Estimating the Moral Hazard Cost of Private Disability Insurance and its Welfare Consequences

Sebastian Seitz*

University of Mannheim

November 29, 2021

Job Market Paper

Current version

Abstract

Although one-third of workers in the USA and Germany contract supplementary private disability insurance (DI), most studies on the design of public DI systems abstract from private DI. Using unique and comprehensive contract data from a major German insurance company and representative survey data, I add novel insights on the interaction between private and public DI by estimating a rich dynamic life-cycle model with private insurance choices. I find that the welfare-improving public DI schedule is less generous in the presence of supplementary private DI. This is a consequence of the additional moral hazard private DI take-up imposes on the public system. I show that while having a private DI market is welfare-improving under the current public DI schedule in Germany, private DI markets can be welfare-reducing for more generous public DI systems as observed in other countries.

JEL codes: D14, G22, J28, I38

Keywords: Disability Insurance, Social Welfare Programs, Labor supply of disabled

* University of Mannheim and ZEW, L7 3-5, 68131 Mannheim, Germany, sebastian.seitz@gess.uni-mannheim.de

I thank Sebastian Siegloch, Eckhard Janeba, Cormac O'Dea, Karen Kopecky, Fabian Kindermann, Arthur Seibold, Karl Schulz, Antonio Ciccone, and seminar participants at the University of Mannheim and the Bonn DSE Summer School for their helpful comments. I am grateful to Christian Hilscher, Andreas Böhm, and Dorian Weber who provided excellent research assistance.

1 Introduction

Disability poses a substantial risk over the life-cycle. One in four adults in the U.S. and Germany experiences a disability spell before reaching the retirement age (Aktuarvereinigung, 2018; CDC, 2020). While individuals may retain some of their initial productivity despite their disability (Borghans et al., 2014; Kostol and Mogstad, 2014), it still persistently limits the amount and intensity of work they can perform, thus greatly reducing lifetime income while resulting in greater medical spending needs, e.g., for care services. To alleviate some of its risk, all OECD countries provide public disability insurance (DI). In addition, individuals can contract supplementary private DI in many countries, which allows them to top-up public benefits. For instance, in Germany 34.7% of all employees in the private sector have private long-term DI.[1]

Despite the size of private DI markets, there is little empirical evidence on their interaction with public DI policies. In this paper, I provide new evidence on this interaction by analyzing how private DI affects the design of public DI policies and by quantifying the underlying labor supply channels. My analysis makes two contributions. First, I extend the existing literature by explicitly modeling the interaction between private and public DI. Although the importance of this interaction between overlapping private and public insurance has been formally shown (Chetty and Saez, 2010; Golosov and Tsyvinsky, 2007; Pauly, 1974), the empirical DI literature largely abstracts from it (the few notable exceptions are mentioned below). I show that private DI substantially alters the welfare implications of public DI policies and thus their optimal design. Second, I show that private DI take-up can generate substantial additional moral hazard costs by increasing retirement at disability onset, adding to the little existing evidence on the moral hazard cost of private DI (Stepner, 2019). I term the additional labor supply distortions from private DI take-up the *moral hazard of private DI* (see, e.g., Chetty and Saez (2010)).[2]

Public DI schedules have to trade off the provision of disability insurance with incentives to continue working if productivity remains sufficiently high despite the disability (Chetty and Saez, 2010; Diamond and Sheshinski, 1995). I study how introducing private DI alters this trade-off and, therefore, the design of welfare-improving public DI. In particular, I examine how the generosity of public DI benefits and screening stringency affect welfare through private DI take-up and labor supply. For example, making public DI less generous reduces the moral hazard from public DI (fewer people retire), but can increase private DI take-up and thus the moral hazard from private DI (more people retire due to greater total transfers). The total moral hazard response (more/fewer claimants) and

[1] U.S.: 35% (Labor Statistics, 2020) UK: 3% of women and 6% of men (Statista, 2019); Austria: 4% of the population (Kaniovski and Url, 2019). Numbers are for the whole population, conditional on being employed in the private sector.

[2] A second commonly studied channel quantifies the moral hazard from asymmetric information about the true health of an applicant (Low and Pistaferri, 2015). Allowing for this channel amplifies the moral hazard cost in my model, so my results constitute a conservative lower bound estimate.

consequently welfare then depend on the relative size of both responses and are a priori unclear.

The size of the moral hazard from private DI depends on the share of individuals purchasing private DI and their individual retirement decision to private DI coverage at disability onset. Thereby, the moral hazard of private DI acts on top of the moral hazard from public DI. To quantify these responses and to make welfare predictions, I build a rich life-cycle model in which people endogenously choose their labor supply, consumption, savings, and private DI coverage. Individuals are subject to disability shocks which persistently lower their labor productivity and qualify them for public and if covered private DI benefits while still maintaining potentially enough of their initial productivity to allow for gainful employment. The model contains a detailed approximation of German social insurance programs and private insurance contracts to precisely quantify the interaction between the different programs.

I calibrate my model using the method of simulated moments, which matches data moments to the corresponding moments simulated from the model. A major challenge for estimating the preference parameters is that one needs data on both private DI take-up in the population and information on the design of individual private DI contracts. I overcome this challenge by combining data from two sources. First, I estimate the private DI take-up in the population from a representative household survey, which has collected this information from 2013 on. These are the key moments in my estimation and my model has to closely match private DI take-up for the whole population and conditional on income quartiles. Second, I use confidential contract data from a major German insurer to approximate the private DI market. This allows me to estimate the replacement ratio, model private DI pricing, and speak to risk heterogeneity in the population. Finally, I use administrative social security records to supplement the two data sets with detailed information on income and occupational risk distributions. Based on the model solution, I study the interaction of private and public DI for revenue-neutral changes in public DI benefit generosity and screening stringency.

My first set of results characterizes welfare-improving public policies in the presence of a private market. I show that the welfare-improving public DI schedule is relatively less generous with private DI compared to the setting with only public insurance. This corroborates the formal results from Chetty and Saez (2010) empirically. Specifically, the results show that in presence of a private DI market, public DI should impose a higher rejection rate or lower public DI benefits relative to the respective policy schedule without private DI. In addition, I find that private DI markets can change the direction of welfare-improving policies: whereas increases in benefit generosity relative to the statutory benefit level are welfare improving absent private DI, benefit reductions provide the larger welfare gains in the presence of private DI.

The change in welfare predictions is explained by the two behavioral channels mentioned above, i.e., private DI take-up and the size of the underlying moral hazard response. On the one hand, private DI take-up changes the total insurance value (private + public DI) and thus the welfare of individuals. On the other hand, private DI take-up distorts the labor supply decision at disability onset as the additional transfer income makes retirement more attractive.[3] Welfare only improves, if the increase in insurance value can offset the fiscal externality from changes in private DI take-up and the resulting retirement decision at disability onset (Chetty and Saez, 2010). Therefore, it matters who starts/stops buying private DI and how sensitive their labor supply choice is to private DI take-up. For instance, the data shows that private DI is concentrated among high-income individuals under the current public DI schedule (Figure 1): 33% of people in the first income quartile purchase private DI compared to 66% in the fourth quartile. However, individuals in the fourth income quartile display a greater moral hazard response to private DI coverage in my model: a greater share of them stays employed at disability onset absent private DI relative to low-income individuals, who are more likely to retire independently of private DI coverage.[4] Hence, I find that the welfare gains are smaller (or might even be negative) for public DI policies, where private DI is concentrated among the high-income individuals (large fiscal externality). Since this happens in the direction of more generous public policies (fewer rejections/more benefits), public DI has to be less generous in the presence of private DI markets explaining the results above.

The second set of results extends the discussion to the question of whether having a dual system, i.e. a private DI market, is always welfare-improving. I answer this question by studying the same policy experiments as above but comparing welfare across private DI availability. The results show that a dual system is always welfare-improving for all considered rejection rates, but there is a substantial range of benefit levels over which having a private DI market is welfare-reducing: a dual system is only welfare-improving for low benefit generosity, for example as under the status quo in Germany, but welfare-reducing for more generous benefits.

Again, these results are explained by the correlation between private DI take-up and income: for more generous public DI benefits, private DI coverage is increasingly concentrated among high-income individuals. Since these individuals are more productive, a greater share of them stays employed absent private DI relative to low-income individuals/non-private DI owners but retires with private DI coverage. Moreover, they pay more taxes and social security contributions, which also entitles them to greater benefits. Taken together, although fewer people purchase private DI, the marginal private DI buyer is more costly to insure in the public system relative to the average individual. Since the greater public program costs need to be financed via the tax system, all individuals have to pay higher contributions to the public DI system and the cost increases offset

[3] Intuitively, the additional private benefits distort the price of leisure: leisure gets cheaper, so people substitute labor force participation for leisure.

[4] This retirement pattern is a consequence of both higher retained productivity levels for high-income individuals as well as the progressivity of the public DI schedule, which provides a higher replacement ratio to low-income types compared to high-income types.

the welfare gains from more generous public DI benefits. In contrast, private DI take-up hardly responds to changes in the rejection rate, thus the moral hazard cost of the private market remains modest. As a result, having a dual system is welfare-improving for all considered rejection rates.

This second analysis offers relevant insights beyond the German setting, as many countries struggle with the sustainability of their public DI programs (Autor and Duggan, 2006). Since these countries often have a supplementary private DI market, my results offer new input to this debate. I illustrate this point for the U.S. and Austria, which have been frequently studied in the public DI literature (e.g. Haller et al. (2020) and Low and Pistaferri (2015)). Applying the respective public DI schedules in my model, I find that both countries implement policies that are most likely too generous. Based on my analysis, they could increase welfare by adapting alternative policies which limit the fiscal externality from private DI: either mechanically by reducing the generosity of public DI or by imposing alternative regulation, e.g., by including private DI income into the means-test at public DI application (see Golosov and Tsyvinski (2006)). However, these results should be interpreted with caution because they are derived under the model calibrated for Germany and need to be verified in the respective settings.

To the best of my knowledge, my work is the first to comprehensively study the interaction between private and public DI in a single framework. Leveraging the confidential contract data to model the private market, I add to the literature on DI by relating the fiscal externality from private DI coverage to welfare-improving public policies. Thereby, I combine insights from the public and private DI literature. More broadly, I also contribute to the dual insurance literature by empirically quantifying its formal predictions in the context of disability insurance.

I extend the empirical DI literature, which has so far abstracted from private DI. My work is most closely related to the literature applying structural (Bound et al., 2004; Chandra and Samwick, 2005; Low and Pistaferri, 2015; Waidmann et al., 2003) and sufficient statistic (Diamond and Sheshinski, 1995; Haller et al., 2020) approaches to characterize welfare-improving or optimal public policies. Applying a model similar to Low and Pistaferri (2015), I show that the interaction between private and public DI has sizeable and economically meaningful consequences for the design of welfare-improving public DI policies. Abstracting from private DI underestimates the moral hazard response to public DI reforms, which leads to the implementation of too generous and sup-optimal public DI schedules. Since the sustainability of public DI programs is usually a key concern in these models (and reality), abstracting from private DI results in too expensive programs. In this sense, my results also add to the dual insurance literature, which has formally characterized the optimal public policies in overlapping insurance settings (Chetty and Saez, 2010; Golosov and Tsyvinsky, 2007; Pauly, 1974). My results empirically corroborate their findings in the context of DI and are similar to the findings of Cabral and Mahoney (2018), who study private and public health insurance of the elderly in the U.S.

Moreover, I add to the small yet growing literature on private DI which has primarily focused on quantifying the moral hazard inherent to private DI coverage or the valuation for public DI in a reduced form fashion. Most closely related to this paper is Stepner (2019), who finds that private short-term DI has increased public long-term DI inflow in Canada by 33% and program cost by 5%, imposing a sizeable fiscal externality. I find a similar response studying the schedule in Germany where private DI coverage reduces the labor supply by 50%. Autor et al. (2014) find that the plan parameters of employer-provided private DI in the U.S. significantly affect DI accession, where longer waiting periods or smaller replacement rations deter claims. I complement their analysis by relating the moral hazard response from private DI to the design of public DI schedules. In contrast, Seibold et al. (2021) use the abolition of own-occupation public DI in Germany and the subsequent increase in private DI take-up, to estimate the willingness-to-pay for public own-occupation DI. Their results show that while privatizing own-occupation DI can be optimal for rational agents, equity concerns and behavioral frictions can still call for a public mandate. Studying (any occupation) public DI in the U.S., Cabral and Cullen (2019) find that social insurance is valued at least at 2.5 its cost using price variation in employer-provided private DI schedules. I complement their work by discussing the interaction between public and private DI for alternative public DI schedules and how it translates into welfare-improving public DI policies.

More broadly, my paper is related to the literature which studies how public DI compensates individuals for working in high-risk jobs (Jacobs, 2020; Michaud and Wiczer, 2018); the incentive effects of public DI on earnings and employment (e.g. Autor et al. (2019), Gelber et al. (2017), Meyer and Mok (2019), Mullen and Staubli (2016), and Ruh and Staubli (2019)), and the productivity of (rejected) claimants (e.g. Borghans et al. (2014), Bound (1989), French and Song (2014), Kostol and Mogstad (2014), and Wachter et al. (2011)).

The rest of the paper is structured as follows. Section 2 introduces the institutional settings of the public disability insurance system as well as the private insurance market in Germany. Section 3 presents the model and section 4 the data. The estimation procedure is detailed in section 5. The estimation results and counterfactual exercises are discussed in sections 6 and 7 respectively. Section 8 concludes

2 Institutional Settings

The German public DI is part of the public pension system since its establishment in the late 19th century. Contributions are made via the payroll taxes for private-sector employees. Since civil servants and self-employed are not subject to social security contributions, they are not entitled to public DI and are not further studied in this paper.

Public pension contributions have to be made for at least 5 years to be eligible for public

DI benefits. Upon meeting this formal criterion, a medical assessment of the work limitation follows: To qualify for public DI, the existing health condition has to be persistent, i.e., is unlikely to improve within the next years[5], and to severely limit labor productivity. An individual is entitled to the full benefit amount in Germany if she cannot work more than 3 hours per day in *any* job independent of her past occupations (similar to the U.S.).[6] Rejections at this stage are common: 44% of all applications are rejected on average. For successful applications, the benefits are computed following the formula for old-age pension benefits adjusting for missing contributions, and discounting for early retirement (see appendix E.4). The average replacement ratio of public DI amounts to 35% of past gross income (see Table 2), while the public DI schedule is progressive. The actual replacement ratios are greater/lower for very low/high incomes because of defined minimum (Social Assistance) and maximum (Social Security Contribution limits) benefits.[7]

In addition to mandatory public DI, individuals can purchase supplementary private DI, which is an individual insurance directly bought from an insurer. As individual insurance, private DI differs along some noteworthy dimensions from mandatory public DI.

First, public DI covers all employees and charges a single (average) price independent of risk, i.e., risk pooling. In contrast, private DI charges risk-based premiums, separating risks into different contracts. The individual disability risk is primarily assessed via the occupation at application. The insurer maps occupations into discrete risk groups based on observed disability risk, e.g. from "1" (best) to "5" (worst), and a higher risk group translates into a higher premium.[8] This premium is expressed as the price to insure €1, so the final price is the product of the risk group specific premium and the contracted benefit. The benefits are freely contractible and designed as an annuity paid until at most the legal retirement age. They are capped at 70% of current gross income with an average replacement ratio of 36% (Table 2).

Second, the occupation-based risk assessment is complemented by a thorough health survey determining whether an individual is insurable. The health survey asks for the applicant's health history as well as diseases running in the family, e.g., cancer or high blood pressure. To confirm the statements, the insurance company can contact the primary physician. Untruthful statements at this stage can lead to loss of insurance coverage after purchase when discovered. Nonetheless, only 4% of all applications get rejected at this stage (GDV, 2016). Thus, I am going to abstract from this in my model later.

Third, the medical work-limitation criterion is less strict in private DI: An individual is disabled if she can no longer work for more than 50% of her usual hours in her previous occupation. This definition assesses disability based on education and past career, thus

[5] Alternatively, the health condition has already existed for 19 months and no improvement has been observed.

[6] Being able to work between three to six hours per day qualifies her for a partial claim, i.e. 50% of a full claim. Since partial claims are constituting less than 10% of all claims in any given year (Bund, 2017), I focus on full claims only.

[7] Both of these features are included in the model in section 3.

[8] See Seibold et al. (2021) for a discussion of priced (risk groups) vs. non-priced risk in private DI.

constituting an *own-occupation* DI. In contrast, public DI not only requires a greater productivity loss of 62.5% but also requires that she is no longer able to work in *any* occupation (independent of education and past career). Consequently, accession to private DI is relatively easier than accession to public DI for a given disability. The fact that rejections from private DI are less common reflects this: Only 11% of claims are rejected for not meeting the health criterion, while most are rejected because people either recovered/died before the first benefit award (11%) or lied in their health survey at application (7%) (GDV, 2014). In my model, I deal with these differences by assuming that the health impairment always meets the minimum criterion for private DI. such that there are no rejections from private DI, whereas rejections from public DI are still possible.
Finally, private and public DI receipts are independent of each other: neither admission nor benefit amount is conditional on getting the other transfer. Thus, private DI coverage is not reduced for public DI receipts as is the case in the U.S. (Autor et al., 2014).

3 Model

My quantitative model concentrates on individual choices with respect to labor supply, consumption, savings and insurance decisions with exogenously given private insurance contracts. Individuals are subject to exogenous income and health shocks. In my analysis I focus on the question how the labor supply response to disability shocks depends on private insurance ownership. Based on these insights, I discuss the implications of this labor supply channel on the design of welfare-improving public disability insurance systems in the presence of private insurance markets.

3.1 The individual problem

An individual lives for a maximum of T periods and works for the first $T_{retire} < T$ periods, while being retired for the rest. In each period, she maximizes her expected life-time utility V_{it} over her choice variables X_{it} conditional on the state variables S_{it}. The choice variables in each period are consumption c_{it}, leisure l_{it} (in retirement always equal to time endowment), and savings for the next period A_{it+1}. At entry into the model, $t = 0$, an individual can choose to purchase a private DI contract: $pDI_0 = 1$ if she buys and zero else.[9] Private DI insures an individual against disability shocks up to the retirement age T_{retire} by paying the premium p_{it} in each period, in which she do not claim. If an individual is hit by a disability shock, she can choose to continue working or to retire

[9]In a robustness exercise I add an intensive margin choice, allowing people to choose from a menu of private DI contracts. In this setting, $pDI_0 \in \{0, 1, ..., L\}$ denotes the chosen contract as specified by the replacement ratio. $pDI = 0$ denotes a zero replacement rate-zero price contract, i.e. no private DI coverage.

($l_{it} = 0$), thus claiming public and, if purchased, private DI. The state variables S_{it} are current assets, A_{it}, income, y_{it}, health status H_{it}, private DI ownership, pDI_{it}, and the individual health risk group, rg_i,[10] and, if an individual is disabled and retired from the workforce, whether or not she was admitted into public DI, DIs_{it}. Finally, an individual faces a mortality risk in retirement, so there is an additional state M_{it} for all $t > T_{retire}$. Formally, an individual maximizes her expected lifetime utility by solving the following problem:

$$\max_{\{c_k, A_{k+1}, l_k\}_{k=1}^{T}, pDI_0} V_{i0} = \sum_{t=0}^{T} \beta^t \mathbf{E}[U(X_{it}; S_{it})] \tag{1}$$

where β denotes the discount factor and U_{it} the period utility function. Expectations are taken with respect to the information available to the individual in period t, i.e. the health and, in retirement, mortality risk (section 3.2), rejections from public insurance (section 3.3), and income risk (section 3.4). I assume that people enter the model at age 25 ($t = 0$), retire at age 65 ($T_{retire} = 39$), and live at most to the age of 95 ($T = 70$).
An individual maximizes V_{it} subject to the intratemporal budget constraint, given her time endowment and the borrowing constraint:

$$\begin{aligned} \frac{A_{it+1}}{1+r} + c_{it} + pDI_{it} * p_{it} &= A_{it} + y_{it} + y^s_{it} - SSC(y_{it}) - SSC(y^s_{it}) - TAX(y_{it} + y^s_{it}) \\ l_{it} &= L - hours_{it} - \theta \mathbf{1}[hours_{it} > 0] \\ A_{it} &\geq 0 \end{aligned} \tag{2}$$

The intratemporal budget constraint requires that each period's expenses are covered by the disposable income in the same period. Expenses include consumption, savings for the next period discounted by the real interest rate net of capital taxes r, and the private insurance premium, which is zero if individuals do not own insurance ($pDI_{it} = 0$) or are currently claiming it ($p_{it} = 0$). Disposable income comprises current savings A_{it}, income y_{it} and spousal income y^s_{it} net of social security contributions $SSC()$ and income taxes $TAX()$, which are modelled according to their actual schedule (see appendix E). Social security contributions are paid individually, while household income is taxed jointly. I describe the income process in section 3.4.
The second constraint in (2) formalizes the individual time constraint. In each period an individual has M hours, which it can spend on working hours, $hours_{it}$, or consuming leisure. The term θ captures the additional disutility from labor force participation $\mathbf{1}[hours_{it} > 0]$, which is estimated in the model. In the data, I only observe whether an individual works full- or part-time, but not the hours. Therefore, I set $hours_{it}$ to 1 if an individual works full-time, to 0.5 for part-time, and to 0 otherwise. I set $M = 3$, as the standard work contract specifies 8 hours a day as full-time work. This implies that a

[10] As discussed in section 2, insurance companies map occupations into discrete risk groups, which capture risk heterogeneity, but also correlates with income, so I add risk heterogeneity as an additional state to my model (see also Michaud and Wiczer (2018)).

full-time worker spends 8 hours working out of 24 hours a day. In mandatory retirement ($t > T_{retire}$), people consume their entire time endowment M as leisure.
The third constraint is the borrowing constraint: Individuals cannot borrow against their future income and thus can only save.
In solving the model, I assume that the per-period utility $U_{it}(X_{it}; S_{it})$ takes the form of CRRA preferences:

$$U(c_t, l_t; H_t) = \frac{(c_{it}^{\kappa} l_{it}^{1-\kappa} e^{-\varphi * \mathbf{1}[H_t = bad]})^{1-\gamma}}{1-\gamma} \tag{3}$$

where γ denotes risk aversion, κ the weight on consumption relative to leisure, and φ expresses the (dis-)utility from disability (Low and Pistaferri, 2015). Intuitively, φ informs us about how individuals would move consumption across health states if fully insured. A positive value of φ implies that people value an Euro of consumption moved from the good health state at more than this one Euro in the bad health state, e.g. reflecting higher needs in the disabled state, thus disability being a 'bad'. The values of γ, κ, φ, and θ (from the time constraint) are estimated from the data below.
While the model accounts for both secondary earners and household composition (via an equivalence scale adjusting consumption), it treats both of these variables as exogenous. In general, it is possible to include the choices of secondary earners into the model, but since I cannot observe them in my data, I abstain from doing so. Moreover, secondary earner's income and labor supply responses to a disability of the primary earner are contested in the literature, which finds positive, negative and no responses to disability shocks (Autor et al., 2019; Gallipoli and Turner, 2009; Lee, 2020).
The model described above has no analytical solution, thus it needs to be solved numerically with the methods detailed in appendix A.

3.2 Health risks

Health and health risk play an integral part in my model. Health directly affects utility and optimal consumption levels via the utility function. In addition, disability reduces labor market productivity via the income process causing people to adjust their labor supply choices. This section discusses the health measure and transition across the health states, while section 3.4 focuses on the implications for income.
I model the health process as a two-state Markov-process: People are either in good health or disabled in any given period t. They move between these states with probability $\Pi_{it}(H_{t+1}; H_t, rg_i)$, which depends on age t, current health H_t, and their risk group rg_i. Since the primary data source for these transitions by the German Actuary Society[11] (Aktuarvereinigung, 1997) only conditions on age, I need to adjust them for risk group

[11] This table serves as a baseline for insurance companies' risk calculations as well, when calculating their risk premia.

specific disability risk.[12] Thus, I estimate the following probit model for disability risk on the discrete risk group rg_{it} on social security registry data:

$$disabled_{it} = \Phi(\zeta_0 + \zeta_1 * rg_{it}) \tag{4}$$

I compute risk group specific adjustment factors for each risk group as the ratio of its predicted disability probability relative to the predicted probability of risk group 3, the mean and median risk group in the population. The transition probability across health states becomes:

$$\Pi_{it}(H_{t+1}; H_t, rg_i) = \pi(H_{t+1} = J|H_t = j) * \frac{dis\hat{a}bled(rg_{it})}{dis\hat{a}bled(rg = 3)} \tag{5}$$

for $J, j = good, disabled$.

Similar to Low and Pistaferri (2015), the process described in equation (5) allows for recovery, so disability is not an absorbing state. While recovery probabilities differ with age (recovery is more likely at younger ages), I assume that recovery probabilities are identical across risk groups because I do not have the power to detect any heterogeneity in recovery probabilities due to small sample sizes.

Finally, the actuarial table ends at age 70 which is less than my maximum age of 95. Therefore, I estimate the transition probabilities for the last 25 years based on a linear regression model with a cubic age polynomial accounting for the non-linear growth of the disability risk at higher ages. I estimate this model based on the last 17 years prior to retirement for both disability risk and recovery probabilities[13].

Besides disability risk, individuals also face mortality risk in retirement, where death is an absorbing state providing zero utility. Individuals survive period t with probability s_{it} conditional on surviving period $t-1$. While people do not die during the working life, I adjust the survival probability for experiencing retirement by computing the probability of dying before the age of 65, so retirement is an uncertain state in itself. The survival probabilities are taken from the mortality table by the German Federal Statistical Office (German Federal Statistical Office, 2016).

Finally, my analysis abstracts from adverse selection as all variation in risk is captured by the observable risk groups and there is no (unobserved) *within* risk group variation. In general, my model can accommodate adverse selection as well, but Seibold et al. (2021) show that in the German private DI market all selection is on observable (priced) risk, i.e. the risk groups, despite some remaining disability risk heterogeneity within each risk group. Given their findings, I control for observable risks via the risk groups but abstract from unobserved within risk-group heterogeneity.

[12] See section 2 for details on the risk group assignment.

[13] This assumption is similar to the one chosen by insurance companies which estimate the risk at higher ages based on a quadratic polynomial on a number of pre-retirement years using a slightly different objective function. The outcomes, however, are close.

3.3 Private and public disability insurance

In this section, I describe how the private and public DI are modelled based on the institutional setting discussed in section 2.

Private DI is characterized by a risk group specific price $ppE(rg_i)$ to transfer one Euro of income into the disability state[14] and a replacement ratio $RR^{private}$. The total premium $p_{it}(rg_i)$ is defined as:

$$p_{it}(rg_i) = ppE(rg_i)RR^{private}Y_{it}(H_{it} = good). \tag{6}$$

Private benefits are defined as a constant fraction $RR^{private}$ of full-time income in good health $Y_{it}(H_{it} = good)$. Subsection 4.2 and 5.2 explain how $RR^{private}$ and $ppE(rg_i)$ are estimated.

The private insurance choice is modelled as a single decision at entry into the model: Individuals can choose to purchase supplementary private DI after observing their risk group and income.[15] They buy insurance if their expected life-time utility with insurance exceeds the expected life-time utility without. Once purchased, individuals cannot withdraw from their initial choice. They pay their risk group specific price $p_{it}(rg_i)$ while working and are entitled to private benefits once disabled and retired from the labor force. The benefit entitlement lasts until they return to the labor force either due to recovery, gainful employment or retirement.

Since there is no data available for Germany which contains information on wages, employment and disability status, I cannot estimate the productivity reduction as e.g. Low and Pistaferri (2015) but have to assume it. In my baseline estimation, I assume that disability shocks are perfectly observable and reduce the labor productivity by 56%, which exceeds the required 50% reduction in productivity for private DI entitlement. Thus, there are also no rejections from private DI. I assess the sensitivity of my results with respect to this assumption in section 7.3.

The public DI system is modelled in a similar fashion characterized by a replacement ratio and a rejection rate. As for private DI, public DI benefits replace a fixed fraction RR^{public} of labor income in good health[16], which I take directly from the data (see section 5.1):

$$benefits^{public} = RR^{public} * Y_{it}(H_{it} = good). \tag{7}$$

Equation (11) in the next section displays the total DI benefit amount. Recall that private and public benefits can be simultaneously claimed without benefit reduction.

In contrast to private DI, rejections of public DI applications are frequent (44% on

[14] Recall from section 2 that prices are linear in the benefit level and thus can be expressed as a 'price-per-Euro'.

[15] Modelling private DI purchase as a once in a life-time decision is motivated by the data: the mean (median) age at purchase is 30.5 (29) years and 75% of people buy before the age of 36.

[16] Note that both the public and private disability insurance benefits replace a fixed ratio of the current labor income. This means none of them is preferable with respect to reducing income fluctuations.

average). As detailed in section 2, reasons for rejections are the failure to meet the minimum contribution period or the minimum health requirement, requiring a 62.5% reduction in labor productivity. I model the rejection probability as a constant term $Prob(DIs_{it} = 0|DIs_{it-1} = 0, H_t = bad)$[17] , where DIs_{it} is a dummy variable that takes the value 1 if individual i is admitted into public DI in period t and zero else. The probability of admission if an individual is in good health is always zero. This implies that there are no false acceptances (healthy people claiming public disability insurance), but only false rejections.

Once admitted to public DI, people cannot be removed from it while still being disabled. People leave public DI either upon recovery, for work, or to retirement, where departure from public DI for the former reasons restarts the admission process upon the next application.

3.4 Income Process

Individuals receive income from three different sources: Labor income, public and/or private disability insurance benefits, and social assistance income if eligible. Individual income is complemented by spousal income which is assumed to exogenous. In this section I describe each income source in detail.

Labor income is modelled as a function of observable characteristics and two i.i.d. shock processes:

$$\log Y_{it} = \beta_0 + \sum_{k=1}^{4} \beta_k * age_{it}^k + \beta_5 \mathbf{1}[hours_{it} = 1] + \beta_6 * \mathbf{1}[rg_{it} = rg]) + \varepsilon_{it} + \epsilon_{it}. \quad (8)$$

Y_{it} denotes the annual income in 10,000 Euros. The reduced-form specification controls for a quartic polynomial in age age_{it}^k for $k = 1, ..., 4$, a full-time dummy $\mathbf{1}[hours_{it} = 1]$, which captures the wage premium from working full-time relative to part-time, and a dummy for the individual risk group $\mathbf{1}[rg_{it} = rg]$). ε_{it} denotes a persistent shock process of income innovations following an AR(1) process (Guvenen, 2009; Low et al., 2010):

$$\varepsilon_{it} = \rho \varepsilon_{it-1} + \eta_{it} \quad (9)$$

where $\eta_{it} \sim N(0, \sigma_\eta^2)$ and ρ denotes the shock persistence. The persistent shock captures time-varying shocks to productivity unrelated to health, e.g. changes in wages due to technological change. In contrast, the transitory shock ϵ_{it} captures period-to-period fluctuations in productivity, such as temporary fluctuations in wage rates. I assume it is normally distributed with mean 0 and variance σ_ϵ^2. The parameters governing the shock processes, $\{\rho, \sigma_\eta^2, \sigma_\epsilon^2\}$, are estimated directly from the data as described in section 5.2.

[17] Low and Pistaferri (2015) model the rejection probability as age dependent. The German pension fund only records the total number of rejections in any given year, I cannot allow for any heterogeneity in this variable.

Controlling for the risk group in the income process is important to link income to risk and selection into private DI coverage (Seibold et al., 2021), which has relevant welfare effects as shown in section 7. As most individuals do not change the risk group over their working life[18], I cannot simultaneously allow for risk groups and individual fixed effects. Therefore, I have decided to estimate the risk-income gradient, given its importance for selection patterns into private DI coverage.

Moreover, I cannot directly control for health in equation (8) because the health status only gets recorded in the data for disability related withdrawals from the labor force. Thus, I either observe benefit receipt (health) or labor income but not both. Instead, I assume that a disability shock reduces individual productivity to 44% of the productivity in good health. The labor income with disability is then also 44% of the income from equation (8).[19]

Spousal income y^s_{it} is modelled as an exogenous source of household income which depends on own age age_{it} controlled for by a quartic polynomial and the partner's log income in good health:

$$y^s_{it} = \beta^s_0 + \sum_{k=1}^{4} \beta^s_k age^k_{it} + \beta^s_5 log(Y_{it}) \tag{10}$$

The specification implies that spousal income is independent of their partner's health status for the reasons mentioned in section 3.1.

After disability onset and retirement from the labor force, an individual can receive income in form of public DI benefits, if admitted, and private DI benefits, if covered by private DI. As described in section 3.3, both benefits replace a given fraction of the full-time labor income in good health, and can be simultaneously claimed:

$$B_{it} = Y_{it} * RR^{public} * \mathbf{1}[DIs_{it} = Admitted] + Y_{it} * RR^{private} * \mathbf{1}[pDI_0 = 1] \tag{11}$$

B_{it} denotes the total benefit received, RR^j the replacement ratio in the public or private DI whose respective values I estimate from the data. $\mathbf{1}[DIs_{it} = Admitted]$ and $\mathbf{1}[pDI_0 = 1]$ are two dummy variables that take the value one if an individual is admitted into public DI and owns private DI respectively.

Finally, the German social security system guarantees a consumption floor SSI for people out of the labor force, either for health reasons or voluntarily. To qualify for SSI, household income has to fall below this level conditional on passing a means test:

$$y_{it} = SSI \text{ if } \{0, B_{it}\} + y^s_{it} \leq SSI \;\; \& \;\; A_{it} \leq \bar{A} \tag{12}$$

[18] You can think of these movements as upward/downward movements within the same broad occupation as well as horizontal movements due to specialization with no effect on the initial risk group mapping

[19] I check the sensitivity of my results with respect to this assumption. My parameter estimates and counterfactual results are robust to imposing a retained productivity of 38.5%, the maximum amount that always qualifies you for public DI receipt.

In retirement each spouse receives a fixed pension which depends on their life-time income. I compute these pension benefits following the legal pension schedule as detailed in E.4. To keep the model tractable, I assume that spouses are of identical age, such that they also retire at the same time[20].

3.5 Why do not all people buy private DI?

Basic economic theory predicts that risk averse individuals should always fully insure themselves if insurance is fair and no other frictions exist. However, as stated in the introduction, only about 34% (50%) of all people (men below 35) in Germany purchase private DI (EVS 2013). Since I estimate my model by matching the average private DI purchases in the male population, it follows that some people do not buy private insurance despite being risk averse. So which channels in my model can generate this behavior?

First, private DI is not actuarially fair, but sold at a mark-up (around 13% to 32%). Second, the exogenous spousal income and social security income (SSI) can make purchasing private DI less attractive, especially for low income individuals. Given their low income, they are more likely to qualify for SSI (they are also more likely to pass the means test), while the supplementary private DI benefits might only offer slightly higher benefits but at the cost of paying the premium in good health. Moreover, the negative correlation between income and risk implies that these high risk individuals have to pay a larger share of their income for insuring €1. Figure 1 provides some descriptive evidence for this channel: private DI take-up is increasing in income quartiles for both the whole population (diamonds) as well as the estimation sample (squares).

I verify the importance of each channel by shutting them down separately. The results (available upon request) show that all margins matter and are of similar significance. Absent any of these channels, private DI purchases are close to full insurance.[21]

4 Data

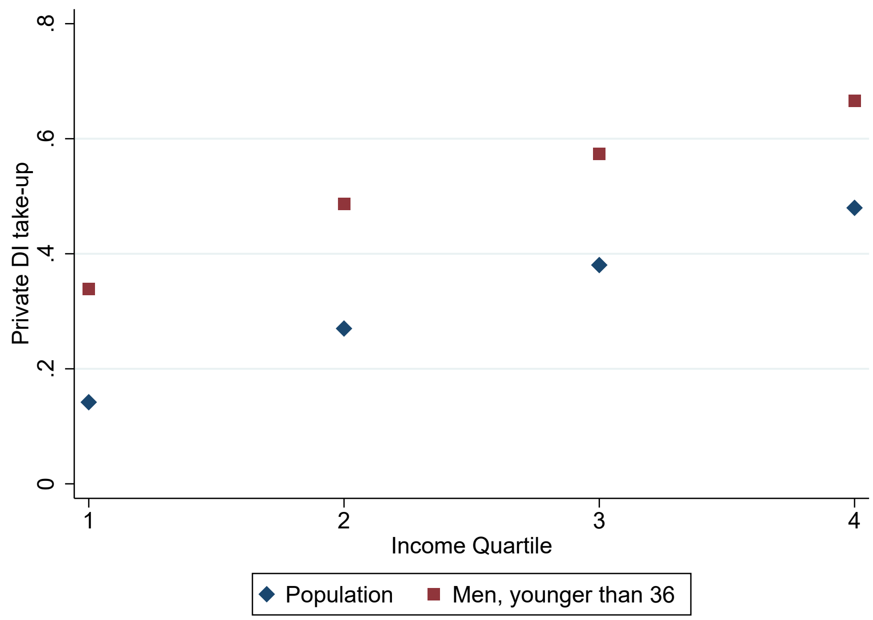
This section describes the data used to estimate the model parameters governing individual choices. My estimation relies on three different data sets with complementary information, each capturing a specific margin of behavior. First, I use four waves of the (German) Income and Consumption Survey (EVS), which contains detailed information on assets and private DI ownership shares. Second, I use proprietary customer data of a major German private insurance company to model the private market. Finally, I use social registry data (SIAB) from the Institute of Employment Research with detailed information on income, program participation, and occupations to model the labor market.

[20]The mean age difference in the data is approximately 2 years, whereby men are older than women.

[21]This is a non-exhaustive list. The points raised here are contained within the model.

Figure 1: Private DI take-up by Income Quartile

The figure below presents the private DI take-up conditional on income quartile for the whole population (diamonds) and the estimation sample (men, 25 to 35 years old; squares). The values are estimated from the EVS 2013 wave.



In all data sets, I restrict the sample to men who are at least 25 years old and are neither retired nor in education. I drop all civil servants and self-employed because they are not eligible for public DI benefits. All monetary values are converted to 2013 prices. Appendix B contains a detailed description of the data set construction.

4.1 Income and Consumption Survey

The Income and Consumption Survey (EVS) is a large representative household level survey conducted every five years by the German Federal statistical office. Participants provide detailed information on income (sources) and their expenditures over a period of three months. Notably, the EVS also contains information on private DI ownership from 2013 on, which is a key moment in the estimation below. Therefore, I use the EVS to estimate the mean and median asset profiles at different ages as well as the mean private DI take-up (unconditional and by income quartile).

I construct the estimation sample by pooling the 1998, 2003, 2008, and 2013 waves and applying the selection criteria discussed above. I drop civil servants and self-employed individuals because they are ineligible for public DI. In addition, all households whose household heads are female, younger than 25 years, still in education or already retired are also dropped. This leaves me with a sample of 87,286 households. Appendix table

B.1 presents some summary statistics.

I use the cleaned sample to generate two sets of moments, which I target in the estimation below. The first set constitutes the key moments in my estimation, the mean private DI ownership in the population and conditional on (gross labor) income quartiles. Since this information is only available from 2013 on, I estimate these moments only on the 2013 wave. Furthermore, due to a public DI reform in 2001, I restrict my sample to individuals who entered the labor market after the reform, i.e. individuals younger than 35 years in 2013 (see Seibold et al. (2021) for details on the reform).

The second set of moments consists of mean and median assets. I estimate the mean and median assets in 3-years age bins for ages 25 to 69 after dropping the top and bottom 1% of the asset distribution similar to Adda et al. (2017). Assets contain all forms of liquid assets, e.g. checking accounts and stocks, plus the value of housing net of liabilities.[22] All four waves are used to estimate these moments. Section 5 provides more information on the estimation procedure.

4.2 Private Insurer Data

I have obtained a novel data set which comprises the universe of contracts from one of the largest German insurance providers.[23] The data contains detailed individual information on demographics, contracts, and health outcomes and is used by the insurer to compute the risk-based premiums. I use this data set to construct the mean replacement ratio, the risk group-occupation mapping, and estimating the prices (by risk group).

A contract still needs to be active as of January 1^{st} 2013 to appear in the data set. I can follow these individuals until January 1^{st} 2018 including all entries and exits during this time as well as various health events. I briefly describe the key variables of interest and cleaning steps here (see Appendix B.2 for the details).

A contract documents basic demographics, such as age, gender, and a detailed occupation title. The latter is the primary input for the applicant's risk assessment, which is mapped into a discrete risk group. In addition, the data contains detailed information on individual annual benefits, the date of purchase, and the expiration date. Between 2013 and 2018, I can also observe disability onset, recovery, death and cancellations.

I add the official occupation codes (2010 version) by job title based on the steps described in Appendix C. This allows me later to export the risk group - occupation mapping to the social security records (SIAB below). Moreover, it allows me to construct predicted individual income by age, occupation, and gender from the 'Verdienststrukturerhebung' (Labor Income Survey). Based on predicted income, I compute the individual replacement ratio as the ratio between benefits and predicted income, which is a key parameter in my

[22] Check the codebook of the Federal Statistical Office for more details on the different types and definitions of assets.

[23] We validate the representativeness of this data set in Seibold et al. (2021).

Table 1: Private DI data: Summary statistics

The table below shows summary statistics for the private insurance data under alternative sample restrictions. Column (1) displays the sample means for the full sample. Column (2) presents the cleaned sample, column (3) the baseline sample for men and column (4) the corresponding estimation sample. The corresponding sample selection criteria is shown in the lower panel. The sample window is 1966 to 2017 in column (1), (2), and (3) and 2001 to 2017 in column (4).

	(1)	(2)	(3)	(4)
Age	40.02	39.84	41.01	43.29
Age: Purchase	29.68	31.54	32.54	34.63
Age: Contract end	62.55	62.79	62.67	65.60
Benefit	16,487.30	17,583.22	19,169,45	20,566.51
Income	52,806.29	51,030.61	56,235.82	59,597.51
Replacement ratio	0.34	0.36	0.35	0.36
Risk group	2.27	2.34	2.34	2.22
Share: Disabled	0.02	0.01	0.01	0.01
Sample selection criteria				
Stand-alone DI	.55	1	1	1
Male	.61	0.57	1	1
Share: Cancel	0.10	0.10	0.10	0
Share: Bought before 2001	0.14	0.01	0.01	0
Share: Age Purchase < 25	0.26	0.18	0.15	0
Share: Miners	0.0002	0.0002	0.0003	0
# Obs.	Confidential	42.1%	24%	99,419

model. Finally, I add prices to the data. Since the data is used in the price calculations, prices are not contained in the data set. Instead, I web-scraped the prices by age and risk group directly from the insurer's website (see Appendix B.2). As the premium is linear in benefit given the risk group (see Section 2), I recover the actual premium by multiplying the web-scrapped prices for insuring a Euro with the reported benefits from the data.

Table 1 reports the summary statistics for different samples in the upper panel and stratifying conditions in the lower panel. Column (1) presents the means for the whole sample before applying any cleaning step. The second column is derived after two cleaning steps. First, I drop all civil servants, self-employed or people in education as I do in the other data sets. I also drop all observations with missing occupation information (see Appendix Table C.6 for details). This leaves me with 80% of the initial sample, whereby 'in-education' and 'missing occupation information' account for about 90% of the dropped observations. Second, I drop all observations that bought their private DI coverage as part of a bundle, e.g. together with life-insurance, as their insurance motive might be different from simply insuring labor productivity. The resulting sample contains 42.1% of the initial sample but looks very similar regarding mean outcomes.

Further restricting my sample to men reduces the sample size to 24% of its initial size. This restriction increases mean benefits and income relative to the full sample, which is mostly driven by the higher average age and the fact that these men are more likely to be academics [not shown]. The replacement ratio, however, is very similar (0.35 vs. 0.34).

Finally, the fourth column contains the estimation sample, which I get by dropping all observations who identify as miners (special public insurance), have cancelled their contract, bought their contract before 2001[24], or were younger than 25 at age of purchase, which is the initial age in my model. The sample consists of 99,419 contracts. Compared to the other samples, this sample has a similar share of disabled and a similar replacement ratio. However, dropping younger individuals translates into higher average age and age at purchase. Given the age gradient in income, these people also have a higher income and insure larger benefits, while the replacement ratio remains constant.

4.3 Social Registry Data

The IAB (Institut fuer Arbeitsmarkt- und Berufsforschung) collects information on the employment related benefit history of each individual in Germany who was in one of the following states between 1975 and 2017: employment, unemployment insurance or social assistance recipient. Since civil servants and self-employed individuals are exempt from social security contribution, they do not appear in this data set. The SIAB is used to estimate the income process, the disability risk probability by risk group, the population risk-group distribution, and the labor supply moments (labor force participation, full-time and part-time shares) for the calibration exercise.

The SIAB is a random 2% sample from this universe of social registry data. It contains the employment and benefit history of 1,875,439 individuals, comprising 66,961,520 spells. The information in this data is relevant for determining unemployment insurance entitlement and benefit level. Hence, the data has comprehensive information on daily wages, occupations, basic demographics (age, gender, citizenship), work arrangement (full-time vs. part-time), industry codes, residency (municipality), and benefit receipt. In addition, the IAB reports the reasons for transitioning employment states including public DI receipt, which allows me to identify these spells in the SIAB data. I use the data to estimate the wage equation (8), the labor market moments, the disability probability by risk group, and the population risk group distribution (see section 5).

I transform the different spells into an annual panel of individual (employment) histories. If spells span several years, I divide them into annual spells. Multiple spells within a given year are ranked according to their timing and I retain only the longest spell in each year. Since my model and estimation sample focuses on the time after the 2001 pension reform, I restrict my sample to spells recorded between 1992 and 2017.[25]

To reflect the annual frequency, I transform daily income into annualized income (2013 Euros)[26]. The income information is third-party reported, so measurement errors are neg-

[24] This is due to a pension reform which changed the incentive to buy private DI and increased coverage substantially. See Seibold et al. (2021) for more details

[25] Including some additional years provides some additional information, especially for people that claim UI or DI after 2001.

[26] Annualized income corresponds to the reported daily income of the retained employment spell multiplied by the number of days in that year.

ligible. However, income in the SIAB is only reported up to the social security contribution limit, thus I impute wages above the contribution limit with a series of Tobit-regressions (see Dauth and Eppelsheimer (2020) for details).
After constructing the panel, I merge the mean, median and mode risk group from the private data to the SIAB by occupation code. If I fail to match an occupation to a risk group from the insurance data, I look up their risk-group mapping in the insurance company's risk table and add their risk-group manually.[27] Overall, I can match all observations with non-missing occupation codes to a risk group, which corresponds to 97.15% of all observations in the raw data and 99.8% in the cleaned sample. Appendix D provides further details on the cleaning steps and the merging process. Based on this mapping I later estimate the risk-group distribution in the whole population as well as controlling for the relationship between income and risk-group.
Finally, I apply the same sample selection criteria as above: I retain all all individuals that are between 25 and 65 years old[28], are not reporting zero income[29], and do not work in non-standard employment forms (e.g. apprenticeship, early retirement,...) or are temporary employees. The final sample then consists of 32 million person-year observations. Appendix table D.1 presents the summary statistics and how the sample selection criteria affect the sample composition.

5 Estimation

I estimate the model described in section 3 following a three-step procedure. First, I take some values from the literature, e.g. tax rates and social security contributions. Second, I estimate some processes outside the model in a reduced form fashion, such as the population risk-group distribution or the income process. Finally, I apply the method of simulated moments (MSM) to estimate the utility parameters of my model by minimizing the weighted distance between the data moments and the corresponding moments simulated in the model.

5.1 Values from the literature

Table 2 displays the parameters I take directly from the literature instead of estimating them alongside their values and source. The first panel shows three model parameters

[27]This can happen due to censoring requirements: If too few observations are within an occupation-risk group cell, this cell is censored in the aggregated insurance data.
[28]In the cleaning step I retain individuals between 20 and 65 years, but drop the ones below 25 in the estimation
[29]Transfer income is also documented and well different from zero. Therefore, zero income spells refer to a special subgroup of "non-eligible" yet documented individuals, which I drop from my analysis, or individuals with missing information.

Table 2: Parameters from literature

The table below shows the parameter values selected outside the model. These parameters include model parameters not estimated in the model, the German tax and benefit schedules, disability and mortality risk, as well as private insurance prices for the different risk groups. Monetary values are deflated to 2013 prices.

Parameter	**Value**	**Source**
Model parameter:		
-Final period T	70 (age 95)	-
-Interest rate r (net-of-tax)	0.0225	-
-β	0.987	-
Tax schedule and social security contributions		
-Income tax schedule	appendix E	Income tax code 2013
-Health, long-term care insurance	0.0775, 0.01025	SSC code in 2013
-pension, unemployment insurance	0.0995, 0.015	SSC code in 2013
Social security contribution income limits		
-Health and long-term care insurance	4000 Euros/month	SSC code in 2013
-Pension and unemployment insurance	5800 Euros/month	SSC code in 2013
Public Benefit programs		
-Social Assistance	6300	Income tax code 2013
-Social Assistance, means test $\bar{A}$	5,000 Euros (per adult)	SSC code in 2013
-Public DI rejection rate	0.44	German Pension Fund
-Replacement ratio (public)	0.35	German Pension Fund
Risk processes		
- Health Transitions	appendix table F.1	German Acturian Society
- Mortality risk	appendix table F.1	German Federal Statistical Office
Annual private DI prices for an annual benefit of €12k, by risk-group		
-Risk-group 1	€353	Company website
-Risk-group 2	€467	Company website
-Risk-group 3	€762	Company website
-Risk-group 4	€1125	Company website
-Risk-group 5	€1736	Company website

I set to specific values commonly used in the literature.[30] The terminal age is set to 95 years corresponding to a final period of $T = 70$. I impose a real interest rate of 3%. Given the linear tax rate of 25% on capital returns, the net-of-tax rate r amounts to 2.25%. I assume that β takes the value 0.987, so people are patient.

The second panel shows the values for the tax and transfer system, which I model according to their statutory rules in 2013 (see Appendix E for details). Household income is assessed jointly based on the income tax schedule in Appendix E. In contrast, social security contributions in the form of payroll taxes are paid individually. The individual payroll tax rates in 2013 were $\{0.015, 0.0995, 0.0775, 0.01025\}$ for unemployment insurance, public pension, health insurance, and long-term-care insurance respectively. Social security contributions are paid up to a fixed income threshold and remain flat for income exceeding these caps. In 2013, these income limits were €5800 (€4000) per month for the

[30] I have verified the robustness of my results with respect to alternative values.

pension and unemployment insurance (the health care and long-term care insurance).[31] In turn, public benefits also remain flat after these thresholds at their maximal amount. The parameters of the public benefit programs are presented in the third panel. I set the consumption floor offered by social assistance (Hartz-IV + additional transfers) to €6300 per year, the statutory values in 2013 (€ 450 per month plus up to € 900 bonus payments).[32] To qualify for social assistance, household income has to be below this value conditional on passing a means test. The means test requires that household assets do not exceed €5,000 per adult. Otherwise households are not eligible.
The public DI system is characterized by two parameters, the replacement ratio and the rejection rate of applications. The replacement ratio is set to 35% of individual gross income, its average from the public pension data (Seibold et al., 2021). The rejection rate is set to its average from 2001 to 2013, which amounts to 44%[33]. Contributions to the public DI system are included in the public pension contributions (cf. section 2).
The health transition probabilities are taken from the disability table provided by the German Acturian Society (Aktuarvereinigung, 1997, 2018). The mortality probabilities come from the mortality tables provided by the German Federal Statistical Office. Appendix table F.1 presents the respective probability and mortality probabilities.
The last panel of table 2 presents the web-scraped prices for private DI by risk group. The prices calculated under the assumption that a 25 year old (healthy) individual purchases insurance until the age of 65 insuring €12,000 per year.

5.2 Parameters estimated outside the model

I estimate the parameters governing (a) the population risk group distribution, (b) disability probabilities by risk group, and (c) the income process from equation (8) in a reduced-form fashion outside the model. The construction of the respective estimation samples is detailed in section 4. If not stated otherwise, the sample window always runs from 2001 to 2017. Table 3 presents the estimated coefficients.
Panel A shows the estimated risk group distribution for men. The assignment to a risk group is based on the insurer's risk group - occupation mapping, where each individual is assigned to a unique risk group. The results reveal substantial heterogeneity in risk. While 5.3% and 19.9% of men work in occupations assigned to the lowest two risk groups, the largest share works in occupations with medium to high disability risk: 28.9% have a job falling into risk group 3, while the large majority (45.3%) work in high risk jobs.

[31]Note, I impose the social security limits for West Germany as the West German population is greater. Imposing the corresponding ones for East Germany are €4900 and €4000 per month has no discernible effects on the results.

[32]As with consumption, this consumption floor is scaled by the equivalence scale to account for household composition.

[33]See https://statistik-rente.de/drv/extern/rente/antraege/tabellen_2015/201512_Rentenantrag_Tabelle03.htm for the data

Table 3: Parameters estimated outside the model

The table below shows parameter values estimated outside the model. Panel A to C is estimated on the subsample of employed or disabled men in the SIAB. Panel A shows the distribution of the discrete risk groups in the population as population shares (sample window: 2001-2017). Panel B displays the predicted disability probabilities by eq. (4) (sample window: 2001-2017). Panel C reports the results from estimating the income equation (8) (sample window: 1999 - 2017).

Parameter	**Value**	**Source**
Panel A: Risk Group Distribution		
Risk Group 1	0.0529	SIAB
Risk Group 2	0.1993	
Risk Group 3	0.2887	
Risk Group 4	0.4534	
Risk Group 5	0.0045	
Risk Group NA	0.0011	
Num. Obs.	4,701,550	
Panel B: Health Risk adjustment		
$Prob(disabled(rg = 1))$	$2.722 * 10^{-4}$	SIAB, eq. (4)
$Prob(disabled(rg = 2))$	$4.227 * 10^{-4}$	
$Prob(disabled(rg = 3))$	$6.476 * 10^{-4}$	
$Prob(disabled(rg = 4))$	$9.787 * 10^{-4}$	
$Prob(disabled(rg = 5))$	$14.592 * 10^{-4}$	
Num. Obs.	4,696,325	
Panel C: Income Process		
β_0	0.7730	SIAB, eq. (8)
β_1 (age)	0.0405	
β_2 (age^2)	-0.0015	
β_3 (age^3)	$2.46*10^{-5}$	
β_4 (age^4)	$-1.91*10^{-7}$	
β_5 (full-time)	0.7921	
β_6^k (risk group):		
2	-0.2035	
3	-0.5412	
4	-0.7253	
5	-0.7558	
σ_η^2	0.0192	
σ_ζ^2	0.1265	
σ_ϵ^2	0.0404	
ρ	0.9459	
Num. Obs.	5,143,326	
Replacement ratio	0.36	contract data

Occupations with the highest disability risk are very rare (0.5%).

Equipped with this risk group assignment, I estimate the probability of experiencing a disability by risk group based on equation (4). Panel B reports the predicted probabilities by risk group. The results show that risk groups and disability risk are positively correlated and that this relationship is not linear: Relative to risk group 1, risk group 3 is 2.3 as likely to become disabled, while risk group 5 is approximately 5.6 times as likely. I plug these values into equation (5) to adjust the average disability probabilities reported by the DAV for heterogeneity in disabiltiy risk by risk group.

Panel C of presents the parameter estimates obtained from estimating the labor income

equation from (8) on the subset of employed men with non-missing occupation information in the SIAB between 1999 and 2017. Based on these estimates I derive the stochastic earnings components as detailed in appendix G following the method described in Guvenen (2009). An important feature of the model is the negative correlation between income and risk group ($\beta_6^k, k = 2, ..5$), which captures two important margins of selection into private DI observed in the data: Low risk (high income) individuals are more likely to own insurance (Seibold et al. (2021) and Figure 1). The correlation between income and private DI ownership plays a central role for the evaluation of alternative public DI systems as it directly relates to the moral hazard response of private DI coverage.
Finally, I assume that people can only purchase one type of contract at baseline characterized by the average replacement ratio observed in the contract data, which amounts to 36% of gross income. In robustness exercises, I include a menu of contracts where people can choose among different replacement ratios, e.g. $\{0.2, 0.25, 0.3, ..., 0.5\}$. Appendix table I.3 presents the parameter estimates.

5.3 Method of Simulated Moments Approach

I estimate the four preference parameters of interest, risk aversion γ, consumption weight κ, (dis-)utility from disability φ, and the fixed cost of labor force participation θ, applying the method of simulated moments approach. This approach minimizes the (weighted) distance between the data moments and the corresponding moments derived from my model given imposed parameter values. I weight each moment by the inverse of its variance, which besides controlling for small sample bias (Altonji and Segal, 1996) also accounts for the different units at which each moment is reported (shares vs. levels). Appendix A.3 provides a more formal description of this method.
The fundamental model parameters are estimated based on the moments presented in table 4, which can be distinguished into three sets of moments: private DI shares, labor supply, and savings rates and assets. Appendix table I.2 presents the each data moment and its weight.

The private DI moments consist of the share of private DI owners in the population and by income quartile. I estimate these moments from the EVS 2013 wave, the first wave to ask for private DI ownership. I restrict my sample to men aged 25 to 35 in 2013 to avoid confounding effects from a pension reform in 2001, which changed the incentives to purchase private insurance (see Seibold et al. (2021).
The second set of moments includes the labor force participation (extensive labor supply margin) and the share of full-time and part-time workers (intensive labor supply margin) at different ages. I estimate these moments from the SIAB pooling the years 2001 to 2017. The estimation sample comprises all men either employed, on social assistance or

Table 4: Moments targeted in the method of simulated moments approach

The table below shows the targeted moments in the estimation step. The first column presents the different group of moments and the second column shows the data sets from which these moments are derived. The third column shows the number of moments contained in each group. See appendix table I.2 for the actual data and simulated moments.

Data Moment	**Source**	**Number moments**
	private DI moments	
Mean ownership	EVS 2013	1
Mean ownership by income quartile	EVS 2013	4
	Labor moments, age 29-53 (every 4yrs)	
Participation	SIAB	7
Full-time	SIAB	7
Part-time	SIAB	7
	Asset moments, age 25-69 (3yrs-bins)	
Mean assets	EVS98 - EVS2013	15
Fraction with below (data) median assets	EVS98 - EVS2013	15
Total Moments		56

on public disability insurance.[34] I take these moments from age 29 to age 53 for every fourth year, for 21 moments in total.

The third set of moments consists of mean and median assets at different ages. The mean and median assets are estimated on the pooled EVS estimation sample described in section 4.1. The assets of ages 25 to 69 are pooled into 3-years age bins, to increase estimation precision. The mean and median asset moments (following French (2005)) are estimated for each of the resulting 15 age bins, for a total of 30 moments.

Before discussing the results, I want to make explicit which moments help to identify which parameter. Risk aversion γ determines the consumption smoothing motive across time and states: A greater value of γ increases the smoothing motive, so people save more. Hence, the asset profiles (mean and median) contribute to its identification. People are willing to work longer hours to increase their consumption, if they value consumption relatively more to leisure, captured by a greater consumption weight κ. Thus, the variation in leisure (full-time, part-time, no participation) helps identifying κ. The fixed cost of labor force participation θ is mainly determined by labor force participation moments: Individuals only participate in the labor force if the compensation from doing so (income which can be used for consumption) exceeds the utility cost of supplying labor. The share of (non-)participants, part-time and full-time shares are informative about this cost. Finally, φ, the (dis-)utility from bad health, governs how people want to move consumption across health states (insurance motive). A greater value of φ raises the value of an additional

[34] I drop individuals on unemployment insurance as my model does not allow for involuntary unemployment spells. Besides, UI benefits are exhausted after one year and people move onto social assistance. Given that my model and thus data is at yearly frequency, only a small fraction of individuals are UI beneficiaries and most individuals re-enter my sample as either employed or on social assistance.

Table 5: Parameters estimated using the method of simulated moments

The table below shows the model parameter estimates obtained from the method of simulated moments. The third column contains the estimated standard errors for each parameter. The fourth column presents the moments that contribute to identifying each utility parameter as discussed in section 5.3.

Parameter	**Value**	**Standard Error**	**Identification**
Risk aversion γ	6.232	0.453	Mean and median assets
Consumption weight κ	0.495	0.003	full-time and part-time shares, LF participation
Labor force participation cost θ	0.161	0.01	LF participation, full-time and part-time shares
Disutility from bad health φ	0.154	0.001	mean private DI, mean and median assets

Euro of consumption in the disabled state thus increasing the demand for formal and informal insurance (assets). Both private disability insurance ownership shares (formal insurance) and asset profiles over the working life (informal insurance) are informative about this parameter.

6 Results

This section presents the estimation results of the preference parameters from the model in section 3. It includes a discussion of the model's performance by evaluating the estimation precision with respect to preference parameters and model fit. Overall, the parameter estimates are in line with values in the literature and precisely estimated. Moreover, the simulated moments match targeted and non-targeted data moments well.

6.1 Estimation Results

Table 5 presents the estimation results. The second column displays the parameter estimates derived from the method of simulated moments and the third column shows the corresponding standard errors for each parameter.[35] The fourth column reports which moment identifies which parameter (see section 5.3). Overall, the parameter estimates are in line with the related literature and precisely estimated. The coefficient of relative risk aversion γ is estimated to be 6.232. Common values found in the related literature on long-term care insurance and pension range from values between 2 to 7 (French, 2005; Jacobs, 2020; Lockwood, 2018). The estimated parameter γ lies at the upper end of this interval. The standard error in the third column of table 5 shows that γ is precisely estimated.

[35]Lockwood (2018) explains the standard error computation in detail in his online appendix.

The consumption weight κ is estimated to be equal to 0.495, which is close to the values found in French (2005) and (Jacobs, 2020) and similar to the one assumed in Low and Pistaferri (2015). This value implies that individuals value consumption and leisure almost equally. The standard error indicates that κ is also precisely estimated (s.e. 0.003). Likewise, the estimate for the labor force participation cost is precisely estimated (s.e. 0.01). The estimated value of 0.161 implies that the labor force participation cost are equivalent to 5.4% of the total time endowment, which is again similar to the values reported in Jacobs (2020) and French (2005). Given that prime-age men in good health exhibit large labor force participation shares (over 90% in the data), it follows that labor force participation cannot be overly costly to them, resulting in this small estimate.

Finally, the disutility of bad health (disability) is estimated to be 0.154 (0.001 standard error). Recall that a positive φ implies that disability is a "bad" given the utility function in eq. (3), so people wish to transfer additional consumption to the bad health state. The parameter estimate lies between the estimates of French (2005) and Low and Pistaferri (2015). An explanation for this is that French (2005) uses a broader measure of bad health[36], which includes also more moderate conditions, thus finding a lower 'penalty'. Low and Pistaferri (2015) focus on low income earners, who might suffer from more severe disabilities compared to the average individual, explaining their higher disutility term.

In appendix table I.3 I show that my estimation results are robust to alternative assumptions by: (a) imposing a lower retained productivity, (b) accounting for selection into employment, and (c) and allowing for a menu of private DI contracts to choose from (intensive margin). In addition, Appendix table I.4 reports the sensitivity of each parameter with respect to the different moments following Andrews et al. (2017).

Summing up, the estimated model parameters are in line with values in the related literature. They are precisely estimated, so the targeted moments carry some information for these parameters[37]. The next subsection presents the model's performance regarding the targeted moments and non-targeted moments, i.e. the in-sample and out-of-sample fit.

6.2 Model Fit

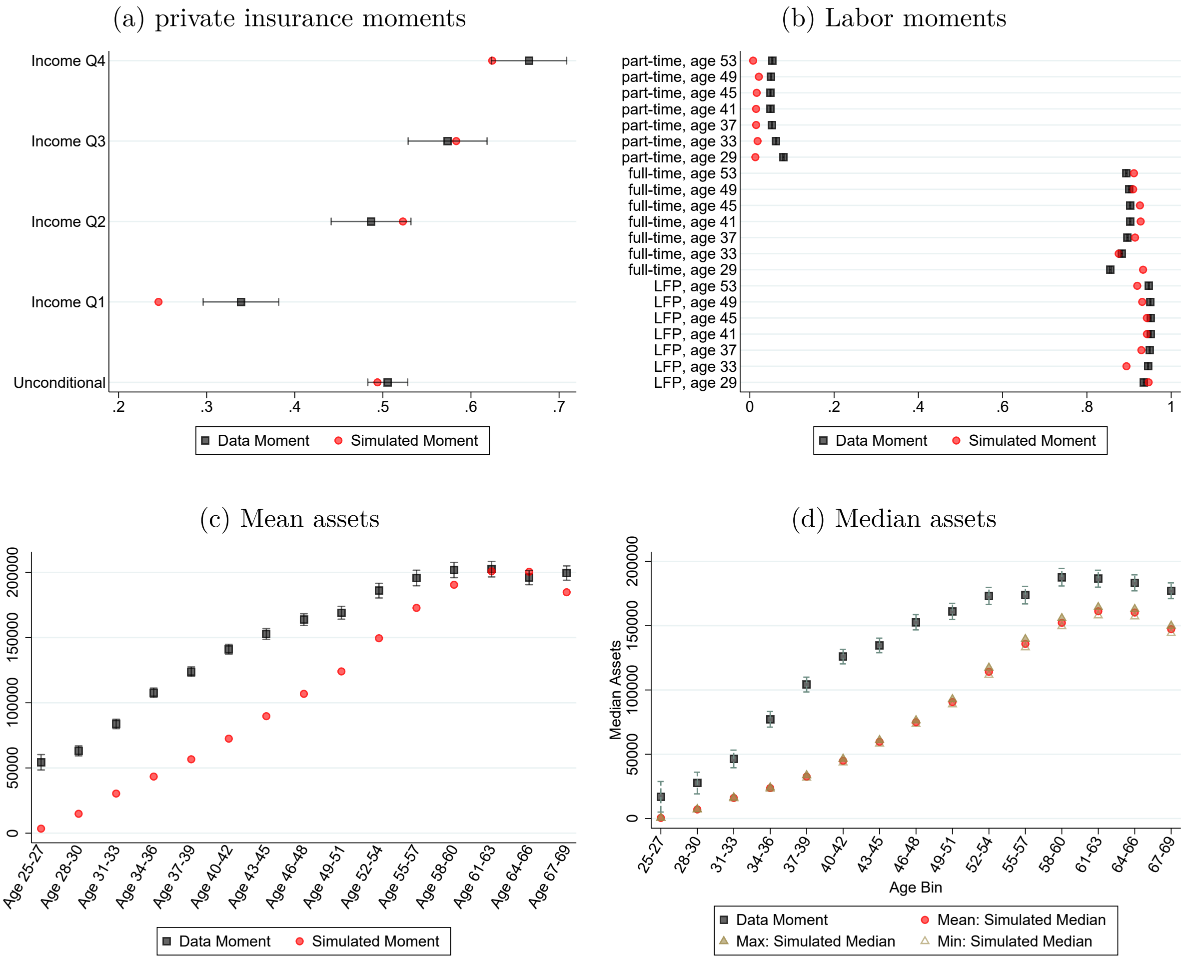
This section evaluates how well the model matches targeted and non-targeted moments, which is informative about the model's performance. By construction, the model should fit targeted moments well as it was estimated on these moments. Matching non-targeted moments corroborates the model's performance by re-producing relationships not used in the estimation. The model matches targeted and non-targeted moments well, which

[36]His measure is based on the answer to the question: "Do you have any physical or nervous condition that limits the type of work or the amount of work that you can do?"

[37]To put it differently, the estimated standard errors imply that the objective function is steep around the optimal values with respect to each parameter. Since small variation in each parameter value produce a substantially lower model fit, this implies that the chosen moments are also informative with respect to the parameters which are to be estimated.

Figure 2: Model fit of data to simulated moments

The figure below presents the in-sample fit of simulated and data moments. The data moments are estimated on the sample of employed men that are at least 25 years of age. Panel (a) displays the private disability insurance moments based on the EVS2013 wave, panel (b) the labor moments estimated on the SIAB, panel (c) and (d) are based on the EVS 98 to 2013 waves and show the mean asset and the median asset profiles over the life cycle respectively. The simulated moments are obtained from 25 populations with 16,000 individuals each. The displayed moments are the average across these populations. The 95% confidence intervals are shown.

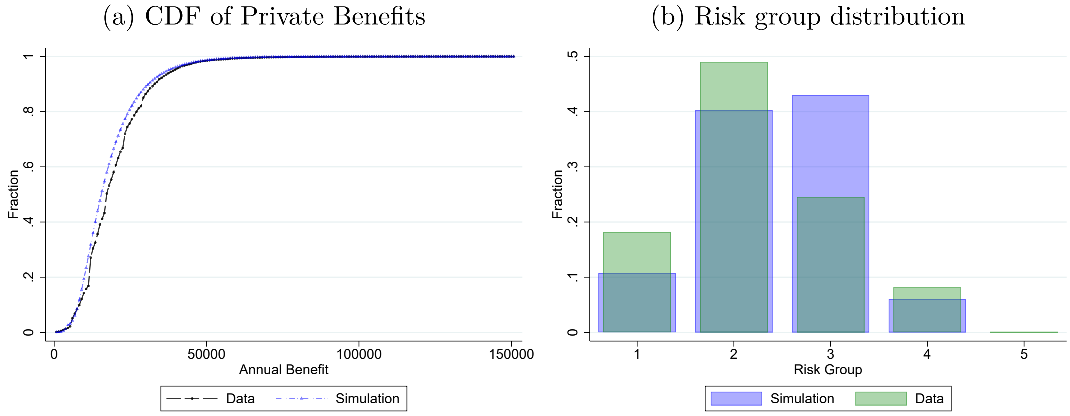


leaves me confident about its performance.

Figure 2 shows the fit between targeted (black) and simulated (red) moments. The standard errors from the data are plotted to speak to precision. The model matches the private DI shares, the key moments of interest, well (Panel a). The simulated moments are close to their data counterparts and the model recovers the positive correlation between income and private DI coverage. The simulated labor supply moments in Panel (b) are also close to the corresponding data moment and they closely track each other. While the model matches the labor force participation well, it generates slightly higher full-time shares at the expenses of too low part-time shares. This is probably a consequence of

Figure 3: Out-of-sample fit of model

The figure below presents the out-of-sample fit of simulated and data moments not targeted in the estimation. The data moments are estimated on the sample of employed men who are at least 25 years of age. Panel (a) shows the cumulative distribution of private DI benefits in the model(blue) and the data (black). Panel (b) shows the risk group distribution of people buying private insurance in the data (green), and in the simulations (25 populations, 16,000 individuals each) (blue). Appendix figure I.1 shows additional out-of-sample fit graphs.



the measurement of part-time as a binary variable in the data and my model instead of hours.[38] Panel (c) and (d) show the model fit for the mean and median asset profiles. The model matches the trends in mean and median assets over the life-cycle well, while there is some discrepancy in the levels. First, this discrepancy is explained by the assumption that people start their life with zero assets, which is not too far off in case of the median; the median level of assets at 25-27 is €16,910 with a confidence interval spanning from €5,000 to €29,000. Second, the asset moments contain net-housing wealth (value housing net of liabilities). In my model, however, I do not separately control for housing, so I cannot match the levels well, especially for ages where most people purchase their first apartment/house. Whereas explicitly modelling the housing decision would increase the model fit, it does not add any conceptual insights to the question at hand: housing wealth is not used to insure against disability risk and in practice banks in Germany often require individuals to have private DI (or life insurance) to secure their housing loans.

Besides closely fitting targeted moments, the model closely matches moments which were not explicitly targeted in the estimation. Figure 3 shows two private market moments conditional on private DI coverage. Panel (a) plots the cumulative distribution function (CDF) of private DI benefits. Despite only offering a single contract with the average replacement ratio of 36%, the model produces a benefit CDF (dotted line) which closely

[38] People in my model can choose to work 20 or 40 hours per week, but part-time work is defined as working 10 to 29 hours. Thus some individuals currently preferring to work 40 hours in my model might move to 29 hours if this option was available.

matches the data CDF (dashed line).

Furthermore, the model produces a risk-group distribution of private DI owners (blue) which is broadly consistent with the data (green) as shown in Panel (b). Private DI coverage is mostly concentrated among low-risk individuals in risk group 1 and 2, whereas very few individuals in risk group 4 (despite being the largest risk group in the population) purchase private DI. The model, however, predicts that too many individuals in risk group 3 and too few in risk group 1 and 2 own private DI. Since all individuals in risk group 1 and 2 purchase private DI, the pattern is explained by too many people being assigned to risk group 3 (and 4).

A possible explanation for this is that I estimate the risk group distribution from the SIAB data based on the occupations held by people between 25 to 35 years of age, i.e. at the early stage of their working life when most people buy private DI. These entry-level jobs are often assigned to a higher risk group, whereas most intermediate and management level jobs are assigned to the next better risk group. In practice, people move up the ranks over their working life which they can report to the insurer to potentially improve their risk group assignment (thus paying less for their insurance). For instance, using the occupation at retirement (old-age or disability) from the public pension data, Seibold et al. (2021) find evidence for this risk improvement over the life-cycle, e.g. risk group 1 in the population increases from 5% to 9.4%, while risk group 4 reduces from 45.4% to 37.6%. My model, however, abstains from such improvements, assuming that the initial contract remains unchanged over the life cycle. Nonetheless, since lower risk groups exhibit the largest moral hazard response at disability onset in my model (see section 7), my results based on the risk group distribution at younger ages provide a conservative lower bound estimate: The moral hazard response would be greater using the risk group distribution at retirement (old-age or disability), calling for even less generous policies.

Appendix figure I.1 presents additional out-of-sample fit graphs with respect to labor supply and income. Again, the model closely fits these non-targeted moments, which leaves me confident about the utility parameters estimated above.

7 Counterfactuals

This section explores the two key questions for changes in benefit generosity and changes in the rejection rate of public DI: How does private DI affect the direction for welfare-improving public DI reforms? For which public DI benefits and rejection rates is having a private market optimal?

To answer these questions and evaluate welfare, I first quantify the behavioral responses to public and private DI coverage. The key question is to determine how private DI coverage distorts the labor supply of people eligible for DI benefits and how selection into private DI coverage varies with the public DI schedule. The resulting cost are weighted against the welfare gains from private and public DI coverage to evaluate overall welfare. All counterfactuals are derived under revenue neutrality by the means of a lump-sum tax levied on all individuals during their working life to balance the government budget. Welfare responses are expressed in terms of consumption-equivalent-variation (CEV), i.e. the constant share of per-period consumption an agent is willing to forgo to move to the new policy regime relative to the baseline. The CEV is computed before any individual uncertainty is revealed ('under the veil of ignorance'). Appendix H contains the details for the computation of the lump-sum tax and the CEV.

Finally, I estimate a partial equilibrium model in which the private market is exogenously given. Characterizing the globally optimal public DI schedule, however, requires larger policy variations which involve estimating general equilibrium effects as well, for instance private firms adjusting their contract menu (prices, risk assessment) in response to public policy changes. Thus, I focus on *local* policy reforms around the observed baseline schedule as is typically done in the literature (see Low and Pistaferri (2015)), keeping the policy environment fixed at its baseline.

7.1 Welfare-improving public DI reforms with private DI

How does private DI affect the design of welfare-improving public DI reforms? Studying alternative public benefit generosity or rejection rates starting from the current German system, I first derive the behavioral responses before relating them to welfare. I compute all results with a private DI market and once without to show how the welfare predictions change conditional on private DI availability. I find that increases in rejection rates are welfare-improving in both scenarios, while benefit increases are only welfare improving without a private DI market.

In this subsection, welfare is normalized at the status quo. While this allows me to infer the direction and size of welfare effects within each scenario (with/without private DI), it does not allow me to compare the welfare across scenarios. This discussion is deferred to section 7.2.

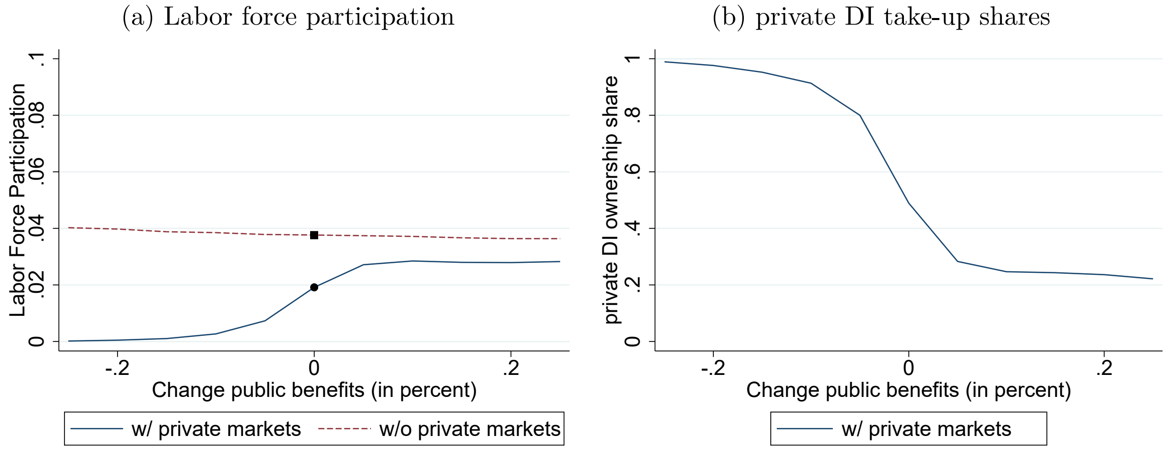
7.1.1 Alternative Benefit Generosity

This counterfactual studies the behavioral and welfare responses to changes in the public benefit generosity between $[-25\%, 25\%]$ around the current public benefit level in Germany. Figure 4 plots the results for the labor force participation (LFP) of the disabled in Panel (a) and the private DI take-up in Panel (b). The status quo level is marked in black, while the solid (dashed) line marks the moments with (without) a private DI market.

I find that private DI coverage reduces the LFP of the disabled across all considered ben-

Figure 4: Labor force participation and mean private DI shares for changes in benefit generosity

The figure below presents the mean labor force participation of disabled individuals (panel (a)) and the mean private DI ownership shares (panel (b)) for alternative public DI benefit generosity. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

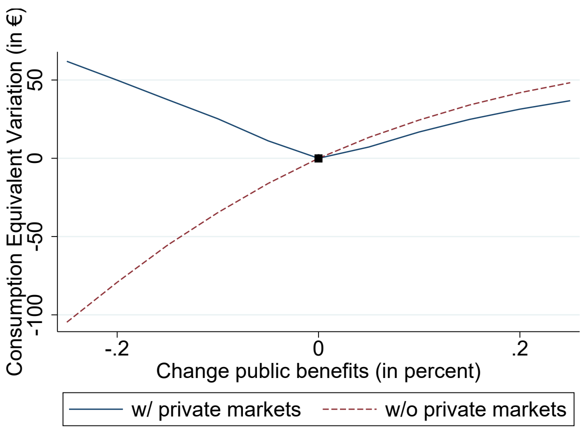


efit changes, which is captured by the gap between the solid and dashed line in Panel (a). For instance, at baseline (in black) the LFP of the disabled is reduced by 50% with private DI coverage. The additional moral hazard inherent to private DI coverage imposes a fiscal externality on the public DI system relative to the scenario without private DI, making it more expensive.

Increasing benefit generosity, the gap in the LFP with and without a private market narrows, while it opens up for less generous benefits. This is driven by the standard LFP response to benefit generosity and the private DI take-up plotted in Panel (b). The LFP of the disabled is decreasing in benefit generosity absent a private DI market, thus the dashed line is downward sloping. The solid line is upwards sloping because fewer people purchase private DI for more generous benefits, such that private DI coverage reduces from 49.5% at baseline to 22% at +25%. As Appendix Figure I.2 shows, people covered by private DI always retire at disability onset while a positive share of them stays employed after removing their coverage. Hence, part of the observed convergence in Panel (a) is explained by fewer people owning private DI. Yet, I find that selection into private DI coverage in benefit generosity is positive on income, e.g. the average income conditional on private DI coverage increases from €36,000 at baseline to €47,000 at +25%. The concentration of private DI coverage among the high-productive types implies that the moral hazard response to private DI coverage of this group is greater (Panel (c) Appendix Figure I.2) and they impose a greater fiscal externality per person on the public system. How do the recorded selection into private DI coverage and moral hazard response affect welfare? Figure 5 plots the welfare gains under the alternative benefit generosity. Without

Figure 5: Consumption - equivalent variation for changes in benefit generosity

The figure below presents the consumption-equivalent variation (CEV) for changes in the benefit generosity. The CEV measures the change in expected life-time utility relative to the baseline level (percentage change = 0) in percent of life-time consumption an agent is willing to forgo to move to the alternative policy. All values are expressed in terms of average (per period) consumption in 2013 Euros. Positive values imply a welfare improvement. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

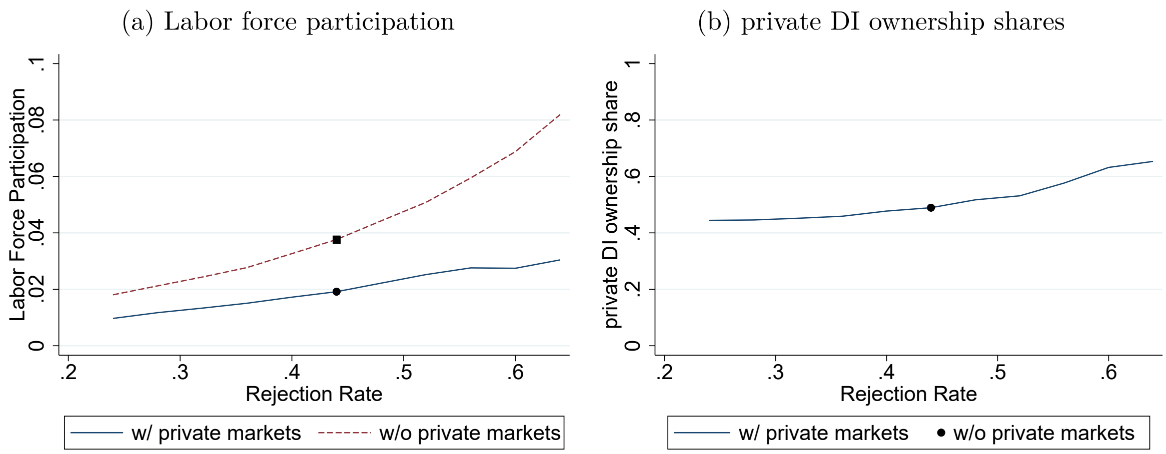


a private DI market (dashed line), welfare unambiguously increases in benefit generosity, such that the increase in insurance value offsets the additional cost from more generous benefits. In contrast, benefit reductions lead to larger welfare gains when a private DI market exists (solid line). For lower benefit generosity, the selection into private DI on income weakens, so the moral hazard response of the marginal buyer is decreasing, as they are less likely to continue working even without private DI coverage (see Panel (e) in Appendix Figure I.2). As a result, the additional fiscal externality remains modest, while the public cutbacks reduce the current program cost. Moreover, more people are covered by private DI and the total insurance value of this group increases substantially offsetting the cutbacks in public DI.[39] Taken together, the welfare gains for benefit reductions are explained by the weakening moral hazard response to private DI coverage and the substantial increase in the total insurance value. Thus, welfare-improving policies with private DI markets are characterized by lower benefit generosity relative to the status quo with private insurance and the scenario without private DI.

[39] The increase in welfare for higher benefit generosity is driven by the higher public insurance value. Yet, fewer people are covered by private DI, such that their total insurance value drops and the welfare gains in this group are smaller. Finally, the people still covered by private DI show the largest moral hazard response to private DI coverage and thus impose a large fiscal externality on the public system, which dampens the total effect. In total, welfare increases, but less so compared to a lower benefit generosity.

Figure 6: Labor force participation and mean private DI shares for changes in screening stringency

The figure below presents the mean labor force participation of disabled individuals (panel (a)) and the mean private DI ownership shares (panel (b)) for alternative public DI rejection rates. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*



7.1.2 Alternative Rejection Rates

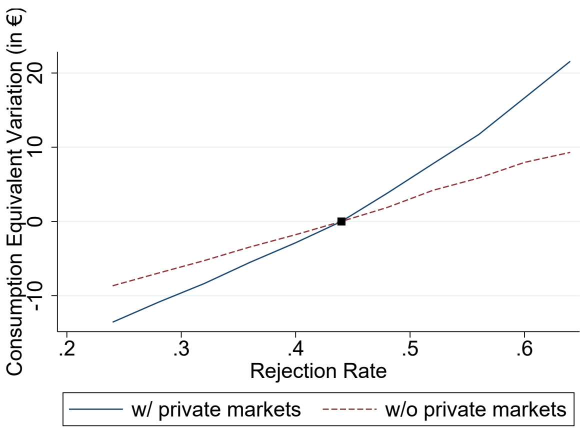
Figure 6 presents the results for labor supply of the disabled in Panel (a) and for private DI coverage in Panel (b) in response to changes in the rejection rate of 24 p.p. around its baseline value of 44%. The baseline value is marked in black, while the solid (dashed) line marks the respective moments for the scenario with(without) private DI.

As before, the LFP of the disabled in Panel (a) is always lower when a private DI market exists. The gap between the solid and dashed line captures the size of the additional moral hazard inherent to private DI coverage, which imposes a fiscal externality on the public DI system relative to the scenario without private DI, increasing the program cost. Increasing the rejection rate and therefore making it harder to claim public DI rises the LFP of the disabled independent of private DI availability. However, without private DI the increase in the LFP is larger, e.g. from 3.9% at baseline to 8.2% at a rejection rate of 64% compared to an increase from 2% to 3% with private DI. Consequently, the gap between the two scenarios opens up and the moral hazard response grows larger.

This is driven by the expansion in private DI coverage for higher rejection rates plotted in Panel (b). Since public DI is harder to obtain, more people rely on private DI coverage to insure against disability, but being covered by private DI these people always retire at disability onset (Appendix Figure I.3 Panel (c) and (e)). At these higher rejection rates, the selection on income into private DI coverage weakens and the marginal buyer's moral hazard response to private DI coverage is smaller. Therefore the additional fiscal externality also remains modest.

Figure 7: Consumption - equivalent variation for changes in screening stringency

The figure below presents the consumption-equivalent variation (CEV) for changes in the rejection rate of applications. The CEV measures the change in expected life-time utility relative to the baseline level (rejection rate = 0.44) in percent of life-time consumption an agent is willing to forgo to move to the alternative policy. All values are expressed in terms of average (per period) consumption in 2013 Euros. Positive values imply a welfare improvement. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*



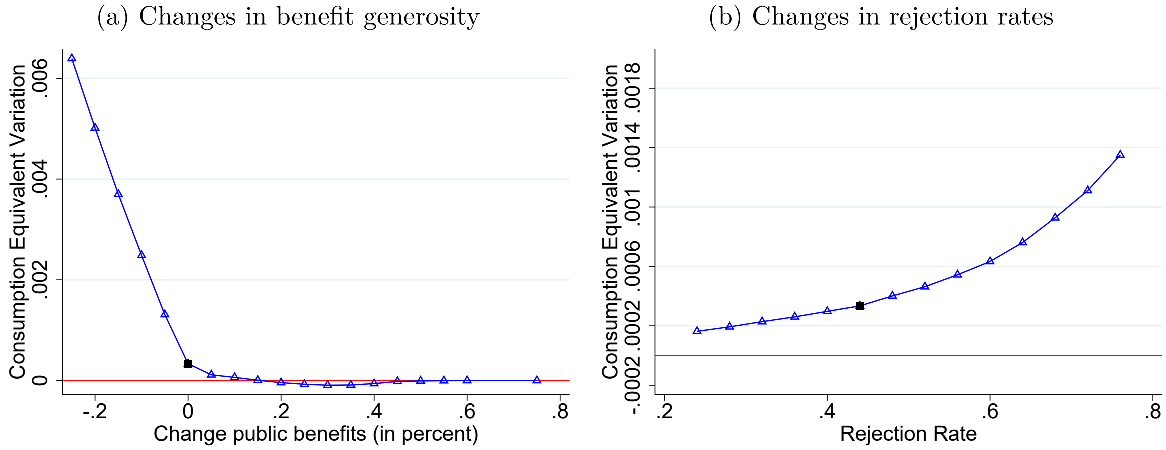
The documented behavior has the following welfare implications summarized in Figure 7. Independent of private DI availability, welfare is increasing in the rejection rate, while the welfare gains are larger with private DI. For instance, people are willing to pay about 0.08% of their consumption per period (€22 on average) to increase the rejection rate to 64% relative to about 0.03% (€9 on average) without a private market. The welfare gains with a private market are increasing in the rejection rate because on the one hand fewer people are admitted into public DI, which given the large fiscal externality from private DI coverage substantially reduces public program cost. On the other hand, more people purchase private DI, recovering some of their lost insurance value. Overall, the total insurance value is decreasing as in expectation people are less likely to be admitted into public DI, but given the large fiscal externality at baseline (the fact that the most productive individuals buy private DI first), the significant cost savings from less public DI claimants still increases welfare. However, note that these increases are small in economic terms and also smaller compared to reforms in the public DI benefit generosity, such that reductions in benefit generosity seem to be the more promising way to increase welfare under the current German schedule.

7.2 Welfare-Effects of private markets

The discussion in the previous section has focused on evaluating the size and direction of the moral hazard response to private DI coverage under alternative policy schedules and

Figure 8: Welfare effects of private DI markets

The figure below presents the consumption-equivalent variation (CEV) for allowing for private DI markets under alternative policy schedules. The CEV is expressed as the percent change of per-period consumption an agent is willing to forgo to have a private market by comparing the expected life-time utility from having a private market to the one without a private market under the same public DI schedule. Positive values imply that private DI markets are welfare enhancing under the considered policy schedule visually presented by the blue line being above the red '0'-line. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*



its effect on welfare. Building on these insights, this section evaluates under which policy schedules private DI markets are welfare-improving relative to only public mandatory insurance. The main discussion is on Germany, but at the end of the section I extrapolate from my findings to the welfare consequences of private DI markets under public DI systems observed in the USA, and Austria[40]. As before, all results are computed under revenue-neutrality by the means of a lump-sum tax.

Figure 8 presents the CEV defined as the percentage of per-period consumption the average agent is willing to forgo to have a private market. It is computed by comparing the expected life-time utility without private markets to the scenario with a private market.[41] Having a private market is welfare improving if the CEV is positive, visually displayed as the blue line being above the red '0'-line.

I find that under the current public schedule (in black), having a private DI market enhances welfare. As Panel (a) shows, having a private DI market is welfare-improving for less generous benefits, whereas it becomes welfare-reducing for more generous public

[40] The choice of countries is motivated by data availability on public DI systems, private DI coverage and by their appearance in research papers

[41] Note the difference to the previous exercise where the comparison was "within a scenario relative to the status quo". Here the status quo is the expected life-time utility with a private market and the comparison is across private DI availability.

DI benefits, before becoming zero at high benefit levels again.[42] The explanation for this pattern is identical to the previous discussion: At higher benefit generosity, private DI coverage is increasingly concentrated among high-income (high-productivity) individuals. This means that fewer people purchase private DI, but the people selecting into private DI coverage display a greater moral hazard response to private DI relative to both the average individual and the average baseline buyer. Since a larger share of them would have stayed employed absent private DI, these people impose a fiscal externality on the public DI system. The fiscal externality reduces the welfare gains from more generous public benefits (higher insurance value) relative to the scenario without a private market, so that not having a private DI market is welfare-improving.[43]

In contrast, Panel (b) shows that having a private DI market is unambiguously welfare-improving for the considered changes in the rejection rate. The black square marks again the baseline rejection rate, at which the CEV is identical to the CEV in Panel (a). Increases in the rejection rate relative to the baseline level enhances welfare. This is again a consequence of the behavioral responses discussed in the previous section. At these higher rejection rates, more people purchase private DI coverage, such that the selection into private DI coverage weakens. It implies that the marginal buyer has a smaller moral hazard response to private DI coverage and the resulting additional fiscal externality from private DI coverage is smaller. Simultaneously, public DI gets harder to claim, thus the number of beneficiaries and the resulting program cost are smaller. Hence, the overall fiscal externality to private and public DI coverage mechanically decreases, while people can recover some of the lost insurance coverage by buying private DI. The latter response is not possible without a private market, so people only benefit from the public program cost reductions. Taken together, having a private DI market is optimal in this case because public program cost decrease while the reduction in insurance value is smaller when private DI is available.[44] Nonetheless, note that the CEV is small in economic terms and relative to the CEV of benefit changes. As before, this implies that changes in the rejection rate might be less effective to increase welfare and policy makers should perhaps focus more on the benefit margin.

While the discussion so far has focused on Germany, an interesting extrapolation exercise is to explore whether having a private DI market is welfare-improving under the policy

[42]This is explained by no one purchasing private DI at these high levels, so the expected life-time utility is identical under both settings.

[43]The argumentation for why private DI is welfare-improving for benefit reductions is similar: At lower benefit generosity, selection into private DI on income weakens/vanishes such that the additional moral hazard remains modest, while a greater share of people owns private DI, thus benefiting from the greater total insurance value. As a result, welfare gains are positive here and having a private market is optimal.

[44]The argumentation for lower benefit generosity is similar: Having a private DI market is still optimal because the response in private DI coverage to rejection rates is small. Hence, the change in the fiscal externality is small, while the welfare gains from easier access to public DI still dominate. However, since the selection into private DI on income worsens at these lower levels (fewer people buy private DI, but advantageously selected), the overall welfare gains from having a private market get smaller.

Table 6: Welfare comparison under alternative public DI systems with and without private markets

The table below presents the welfare change and the share of private DI ownership under the policy regimes characterized by a rejection rate and the replacement ratio observed in Germany (baseline), Austria, and USA. The welfare change is measured in terms of consumption-equivalent-variation, the percentage of per-period consumption an individual is willing to give up to have a private market. Positive values imply that private markets increase expected life-time utility. The CEV is reported in the fourth column. The fifth column displays the private DI ownership share as predicted by the model and the share observed in the data in parenthesis.

Country	**Replacement Ratio**	**Rejection Rate**	**Welfare Change (in percent)**	**private DI ownership share model (data)**
Germany	35%	44%	0.0183	0.4939 (0.5055)
USA	44%	44%	-0.0044	0.2214 (0.35)
Austria	56%	53%	-0.0007	0.0132 (0.04)

schedules observed in other countries. This discussion is motivated by two observations. First, many countries have a private DI market, whose size, however, varies considerably. For instance, the market is large in Germany (50.5%, own calculations) and the USA (35%, Labor Statistics (2020)), but small in Austria (4%, Kaniovski and Url (2019)). Second, many countries offer a greater income replacement compared to Germany, e.g. 44% in the US, 56% in Austria, and 70-75% in the Netherlands.[45] Figure 8 shows, having a private DI market is welfare-reducing under most of these replacement ratios. Hence, I extend the analysis to other countries to illustrate pathways for welfare-improving reforms.

I proceed as follows: I impose the rejection rates and replacement ratio observed in Germany, the USA and Austria[46], while keeping all other distributions fixed (e.g. income distribution, disability risk, risk group distributions,...). I then compute the CEV for having a private market. The fourth column in Table 6 reports the results. The second and third column report the replacement ratio and the rejection rate respectively. In the final column, I display the private DI ownership share as predicted in my model and the observed one in parenthesis.

The results show that having a private DI market only increases welfare under the current schedule in Germany (CEV = 0.0183%). In the USA and Austria having a private DI market reduces welfare relative to the scenario of not having private DI by -0.0044% and -0.0007% respectively. This is a consequence of the behavioral responses discussed throughout this paper. For instance, the USA pays about 25.7% more generous public benefits compared to Germany, while the rejection rate is identical. From the discussion above, we know that at these higher benefit levels, fewer people own private DI (0.2214)

[45]The variation in rejection rates is much smaller and more comparable across countries.

[46]These are the only countries for which I could find both information on private DI coverage and the public DI schedule. The values are taken from Autor et al. (2014) and the BLS (Labor Statistics, 2020) for the USA and from Haller et al. (2020) and Kaniovski and Url (2019) for Austria.

and selection into private DI coverage becomes increasingly advantageous on income. The resulting fiscal externality dampens the welfare gains from greater public insurance coverage with a private DI market, so the expected welfare without a private DI market is greater. The same reasoning applies to Austria, but since fewer people own private DI here (only 1.32% of the population) the welfare losses due to the additional fiscal externality of private DI is smaller, albeit still negative because of the still advantageous selection on income.

Summing up, since most countries have a public DI schedule that is more generous than Germany, so their benefit generosity is to the right of the black square in Panel (a) of figure 8, they could arguably improve welfare by either altering their public DI schedule or by taking means to reduce the fiscal externality stemming from private DI coverage. While this section focused on the most intuitive but also controversial approach, banning private DI markets, there are certainly alternative policy instruments available, which allow for both having a private DI market and high public DI benefits. For example, public DI could include a means-test similar to social security income reducing public benefits if private benefits are paid, which would work similar to a tax on private benefits. This idea includes common concepts such as opt-out insurance (infinite tax rate reducing public benefits to €0 for the first €1 of private benefits) or secondary payer insurance, which replaces a maximum amount of income, e.g. 50% and public insurance only tops up the private benefits to this level. Moreover, the results are derived under rather strong assumptions on the distributions and under the unique German setting, where private DI is an individual insurance as opposed to an employer-provided benefit. Thus, I consider the analysis as illustrative and leave it to future research to answer these questions in the respective country-specific context.

7.3 Robustness Exercises

The results above are derived under the baseline specifications and assumptions. In Appendix I I show that these results are not sensitive to the chosen retained productivity and the inclusion of an intensive private DI margin (a menu of private DI contracts to choose from). Appendix Table I.3 presents the corresponding parameter estimates. Appendix Figures I.4 and I.5 show the welfare effects of changing the benefit generosity and the rejection rate respectively. Appendix Figures I.6 (benefit generosity) and I.7 (rejection rates) show the corresponding behavioral responses. Finally, Appendix Figure I.8 shows the change in welfare under alternative public DI policies after shutting down the private market. The welfare effects, behavioral labor supply responses, and private DI take-up responses are qualitatively and quantitatively close to the baseline results.

8 Conclusion

Although private DI markets exist in many countries to top up public DI benefits, there is little empirical evidence on their interaction with public DI policies. In this paper, I provide novel evidence on this interaction by analyzing how private DI alters the design of public DI schedules and quantifying the underlying labor supply channels. My results highlight the importance of accounting for these channels. The additional moral hazard from private DI take-up is sizeable and has economically meaningful consequences for the design of welfare-improving public policies: in the presence of private DI, welfare-improving public DI schedules are less generous, characterized by either higher rejection rates or less generous benefits. Comparing welfare across private DI availability, I show that the same fiscal externality explains why having a private DI market is only welfare-improving for low benefit generosity as observed in Germany. Under more generous public DI policies, however, having a supplementary private insurance market may be welfare-reducing. I illustrate this for the U.S. and Austria, which both have a private DI market. Imposing their respective public DI schedule in my model, I find that both countries could improve welfare by making public DI less generous or by regulating private DI more.

My findings have practical relevance. Public DI systems have come under financial pressure in recent years due to a rising number of beneficiaries and cost (Autor and Duggan, 2006), and both policymakers and academics have discussed ways to reform the system. My results provide novel input to this debate. Since private DI markets exist in many countries and are often large, abstracting from them can result in a sizeable fiscal externality increasing public program costs. This adds additional strain to the public programs, further threatening their sustainability. Hence, the discussion on how to reform public DI should account for private DI markets.

While my analysis takes the first step into modeling the relationship between private and public DI, focusing on the insurance-incentive trade-off, more research is needed to better understand this interaction, especially with other government programs or under equity concerns. For instance, future studies could analyze the effectiveness of programs aimed at incentivizing public DI claimants to re-enter the labor force in the presence of private DI (Kostol and Mogstad, 2014; Ruh and Staubli, 2019).

References

Adda, Jerôme, Christian Dustmann, and Katrien Stevens (2017). "The Career Costs of Children". In: *Journal of Political Economy* 125.2, pp. 293–337.

Aktuarvereinigung, DAV Deutsche (1997). *Neue Rechnungsgrundlagen für die Berufsunfähigkeitsversicherung DAV 1997*. DAV-Mitteilung. URL: https://books.google.de/books?id=XskKvwEACAAJ.

— (2018). *Überprüfung der Angemessenheit der DAV 1997 I als Reservierungstafel für Berufsunfähigkeitsversicherung*. DAV-Mitteilung.

Altonji, Joseph and Lewis Segal (1996). "Small-Sample Bias in GMM Estimation of Covariance Structures". In: *Journal of Business and Economic Statistics* 14.3, pp. 353–366.

Andrews, Isaiah, Isaiah Gentzkow, and Jesse M. Shapiro (2017). "Measuring the Sensitivity of Parameter Estimates to Estimation Moments". In: *Quarterly Journal of Economics* 132.4, pp. 1553–1592.

Autor, David, Mark Duggan, and Jonathan Gruber (2014). "Moral hazard and claims deterrence in private disability insurance". In: *American Economic Journal: Applied Economics* 6.4, pp. 110–141.

Autor, David, Andreas Ravndal Kostol, Magne Mogstad, and Bradley Setzler (2019). "Disability Benefits, Consumption Insurance, and Household Labor Supply". In: *American Economic Review* 109.7, pp. 2613–2654.

Autor, David H. and Mark G. Duggan (2006). "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding". In: *Journal of Economic Perspectives* 20.3, pp. 71–96.

Borghans, Lex, Anne C. Gielen, and Erzo F. P. Luttmer (2014). "Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform". In: *American Economic Journal: Economic Policy* 6.4, pp. 34–70.

Bound, John (1989). "The Health and Earnings of Rejected Disability Insurance Applicants". In: *The American Economic Review* 79.3, pp. 482–503.

Bound, John, Julie Berry Cullen, Austin Nichols, and Lucie Schmidt (2004). "The welfare implications of increasing disability insurance benefit generosity". In: *Journal of Public Economics* 88, pp. 2487–2514.

Bund, Deutsche Rentenversicherung (2017). *Rentenversicherung in Zeitreihen*. DRV Schriften.

Cabral, Marika and Mark. R. Cullen (2019). "Estimating the Value of Public Insurance Using Complementary Private Insurance". In: *American Economic Journal: Economic Policy* 11.3, pp. 88–129.

Cabral, Marika and Neale Mahoney (2018). "Externalities and Taxation of Supplemental Insurance: A Study of Medicare and Medigap". In: *American Economic Journal: Applied Economics* 11.2, pp. 37–73.

CDC (2020). *Disability Impacts All of US*. Tech. rep. Accessed: July 23, 2021. Centers of Disease Control and Prevention. URL: https://www.cdc.gov/ncbddd/disabilityandhealth/infographic-disability-impacts-all.html.

Chandra, Amitabh and Andrew A. Samwick (2005). *Disability Risk and the Value of Disability Insurance*. NBER Working Paper No.11605.

Chetty, Raj and Emmanuel Saez (2010). "Optimal Taxation and Social Insurance with Endogenous Private Insurance". In: *American Economic Journal: Economic Policy* 2.2, pp. 85–114.

Dauth, Wolfgang and Johann Eppelsheimer (2020). "Preparing the sample of integrated labour market biographies (SIAB) for scientific analysis: a guide". In: *Journal of Labour Market Research* 54.10.

Diamond, Peter and Eytan Sheshinski (1995). "Economic aspects of optimal disability benefits". In: *Journal of Public Economics* 57, pp. 1–23.

Eisenhauer, Philipp, James J. Heckman, and Stefano Mosso (2015). "Estimation Of Dynamic Discrete Choice Models By Maximum Likelihood And The Simulated Method Of Moments". In: *International Economic Review* 56, pp. 331–357.

German Federal Statistical Office (2016). *Sterbetafeln*. data retrieved from 'Statistischen Bibliothek', https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00057034, accessed August 8, 2021.

Finkelstein, Amy, Nathaniel Hendren, and Erzo F.P. Luttmer (2019). "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment". In: *Journal of Political Economy* 127.6, pp. 2836–2874.

French, Eric (2005). "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour". In: *Review of Economic Studies* 72.2, pp. 395–427.

French, Eric and Jae Song (2014). “The Effect of Disability Insurance Receipt on Labor Supply”. In: *American Economic Journal: Economic Policy* 6.2, pp. 291–337.

Gallipoli, Giovanni and Laura Turner (2009). *Household Responses to Individual Shocks: Disability and Labor Supply.* FEEm Working Paper No.97.2009.

GDV, Gesamtverband Deutscher Versicherer (2014). *Versicherungsunternehmen wollen leisten!* Report. URL: https://www.gdv.de/de/themen/news/-versicherungsunternehmen-wollen-leisten---15994.

— (2016). *So werden Kunden gegen Berufsunfähigkeit versichert.* Report. URL: https://www.gdv.de/de/themen/news/so-werden-kunden-gegen-berufsunfaehigkeit-versichert-11226.

Gelber, Alexander, Timothy J. Moore, and Alexander Strand (2017). “The Effect of Disability Insurance Payments on Beneficiaries’ Earnings”. In: *American Economic Journal: Economic Policy* 9.3, pp. 229–61.

Golosov, Mikhail and Aleh Tsyvinski (2006). “Designing Optimal Disability Insurance: A Case for Asset Testing”. In: *Journal of Political Economy* 114.2, pp. 257–279.

Golosov, Mikhail and Alex Tsyvinsky (2007). “Optimal Taxation with Endogenous Insurance Markets”. In: *The Quarterly Journal of Economics* 122.2, pp. 487–534.

Guvenen, Fatih (2009). “An Empirical Investigation of Labor Income Processes”. In: *Review of Economic Dynamics* 12.1, pp. 58–79.

Haller, Andreas, Stefan Staubli, and Josef Zweimüller (2020). *Designing Disability Insurance Reforms: Tightening Eligibility Rules or Reducing Benefits.* IZA Discussion Paper No.13539.

Hilmes, Christian (2019). *Warum Versicherer jeden fünften BU-Antrag ablehnen.* report. URL: https://www.dasinvestment.com/berufsunfaehigkeitsversicherung-warum-versicherer-jeden-fuenften-bu-antrag-ablehnen/.

Jacobs, Lindsay (2020). *Occupations, Retirement, and the Value of Disability Insurance.* Working Paper.

Kaniovski, Serguei and Thomas Url (2019). *Die Auswirkung dauernder Berufsunfähigkeit auf das erwartete Lebenseinkommen in Österreich.* policy report. Österreichisches Institut für Wirtschaftsforschung.

Kostol, Andreas Ravndal and Magne Mogstad (2014). "How Financial Incentives Induce Disability Insurance Recipients to Return to Work". In: *American Economic Review* 104.2, pp. 624–55.

Labor Statistics, Bureau of (2020). *The Economics Daily, Employee access to disability insurance plans*. Tech. rep. Accessed: July 19, 2021. U.S. Department of Labor. URL: https://www.bls.gov/opub/ted/2018/employee-access-to-disability-insurance-plans.htm.

Lee, Siha (2020). *Spousal Labor Supply, Caregiving, and the Value of Disability Insurance*. Working Paper.

Lockwood, Lee (2018). "Incidental Bequest and the Choice to Self-Insure Late-Life Risks". In: *American Economic Review* 108.9, pp. 2513–2550.

Low, Hamisch and Luigi Pistaferri (2015). "Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off". In: *American Economic Review* 105.10, pp. 2986–3029.

Low, Hamish, Costas Meghir, and Luigi Pistaferri (2010). "Wage Risk and Employment Risk over the Life Cycle". In: *American Economic Review* 100.4, pp. 1432–1467.

Meyer, Bruce D. and Wallace K.C. Mok (2019). "Disability, earnings, income and consumption". In: *Journal of Public Economics* 171, pp. 51–69.

Michaud, Amanda and David Wiczer (2018). "Occupational hazards and social disability insurance". In: *Journal of Monetary Economics* 96, pp. 77–92.

Mullen, Kathleen J. and Stefan Staubli (2016). "Disability benefit generosity and labor force withdrawal". In: *Journal of Public Economics* 143, pp. 49–63.

Nelder, John and Roger Mead (1965). "A simplex method for function minimization". In: *The computer journal* 7.4, pp. 308–313.

Pauly, Mark V. (1974). "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection". In: *The Quarterly Journal of Economics* 88.1, pp. 44–62.

Ruh, Philippe and Stefan Staubli (2019). "Financial Incentives and Earnings of Disability Insurance Recipients: Evidence from a Notch Design". In: *American Economic Journal: Economic Policy* 11.2, pp. 269–300.

Seibold, Arthur, Sebastian Seitz, and Sebastian Siegloch (2021). *Privatizing Disability Insurance*. Unpublished Manuscript.

Statista (2019). *Share of respondents with disability income insurance in addition to compulsory social security in the United Kingdom (UK) in 2017, by gender*. Tech. rep. Accessed: July 21, 2021. Statista. URL: https://www.statista.com/statistics/681261/individuals-with-disability-income-insurance-in-the-united-kingdom/.

Stepner, Michael (2019). *The Long-Term Externalities of Short-Term Disability Insurance*. Working Paper.

Tauchen, George (1986). "Finite state markov-chain approximations to univariate and vector autoregressions". In: *Economics Letters* 20.2, pp. 177 –181.

Wachter, Till von, Jae Song, and Joyce Manchester (2011). "Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program". In: *American Economic Review* 101.7, pp. 3308–29.

Waidmann, Timothy, John Bound, and Austin Nichols (2003). *Disability benefits as social insurance: tradeoffs between screening stringency and benefit generosity in optimal program design*. Michigan Retirement Research Center, WP no.2003-042.

A Appendix: Numerical Methods

This appendix provides the details on the numerical approaches applied to estimate the preference parameters of interest. To this end, it first discusses the solution approach to the individual problem and associated modelling choices. Next, it describes how the individual profiles are simulated based on the model solution. Finally, I come back to the method used on how to estimate the preference parameters based on the Method of Simulated moments approach.

A.1 Solution

The model needs to be solved numerically as no analytical solution to the problem described in section 3 exists. Therefore, I apply a backwards iteration approach: By backwards iterating on the value function starting in the final period of the model, I obtain the value of the value function for that period which I can then use to solve the maximization problem in period $T-1$, and so on. Formally, the individual decision problem from eq. (1) in $T = 60$ simplifies to the following problem because death occurs with certainty in the next period leaving the individual with zero utility:

$$V(S_T) = \max_{c_T, A_{T+1}} U(c_T, M) \tag{A.1}$$

where S_T is the set of state variables at time T. Since the per-period utility function $U(\circ)$ is given (eq. 3), I can derive the policy functions $c_T(S_T)$ and $A_{T+1}(S_T)$ which maximizes the value function $V(S_T)$ for any given values of state variables S_T. As detailed below, the maximization method relies on discretized state space grids, so I only solve this problem for this subset of the state space. To obtain the value of $V(S_T)$ at any point in S_T including off-grid points, I need to apply an approximation approach, which is also detailed below. This approach then yields the approximation $\hat{V}(S_T)$, which I use to derive the policy functions for $c_{T-1}(S_{T-1})$ and $A_T(S_{T-1})$ by solve the decision problem in period $T-1$:

$$V(S_{T-1}) = \max_{c_{T-1}, A_T} U(c_{T-1}, M) + s_{T-1} * \beta * \hat{V}(S_T|S_{T-1}) + (1 - s_{T-1}) * 0 \tag{A.2}$$

where s_t denotes the survival probability conditional on having survived till period t.
This approach is repeated until period $t = 0$ is reached. Note that for all ages below 65 (t = 40, the legal retirement age) individuals additionally need to choose their labor supply. Furthermore, the state space changes: For $t < 40$ I drop the survival probability but instead include income risk into the model (transitory and persistent shocks). Moreover, during working life it matters whether people purchased private disability insurance in period $t = 0$. I compute the value functions for this initial choice separately. The

policy function with respect to private insurance ownership is then derived by comparing the expected life-time utility function under each decision: Individuals purchase private insurance if and only if the value function associated with private purchases is greater than the utility function without conditional on being able to pay for insurance.
To solve this model as described here, I have to make some choices regarding (a) discretization of the state space, (b) integration over stochastic variables, (c) approximation of the value function at each point of the state space, and (d) the implications for optimization.

(a) Discretization of the state space

There are six state variables in my model: current assets, persistent income shock realization, transitory income shock realization, health shock realization, individual risk group, and (if disabled) public DI admission decision. The first three variables are continuous, thus they need to be discretized for my model. Assets are discretized by placing them on an equidistant grid with 49 grid points. The minimum of this grid is set to 0 (borrowing constraint), while the maximum depends on the period t. It is equal to the minimum of either the maximal possible income and individual can earn, thus restricting the asset grid to the feasible asset set, or €2,000,000 which corresponds to 10-times the average savings at retirement age.. The continuous stochastic processes are discretized using the Tauchen method (Tauchen, 1986). The grid consists of 15(9) equally spaced grid points for the persistent(transitory) shock, which are assumed to be normally distributed. Thereby, the persistent shock process accounts for path-dependency. The three remaining state variables are already discrete: health shock realizations and the public DI admission decision are binary distributed, while the risk group consists of 5 mutually exclusive realizations. The three control variables[47] in my model, savings, labor supply, and the insurance decision (only in t=0), also need to be discretized. The latter two are already discrete, so no further steps are necessary. The savings decision, however, is continuous. Yet no discretization is needed because the optimal savings choice given all other variables is obtained by maximizing the individual problem in each period over the choice of savings.

(b) Integration over stochastic values

Solving the individual maximization problem requires to evaluate the expected utility by integration over the four stochastic variables. These shocks are the persistent and transitory income shocks, the health shocks, and the public DI admission decisions during the working life and health as well as survival shocks during retirement. All of these shocks are discrete: Health, survival, and public DI admission shocks are already binary random variables, while persistent and transitory income shocks are discretized using the

[47]Note that consumption as a control variable is redundant as it is pinned down by the labor supply, insurance purchase and savings decision in every period via the budget constraint.

Tauchen method (Tauchen, 1986) mentioned above. Consequently, the integration of the value function over the discrete realizations of these stochastic outcomes is equivalent to computing the weighted sum over the value functions at the respective realizations. The weights correspond to the probability of each realization.

(c) Approximation of the value function

The results of the individual optimization problem are only derived for the subset of the discretized state-space. However, solving the problem requires to evaluate the value function for the entire state space. To this end the value function is approximated at these off-grid points by applying multidimensional spline-evaluation for equi-distant grids.[48]

(d) Optimization

I solve the problem separately for each private insurance purchase decision. For each point of the discrete state space, I compute the optimal decision rules conditional on (not) having purchased private disability insurance. In addition, I compute the optimal savings choice within each period separately for each labor supply decision. The resulting decision problem is then continuous in assets and solved using the Brent-Method. Next, I compare which labor supply - asset choice maximizes the value function in that period (at fixed state-space points). The maximizing pair defines the policy functions (labor, assets) and value function for this state space point.

A.2 Simulation

After deriving the optimal decision rules for consumption $c_t(S_t)$, assets $A_{t+1}(S_t)$, leisure $l_t(S_t)$, and private DI purchases, I simulate the decisions of 16,000 households. I follow Eisenhauer et al. (2015) and simulate 25 different data sets to reduce the idiosyncratic errors introduced into the model by drawing from random distributions. The simulated moments are then computed by averaging the respective moments across runs.
Within each run, I simulate the behavior of each individual as follows:

1. I initialize the simulations by setting all decision paths to zero (consumption, purchase decision, assets, labor supply). Individuals start their "life" in good health and with zero assets.

2. I then draw the shock realizations (health, persistent and transitory income, public DI admission, survival) for all individuals in each period from the corresponding

[48]The routine for this is provided by Fehr and Kindermann https://www.ce-fortran.com/toolbox/

distributions, which is normal for continuous variables and uniform for binary variables. Likewise, I draw the risk group realization from a uniform distribution. Based on the draws from the probability distribution, I map the realizations of the continuous variables, transitory and persistent income shocks, in the corresponding outcome (income). Consequently, I compute the continuous gross income that follows from the deterministic income process (eq. (8)) and the shock realizations.
For the discrete outcomes health, survival, public DI admission, and risk group distribution, I assign an individual to the a certain outcome, if the shock realization does not exceed the probability of being in said state, e.g. I assign an individual to the outcome "good health" (conditional on good health before) if the shock realization does not exceeds the risk group specific probability of being in good health.

3. After initializing the decision paths as well as computing the state variable realizations, I start the simulation by determining whether people purchase private DI at age 25. For this purpose, I evaluate the policy function given the individual's assets and their persistent and transitory shock realization using a spline evaluation for equi-distant grids. If the resulting evaluation is exceeds 0.5, the individual buys private DI. This initial decision then determines which policy functions apply for the rest of their life.

4. The remaining decision profiles for $t = 0, ..., 60$ are computed by repeating the following steps:

    (a) Given the risk group and the current health status, I first simulate individuals labor supply decision which pins down their gross income. Again, I apply a spline evaluation for equi-distant grids given the current assets and income shock realizations to interpolate the labor supply policy function. I then assign the individual to its nearest neighbor (in absolute values) labor supply. Based on the labor supply decision, I compute spousal income, tax liability and, conditional on bad health, benefit receipt. I then pool all these incomes to compute the disposable income (income net of taxes and social security contributions). This step is ignored in retirement as people are forced to consume their entire leisure endowment.

    (b) I compute savings (and by the property of the budget constraint consumption). Again, I apply the same spline interpolation approach conditional on current assets and income shock realizations. Since assets are continuous, no further adjustment is needed except for verifying that this amount of savings is feasible (so the optimal assets do not exceed current savings plus disposable income).

    (c) Finally, consumption is computed as the difference between disposable income, this periods savings and the price of private DI (if purchased and not in bad health).

A.3 Estimation of preference parameters

The preference parameters of interest risk aversion γ, consumption weight κ, (dis-) utility from bad health φ, and labor force participation cost θ are estimated via the Method of Simulated Moments approach. This is a GMM approach which minimizes the weighted distance between a set of data moments (depending on the true parameters denoted by index 0) and the corresponding simulated moments derived in the model which takes the preference parameters as arguments. Let **G** denote the difference between the data moments and the simulated moments:

$$G(\gamma, \kappa, \varphi, \theta) = \Sigma^{data}(\gamma_0, \kappa_0, \varphi_0, \theta_0) - \Sigma^{sim}(\gamma, \kappa, \varphi, \theta) \tag{A.3}$$

where $\Sigma^j, j = \{data, sim\}$, is an $Nx1$ vector of the stacked moment conditions. There are two types of moment conditions: mean comparisons and median comparisons. The mean comparisons compare the difference in data and simulated means (M_t and $\hat{M}_t$), while the median conditions are computed following French (2005):

$$\begin{aligned} &M_t - \mathbf{E}[\hat{M}_t(\gamma, \kappa, \varphi, \theta)] = 0 \\ &0.5 - \mathbf{E}[\mathbf{1}[A_{ia} \leq median(\hat{A_{ia}}(\gamma, \kappa, \varphi, \theta))] = 0 \end{aligned} \tag{A.4}$$

A_{it} denotes the asset of individual i in age bin a in the data. $median(\hat{A_{ia}}(\hat{\Upsilon})$ is defined as the median of assets at age bin a from the simulated asset profiles $\hat{A_{ia}}(\hat{\Upsilon})$. Finally $\mathbf{1}(\dot{)}$ denotes an indicator function that takes the value 1 if the assets from the data are below the median assets in the simulations. The corresponding data moment is 0.5, i.e. 50% of all assets in the data are below the median assets from the data.
The optimal preference parameters are then determined by solving:

$$\min_{\gamma,\kappa,\varphi,\theta} G(\gamma, \kappa, \varphi, \theta)' W G(\gamma, \kappa, \varphi, \theta) \tag{A.5}$$

where W denotes the weighting matrix.
I use the inverse of the variance matrix as the weighting matrix and not the optimal weighting matrix, which has to be shown to have poor small sample properties (Altonji and Segal, 1996). Using the inverse variance matrix also has the advantage that it automatically controls for differences in units (shares vs. levels). The variance matrix is estimated directly from the data via bootstrapping. To assigns more weight to the private DI moments, the key moments in my estimation, I modify the inverse variance matrix to become a block-weighted matrix (cf. Finkelstein et al. (2019)). This modification is needed because I only observe the private DI ownership shares in a single wave of the EVS, while the sample size for the mean moments is 4 times (4 waves pooled) and the sample size for the labor supply moments (SIAB) almost 20 times as large. Hence, absent any re-weighting, the method of simulated moments approach assigns the greatest weight to the labor supply moments (most precisely estimated) at the cost of matching

the private DI moments less well. Since they are the key moments in my model using the block-weighting approach then ensures that there is still enough weight put on them without ignoring the information on precision contained in the variances[49].

I compute the solution to the GMM method using the Nelder-Mead simplex algorithm (Nelder and Mead, 1965). I initialize the algorithm by randomly drawing 150 different parameter combinations from the parameter space. The starting value is then a convex combination of the parameter values returning the two smallest function values. To increase precision, I do this for three different sub-spaces (especially with respect to gamma) and repeat the exercise several times (at least 3 or 4 times), always including the previously found optima as values in the new search. All of this leaves me confident that the algorithm really finds the global minimum.

B Appendix: Data

To estimate the fundamental parameters of my model, I draw on three different data sets: the (German) Income and Consumption Survey (*Einkommens- und Verbrauchsstichprobe*, EVS), German administrative register data from the history of social security records (*SIAB*), and a proprietary data set from major German private insurance company comprising their existing contracts from as of January 1^{st}. This Appendix contains a detailed discussion of the sample construction and cleaning procedure for each data set (for short summary, see section 4).

B.1 Income and Consumption Survey

The Income and Consumption Survey (*Einkommens- und Verbrauchsstichprobe*, EVS) is a large representive household level survey conducted by the German Federal Staticial Office every 5 years. It is a repeated cross-section with a sample size of approximately 60,000 private households. Since participation is not compulsory, the actual sample sizes varies across waves. To account for this, sample weights on basis of the Microcensus are constructed and all numbers presented here are weighted. In this paper I use the 1998, 2003, 2008, and 2013 waves, which have between 42,000 to 49,000 participants.

The EVS contains detailed information on household's income sources, expenditures, and some basic demographics of each household member. Households are asked to document their total income from all sources (e.g. labor, transfer, capital, sales of property,...) as well as expenditures (e.g. consumption goods, durable goods, housing, health, insurance, loans,...) over a period of three months. To account for household composition, I construct separate identifiers for spouses and children, which I use to construct the modified OECD

[49] I re-weight the moments by dividing the asset and labor market moments by their respective number of moments, so 21 for labor market moments and 51 for the asset moments.

Table B.1: EVS: Summary Statistics

The table below presents the mean of selected variables across different sample selection steps. The first column shows the means for the cleaned sample, while the second column shows the means for the estimation sample. Since private DI ownership is only available in 2013, the shown means are only computed based on the 2013 EVS wave. Monetary values expressed in 2013 prices.

	Cleaned Sample	Estimation Sample
Gross labor income (€/year)	22,672	23,396
Assets (€)	150,265	170,810
Median assets (€)	69,482	98,509
private DI owners	0.24	0.25
private DI owners, 25-35 years old	0.45	0.51
Age	51.13	52.79
Family size	2.20	2.39
Male household heads	0.76	1
# Obs.	112,918	87,286

equivalence scale converting household consumption to individual consumption.

I construct the estimation sample by imposing the following restrictions across all waves. First, I drop all self-employed and civil servants because they are not covered by the social security system. Consequently, they are also not eligible to public DI benefits. Second, household heads that are younger than 25 and people who are still in training or education are dropped as my model focuses on choices of the working life after completing education. Finally, I restrict my sample to male household heads, which is still the prevalent family model in Germany (76% of all respondents in the EVS). The cleaned (estimation) sample has a sample size of 112,918 (87,286) observations. Table B.1 presents relevant summary statistics.

I estimate two sets of moments from the EVS which I use in my methods of simulated moments approach. First, I compute the mean private disability insurance (DI) ownership overall and by income quartile in 2013. I use "gross labor income from employment" as the conditioning income variable, because private disability insurance insures against health-related labor productivity shocks. Since private DI ownership is only elicited from 2013 on, I am restricted to this wave. Furthermore, due to a public pension reform in 2001 which changed the public DI system for people born 1961 and later, I restrict my sample to individuals who entered the labor market after the reform, i.e. individuals younger than 35 years in 2013. As the share of private DI owners in table B.1 shows, private DI coverage increased greatly among the cohorts who lost their coverage in 2001. Seibold et al. (2021) study the effects of the reform on the private DI market in a related paper. Second, I use all four waves to estimate mean and median asset by age bins. To this end, I pool the data sets and deflate all prices to 2013 Euros using the CPI.[50] I estimate the mean and median assets in 3-years age bins for ages 25 to 69 after dropping the

[50]The prices in 1998 are still in "Deutsche Mark" values, so I first convert them to Euros and then deflate them.

top and bottom 1% of the household net income and asset distribution following Adda et al. (2017). Assets are defined as liquid assets (savings accounts, home loan and savings contracts, stocks, private loans, annuities, and 'other' liquid assets) and the net value of housing, i.e. the value of housing net of liabilities (mortgage, credits/loans). This corresponds to the asset definition suggested by the Federal statistical office (see 'EVS 2013 Codeverzeichnis' [German only]).

B.2 Private Insurance Data

Modelling the private insurance market requires information on specific contract details such as prices, insurance sums, contract duration, occupational information (sorting into insurance), and the risk assessment on behalf of the insurer. No publicly available data set has these required information. Instead, firm-level micro data on their customers is required to speak to these points.

For modelling the private insurance market, I have obtained the customer data of a major German insurance company, which is among the ten largest insurers. The data comprises all private DI contracts that still have existed as of January 1^{st} 2013 or have been purchased thereafter up to 2018. The insurance company uses this data for evaluating their risk assessment and pricing strategy, i.e. as the basis for their daily business operations. The data set has detailed records on demographics, contract details, and health outcomes. The demographic information recorded comprises age, gender, and detailed occupation titles (based on official occupation titles as used by the Unemployment Agency and the Federal Statics Office), which are primarily used to assess risk and price contracts. The risk group assignment of each individual is contained in the data alongside other contract details such as insurance type (pure DI vs. bundled with life-insurance), annual benefits, date of contract purchase, expiration, final payments. Furthermore, the dates of health outcomes and cancellations are reported between 2013 and 2018.[51] The health outcomes consist of the date of entry into disability, date of recovery, and date of death. All dates are reported at the month-year level.

To enable matching aggregated information from the private data with the IAB data, I add occupation classification codes to the private data, based on the recorded occupation titles. I propose two different strategies to match occupation titles to occupation codes. The first approach involves matching the occupation titles from the contract data to the risk table used by the insurance company for risk-assessment. I call this approach "string matching" and I describe it in detail in appendix C.1. The second approach matches the occupation title from the insurance data to the occupation title - code pair in the occupation code handbook published by the German Unemployment Agency. Unfortunately, string matching is not feasible in this case due to different naming conventions in the

[51]Except disability spells that started before 2013 and no recovery has been reported

insurance data. Thus, I searched line-by-line for each occupation title and match them accordingly, hence I refer to this as "line-by-line" matching. Appendix C.2 explains the procedure. The results in the paper based on the 'line-by-line' matching, as I can match more occupations to an occupation code. However, both procedures produce a large overlap as Appendix Table C.1 shows and they are therefore robust to either assignment.

Next, I add two variables I need to estimate my model, replacement ratios and prices. The replacement ratio is defined as the ratio of annual benefits to annual income. However, the annual income is not documented by the insurer, so I estimate the predicted income from the "Verdienststrukturerhebung 2014" (Labor Income Survey), a large cross-section survey conducted by the German Federal Statistics Office which contains detailed information on employment and income. Since the employer completes the survey, income is third-party reported and draws on the same source as the social security records (so little measurement error), while not being top-coded. I apply the same sample selection criteria as throughout my analysis (no civil servants, older than 25, not in education or training) to estimate predicted income by regressing annual income on a quartic age polynomial, a gender dummy, a full-time dummy, and a full set of occupation code classification dummies. Based on these estimated coefficients, I then predict the income for each individual in the insurance data, again conditional on their age, gender, working full-time, and their occupation code. The replacement ratio is then the ratio between the benefits and the predicted income.

Prices are another key variable in my analysis, which are not contained in the data set directly. However, since prices are publicly available at the insurer's website, I web-scrape them for each risk group directly from the website in 2020. I elicit the prices for identical contracts varying only the risk group by assuming that in individual seeks to insure 1,000 Euros from the age 25 to 65 (contract duration 40 years). As the insurance premium is linear in benefits conditional on risk-group assignment and contract duration, I generate the price to insure one Euro by dividing the resulting prices by 1,000. The insurance premium variable is then the product of this price per insured euro and the insurance sum I observe in the data. Appendix section C.3 presents the prices by risk group and the imposed assumptions to elicit them before comparing them to prices of other insurers for 2020/2021.

I clean the sample by dropping all civil servants, self-employed, and people in education. I can identify these people based on their reported occupation titles, e.g. "Entrepreneur" or "tax attorney (self-employed"). Besides, I drop all observations with missing occupation information or observations for which I failed to find the corresponding occupation code (175 in total). This also includes students who do not state their major, as no assigment to an occupation code is feasible.[52] Overall, I can assign 80% of the sample to an occupation code and the most common reason for failing to do so is "missing occupation

[52]Note that for some majors occupation codes exist. Thus, I could assign those students to an existing occupation code and retained them.

information" or being a "student" (90% of all failures).
Moreover, the insurance company sells two types of disability insurance: disability insurance as a stand-alone product and as part of a package (usually together with life-insurance). Since the focus of this paper is on insurance motives of labor productivity and the motives for purchasing private DI together with life-insurance are potentially different from purchasing a stand-alone DI contract, I drop the former contract types from my analysis. Likewise, I drop all individuals that ever cancel their insurance contract to focus on the group that keeps their insurance. In addition, I have to drop all miners, who are covered by a special public DI program, and people, who bought their private DI before 2001 due to a major pension reform that removed private DI coverage for people younger than 41 in 2001 (see Seibold et al. (2021) for discussion). Finally, I apply the same selection criteria as in the other data sets, by only retaining men who purchased their private DI contract after turning 25, which is the starting age in my model

B.3 Social Security Register Data

The IAB (Institut fuer Arbeitsmarkt- und Berufsforschung) collects information on the employment and labor market related benefit history of each individual in Germany who was in one of the following states between 1975 and 2017: employment, unemployment insurance beneficiary, social assistance recipient. Individuals working in a "mini-job" (defined as earning below a certain minimum threshold, currently 450 Euros per month) or taking part in job-retraining appear since 1999. Civil servants and self-employed are exempt from social security contributions, so they do not appear in this data set.
The SIAB is a random 2% sample drawn from the universe of these social security records. It contains the employment and benefit history of 1,875,439 individuals, comprising 66,961,520 spells. The information collected in this data set is relevant for determining unemployment insurance entitlement and benefit level. Hence, the data set has comprehensive information on the daily wage, the occupation title and classification (2010 version), some demographics (age, gender, citizenship), Work arrangement (full-time vs. part-time), sector of employer, residency (municipality), and benefit receipt. In addition, the IAB reports the reasons for transitioning employment states including public DI receipt, which allows me to identify these spells in the SIAB data. I use the data to estimate the wage equation (8), the labor market moments, the disability probability by risk group, and the population risk group distribution (see section 5).
I transform the different spells into an annual panel of individual (employment) histories. If spells span several (calendar) years, I divide them into annual spells, e.g. if a spell lasts from May 2011 to May 2012, I create two spells, one from May to Dec. 31^{st} 2011, and the other from Jan. 1^{st} 2012 to May 2012. Multiple spells within a given year are

ranked according to their timing. I retain only the longest spell in each year.[53] Since my model and estimation sample focuses on the time after the 2001 pension reform, I restrict my sample to spells recorded between 1992 and 2017. I include the years 1992 to 2000 because they provide some additional information, especially for people that claim UI or DI after 2001.

To reflect the annual frequency, I transform daily income into annualized income (2013 Euros). The annualized income corresponds to the reported daily income of the retained employment spell multiplied by the number of days in that year. The income information is third-party reported, so measurement errors are negligible. However, income in the SIAB is only reported up to the social security contribution limit, thus I impute wages above the contribution limit with a series of Tobit-regressions (see Dauth and Eppelsheimer (2020) for details).

After constructing the panel, I start cleaning the data set. Appendix D provides further details on the cleaning steps and the merging process. Here I provide a brief overview over the steps taken. In an initial cleaning step I only retain spells related to employment, unemployment, non-participation and health-related departures. Some spells are recorded twice in the data set, because they originate from different sources. I delete one of these spells, whereby I retain the more detailed spell or the health-related spell. Before I can merge the risk group mapping from the private data by occupation code to the SIAB, I need to deal with spells which have missing occupation information, e.g. social security spells. I assign the individual mode occupation code to these spells.

After dealing with missing occupation spells, I merge the mean, median and mode risk group from the private data to the SIAB by occupation code. If I fail to match an occupation to a risk group from the insurance data, I look up their risk-group mapping in the insurance company's risk table and add their risk-group manually. This can happen due to censoring requirements: If too few observations are within an occupation-risk group cell, this cell is censored in the aggregated insurance data. Overall, I can match all observations with non-missing occupation codes to a risk group, which corresponds to 97.15% of all observations in the raw data and 99.8% in the cleaned sample. Based on this mapping I later estimate the risk-group distribution in the whole population as well as controlling for the relationship between income and risk-group.

Finally, I apply the same sample selection criteria as above: I retain all all individuals that are between 25 and 65 years old[54], are not reporting zero income[55], and do not work in non-standard employment forms (e.g. apprenticeship, early retirement,...) or are temporary employees. The final sample then consists of 32 million person-year observations.

[53] I tried other common 'retention' criteria, such as the spell with the largest income or weighting by spell duration. The results are insensitive to this choice, so I went with the initial strategy.

[54] In the cleaning step I retain individuals between 20 and 65 years, but drop the ones below 25 in the estimation

[55] Transfer income is also documented and well different from zero. Therefore, zero income spells refer to a special subgroup of "non-eligible" yet documented individuals, which I drop from my analysis, or individuals with missing information.

Table C.1: Comparison between both Occupation Title to Code Mapping Strategies

Flag	Number of Observations	Percent
Perfect Overlap	-	73.06
Different Assignment	-	6.11
Only Line-by-Line Assignment	-	0.0
Match Line-by-line, not contained in Risk Table	-	1.22
Only Risk Table Assignment	-	0.23
Both: No Assignment	-	0.61
No Match Line-by-Line, not contained in Risk Table	-	18.76
Total	Confidential	100

The table presents the overlap in occupation code assignments based on the "String Matching" relying on the company's risk table and the "Line-by-Line" matching.

Appendix table D.1 presents the summary statistics and how the sample selection criteria affect the sample composition.

C Appendix: Occupation Code Assignment

As explained in appendix B.2, the private insurer's data only records people's occupation by title. However, in the public data, the occupations are only recorded by their occupation code. Therefore, I map each occupation title in the private insurance data to the corresponding occupation code (2010 version) as specified in the handbook of occupation titles published by the German Unemployment Agency.
I apply two different approaches to assign the occupation code: (i) "String Matching" based on the insurer's risk table mapping occupation titles to risk groups and (ii) "Line-By-Line" matching where I search for each occupation title the corresponding occupation code by hand in the official handbook. I employ both approaches as "String-Matching" allows me to observe more information on how the insurance evaluates risks and prices them, while the latter approach allows me to match more occupation titles to the respective occupation code.

Appendix table C.1 shows that 72 percent of the sample receive the same occupation code under both approaches and only 7 percent are assigned different codes. The main reason for the latter is that the risk table is more aggregated then the actual occupation information from the contract data. Consequently, the "Line-by-Line" approach can match at a finer level. Likewise, 1.31 percent of contracts receive only a occupation code in the "Line-by-Line" but the corresponding occupation titles are not contained in the risk table. Finally, the last row of table C.1 yields 18.76 percent of observations for whose "occupations" no matching occupation code can be found. This number corresponds to the unmatched occupations under the "Line-by-Line" approach and table C.6 in section

C.2 displays the underlying reasons.[56] Taken together, both methods produce similar mappings, thus the results are robust to the choice of either mapping.

C.1 String Matching

The first procedure is based on the insurance companies occupation-to-risk-group mapping. The company uses a table where each row corresponds to an occupation title (of any occupation that ever applied for an disability insurance contract) and assigns this occupation to a risk group. I match the contained occupation titles to their codes based on the German Unemployment Agency's official mapping. Due to differing naming conventions, string matching is not feasible and I assign the occupation titles to the corresponding codes by hand. I create a flag to control for conflicts in this assignment (assignment not unique, old occupation title,...). Since the insurance company draws on the same source for classifying occupations and periodically updates it, the flag is empty here.
After adding the occupation codes to this table, I merge the table to the contract data based on occupation titles (string matching). In this first step, I can match already 78 percent of all contracts to their corresponding occupation code.[57] To match the remaining 22 percent, I check the data row-by-row why the matching failed. I resolve these conflicts by applying the following approach:

1. If the job title from the contract data is not contained in the risk table (for example change of naming convention), I search for it in the job classification table provided by the German Unemployment Agency. I retrieve the corresponding occupation code and search in the insurer's risk table for a match. If a match is produced, I check if the occupation titles and descriptions are similar. If they are, I store the occupation title as used by the insurer in a new variable.

2. If neither the job title nor the associated occupation code are contained in the insurer's table, I apply a "nearest neighbor" approach by checking for slight variations of the occupation code in the risk table. I proceed as follows:

   (a) Is there an occupation whose occupational code only differs in the 5th digit? If yes, use that occupation's title and store it in a new variable, conditional

[56] The "String Matching" approach is able to match 0.23 percent of occupations which are later identified as self-employed individuals. Theoretically, these occupations could also be matched in the "Line-by-Line" method. Since self-employed individuals, however, are not eligible for public DI receipt, I have decided to not assign them any occupation code and rather mark them as self-employed. Also I was only able to find the occupation code for roughly 20 percent of the self-employed, which is why I later forced them to "NA".

[57] Approximately 27% of all contracts can be matched to their corresponding occupational code. I can match another 51 percent controlling for case sensitivity, spelling errors, the treatment of (ä, ö, ü), or additional information the insurance collects which matter for the risk assessment but not for the occupational classification, e.g. share of office work, exposure to hazardous chemicals, etc.

on these occupations being almost identical (e.g. different levels of managerial positions receive different digits).

(b) If (a) does not produce a match, I check if the risk table contains any occupation whose first 3 and final digit are identical to the occupation code of the unsuccessfully merged occupation. These differences can occur based on very narrow specialization, for example gardeners growing fruits (code: 12112) versus flowers (code 12122) differ in their 4th digit, yet both classify as gardeners (code 12102). If I can match occupational code (first 3 + final) and title successfully, I store the occupation title in a new variable.

(c) If (b) does not produce any match, I check for existing neighbors with respect to variations in both the 4th and 5th digit. These cases can arise for special occupations which are pooled into one general term, for example "Ausbildungsmeister" (apprentice trainer/mentor) is not contained as an extra occupation but the first three digits of its occupational code coincide with "Master of Education". Again, if I am able to find a matching occupational code with a similar occupational title or educational background, I store this matched occupational title in a new variable.

If I am unable to match a job based on its "nearest neighbor", I assign the value "NA" to it, indicating the failure to match it (given the next two steps also do not yield any match).

3. Some people state foreign occupation titles, which I match to their German equivalent (official conversion). This occurred for only two occupational titles.

4. There are six occupations for which the insurance company treats as identical despite having different occupation codes (called synonyms by the insurance). I refer to these occupations as "by insurers discretion".

I create a flag to mark each of these different steps. Appendix table C.2 summarizes the final distribution of this flag. As aforementioned, 78 percent of contracts are perfectly matched or after correcting for minor mistakes. I can match another three percent based on steps 2.) to 4.), so it is very unlikely that our assignment strategy biases our results systematically. Finally, I am unable to match roughly 20% of the contracts to the risk table or some occupation code.

Table C.3 presents the reasons for the matching failure. The most common reason is that people are still in education, training or high-school so that they still have to decide on an occupation. This accounts for 67 percent of all failures. Another 29 percent cannot be matched due to missing values in the occupation variable. The remaining 4 percent are either self-employed individuals, home producers, or people in-between jobs (unemployed, interns, community service,...). Note that these occupations also cannot be

Table C.2: Flags for Matching Procedure (String Matching)

Flag	Number of Observations	Percent
Perfect String Match	-	27.64
Correction of minor mistakes	-	51.03
By neighbor (5th digit)	-	0.48
By neighbor (4th digit)	-	0.62
By neighbor (4th and 5th digit)	-	0.07
Foreign title	-	0.12
Insurer's discretion	-	0.04
Discretion (researcher)	-	0.02
Not matched	-	19.98
Total	Confidential	100

The table presents the distribution of the flag indicating how occupations contained in the risk table and the contract data were matched.

Table C.3: Reasons for Matching Failure (String Matching)

Flag	Number of Observations	Percent
In-Training/Education	-	41.17
High-School Student	-	21.57
Missing Occupation Title	-	26.88
Self-Employed	-	3.74
Occupation: Employee, Home Producer	-	0.29
Community/ Military Service	-	0.17
Intern	-	0.02
Unemployed	-	0.02
Unable to find matching occupation	-	0.04
Occupation not in risk table	-	6.10
Total	Confidential	100

The table presents the distribution of the occupation titles that could not be matched in the string matching (risk table) approach. The total corresponds to the category "Not matched" from table C.2.

matched based on the "Line-by-Line" matching in section C.2.Nonetheless, 6.1 percent of individuals work in an occupation that is not contained in the risk table. About 55 percent of these observations are military personal, which in the past could purchase private DI, but recently are in a separate insurance market. This poses no problem to our analysis, as military personal are not subject to the public disability insurance system and we drop them later anyways. The remaining 45 percent of "unmatched" occupations cannot be matched despite our best efforts. However, since they constitute less than one percent of all successful matches, they do not bias our estimation results.

C.2 Line-by-Line matching

The second approach tries to improve upon the first by directly matching the occupation titles from the contract data to the corresponding job classification code from the handbook of job classifications provided by the German Unemployment Agency. As before, string-matching is not feasible, thus I match each occupation by hand. I create a flag that documents the source for each match. Since I am not able to match all occupation titles uniquely to a 2010 occupation code, e.g. because the 1988 occupation classification job title is reported, I generate an additional variable that reports whether a match was unique or not. I then proceed as follows:

1. First, I search for the occupation code by occupation title in the 2010 handbook of the German Unemployment Agency. I assign the corresponding occupation code only for precise matches with respect to occupation titles.

2. If I am unable to find the occupation title (or a precise match) I turn to older versions of these tables, the 1992 and 1988 versions. In those tables, I search again for the old occupation code by occupation title. For precise matches, I extract the old occupation code and searched for its mapping into the 2010 code in the transformation tables provided by the German Unemployment Agency.
   However, these matches are not necessarily unique because the 2010 version is more detailed. Hence, I applied the following steps:

   (a) If the occupation title/description rules out certain matches based on the old code, I drop them, e.g."Stukkateur" (mason) has the code "4810" in 1988, which is associated with 4 possible 2010 codes, two of which I rule out as they refer to "Stukkateur-Meister" (mason master), because they are a separate category, even in the insurance data. From the remaining two, one was referring to "carpenters", which I could rule out. The remaining one is the unique match.

   (b) Some old occupation titles contain further descriptions, often in brackets behind the actual title. I use this additional information to look for a match in the 2010 handbook and compare the resulting code with the one obtained from the transformation table. If they match, I treat them as a unique match, e.g. "Sicherheitsberater" (Work Safety expert) has different potential matches (work areas), but in the 1988 version, there is only one "Sicherheitsberater" without any additional terms (the default occupation, so to speak) which clearly identifies this occupation as an engineer. Only one of the listed occupations refering to "Sicherheitsberater" in the 2010 handbook is an engineer, so the match is unique.

   (c) If still several candidate occupations remain after steps (a) and (b), I document all possible candidates with their occupation codes (see table C.6).

3. If I am unable to match an occupation on the handbooks from 1988 to 2010, I apply an internet search where I search for "occupation title + KldB". [58] Often these

[58] "KldB" is the German abbreviation for "Klassifikation der Berufe", which translates as job classification system.

Table C.4: Flags for Matching Procedure (Line-by-Line)

Flag	**Number of Observations**	**Percent**
Perfect Match	-	69.84
Old Job Title	-	9.41
By neighbor (5th digit)	-	0.14
By neighbor (4th digit)	-	0.00
By neighbor (4th and 5th digit)	-	0.00
Foreign title	-	0.02
Insurer's discretion	-	0.18
Researcher's Discretion	-	0.80
Not matched	-	19.61
Total	Confidential	100

The table presents the distribution of the flag documenting the "by hand" matching approach.

occupations can be found on the website of the German Unemployment Agency. I provide the link to these web-pages in my code.

4. If none of the above returns a precise match, I report "NA" for the occupation code.

Appendix Table C.4 reports the distribution of matches. 69.84 percent of observations could be directly matched and an additional 9.41 percent of observations via their old occupation title. The contribution of all other procedures are negligible. 19.61 percent of contracts could not be assigned to an occupation code.

Table C.5 presents the distribution of occupations matches with respect to whether a match was unique or several potential occupation codes are applicable for the same occupation title. 70.2 percent of the sample could be uniquely matched (4,543 occupations in total). From the remaining 29.8 percent, 7.7 percent had two competing occupation codes (162 occupation titles), while 2.5 percent had even 3 or more competing codes (154 occupation titles). As in table C.4, 19.6 percent of observations could not be matched to any occupation code.
The main reason for multiple occupation codes is the updating of the occupation codes in 2010, which were more diversified then the previously used codes (1992, 1988). Contracts contain the occupation title of the respective year of purchase, implying that all contracts before 2010 used the 1988/1992 codes. These occupations still exist in the 2010 version, but sometimes where split into different "specializations". On a whole, 80.5 percent of the "non-unique" matches are due to the differentiation into specializations within an occupation. The remaining 19.5 percent are explained by the insurance company summarizing several similar occupations with differing codes into one occupation.[59] The differences in the 2010 occupation codes, however, are minor and the results are robust to interchanging the codes.[60]

[59]For example, "Steuerassistent/Steuerfachgehilfe" are one occupation group in the insurance data but

Table C.5: Distribution of Occupation Title to Occupation Code Mapping - Uniqueness of Match

Flag	Number of Observations	Percent
Unique Match	-	70.20
Two Candidates	-	7.71
Three Candidates	-	2.48
No Match	-	19.60
Total	Confidential	100

The table presents the distribution of the uniqueness of matches.

Finally, table C.6 explores the reasons for the failures to match the occupations. The most common reason for matching failures is that individuals are currently "out-of-employment", either because they are unemployed, not participating in the labor force or because they are still in training, education or high school and have yet to choose an occupation, thus no occupation code can be assigned to these individuals. They account for 66.9% of all matching failures. 27.4% of all matching failures are due to missing or corrupted[61] occupation information. In 0.6% of cases people stated "Employee", "Worker" or "Home Producer" as their primary occupation, which is not specific enough to allow for any match. Likewise, I am unable to match most self-employed individuals to their respective occupation code as the data often refers to them as "entrepreneurs" or "self-employed". They account for another 4.9% of matching failures. Finally, I am unable to match 174 observations (0.2%) reporting a "specific occupation" to any classification code. Since the occupation title stated does not exist in the occupation code handbook published by the German Unemployment Agency, it is very likely that they are own creations either by the insurance holder or by the insurance company.[62]

C.3 Private DI market - price comparison

This section presents the price of each risk group across different insurance companies in 2021. The objective behind this price comparison is to show that the insurance provide whose data we are using offers comparable contracts to other insurance companies, thus being representative for the market as a whole. See Seibold et al. (2021) for a more thorough discussion and validation of this point.

In this price comparison we proceeded as follows. First, we select the insurance companies

correspond to two different occupation codes, "72303" and "72302" respectively.

[60] More than half of the observations with two candidates are "engineers" (Ingenieur o.n.A.). The 1988 occupation codes allowed for "not stating a sub-field of engineering". This was abolished in the 2010 version and all engineers must provide their field of specialization, such as mechanics, electrical engineering, etc. The formerly "engineers" are now either belonging to occupation code "27104" or "27304".

[61] Stated as "unable to match to occupation" in the data

[62] In most cases a "related" job exists in the sense that parts of the occupation title appear in other occupations as well, but no unique match can be created.

Table C.6: Flags for Matching Procedure (Line-by-Line)

Flag	**Number of Observations**	**Percent**
Stay-At-Home Parent	-	2.79
Missing Occupation Title	-	27.39
Occupation: Employee, Home Producer	-	0.59
Community/ Military Service	-	0.17
Intern	-	0.02
Unemployed	-	0.01
In-Training/Education	-	41.96
High-School Student	-	21.99
Self-Employed	-	4.93
Unable to find matching occupation	-	0.15
Total	Confidential	100

The table presents the distribution of the occupation titles that could not be matched in the line-by-line approach. The total corresponds to the sum of "Self-Employed" and "Not matched" from table C.4.

we want to include in our search. To be included, we require that the insurance company has sold at least 100,000 contracts. This leaves us with 13 companies for a total of 9.38 million contracts. Second, we then search for each company by name for online information on their pricing, which usually comes in the form of an online calculator tool. This calculator then generates a price offer based on the entered information (see below). ´ Only 7 companies offer such an online tool, but they account for 67.4% of the market, so I am confident that our results hold even for the companies that up till now do not offer such an online tool. Third, we then compute the prices for 12 to 18 selected occupations from each risk-group[63]. To this end, we enter the required information into the online tools as follows:

- Age at purchase: 25 years
- Age at contract end: 60, 65 and 67 years
- Benefits: €1,000/month
- Highest (occupational) degree (if requested): Either explicitly stated or obvious, e.g. doctor has an university degree
- Share of working hours spend in office: Explicitly stated, but thresholds differed somewhat from company to company, e.g. $> 75\%$ or >80%. For construction workers we picked the minimum, for white-collar workers (if not otherwise indicated) the maximum).
- Number of subordinates: Default value set to 0, except for management positions where I picked both zero and the maximal available value

[63] Number of occupations included depends on the actual number of people with this occupation in each group. We provide the list of occupations upon request.

- Self-employed or civil servant: No
- Nationality: German
- Smoker, drinker, other addictions: No
- Dangerous hobbies: No

In short, I elicit the prices for a 25 year-old employed and healthy individual who wants to insure himself/herself against disability until the age of 60/65/67. I use three age cutoffs because not all insurance companies offer contracts up to the age of 67 for high risk-groups while some insurance companies only insure people up to age 67.

The results are presented in table C.7. Risk groups are entered row-wise, while each column corresponds to the insurance premium of this group charged by a distinct insurer. Overall, prices are very similar for the same risk groups across insurance companies, especially for the (better than) average risk group 1,2 and 3. This relationship is independent of the final contract age. In contrast, risk groups 4 and 5 show considerably more variation in prices across companies and also availability at different ages. The underlying explanation is that insurers become more restrictive regarding the occupations they insure. Even if they allow certain occupations in, they often impose alternative contract end ages for these occupations. For instance, the insurer in the first column accepts the most occupations while not imposing any additional age restrictions. In contrast, the one in the second column, rejects several occupations still insured by the first insurer or only allows them to buy insurance up to the age of 63. As the empty cells show, some insurers even refuse to offer contracts to high-risk occupations up to age 67. Taken together, the price variation in risk group 4 and 5 is explained by two composition effects: Insurers differ with respect to the occupations they insure (occupational composition) and the contracts they offer (contract menu differs). Hence, the first panel of table C.7 shows the most complete comparison for each risk group across insurance companies. The other two panels show that even if certain occupations and thus risk groups are no longer insured, the insurance companies do not target the good risk groups more.

I conclude that insurance companies offer comparable contracts across the different risk groups. They do not target certain risk-groups (or occupations), especially the better risks. Thus, each risk-group should be close to indifferent from whom to buy insurance, conditional on not being rejected by the insurer. It follows that our insurance company is representative for the market with respect to the offered contracts and we do not expect that they attract different customers relative to its competitors.

Table C.7: Private DI price comparison

The table below shows the price comparison across insurance companies for different ages at contract end. The prices have been computed under the assumption that a 25 year old, employed and healthy person wants to insure himself/herself with €1000/month against disability until the respective contract end. All prices in €/month.

Risk-Group	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Contract end at age 60							
1	23.241	18.862	22.435	22.423	-	-	19.089
2	27.812	21.433	24.821	27.272	-	-	23.613
3	46.301	44.23	37.534	44.18	-	-	39.618
4	84.286	76.971	64.022	69.773	-	-	63.314
5	90.142	96.971	80.49	94.379	-	-	71.409
Contract end at age 65							
1	32.664	28.612	32.367	31.59	-	-	28.662
2	39.236	32.23	35.755	38.448	-	-	35.538
3	65.509	63.735	53.064	63.09	-	-	59.329
4	119.612	105.905	90.372	99.555	-	-	82.597
5	125.622	132.174	107.489	139.704	-	-	94.763
Contract end at age 67							
1	37.253	33.871	37.820	37.075	33.253	31.931	33.819
2	44.846	38.049	41.747	45.023	39.894	41.529	40.859
3	74.856	75.736	69.693	69.801	68.117	68.835	69.715
4	136.785	107.266	107.193	-	118.806	114.315	94.889
5	147.393	-	-	-	132.594	139.844	110.417

D Data Appendix - SIAB

This appendix provides further information on the construction of my sample in the SIAB data as well as on the matching procedure with the private data. Finally, I present some summary statistics from this sample construction at the end of this appendix section.

D.1 Cleaning the data

While I write separate cleaning files for each analysis, the ones for Labor Moments, DI probabilities and the risk group distribution are actually identical. The cleaning file for the income process is a subset of the this file. Since I only retain employment spells for constructing the income estimation sample, I do not need to assign an occupation code for non-employment spell with missing occupation information. In turn, I need to deal with top-coded income spells in the income estimation. However, all these files follow the same general structure and they are all using the panel version of the data (so after I transformed the initial spell data into an annual panel). I point out the respective differences as they come.

1.) I identify different spells of interest, such as regular employment, unemployment, social assistance claims, and so on. I then drop all observations that do not fall into one of the following categories: "employment", "unemployment", "social assistance recipient", "non-participation" or "health related absences".

2.) Some spells are recorded twice in different systems, such that these (almost) identical spell appear also twice in the data. I delete these duplicates whereby I give precedence to health-related causes, e.g. a disabled person can also appear as "non-participating" in one source and "disabled" in the other. The ordering of these two spells is random, such that I retain only the spell related to the "health-state". Finally, even the same disability spell can be recorded in two different sources. In that case I pick the source with a more detailed "health state" description.

3.) I construct a variable that measures the spell duration within a given calendar year for each individual and spell, so spells spanning several years are appearing in each calendar year with the respective number of days. Based on this annual spell duration variable, I define the dominant employer for each individual as the longest spell in this calendar year.
In the income process cleaning file I also compute the annualized income for these dominant employer spells by multiplying the daily wage with the number of days in that year.

4.) I prepare the SIAB data for the merging procedure, which I explain below in detail. To this end, I need to assign an occupation code to spells with missing occupation information, e.g. social assistance spells or disability spells. I apply two methods:

- I assign the mode occupation of the individual to these spells.
- I assign the last observed occupation to the missing spell.

With respect to the results, the methods produce similar estimates, albeit the mode method is a little bit less sensitive and matches slightly more occupations. Hence, I choose this method as my baseline setting.
For the income process cleaning, this step is skipped: Since I restrict my sample to the employed individuals, they all have an occupation code and no further assumption is needed.

5.) Described below: I merge the private data (occupation code to risk-group mapping) to the SIAB.

6.) I adjust the income variable for inflation by dividing the income by the CPI.

6.a) *(Only for income process estimation)* I use the code described in Dauth and Eppelsheimer (2020) to impute the wages for top-coded spells, after adjusting the underlying model to my setting. Thus, I first assign each individual to a unique risk group based on the same method used later in the estimation (including same seed value). I then estimate the wages for the top-coded spells based on the adjusted code and store them in a separate variable.

7.) I produce some initial summary statistics (tables below) and then drop observations meeting one of these criteria:

- People that are older than 65 (retired) or younger than 20 (in education) [797,990 spells]
- Individuals that are temporary workers only ("Leiharbeiter"). [333,753 spells]
- Military personal (civil servants that sometimes appear in the SIAB) [4,329 spells]
- People with zero income [549,345 spells]

Appendix table D.1 presents some summary statistics of interest for the whole sample and after the imposing the sample selection criteria discussed in this section.

Table D.1: Sample Restriction and Composition

The table below shows the composition of the sample under different sample restriction criteria. Column (1) displays the sample means for the full sample of either employed, unemployed, non-participating or social security beneficiaries. Column (2) presents the baseline sample after imposing the sample selection criterion (as shown in the table) and column (3) shows the baseline sample conditional on matching the occupation to a risk group from the private company data. Column (4) presents the sample means for the subsample of employed individuals and column (5) the sample means for occupation codes successfully merged to a risk-group. The sample window is 1992 to 2017.

	(1)	(2)	(3)	(4)	(5)
Age	39.15	39.63	39.65	40.54	40.76
Spell-duration	187.26	189.71	188.48	297.66	301.63
daily wage	73.12	73.72	74.02	88.56	89.93
Male	0.5532	0.5518	0.5558	0.5536	0.5489
Share employed	0.6253	0.6299	0.6448	1	1
Share part-time	0.1289	0.1298	0.1329	0.1964	0.1971
Share full-time	0.4962	0.4998	0.5116	0.8032	0.8029
Share unemployed	0.1945	0.1956	0.19	-	-
Share social assistance	0.031	0.0319	0.0308	-	-
Share non-participation	0.1321	0.1251	0.1172	-	-
Share public DI	0.0005	0.0005	0.0005	-	-
Sample selection criteria					
Share: Occ. merged from risk-table	0.9715	0.9759	1	.9727	1
$19 < age < 66$	0.9765	1	1	0.9894	1
Temporary worker	0.01	0	0	0.01	0
Military personal	0.0001	0	0	0.00002	0
# Obs.	33,952,157	32,816,085	32,024,456	14,824,126	14,128,622

D.2 Merging Risk Groups to Public Data

As discussed in appendix C, the private insurer records the occupation title, which are more disaggregated than the occupation codes. Since the SIAB only reports the occupation codes, I need to assign the respective occupation code to each occupation title in the private data before I can merge the risk group mapping to the SIAB. Yet, several occupation titles can share the same occupation code despite falling into different risk groups. I compute four different statistics to account for this: the mean (baseline), median and mode (min and max) risk group by occupation code, at the 5 digit, 4 digit and 3 digit occupation code level.[64]. In my baseline, I use the mean risk group and assess the robustness of this exercise by using the other assignment strategies. Note that in general the mapping of occupation code to risk group will not be discrete anymore, so I later need to discretize them again.

Next, in order to ensure a large overlap even for periods in which an individual is out of the labor force (social assistance, disability) when merging the aggregated data to the SIAB, I have to assign an occupation to spells for which no occupation code is reported. As mentioned in the previous subsection, I choose two approaches to deal with those spells: (a) I assign each individual their mode occupation code or (b) I assign them their last observed occupation code. Reassuringly, the results are robust to both approaches because people hardly change their occupation[65], so that I choose the mode - approach as the baseline approach.

After this preparation, I merge the aggregated data from the private insurance company to the SIAB panel based on the occupation codes. Thereby I proceed as follows:

1. I merge all occupations based on their 5-digit occupation codes. If a cell in the private data had less than 3 observations for the 5-digit code, the corresponding risk group had to be censored (set to missing). In that case, I replace the corresponding risk group with the risk group based on the 4(3)-digit occupation code given that those cells are non-missing.

2. Some 5-digit occupation codes from the public data are not contained in the private data. To get a chance at matching them to a risk group, I check whether I can assign them based on the 4-digit occupation codes (3-digits plus skill level [fifth digit]). If possible, I assign the corresponding risk group to these occupations. As before, if certain cells are censored due to small cell sizes, I attempt to match them based on the corresponding 3-digit occupation codes.

3. I check whether I can increase the overlap by matching the remaining unmatched occupation codes based on the 3-digit codes (first 2 plus final digit). Successful

[64] 3(4) digit code refers to the combination of the first 2(3) numbers plus the final digit recording the skill level.

[65] This is precisely the reason I cannot include fixed effects in the labor income estimation equation 8

matches receive the corresponding risk-group.

4. Finally, I check all the occupations codes by hand which had no match with occupation codes from the private insurance data. I proceed by looking up the corresponding occupation titles and searching for them in the insurer's risk group occupation mapping table.

Following these steps, I am able to match all occupation codes to a risk group. Unsuccessful matches only occur when no occupation code is recorded (across all spell of an individual), affecting 2.8% of all spells.
Before discussing the summary statistics which document the merging success below, I want to point out again that both the mean and median are no-longer discrete. While some computations allow for using continuous risk groups (e.g. income regression, disability probabilities by risk group), it is still sensible to discretize the risk groups again. I discretize the risk-group - occupation mapping using the following two approaches. First, I assign each individual to the lower risk group if the mean (median) risk-group is less or equal $x.5$, e.g. if an occupation has the mean risk group 1.49 then it falls into risk group 1, but for mean risk group 1.5 it would be assigned risk group 2. Second, I assume that individuals are uniformly distributed on the interval between the two nearest integers around the mean (median). Drawing a number from an uniform distribution over this interval, I assign an individual to the larger risk group if the draw is larger than $1-$ (RG - next smaller integer), which is the probability of falling below the mean. For example, let the mean be 1.6, then I assume that the probability of falling into risk-group 1 (next smaller integer) is equal to $(1-(1.6-1)=0.4)$. Again, both approaches deliver similar results, but the second approach tends to put more mass on smaller risk groups (groups 1 and 5). Hence, I use the probabilistic assignment as my baseline method.

E Appendix: German Institutional Setting

E.1 Income Taxation

In the following, I discuss the German income tax code in its version from 2013 (for singles). Compared to the 2020 tax code, the same tax rates apply to today, only the tax brackets have shifted upwards to account for inflation and wage growth.
The German tax code consists of five tax brackets with increasing marginal tax rates in each bracket. The marginal tax rates range from 14% at the bottom to 45% at the top. The first tax bracket ranges from zero to the tax-free allowance, which was 8130 Euros per year. This income is not taxed.
The second bracket ranges from the tax-free allowance to 13,469 Euros of annual income.

The tax liability in this bracket is computed by the following formula to ensure the continuity of the tax schedule, where y_{it} refers to annual labor income:

$$Liability = (933.70 * \frac{(y_{it} - 8130.00)}{10,000} + 1,400.00) * \frac{(y_{it} - 8130.00)}{10,000}$$

whereby this formula ensures (a) the continuity of the tax liability at the bracket limit and (b) that the marginal tax rates are increasing in income.
The third bracket ranges from $13,470$ to $52,881.00$ Euros per year. Again, the continuity of the tax liability is ensured by applying the following formula:

$$Liability = (228.47 * \frac{(y_{it} - 13,469.00)}{10,000} + 2,397.00) * \frac{(y_{it} - 13,469.00)}{10,000} + 1,014.00$$

Household income falls into the fourth tax bracket if it exceeds $52,881.00$ but not $250,730.00$ Euros per year. Starting at this bracket, the German tax code simplifies, as individuals pay a linear tax:

$$Liability = 0.42 * y_{it} - 8,196.00$$

where the subtraction of $8,196.00$ ensures the continuity of the tax schedule.
The last bracket contains the income exceeding $250,730.00$ Euros per year. People pay here a marginal tax rate of 45% and the tax formula again looks as follows:

$$Liability = 0.45 * y_{it} - 15,718.00$$

The resulting income tax liability is always rounded down to the next integer.
The income of married couples is assessed jointly. Their incomes are pooled and then divided by two. The resulting expression is entered into the tax formula and the tax liability is computed. Finally, the resulting tax liability is multiplied by two, which is then the final household income tax liability. Given the progressivity of the tax schedule, this tax liability is always less or equal then the tax liability for separate assessment. Since joint household taxation is the default setting for married couples and they need to explicitly opt for separate taxation, most households opt for this arrangement.
Moreover, each household in Germany has to pay an additional tax called "Solidaritaetszuschlag" (solidarity surcharge) which amounts to 5.5% of the income tax liability (not income). I also take that special tax into account when computing the tax liability.
Pension and public DI benefits are tax-free to a certain percentage of total benefits, while the remained is subject to the standard income tax schedule. The fraction of your gross (DI) pension, which is tax-free, depends on the year you first claimed pension. This year is "fixed" in the sense that the pension tax treatment does not change thereafter. For example, someone receiving a pension from 2005 or earlier has a tax-free pension allowance of 50%. After that (until 2020) it reduces by 2% each year, so someone receiving his first pension payment in 2013 will only have a tax-free allowance of 34% (50-(2013-2005)*2).

Starting 2020 the reduction is 1% per year until 2040 when the entire pension is subject to taxation. The fraction, which is subject to income taxation is entered in the above formula. However, most of the pensions are rather small so that almost all of them (more precisely the share subject to income taxes) are within the first two brackets.
Likewise, only a fraction of the private disability insurance benefits is subject to income taxation. As discussed in section E.3 below, this fraction is positive correlated with the remaining time of the benefit receipt: A longer payment period implies a higher taxable fraction. This fraction is then plugged into the income tax formula discussed here to determine the income tax liability. In case of simultaneous receipt of private and public benefits, the taxable income from both sources is pooled to determine the income tax liability.

E.2 Social Security Constributions

In contrast to income taxes, social security contributions are paid at an individual level. The reason is that social security contributions are split between the employer and the employee each paying half of the contribution. These contributions are immediately deducted from the gross wage and paid by the employer to the respective fund or insurance company.
Employed individuals pay social security contributions (total numbers in brackets) to the pension fund (18.9%), to the unemployment insurance (3%), healthcare insurance (15,5%), and nursing (long-care) insurance (2.05%), which amounts to roughly 40% of gross wages, 20% paid by the individual themselves. Note, that while the former two contributions are only paid by employed individuals, the later contributions have to be paid by everyone including pensioners and people on (public/private) DI. Thereby, they pay social security contributions on their total benefits and not only the taxable fraction. Therefore, the taxable fraction only matters for income tax treatment but not for social security contributions.
Social security contributions in Germany are capped at a maximum contribution limit. For pension and unemployment insurance contributions this cap is 5,800 Euros of monthly income in 2013, and roughly 4000 Euros per month for the healthcare and nursing insurance contributions. After exceeding these caps, individuals always pay the maximum contribution, but they do not increase in income anymore. Again, these earning caps change on an annual base (usually shifting upwards).
Finally, an important difference between public and private disability insurance receipt is that people receiving only private benefits have to pay the full health and nursing insurance contributions whereas they only have to pay half (like employed individuals) when being on public benefits. Individuals receiving benefits from both insurances only have to pay health and nursing insurance contributions for their public benefits. My code

accounts for all of these distinct cases.

E.3 Annuity Taxation

Private DI benefits are treated as a special form of annuity and are taxed accordingly. Thereby, the amount of taxable income depends on the remaining time of the contract duration. Table E.3 depicts these fraction of annuity income that is subject to income taxation for different remaining terms.

Table A1: Taxable Fraction of Annuity income

Remaining Term	**Taxable Fraction**
0	0
1	1
5	7
10	13
15	17
20	21
30	30
35	35

This table shows the relationship between remaining terms of an annuity and the fraction of its benefits that are subject to income taxes. Greater remaining terms are associated with higher taxable fractions and the relationship is almost linear.

The income tax code is then applied to the so determined taxable income and the taxable fraction of the private DI benefits are treated just as regular labor income. As for public DI or pension benefits, social security benefits have to be paid on gross benefits, which implies the total private DI benefits and not only the taxable fraction. Thereby the distinction mentioned in the previous section applies: Individuals only receiving private insurance benefits have to pay the full amount (employer + employee contribution) for health and nursing insurance. In contrast, public DI recipients only pay the employee contribution (half the amount).

E.4 Public DI and Pension Formula

The formula for computing public DI and pension benefits consists of four factors: The sum of actual and hypothetical pension points, the pension value, the discount factor, and the claim size. I will explain each of these components separately below.
The first factor is the sum of the actual pension points, $actPP$, and hypothetical pension

points, $hypPP$. Initially, the actual pension points are computed as the ratio between individual income and average income (monthly or annually is irrelevant). For incomes above the earnings threshold y_{it}^{max}, set at €5800/month in 2013, the actual pension points are the ratio of the earnings threshold to average income $\bar{y}$:

$$actPP_{it} = \begin{cases} \frac{y_{it}}{\bar{y}} & \text{if } y_{it} \leq y_{it}^{max} \\ \frac{y_{it}^{max}}{\bar{y}} & \text{else} \end{cases}$$

The hypothetical contributions are computed on a monthly base. The law postulates that an individual would have earned pension points according to the monthly average across all the years (s)he has contributed to the pension system. These hypothetical monthly points are then multiplied by (62*12), the cutoff age in months, minus 12*T^k, the age at which the disability occurred[66]

$$hypPP_i = \left(\frac{1}{(T^k - T^0) * 12} * \sum_{j=0}^{T^k - T^0} actPP_{ij}\right) * (62 * 12 - T^k * 12) \tag{E.1}$$

where T^0 is the age at which an individual entered the labor force. For all years above 62, no pension points can be earned. Summing the actual and hypothetical pension points completes the first step.

The second step determines the discount factor $Disc_{it}$, which adjusts the pension benefits for claiming them before reaching the legal retirement age. The pension benefits are reduced by 0.3% for each month an individual claims before the age of 63 years and 7 months. The maximal reduction is 10.8%. Hence, the discount factor is:

$$Disc_{it} = 1 - \min\{0.108, (63 * 12 + 7 - T^k * 12)\} \geq 0.892 \ \forall t \tag{E.2}$$

The third factor is the pension value $PensVal_{it}$, which is just a Euro valued multiplier translating the product of the factors into a Euro-valued benefit. It depends on the state of residence of the claimant. The distinction is made between East and West Germany to account for differences in living expenses. This factor was 28.14 (25.74) Euros for West (East) Germany in 2013. In this paper, I abstract from such distinctions.

Finally, an adjustment for the severity of the work impairment is made: People deemed as fully work-impaired receive a full claim, $HM_{it} = 1$, while those awarded a partial claim receive $HM_{it} = \frac{1}{2}$.

The complete formula then looks like this:

[66] Assuming a hump-shaped earnings profile, this explains why the replacement rate is lower for individuals that claim public DI at earlier ages: The average of their past income is lower and they forgo the higher incomes at later points in their careers.

$$DIb_{it} = (\sum_{j=0}^{T^k - T^0} actPP_{ij} + hypPP_i) * Disc_{it} * PensVal_{it} * HM_{it} \tag{E.3}$$

The Pension System

The pension benefits are computed in a similar fashion as the public DI benefits. In fact, they both apply the same formula with exactly the same factors. The only difference is that there is no "partial claim" and no discounting as long as benefits are not claimed before the legal retirement age. Hence, $HM_{it} = 1$ and $Disc_{it} = 1$ in (E.3).

Likewise, pension benefits are subject to the same tax treatment as public DI receipts. Besides, the same rules for earning additional labor income apply, which I ignore for the same reasons as in the case of public DI receipts.

Finally, a special case occurs when some claiming public DI reaches the legal retirement age. As aforementioned, the benefits for public DI are computed once and are then not re-adjusted. The exception from this rule is when transforming the DI pension into a classical old age pension. In this case, the benefits are re-computed and the DI receipts are treated as contributions to the system. This increases the pension claims in general: First, the partial factor drops (ignored in my model). Second, the discount factor is increased (if less than one) to one. And last, treating your DI income as labor income tends to increase the sum of pension points compared to the computed average. Hence, people see their income rise upon entering retirement. My model accounts for this by recomputing pension benefits upon entering retirement, while keeping them constant over the claiming period.

F Appendix: Health transition probabilities and mortality risk

Table F.1 presents the mortality risk and health transition probabilities on which my computations are based. The mortality risk is taken from the mortality probability table provided by the German Federal Statistical Office (table).

The health transition probabilities are based on the disability risk and recovery probability tables as provided by the German Acturian Society (DAV). The first table was computed in 1997 and its values are contained in columns 3 and 4. Since the table is reassessed periodically, I also include the updated values for 2018. The results, however, are robust to the choice of year. The values for 1997 can be found in Aktuarvereinigung (1997), table 1a and table 10a (average by row). The values in column 5 and 6 are taken from Aktuarvereinigung (2018), which shows that there are hardly any changes compared to 1997.

Table F.1: Mortality Risk for men in Germany, observation period 2011-2013

Age	Mortality Risk	DAV1997 disability	DAV1997 recovery	DAV2018 disability	DAV2018 recovery
Age	**Mortality Risk**				
25	0.0005265	0.002807	0.1274183		
26	0.00054558	0.002807	0.1225644		
27	0.0005348	0.002807	0.1178347		
28	0.00056253	0.002807	0.1129593		
29	0.00061266	0.002807	0.1075248		
30	0.0006439	0.002807	0.1017475		
31	0.00067566	0.002807	0.095844		
32	0.00072405	0.002807	0.0900313		
33	0.00073322	0.002807	0.0840878		
34	0.00077233	0.002807	0.0778347		
35	0.00079924	0.0023012	0.0715534		
36	0.00085893	0.0024604	0.0655245		
37	0.00092543	0.0026587	0.0600292		
38	0.00103712	0.0028520	0.0548864		
39	0.00114203	0.0030383	0.0498151		
40	0.00125967	0.0032306	0.0449464		
41	0.00134366	0.0034725	0.0404114		
42	0.00151628	0.0037716	0.0363407		
43	0.001703	0.0041007	0.0326225		
44	0.00190832	0.0044404	0.0290863		
45	0.00214117	0.0047767	0.0257788		
46	0.00239826	0.0051541	0.0227454		
47	0.00264757	0.0056249	0.0200321		
48	0.00299514	0.0062273	0.0175387		
49	0.00342064	0.0070534	0.0151724		
50	0.00378938	0.0081259	0.0129902		
51	0.00435519	0.0095007	0.01105		
52	0.00488015	0.0112013	0.0094093		
53	0.00534974	0.0132062	0.0079835		
54	0.00601979	0.0155535	0.0066684		
55	0.00657967	0.0182793	0.0054917		
56	0.00709137	0.0213377	0.0044811		
57	0.0080032	0.0246920	0.0036643		
58	0.00878771	0.0282059	0.0029725		
59	0.00952787	0.0317913	0.0023365		
60	0.01036533	0.0353828	0.0017826		
61	0.01123395	0.0403322	0.0013381		
62	0.01193909	0.0454595	0.0010296		
63	0.0127597	0.0510343	0.0010296		
64	0.01422552	0.0570642	0.0010296		
65	0.01512419	0.0635517	0.0010296		
Retirement					
66	0.01645711				
67	0.01755892				
68	0.01950847				
69	0.02082376				
70	0.02213184				
71	0.024409				
72	0.02660199				
73	0.0291149				
74	0.03184282				
75	0.03568586				
76	0.04010634				
77	0.04504754				
78	0.05020154				
79	0.05618894				
80	0.06296475				
81	0.07138584				
82	0.08034131				
83	0.08966652				
84	0.09918926				
85	0.10976565				
86	0.12332763				
87	0.13689281				
88	0.15159255				
89	0.16942819				
90	0.18562917				
91	0.21345515				

continued

Table F.1 continued

Age	Mortality Risk	DAV1997 disability	DAV1997 recovery	DAV2018 disability	DAV2018 recovery
92	0.22875864				
93	0.24681616				
94	0.25110654				
95	0.28847579				
96	0.31934597				
97	0.33600214				
98	0.35797005				
99	0.37769352				
100	0.39896378				

Table F.1: The table displays the mortality risk by age for men based on the values from 2011-2013. The table can be accessed via https://www-genesis.destatis.de/genesis/online#astructure. The last four columns present the disability and recovery probabilities based on the tables published by the German Acturian Society in 1997 and 2018

G Appendix: Estimation of stochastic earnings components

The earnings process described in equation (8) is governed by two i.i.d. stochastic processes: An AR(1) persistent shock (ε_{it}) and a transitory income shock (ϵ_{it}). The AR(1) process depends on the persistence term ρ and the innovation variance σ_η^2 and the transitory shock only on its innovation variance σ_ϵ^2.

I estimate these terms using the methods detailed in Guvenen (2009) and Low et al. (2010) by minimizing the distance between data moments and their theoretical counterparts using the metric described below. The data moments used are estimated variance-covariance matrix ($\hat{\Sigma}$) of the residuals obtained from estimating equation (8) on the SIAB. Their theoretical counterparts are the variance-covariance matrix (Σ) of the sum of the persistent and transitory error component ($u_{it} = \varepsilon_{it} + \epsilon_{it}$) from the same equation. Before discussing the actual estimation procedure in detail, I briefly want to make the theoretical moments more explicit.

Maintaining the assumption that ε_{it} and ϵ_{it} are i.i.d., the variance and covariance of u_{it} is then defined as (dropping the i index for clarity of presentation):

$$var(u_t) = var(\varepsilon_t) + \sigma_\epsilon^2 \tag{G.1}$$

$$\begin{aligned} cov(u_t, u_{t+j}) &= cov(\varepsilon_t + \epsilon_t, \varepsilon_{t+j} + \epsilon_{t+j}) \\ &= cov(\varepsilon_t, \varepsilon_{t+j}) + cov(\epsilon_t, \epsilon_{t+j}) \end{aligned} \tag{G.2}$$

Given the transitory nature of ϵ_t, $cov(\epsilon_t, \epsilon_{t+j}) = 0, \forall j > 0$ and $cov(u_t, u_{t+j}) = cov(\varepsilon_t, \varepsilon_{t+j})$ On the contrary, the persistent shock's variance and (auto-) covariance are time dependent as captured by the persistence term ρ (I define them recursively later):

$$var(\varepsilon_t) = \rho^2 var(\varepsilon_{t-1}) + \sigma_\eta^2 \tag{G.3}$$

$$cov(\varepsilon_t, \varepsilon_{t+j}) = \rho cov(\varepsilon_t, \varepsilon_{t+j-1}) \tag{G.4}$$

where $cov(\varepsilon_t, \varepsilon_{t+1}) = \rho var(\varepsilon_t)$.

Finally, I need to impose an assumption for the persistent shock's initial variance $var(\varepsilon_0)$. While I could impose $var(\varepsilon_0) = \sigma_\eta^2$, I follow the literature that commonly imposes a more flexible assumption: $var(\varepsilon_0) = \sigma_\zeta^2$ with $\sigma_\zeta^2 \neq \sigma_\eta^2$.

Taken together, the elements defined in equations (G.1) and (G.2) with the subsequent definitions define the theoretical variance-covariance matrix $\Sigma(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2)$. For the estimation, I stack the elements of Σ and $\hat{\Sigma}$ into a $Nx1$ vector $\mathbf{vec}(\Sigma)$, where N corresponds to the number of included moment conditions. Let $\mathbf{G}$ denote the difference between the data and theoretical moment vector taking the parameters $(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2)$ as arguments:

$$\mathbf{G}(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2) = \mathbf{vec}(\Sigma)(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2) - \mathbf{vec}(\hat{\Sigma}) \tag{G.5}$$

The stochastic components are then estimated by solving the following problem applying standard GMM methods (Guvenen, 2009):

$$\min_{\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2} \mathbf{G}(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2)' \mathbf{W} \mathbf{G}(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2) \tag{G.6}$$

where $\mathbf{W}$ denotes the weighting matrix. I choose the identity matrix $\mathbf{W} = \mathbf{I}$ following Altonji and Segal (1996). The resulting parameter estimates are reported in table 3.

H Appendix: Computation of counterfactuals

This appendix presents in greater detail how the counterfactuals are solved. I compute the counterfactuals for changes in the benefit level and the rejection rates separately. The changes are centered around the respective baseline values (0.44, 0) and I include changes of 24pp for the rejection rate and 25% for the benefit level around the baseline level. The same counterfactuals are computed with and without a private DI market.

In computing these counterfactuals, I have to impose some assumption on the government revenue and on how welfare is measured.

Revenue neutrality

I impose revenue neutrality in all counterfactual exercises, meaning that the government revenue kept constant relative to baseline. Since the policy changes lead to mechanical and behavioral responses, the government revenue (income - cost) is different under each counterfactual studied. To balance the government budget relative to its baseline level, I levy a lump-sum tax rate which individuals pay in every state of the world until retirement. I choose to levy a lump-sum tax as it has the desirable property of being

non-distortionary.
Formally, the budget-balancing lump-sum tax is computed as:

$$LS = \frac{\hat{R} - R_0}{N_s} * \frac{(1+r)^{T_{retire}} * r}{(1+r)^{T_{retire}} - 1} \tag{H.1}$$

where $\hat{R}$ denotes the government revenue under the new policy regime, while R_0 refers to the baseline revenue level. N_s is the number of simulated individuals and r denotes the real interest rate after taxes. The resulting lump-sum tax is paid constantly until retirement.
The revenue-neutral lump-sum tax rates is determined by minimizing the distance between the simulated tax rate in two subsequent runs, in other words by iterating over the lump-sum tax. While non-distortionary, lump-sum taxes still affect the optimal decisions by altering the budget constraint. Since the lump-sum tax in the current run balances the budget from the previous run, it can still induce large changes in individual decisions and thus the government revenue. Therefore, the program searches for the lump-sum tax rate (and thereby the government revenues) for which the behavioral changes in two subsequent runs are negligible, implying that at this lump-sum tax rates people will no longer change their behavior[67].

Consumption equivalent variation

After solving for the revenue-neutral lump-sum tax rate as described above, the program computes the consumption-equivalent variation (CEV) and stores the simulated decision paths of 16,000 individuals. These are the same individuals under each counterfactual meaning that they have identical shocks (income, health, mortality,...) and risk-groups and only the policy environment changes across simulations.
The CEV is computed by comparing the expected life-time utility under the baseline policies to the life-time utility under the new policy regime prior to the realization of any risk including learning about ones risk group (under the veil of ignorance). The CEV is defined as the (constant) fraction of life-time consumption (α) an individual is willing to forgo in each period under the new policy to receive the same expected life-time utility as at baseline (V_0) under the new policy regime ($\hat{V}$):

$$\hat{V}((1-\alpha)c, l) = V_0(c, l) \tag{H.2}$$

Assuming that the per-period utility function has a CRRA-form, an analytical solution for this expression exists for $\gamma > 1$, where α is defined as:

[67] A test to verify that this approach works is to see whether the program returns a zero lump-sum tax rate in the baseline case. Re-assuringly it does.

$$\alpha = 1 - (\frac{V_0(c,l)}{\hat{V}(c,l)})^{\frac{1}{\kappa*(1-\gamma)}} \tag{H.3}$$

Since my estimation for $\gamma > 1$, I use this formula to compute α. Thereby, a value for α greater 0 implies that individuals are willing to give up consumption to move to the new policy regime. To put it differently, the reform is welfare improving. Vice versa, negative values of α imply that the reform is welfare reducing relative to the baseline.

Finally, when computing the valuation for the second counterfactual, i.e. when exploring under which policy schedule having a private market is welfare-improving, I compare the expected utility with private DI markets to the expected utility without a private market. In terms of the eq. (H.3) this means V_0 ($\hat{V}$) corresponds to the expected utility without (with) a private DI market. Thus, α in this case measures the valuation for having a private market and a positive value implies that a private market is welfare improving.

Return to section 7

I Appendix: Additional Tables and Graphs

Table I.1: Comparison of private and public DI

The table below compares the characteristics across private and public DI in Germany. All prices and benefits are expressed at 2013 values.

Parameter	**Public**	**Private**
Eligibility		
Formal Criterion	Contributions to public pension system for 60 months	Purchased contract
Health Criterion	Unable to work (a) more than 3 hours per day (full claim) or (b) for $3 \leq$ hours per day < 6 (partial claim)	Unable to work for more than 50% of usual work hours
Occupations used for assessing retained productivity	Any occupation	only previous occupation (own-occupation)
Rejected claims	44% (2001-2011, DRV)	30% (GDV, 2014)
Rejected applicants	-	4% (GDV, 2016)
Benefits		
Benefit computation	Pension formula (with discounting for early claiming)	Freely contractible
Average replacement ratio ($\frac{benefit}{grossincome}$)	35%	36%
Maximal benefit	€2320/month	70% of gross wage
Prices		
Price	pension contribution: 9.45% of gross income, up to monthly gross wage of€5800 then maximum contribution: €548/month	3.47ct. - 1.305ct. (see table 2) per € insured

Notes: The most common reason for a rejection in private DI at the claiming stage is that the degree of disability is too small (42% of cases), followed by customers not responding (18%) or not providing the required documents on time (13%) (Hilmes, 2019). At the application stage, only 4% of all applicants are rejected by the insurance company, while 5% of offers are rejected by the customers. 75% accept the standard offer, and the remaining 16% accept an offer with some additional conditions, e.g. exclusion of pre-existing health condition (GDV, 2016).
The years at DRV are continuously updated. Earlier years are available upon request to the DRV.

Table I.2: Targeted data moments, Variances (weights) and Simulated Moments from the Model

The table below shows the estimated values for the targeted moments, their variance, and the corresponding simulated moments from the model estimation step. The final column also shows the standard errors for each moment from the data, which provides additional information on the precision of the model. Note that no standard error can be defined for the median (by definition, standard error is of the mean). Finally, recall that the difference between the data moment and the simulated moment in the estimation step is weighted by the inverse of the variance. Abbreviations are as follows: pDI = private DI, LFP = Labor Force Participation, FT = Full-time, PT = Part-Time.

Moment	**Data**	**Variance**	**Simulation**	**Standard Error**
mean pDI	0.50552	1.65* 10^{-4}	0.4939	0.011546
mean pDI, 1^{st} inc. quartile	0.33888	5.94 * 10^{-4}	0.2453	0.02188
mean pDI, 2^{st} inc. quartile	0.48659	6.63 * 10^{-4}	0.5227	0.023104
mean pDI, 3^{st} inc. quartile	0.573472	5.37 * 10^{-4}	0.5833	0.022862
mean pDI, 4^{st} inc. quartile	0.66588	5.75 * 10^{-4}	0.6241	0.021803
LFP, age 29	0.93525	5.05 * 10^{-7}	0.9465	0.000734
FT, age 29	0.855579	8.81 * 10^{-7}	0.9333	0.001025
PT, age 29	0.079672	5.6 * 10^{-7}	0.0132	0.000777
LFP, age 33	0.94513	7.43 * 10^{-7}	0.8939	0.00064
FT, age 33	0.882776	8.36 * 10^{-7}	0.8754	0.00089
PT, age 33	0.062351	4.68 * 10^{-7}	0.0184	0.00066
LFP, age 37	0.94907	4.31 * 10^{-7}	0.9293	0.00059
FT, age 37	0.896408	5.28 * 10^{-7}	0.9143	0.00081
PT, age 37	0.052661	3.52 * 10^{-7}	0.015	0.00059
LFP, age 41	0.951684	5.27 * 10^{-7}	0.9424	0.00056
FT, age 41	0.902678	6.18 * 10^{-7}	0.9276	0.00077
PT, age 41	0.049006	2.79 * 10^{-7}	0.0149	0.00055
LFP, age 45	0.951424	5.97 * 10^{-7}	0.9423	0.00056
FT, age 45	0.902466	6.72 * 10^{-7}	0.9259	0.00076
PT, age 45	0.048957	2.88 * 10^{-7}	0.0164	0.00055
LFP, age 49	0.950668	3.89 * 10^{-7}	0.9313	0.00057
FT, age 49	0.900232	6.12 * 10^{-7}	0.9098	0.00079
PT, age 49	0.050436	2.65 * 10^{-7}	0.0215	0.00057
LFP, age 53	0.946916	6.34 * 10^{-7}	0.9195	0.00063
FT, age 53	0.893663	8.63 * 10^{-7}	0.9117	0.00086

continued

Table I.2 continued

Moment	**Data**	**Variance**	**Simulation**	Standard Error
PT, age 53	0.053254	$3.6 * 10^{-7}$	0.0078	0.00062
Mean Assets, age 25-27	54352.71	10600000	3461.39	3008.984
Mean Assets, age 28-30	63104.58	5759191	14845.27	1942.597
Mean Assets, age 31-33	83752.07	5970413	30359.74	1842.467
Mean Assets, age 34-36	107676.00	4444191	43390.04	1838.578
Mean Assets, age 37-39	123851.20	4931617	56647.37	1862.858
Mean Assets, age 40-42	141030.00	6029934	72458.05	1930.322
Mean Assets, age 43-45	152744.80	6329157	89696.98	2105.564
Mean Assets, age 46-48	163755.10	7277397	106850.83	2300.355
Mean Assets, age 49-51	169032.00	8348281	124036.24	2505.975
Mean Assets, age 52-54	186003.60	12800000	149498.19	2866.231
Mean Assets, age 55-57	195703.90	13200000	172763.54	3050.546
Mean Assets, age 58-60	201794.10	12700000	190500.01	2978.23
Mean Assets, age 61-63	202461.00	12900000	200801.71	3035.444
Mean Assets, age 64-66	195975.40	10900000	200484.03	2760.561
Mean Assets, age 67-69	199461.30	10100000	184773.67	2807.358
Median Assets, age 25-27	0.5	$2.3 * 10^{-4}$	0.1496	-
Median Assets, age 28-30	0.5	$1.21 * 10^{-4}$	0.2804	-
Median Assets, age 31-33	0.5	$9.47 * 10^{-5}$	0.314	-
Median Assets, age 34-36	0.5	$6.43 * 10^{-5}$	0.2965	-
Median Assets, age 37-39	0.5	$5.96 * 10^{-5}$	0.2888	-

continued

Table I.2 continued

Moment	**Data**	**Variance**	**Simulation**	Standard Error
Median Assets, age 40-42	0.5	$7.27 * 10^{-5}$	0.2912	-
Median Assets, age 43-45	0.5	$5.56 * 10^{-5}$	0.3201	-
Median Assets, age 46-48	0.5	$7.07 * 10^{-5}$	0.3333	-
Median Assets, age 49-51	0.5	$7.62 * 10^{-5}$	0.3465	-
Median Assets, age 51-53	0.5	$7.94 * 10^{-5}$	0.3935	-
Median Assets, age 54-56	0.5	$9.41 * 10^{-5}$	0.4246	-
Median Assets, age 57-59	0.5	$8.48 * 10^{-5}$	0.4297	-
Median Assets, age 60-62	0.5	$8.33 * 10^{-5}$	0.4607	-
Median Assets, age 63-65	0.5	$8.84 * 10^{-5}$	0.4615	-
Median Assets, age 67-69	0.5	$6.39 * 10^{-5}$	0.4513	-

Table I.3: Robustness of parameter estimates to model assumptions

The table below shows the model parameter estimates derived under different assumptions relative to the baseline model. The second column presents the baseline estimates (retained productivity = 0.44, no intensive margin, no control for selection into employment). The third column shows the results for a retained productivity of 38.5% (the requirement for public DI). The fourth column shows the estimation results if people can choose from six different private DI contracts, i.e. six different replacement ratios [0.25, 0.3, 0.35, 0.4, 0.45, 0.5]. The fifth column shows the results controlling for selection into employment following French (2005).

Parameter	**Baseline**	**Retained productivity 0.385**	**Adding intensive margin**	**Selection into employment**
Risk aversion γ	6.232	6.020	4.997	6.334
Consumption weight κ	0.495	0.511	0.552	0.498
Labor force participation cost θ	0.161	0.230	0.372	0.202
Disutility from bad health φ	0.154	0.160	0.131	0.159

Table I.4: Parameter sensitivity to targeted moments

The table below shows the sensitivity of each utility parameter estimate with respect to the moments used in the method of simulated moments approach. They are computed following the method detailed in Andrews et al. (2017). The values shown document how mis-measuring a given moment would alter the parameter estimate by δ = value $\times$ measurement error, such that the correct value would be parameter + δ. Besides, the sensitivity estimates are informative about the relative importance of each moment for identifying the respective parameter. Abbreviations are as follows: pDI = private DI, LFP = Labor Force Participation, FT = Full-time, PT = Part-Time.

Moment	γ	κ	pc	hc
Mean pDI	-2.54903	.0219569	.0225482	.0275698
q0	-7.65527	-.0276607	-.127524	.0309264
q25	1.9468	-.0068154	.0016337	-.0062479
q50	-9.95613	-.0438408	-.182596	.0374958
q75	12.0435	.102582	.32634	-.0281645
LFP1	.262426	.0430078	-.0074368	-.0182088
FT1	.355425	.0266418	.0002969	-.0113023
PT1	-.368713	-.0075447	-.0068999	.003232
LFP2	12.3226	.102756	.268271	-.048481
FT2	11.4131	.0950641	.247798	-.045078
PT2	-1.97828	-.0163166	-.0419482	.0080747
LFP3	-.577453	.0419904	-.0285609	-.0177709
FT3	10.7244	.0515493	.258681	-.0223642

continued

Table I.4 continued

Moment	γ	κ	pc	hc
PT3	-19.7093	-.0487671	-.490431	.0216155
LFP4	-54.781	-.0322087	-1.40778	.0131861
FT4	-41.7706	-.0074032	-1.07828	.0030262
PT4	8.57553	-.0302833	.230365	.0124075
LFP5	-1.85278	.024898	-.0562203	-.0106086
FT5	-.547897	.0307986	-.0235947	-.0128501
PT5	-1.24586	-.0389097	-.021254	.0159314
LFP6	-6.76309	.0178695	-.181354	-.0078199
FT6	.807806	.04866	.0037707	-.0209765
PT6	-9.36479	-.0811681	-.213814	.0348576
LFP7	-1.3979	.0018453	-.0365487	-.0007549
FT7	.547176	.013135	.0095081	-.0056668
PT7	-2.42637	-.0194876	-.055582	.0084525
Mean1	4.94e-08	1.12e-10	1.22e-09	-5.40e-11
Mean2	4.53e-07	1.05e-09	1.12e-08	-5.06e-10
Mean3	8.84e-07	2.08e-09	2.18e-08	-9.99e-10
Mean4	1.51e-06	3.62e-09	3.71e-08	-1.74e-09
Mean5	1.51e-06	3.70e-09	3.70e-08	-1.77e-09
Mean6	1.34e-06	3.36e-09	3.29e-08	-1.61e-09
Mean7	1.34e-06	3.46e-09	3.29e-08	-1.66e-09
Mean8	1.19e-06	3.18e-09	2.91e-08	-1.52e-09
Mean9	1.04e-06	2.92e-09	2.56e-08	-1.39e-09
Mean10	8.48e-07	2.35e-09	2.08e-08	-1.11e-09
Mean11	9.99e-07	2.74e-09	2.45e-08	-1.30e-09
Mean12	1.21e-06	3.32e-09	2.97e-08	-1.56e-09
Mean13	1.35e-06	3.73e-09	3.31e-08	-1.74e-09
Mean14	1.75e-06	4.87e-09	4.31e-08	-2.25e-09
Mean15	2.27e-06	5.88e-09	5.59e-08	-2.72e-09
Median1	.0197815	.0000427	.0004899	-.0000204
Median2	.29057	.0007325	.0071318	-.0003539
Median3	.348745	.0008747	.0085805	-.0004166
Median4	.42126	.0011197	.0103141	-.0005393
Median5	.576927	.0014818	.0141623	-.0007113
Median6	.408869	.0010198	.0100964	-.0004764
Median7	.536233	.001315	.0132402	-.0006175
Median8	.209686	.0005916	.0051118	-.0002867
Median9	.223725	.000638	.0054306	-.0003152
Median10	.209561	.0005898	.0051282	-.0002799
Median11	.354117	.0009338	.0087256	-.0004342
Median12	.339492	.0009241	.0083474	-.0004305

continued

Table I.4 continued

Moment	γ	κ	pc	hc
Median13	.412423	.0011592	.0101075	-.0005438
Median14	.299978	.0007982	.0073854	-.00037
Median15	.610452	.0016758	.0149768	-.0007854

Figure I.1: Out-of-sample fit of model

The figure below presents the out-of-sample fit of simulated and data moments not targeted in the estimation. The data moments are estimated on the sample of employed men who are at least 25 years of age. Panel (a) shows the cumulative distribution of private DI benefits in the model(blue) and the data (black). Panel (b) shows the risk group distribution of people buying private insurance in the data (green), and in the simulations (25 populations, 16,000 individuals each) (blue). Panel (c) is based on the SIAB and shows the profile of full-time and part-time work between age 25 to 60. Targeted moments are marked in red. Panel (d) shows the mean income by age for the baseline sample from the data (black) and the simulations (blue).

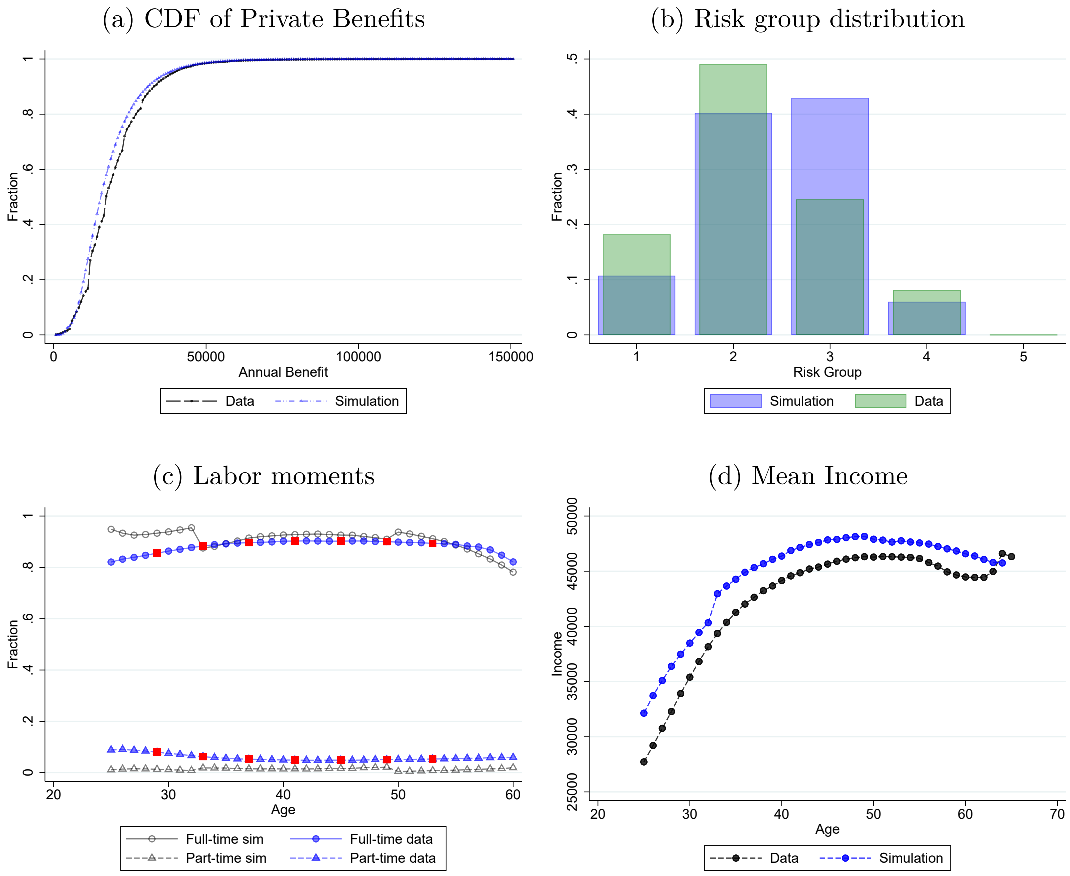


Figure I.2: Labor Supply by private DI coverage at baseline - Benefit Generosity Changes

The figure below presents the labor supply response for disabled individuals under alternative benefit generosity levels conditional on their private DI coverage at baseline. Panel (c) to (f) further condition on whether people continue to buy private DI or stop buying. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

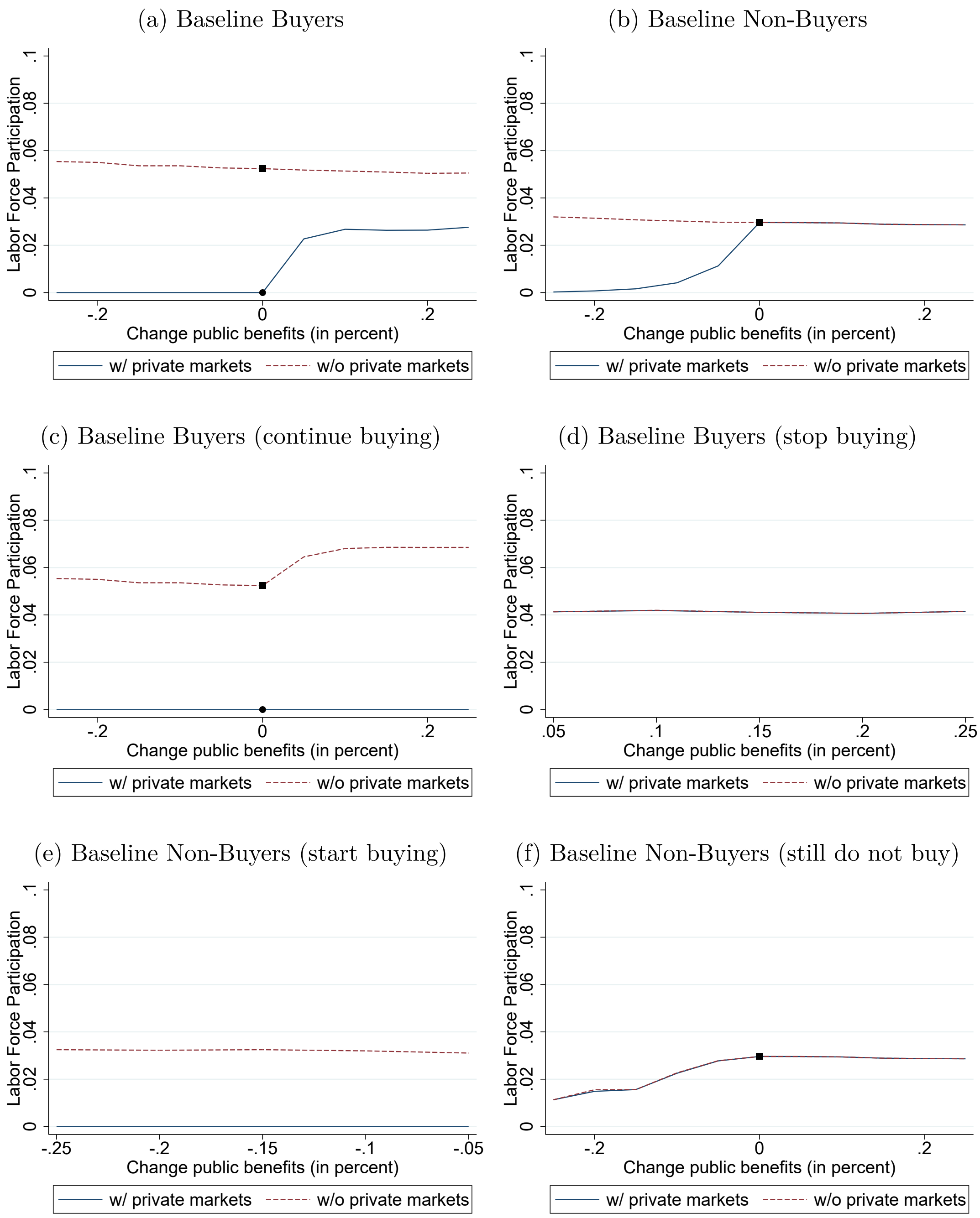


Figure I.3: Labor Supply by private DI coverage at baseline - Rejection Rate Changes

The figure below presents the labor supply response for disabled individuals under alternative rejection rates conditional on their private DI coverage at baseline. Panel (c) to (f) further condition on whether people continue to buy private DI or stop buying. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

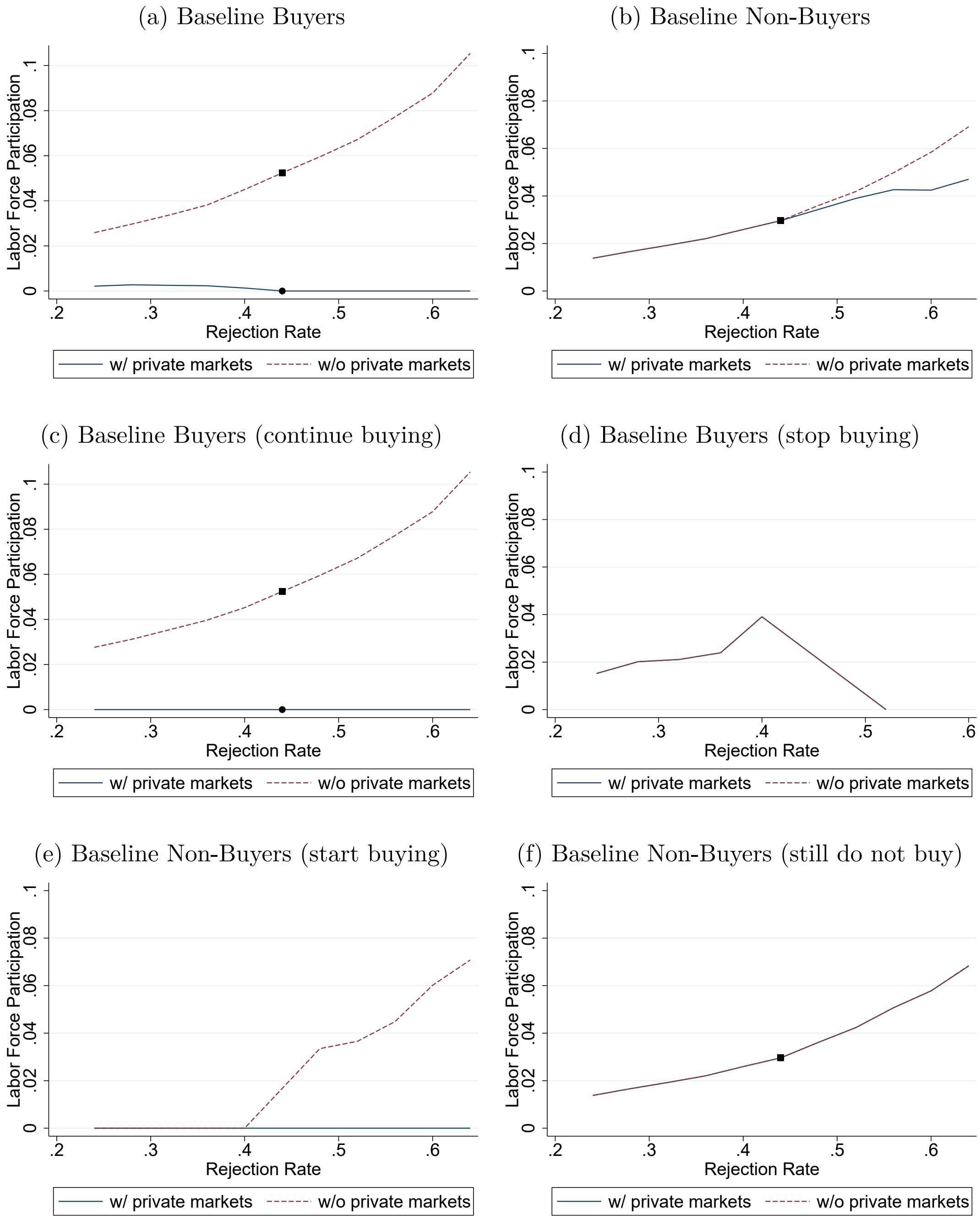
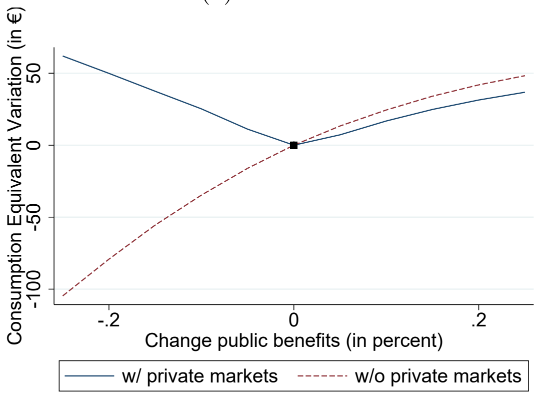


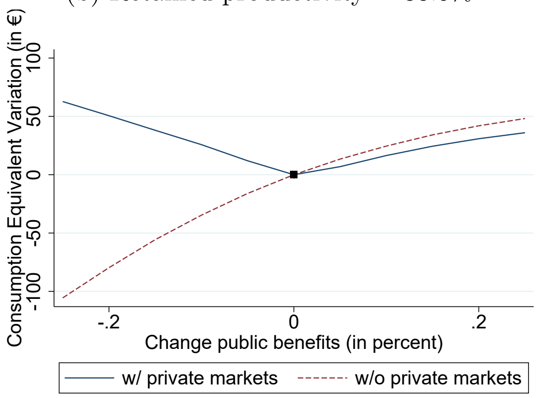
Figure I.4: Consumption - equivalent variation for changes in benefit generosity

The figure below presents the consumption-equivalent variation (CEV) for changes in the benefit generosity for the baseline specification (a), for a smaller retained productivity (b), and when an intensive margin is added to the problem (c). The CEV measures the change in expected life-time utility relative to the baseline level (percentage change = 0) in the percentage of life-time consumption an agent is willing to forgo to move to the alternative policy. Positive values imply a welfare improvement. All values are expressed in 2013 Euros. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

(a) Baseline



(b) Retained productivity = 38.5%



(c) Added intensive margin

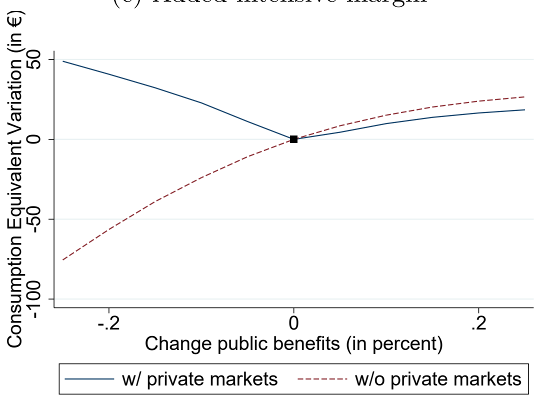
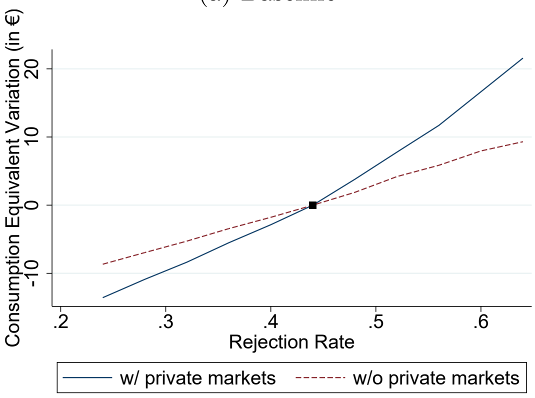


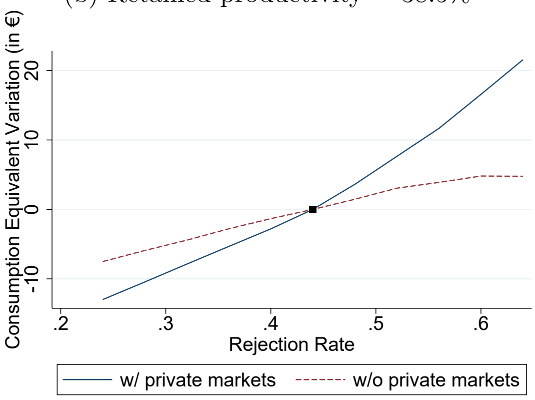
Figure I.5: Consumption - equivalent variation for changes in rejection rates

The figure below presents the consumption-equivalent variation (CEV) for changes in the rejection rate for the baseline specification (a), for a smaller retained productivity (b), and when an intensive margin is added to the problem (c). The CEV measures the change in expected life-time utility relative to the baseline level (percentage change = 0) in the percentage of life-time consumption an agent is willing to forgo to move to the alternative policy. Positive values imply a welfare improvement. All values are expressed in 2013 Euros. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

(a) Baseline



(b) Retained productivity = 38.5%



(c) Added intensive margin

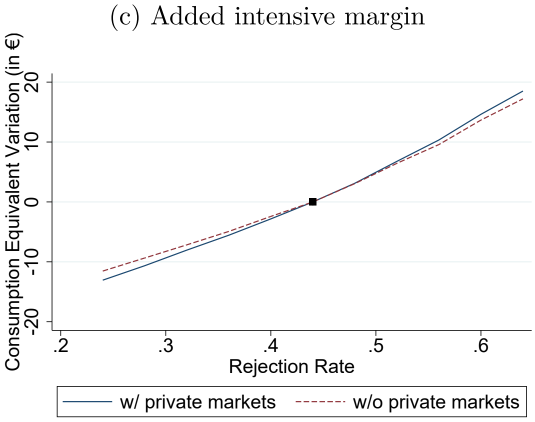
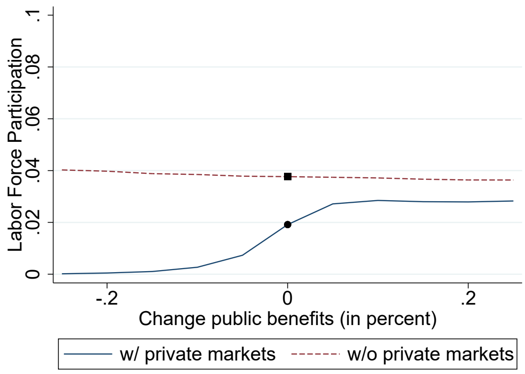


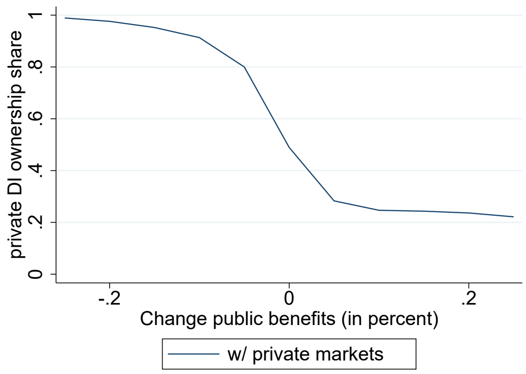
Figure I.6: Labor force participation and mean private DI shares for changes in benefit generosity

The figure below presents the mean labor force participation of disabled individuals and the mean private DI ownership shares for alternative public DI benefit generosity. Panel (a) and (b) show the baseline results from the main text for the mean LFP and mean private DI shares respectively. Panel (c) and (d)/ Panel (e) and (f) present the results for the mean LFP and mean private DI shares under lower retained productivity/ when adding an intensive private insurance margin . The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

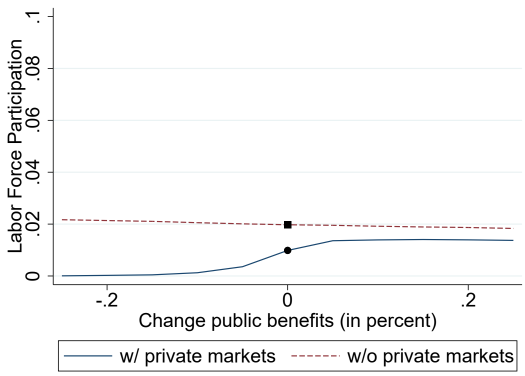
(a) Baseline, LFP



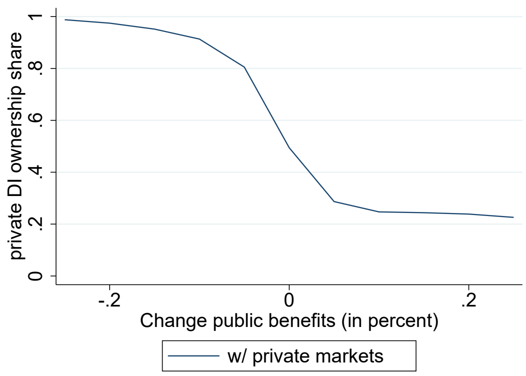
(b) Baseline, private DI shares



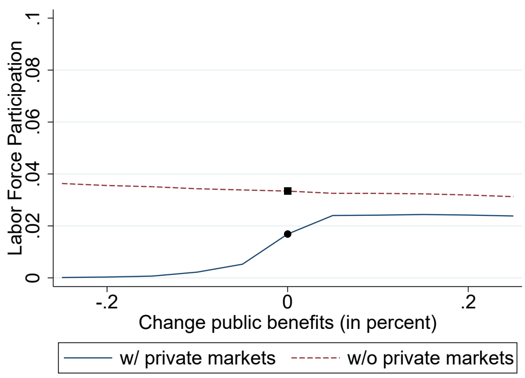
(c) Retained productivity = 38.5%, LFP



(d) Retained productivity = 38.5%, private DI shares



(e) Added intensive margin, LFP



(f) Added intensive margin, private DI shares

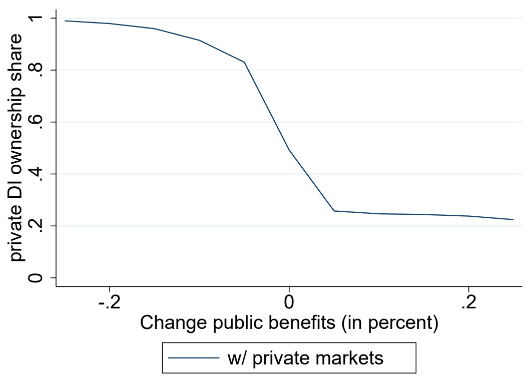


Figure I.7: Labor force participation and mean private DI shares for changes in the rejection rate

The figure below presents the mean labor force participation of disabled individuals and the mean private DI ownership shares for alternative public DI rejection rates. Panel (a) and (b) show the baseline results from the main text for the mean LFP and mean private DI shares respectively. Panel (c) and (d)/ Panel (e) and (f) present the results for the mean LFP and mean private DI shares under lower retained productivity/ when adding an intensive private insurance margin . The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

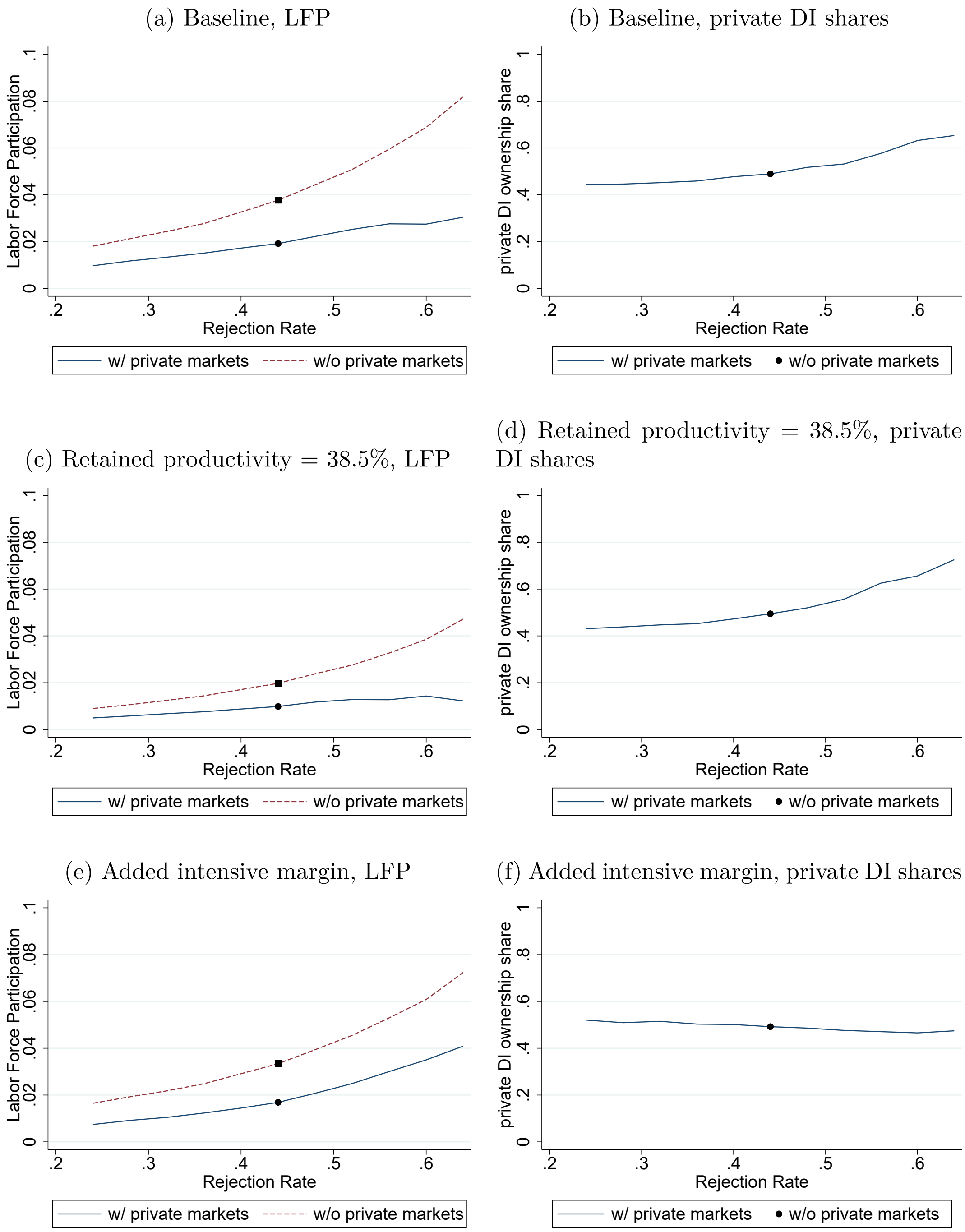
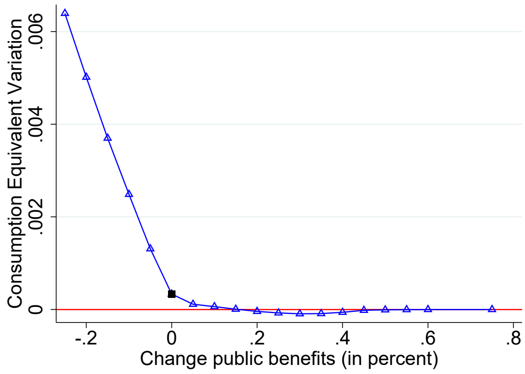


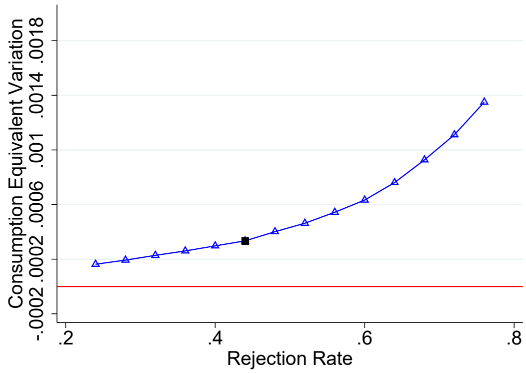
Figure I.8: Welfare effects of private DI markets

The figure below presents the consumption-equivalent variation (CEV) for allowing for private DI markets under alternative policy schedules. The first (second) column depicts the CEV for changes in public benefit generosity (rejection rates). The first row shows the results derived under the baseline model, while the second and third row show the results for lower retained productivity and with an intensive margin respectively. The CEV is expressed as the percent change of per-period consumption an agent is willing to forgo to have a private market by comparing the expected life-time utility from having a private market to the one without a private market under the same public DI schedule. Positive values imply that private DI markets are welfare enhancing under the considered policy schedule visually presented by the blue line being above the red '0'-line. The results are computed for a population of $N = 16,000$ *individuals and under revenue-neutrality.*

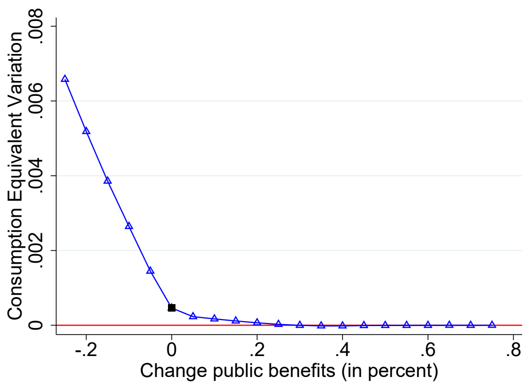
(a) Changes in benefit generosity, baseline



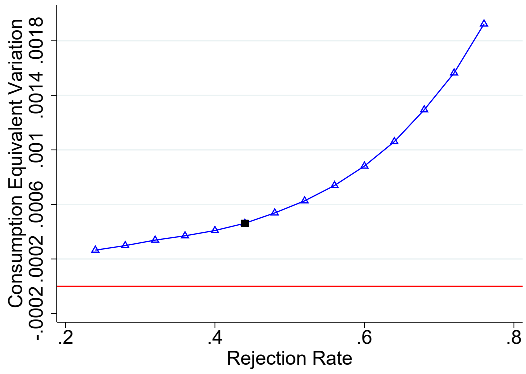
(b) Changes in rejection rates, baseline



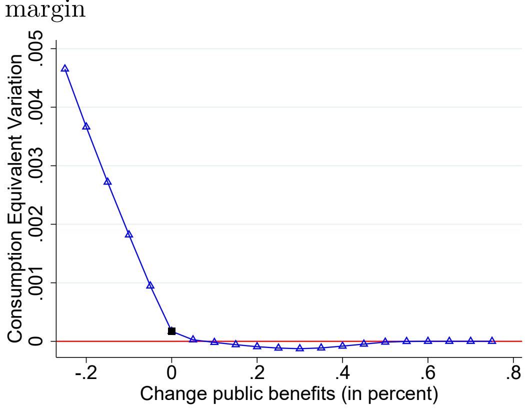
(c) Changes in benefit generosity, retained productivity = 38.5%



(d) Changes in rejection rates, retained productivity = 38.5%



(e) Changes in benefit generosity, intensive margin



(f) Changes in rejection rates, intensive margin