57)$. We can estimate each element on the right-hand side by the corresponding sample mean: $(1/N_{T,B57}^{a}) \cdot \sum_{i \in (T,B57)} X_{i} \cdot I(A_{P1} = 1) + (1/N_{C,A57}^{a}) \cdot \sum_{i \in (C,A57)} X_{i} \cdot I(A_{P1} = 1)$.

**Complier Analysis for RSA 59.** Table W.7 shows the population shares and average characteristics of marginal, always, 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 D.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 (lower) a higher work capacity than always (never) applicants and enrollees.

---

46Formally, the independence assumption is equal to $A_{P0}, A_{P1} \perp Z \mid a, c$ and the parallel trend assumption is equal to $E(A_{P0}|A57,T)-E(A_{P0}|B57,T) = E(A_{P0}|A57,C) - E(A_{P0}|B57,C)$.  

19
Table W.7: Characteristics of DI Applicants and DI Recipients, RSA 59

<table>
<thead>
<tr>
<th>A. Applicants</th>
<th>Marginal (M)</th>
<th>Always (A)</th>
<th>Difference M-A</th>
<th>Never (N)</th>
<th>Difference M-N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share in population (%)</td>
<td>0.033***</td>
<td>0.102***</td>
<td>-0.069***</td>
<td>0.866***</td>
<td>-0.833***</td>
</tr>
<tr>
<td>Sick Leave at age 56 (%)</td>
<td>-0.88</td>
<td>8.38***</td>
<td>-9.26***</td>
<td>0.92***</td>
<td>-1.80**</td>
</tr>
<tr>
<td>Unemployed at age 56 (%)</td>
<td>10.65***</td>
<td>25.53***</td>
<td>-14.88***</td>
<td>4.26***</td>
<td>6.40***</td>
</tr>
<tr>
<td>Employed at age 56 (%)</td>
<td>84.83***</td>
<td>62.80***</td>
<td>22.03***</td>
<td>87.15***</td>
<td>-2.31</td>
</tr>
<tr>
<td>Avg. annual earnings (euro)</td>
<td>41,392***</td>
<td>41,054***</td>
<td>339</td>
<td>46,347***</td>
<td>-4,955***</td>
</tr>
<tr>
<td>Blue-collar (%)</td>
<td>85.87***</td>
<td>80.34***</td>
<td>5.53***</td>
<td>53.89***</td>
<td>31.98***</td>
</tr>
<tr>
<td>Musculoskeletal (%)</td>
<td>74.93***</td>
<td>36.37***</td>
<td>38.57***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental (%)</td>
<td>4.55***</td>
<td>17.83***</td>
<td>-13.27***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (%)</td>
<td>20.51***</td>
<td>45.80***</td>
<td>-25.29***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Enrollees</th>
<th>Marginal (M)</th>
<th>Always (A)</th>
<th>Difference M-A</th>
<th>Never (N)</th>
<th>Difference M-N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share in population (%)</td>
<td>8.00***</td>
<td>3.82***</td>
<td>4.19***</td>
<td>88.18***</td>
<td>-80.17***</td>
</tr>
<tr>
<td>Sick Leave at age 56 (%)</td>
<td>6.40***</td>
<td>9.90***</td>
<td>-3.50***</td>
<td>0.83***</td>
<td>5.58***</td>
</tr>
<tr>
<td>Unemployed at age 56 (%)</td>
<td>23.20***</td>
<td>16.92***</td>
<td>6.28***</td>
<td>4.68***</td>
<td>18.52***</td>
</tr>
<tr>
<td>Employed at age 56 (%)</td>
<td>70.05***</td>
<td>64.31***</td>
<td>5.74***</td>
<td>86.79***</td>
<td>-16.74***</td>
</tr>
<tr>
<td>Avg. annual earnings (euro)</td>
<td>40,852***</td>
<td>42,207***</td>
<td>-1,355***</td>
<td>46,231***</td>
<td>-5,379***</td>
</tr>
<tr>
<td>Blue-collar (%)</td>
<td>87.69***</td>
<td>72.73***</td>
<td>14.96***</td>
<td>54.24***</td>
<td>33.45***</td>
</tr>
<tr>
<td>Musculoskeletal (%)</td>
<td>57.28***</td>
<td>24.91***</td>
<td>32.37***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental (%)</td>
<td>7.43***</td>
<td>25.18***</td>
<td>-17.76***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (%)</td>
<td>30.57***</td>
<td>55.02***</td>
<td>-24.45***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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-59 increase. Earnings are reported in 2018 Euro. Avg. annual earnings measures the average earnings of the best 15 years. Levels of significance: *1%, **5%, and ***1%. 
W.2 Pre-57 Applicants: Representative for Always Applicants?

Table W.8 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 6,000 pre-57 applicants in the control group (RSA 57) and around 3,000 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 W.8: Summary Statistics, Pre-57 Applicants

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

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). RSA 57 comprises men born between December 1953 and November 1955, RSA 58 comprises men born between December 1955 and November 1956, and RSA 59 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 W.23: Applicants at 57 versus Pre-57 Applicants, Treated Cohorts Only

(a) DI Application Rate

(b) DI Award Rate

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 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 W.24: Pre-57 Applicants versus Age-57 Applicants

(a) Employment

RSA 58

RSA 59

(b) Other Benefit Receipt

RSA 58

RSA 59

(c) Tax Revenues

RSA 58

RSA 59

(d) Other Benefit Payments

RSA 58

RSA 59
W.3 Relation between Insurance Loss and Income Loss

Income Losses of DI Eligibility Rules and DI Benefits. As an additional interesting exercise, we look at the (unbounded) income losses of the two policy instruments by income quintile. Figure W.25 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 1,000 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. The 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.

Figure W.25: Effect of Stricter DI Eligibility Rules, by Income Bins

(a) DI Benefits

(b) Earnings

(c) Other Benefits

(d) Total Income

Notes: The Figure plots the estimated income losses by income quintile for the RSA increases by income source. In contrast to Figure 6, we do not implement the bounds on the income loss. 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.

We do the same exercise for lower DI benefits in W.26. 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 DI benefit cut.

In Figure W.27 we present the normalized income losses by income quintile for the RSA increases and benefit generosity from the unbounded income losses from Figures W.25 and W.26.\footnote{The normalized income loss is simply the the total income loss by quintile in Panel (d) in Figures W.25 and W.26 divided by the respective mechanical fiscal effect from Table 3.} Figure W.27 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 larger in all quintiles than that of stricter eligibility rules (the red line is above the blue line, except for the third quintile where the losses are similar).

If the objective function of the planner would not take disutility of work into account, i.e. $\theta_{t,t'}$ would not show up in (D.37), we would have the following relationship...
between the insurance loss and the normalized income loss

\[ \frac{\Delta V_\theta}{\lambda \mathbb{E}[M_\Delta(\theta_*^s)]} \leq \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t \frac{v'(c_{i,t})}{\lambda} \left( \frac{\Delta D_{i,t}^W (b_{i,t} - (w_{i,t} - \tau_{i,t})) + \Delta D_{i,t}^Z (b_{i,t} - z_{i,t})}{\mathbb{E} [M_\Delta(\theta_*^s)]} \right) \right]. \]

This bound corresponds to Figure W.27 where we estimate the normalized income loss by income quintile (taking actual labor earnings into account instead of potential benefits as in (D.41)). Hence, a policy maker who is 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 benefits as it has lower costs (lower income losses across the income distribution), while it creates a higher fiscal cost reduction (higher multiplier).

Figure W.27: 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 W.25 and W.26 divided by the respective mechanical fiscal effect from Table 3. The normalized income loss measures the quintile’s income reduction for a 1 Euro mechanical reduction in fiscal spending.

W.4 Insurance Loss of Stricter DI Eligibility Rules

Implementing the insurance value (the rhs of 8) poses several challenges. First, the utility loss is expressed in differences in utility levels rather than in marginal utilities. Second, the insurance value also depends on the abstract quantity \( \theta \). We tackle this challenges by deriving bounds of the insurance value that do not depend on the unobserved disability level \( \theta \). 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.\(^{48}\) In the following we discuss the derivation of the bounds and the implications of our assumptions. For this

\(^{48}\)We only observe transfers and incomes in our data and cannot measure consumption.
define the insurance value as
\[
\Delta V = \frac{1}{\mathbb{E}[M(\theta^*)]} \mathbb{E} \left[ \sum_{t=1}^{T-1} \beta^t \left( M_{W_{i,t}} \left[ v(c_{i,t}^D) - \left( u(c_{i,t}^W) - \theta_{i,t} \right) \right] + M_{Z_{i,t}} \left[ v(c_{i,t}^D) - v(c_{i,t}^Z) \right] \right) \right]
\]  
(W.57)

In the following we derive lower and upper bounds on this insurance value \( \Delta V \).

**Upper Bound Insurance Value.** The social welfare benefits act as safety net. An agent cannot do worse than being on welfare benefits in all periods. The insurance loss can therefore not be larger than
\[
\Delta V \leq \frac{1}{\mathbb{E}[M(\theta^*)]} \mathbb{E} \left[ \sum_{t=1}^{T-1} \beta^t \left( M_{W_{i,t}} + M_{Z_{i,t}} \right) \left( v(c_{i,t}^D) - v(c_{i,t}^Z) \right) \right].
\]  
(W.58)

This bound assumes that individuals who are screened out are all on welfare benefits. Individuals who decide to work at some points can only do better than being on welfare benefits in all periods and 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
\[
\Delta V \geq \frac{1}{\mathbb{E}[M(\theta^*)]} \mathbb{E} \left[ \sum_{t=1}^{T-1} \beta^t \left( M_{Z_{i,t}} \left( v(c_{i,t}^D) - v(c_{i,t}^Z) \right) \right) \right].
\]  
(W.59)

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 (W.58) we make four assumptions. First, we measure the insurance loss relative to an increase in resources while employed (\( \lambda = \mathbb{E} \left[ \sum_{t=0}^{T-1} \beta^t u(c_{i,t}^W) \right] \) where \( c_{i,t}^W \) is the consumption level of working individuals). This is the standard measure of 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 \( (c_{i,t}^D = b_{i,t}, c_{i,t}^Z = z_{i,t}, c_{i,t}^W = w_{i,t} - \tau_{i,t}) \), 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

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

(W.60)

and

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

(W.61)

We calculate $\mathbb{E}[\sum_{t=0}^{T-1} u'(w_{i,t} - \tau_{i,t})]$ from the data for a given value of risk aversion. We then estimate the mechanical fiscal effect $\mathbb{E}[M(\theta^*)]$ in Section 6. In (W.60) 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 $\gamma$ 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 $\mathbb{E}\left[ \sum_{t=s}^{T-1} (M_{W_{i,t}} + M_{Z_{i,t}}) \frac{1}{1-\gamma} ((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}) \right]$.

To estimate $\mathbb{E}\left[ \sum_{t=s}^{T-1} M_{Z_{i,t}} \frac{1}{1-\gamma} ((b_{i,t})^{1-\gamma} - (z_{i,t})^{1-\gamma}) \right]$ in (W.61) 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 an insurance loss if $M_{Z_{i,t}} = 1$). We then calculate for a given risk aversion $\gamma$ the
utility $u_{it} = \frac{1}{1-\gamma}(l_{it})^{1-\gamma}$ and run our DiD strategy on this outcome variable $u_{it}$.

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 W.28. 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 W.28: Stricter Eligibility Rules

(a) RSA 58

(b) RSA 59

Notes: Figure plots the LHS and the upper and lower bounds of the RHS of inequality (8) 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.

W.5 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

$$
\mathbb{E} \left[ \sum_{t=s}^{T-1} (b_{it})^{-\gamma} \right] / \mathbb{E} \left[ \sum_{t=0}^{T-1} (w_{it})^{-\gamma} \right]
$$

(W.62)
We can directly calculate this for different values of risk aversion based on the pre-reform benefit levels. Figure W.29 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 W.29: Welfare Effects Benefit Generosity, Men

(a) Ages 57-60

(b) Ages 30-56

Notes: Figure plots the LHS and RHS of inequality (9) 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.

W.6 Optimal Policy Mix

For the gradient we can use the implementation from above to express $\gamma$ and $\sigma$ from equation (S.54) as a function of risk aversion. For $\gamma$ 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 W.30 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 $\frac{d\sigma^*}{db} \approx 0.1-0.2$). 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 W.30: Optimal Policy Mix: Gradient

Notes: Figure plots the optimal combination of changing benefit generosity and eligibility, i.e. the direction $\gamma$ from equation (S.54). The red lines indicate the critical risk aversion values for optimal DI benefits and eligibility rules.