# Essays in Public Economics and Labor Economics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Sebastian Seitz

im Frühjahrs-/Sommersemester 2022

| | |
|---|---|
| Abteilungssprecher | Prof. Volker Nocke, PhD |
| Referent | Prof. Dr. Eckhard Janeba |
| Koreferent | Prof. Dr. Sebastian Siegloch |

| | |
|---|---|
| Tag der Verteidigung | 9. Mai 2022 |

# Acknowledgements

I want to thank my advisors Eckhard Janeba, Sebastian Siegloch and Cormac O'Dea for their support and guidance throughout my time as a graduate student. I am especially grateful for the many insightful discussions about each chapter in this thesis and my other research projects. My work greatly benefited from their comments, suggestions and experience.

This thesis is the result of many stimulating collaborations with my co-authors. In particular, I want to thank Andreas Gulyas, Sourav Sinha, Arthur Seibold, and Sebastian Siegloch for many inspiring discussions. In addition, I have benefited greatly from my close and long-standing interaction with the "Social Policy and Redistribution" group at the ZEW under the leadership of Sebastian Siegloch and my affiliation with the CRC224 Transregio hosted by the University of Bonn and the University of Mannheim. I thank the University of Mannheim's Graduate School of Economics and Social Sciences for financial support.

Moreover, I want to thank the amazing people I had the pleasure meeting during my graduate studies. Lukas, and Daniel, thank you for the great time I had at the chair and at diverse occasions outside the office. Max, Karl, and Nils, thank you for your close friendship and the many great times we spent together. Alexander, Tomasz, David, thank you for the football matches and the after-match analysis. Lukas, Dominik, and Florian thank you for your friendship and support during my Bachelor's, Master's, and graduate studies. Beatriz, thank you for your close friendship and support. Michaela, thank you for all the advice regarding research and life.

Finally, I want to thank my family for their lifelong support, especially my parents Angela and Wolfgang, my siblings Bernardin and Ann-Christine, and my grandmother Ursula on whom I could always count. Akari, thank you for always supporting me and the great times we spent together on our (little) adventures outside the office. I am

grateful for always having by my side and being able to talk to you about anything. You made all of this possible.

# Contents

# List of Figures

# List of Tables

# Preface

Inequality has been rising in many countries for instance in income (Atkinson, Piketty, and Saez, 2011) or between genders (Blau and Kahn, 2017). Simultaneously, social security systems providing insurance against economic hardships struggle to maintain their sustainability following a rise in the number of recipients and cost (Autor and Duggan, 2006; OECD, 2010). Hence, policymakers face pressure to address these issues, which relates to a classical issue in public economics: the design and evaluation of policies aimed at alleviating inequality and providing social insurance. Designing effective policies requires a firm understanding of the underlying mechanisms, e.g., the causes of inequality, as well as individuals' responses to policy changes.

In this thesis, I show that designing effective policies is not always an easy task. Since there are many causes of inequality as well as different individual responses to policies, a careful evaluation is crucial to progress our understanding of the relationship between policies, individual behavior, and inequality. I investigate these relationships in the three chapters of this thesis, which can be read independently. The first chapter studies the effect of pay transparency legislation, a recent popular policy to address the remaining gender wage gap, on female wages and the gender wage gap. The second chapter investigates the interaction between private and public disability insurance and what this interaction implies for the design of welfare-enhancing public policies. The third chapter investigates whether certain aspects of the public disability insurance system can be efficiently privatized. I discuss the three chapters in greater detail below.

The gender wage gap has been converging in the 1990s and early 2000s but has remained roughly stable for most of the 2010s (Blau and Kahn, 2017). A popular explanation for this stagnation is that women lack precise information about their employers' remuneration schedules, preventing them from asking for equal pay. To address this issue, pay transparency legislation has recently received widespread attention, and variants of it have been introduced in Finland, Sweden, Norway, Denmark, Austria, the

UK, Germany, Iceland, and the U.S. In Chapter 1, which is joint work with Andreas Gulyas and Sourav Sinha, I investigate the effect of the Austrian pay transparency legislation on wages and the gender wage gap. Using Austrian social security data and the size threshold introduced by the law for producing internal pay reports, I show that the policy had no discernible effects on male and female wages, thus leaving the gender wage gap unchanged. In addition, I do not find any evidence for wage compression within establishments. I discuss several possible explanations for why the reform was ineffective in narrowing the gender wage gap.

Overlapping private and public insurance interact in important ways, as has been shown for instance in the context of health insurance (Cabral and Mahoney, 2018; Chetty and Saez, 2010). However, despite the existence and size of private disability insurance markets, there is little empirical evidence on their interaction with public disability insurance and what this implies for the design of public disability insurance programs. I address this gap in Chapter 2, where I evaluate the effect private disability insurance take-up has on the design of welfare-enhancing public policies. To answer this question, I build a rich dynamic lifecycle model with private insurance choices. I estimate the model using novel and comprehensive contract data from a major German insurance company together with a representative household survey and social security records. I derive two results: first, I show that welfare-enhancing public disability insurance programs have to be less generous in the presence of private insurance markets as private disability insurance take-up leads to additional withdrawal from the workforce at disability onset. Second, I show that having a dual insurance system, i.e. private plus public disability insurance, might be welfare-reducing if the public system is too generous.

While the previous chapter focuses on the change in labor supply incentives at disability onset due to private disability insurance take-up, Chapter 3, which is joint work with Arthur Seibold and Sebastian Siegloch, investigates whether certain aspects of the public disability insurance system can be efficiently privatized. In particular, I exploit the abolition of own-occupation public disability insurance for certain birth cohorts in Germany in 2001 to estimate the crowding in response to private own-occupation disability insurance and ultimately welfare. I find that despite significant crowding-in effects, overall private disability insurance take-up remains modest. Studying selection, I find that private DI tends to be concentrated among high-income, high-education, and low-risk individuals but I find no evidence of adverse selection on unpriced risk. Applying a revealed preference approach, I estimate individual valuations for (private) own-occupation disability insurance. I find that own-occupation disability insurance can be efficiently privatized for rational agents. However, distributional concerns, as well as individual risk misperceptions, can justify a full public mandate.

In conclusion, this thesis underscores the many important interactions between causes of inequality and policies aimed at alleviating it. Chapter 1 shows that without a good understanding of the drivers of inequality, policies fail to achieve their goals. Likewise, Chapter 2 shows how abstracting from important interactions between private and public disability insurance can lead to the implementation of expensive and suboptimal (or even welfare-reducing) policies. Since public disability insurance programs in many countries struggle to maintain their sustainability, not accounting for private disability insurance further threatens their sustainability, which has important implications for the success of reforms. Finally, in Chapter 3 I study a specific reform aimed at reducing public disability insurance program cost: the privatization of own-occupation disability. I show that in this specific context, privatization is efficient. However, distributional concerns and risk misperception can justify the implementation of a public mandate as especially low-income, high-risk individuals purchase too little private insurance.

The three chapters of this thesis show that there are many relevant and non-trivial interactions between causes of inequality and policies which might not be obvious at first. Since incomplete accounting for these relationships can lead to erroneous policy conclusions, it is of utmost importance to progress our understanding of these relationships and consider them when designing policy.

# Chapter 1

# Does Pay Transparency Affect the Gender Wage Gap? Evidence from Austria

Joint with Andreas Gulyas and Sourav Sinha.

## 1.1 Introduction

Gender disparity in earnings is a persistent feature of labor markets around the world. Women earn about 23% less than men in the US, 20% in Austria, and 15% on average across the European Union.[1] There is an ongoing debate among academics, policy makers, as well as the general public about the reasons behind the gender wage disparity and about the best policy instruments to close the gap.[2]

One policy instrument that has recently received widespread attention is some form of pay transparency legislation, whereby firms are required to provide information on pay disparities between genders. Proponents of transparency argue that the lack of information on pay sustains the gender gap and transparency helps women to challenge discriminatory pay schedules.[3] However, critics worry about administrative costs and that men might use the information revealed by transparency more actively than

[1] Eurostat, 2018

[2] see Blau and Kahn (2017) for a review.

[3] For example, the European Commission writes in the Factsheet on Pay Transparency (2019): "[...] the effective enforcement of the right to equal pay [...] for women and men remains a major challenge, partly because of a lack of information on pay."

women, further widening the gender pay gap instead. Nevertheless, these policies have garnered widespread attention among policy makers and variants of it have been introduced in Finland, Sweden, Norway, Denmark, Austria, the UK, Germany, Iceland, and the United States.[4] Despite its recent introduction in many countries, the causal evidence of transparency laws on wages and the gender wage gap is scarce. This paper studies the Austrian transparency law to fill this gap.

The Austrian transparency law was rolled out in phases, starting off with the largest firms in 2011. Over the next three years smaller firms were brought under coverage, and by 2014 all firms with more than 150 employees were required to publish and update income reports every second year. These reports must contain annual gross income, itemized by gender and occupation groups as defined in the respective collective bargaining agreements. However, wage reports are company secret and not public information. Using the universe of Austrian social security records, we exploit the size-based cutoff rule and employ an event-study design to estimate the causal effects of pay transparency on wages and the gender wage gap.

In our baseline specification we focus on a narrow window around the lowest cutoff to make the control group as comparable to treated establishments as possible. We do not find evidence that transparency has any discernible effect on the gender wage gap. The point estimate is close to zero, precisely estimated, and we can rule out that the policy narrowed the gender wage gap by more than 0.4 percentage points. When we study the effects on wages of men and women separately, we again do not find any statistically or economically significant effects. Therefore, transparency seems to have failed in its twin objectives of reducing the gender pay gap and boosting female earnings. We show that this conclusion holds under a number of alternative specifications using different control variables and alternative sample restrictions on top-coding, firm size windows, and compliance with treatment assignment. We further consider the full roll-out of the policy across all firm size groups and show that transparency did not affect the gender wage gap in large firms either.

While pay transparency does not affect average wages, it could potentially lead to

https://ec.europa.eu/info/sites/info/files/factsheet-pay_transparency-2019.pdf

[4] In the United States, during President Obama's tenure, the Equal Employment Opportunity Commission (EEOC) proposed changes which would have required firms with more than 100 employees to provide annual reports on gender pay gap, to the Department of Labor. This move was subsequently rolled back by President Trump. See: Obama EEOC Action on Pay Data collection).

wage compression within establishments. Yet again, we find no evidence for this. The variance of log-wages within treated establishments evolves in tandem with the control group, with no discernible effect of the policy. Furthermore, we do not find heterogeneous effects for workers earning below or above the establishment-level gender-specific median wage.

Why does pay transparency not affect the gender pay gap and wage setting in general? Surveys of worker representatives and work councils reveal that compliance was universal and a majority of respondents found the reports informative and useful.[5] Therefore, imperfect implementation seems an unlikely explanation.

Our data does not allow us to definitively pin down the reasons behind the lack of policy effects. Nevertheless, we highlight several possible channels why the policy might not be effective in narrowing the gender wage gap. First, wage reports are legislated to be company secret, and therefore can only affect within-firm wage differences. Without wage reports being public, they cannot affect the differential sorting patterns of men and women, which we show to be a major contributor to the gender wage gap in Austria.[6]

In addition, it could be that within firms, the pay gap between men and women in the same occupation is too small for the employee to initiate a renegotiation.[7] Alternatively, workers might lack the bargaining power to renegotiate wages, since firms are not required to act upon unequal firm pay policies. If workers have low bargaining power but feel unfairly compensated, we would expect them to have lower job satisfaction and higher quit rates (Card et al., 2012, Rege and Solli, 2015, Dube, Giuliano, and Leonard, 2019). In Austria we find that pay transparency leads to a reduction in separation rates in treated firms. This is perhaps indicative that transparency alleviated previously held concerns about unfair pay schedules among workers and increased their job satisfaction.

Our work contributes to a small literature studying the effects of transparency in very specific labor markets, which typically documents unintended consequences of such policies. Schmidt (2012) and Mas (2016) show that mandated disclosure of CEO

---

[5] Arbeiterkammer, 2014

[6] International evidence also points towards the importance of sorting for the gender wage gap (Card, Cardoso, and Kline, 2016; Morchio and Moser, 2019).

[7] Our data lacks detailed occupation information, and thus we cannot compute the within-establishment, within-occupation gender wage gap.

compensation leads to 'ratcheting' effects, whereby CEOs who earned below the average, received a pay raise. Using a field experiment in an online labor market, Cullen and Pakzad-Hurson (2019) document that transparency led to overall wage reductions. Baker et al. (2019) show that a public sector salary disclosure law for university faculty in Canada reduced the gender wage gap, though partly by lowering male wages.

Our paper is one of the first to document the effects of a broad introduction of pay transparency. The most closely related studies are Bennedsen et al. (Forthcoming), Duchini, Simion, and Turrell (2020) and Blundell (2020), which analyze similar policies in Denmark and the UK. These studies show that similar to Austria, pay transparency in both countries failed to achieve its goal of increasing female wages. However, in contrast to our study, they find that transparency moderately depressed male earnings, and thus slightly narrowed the gender wage gap. We argue that transparency policies can potentially have a larger impact on the gender wage gap if the wage reports are public information. This can guide women in their job search towards more equitable and higher paying firms. This could be one of the reasons why the UK reform, which makes gender wage gaps public information, was more successful compared to Austria in closing the gender wage gap.[8]

More broadly, our work is related to the literature which studies the effects of information about relative earnings on behavioral and labor market outcomes: municipal salary disclosure on pay compression among city managers (Mas, 2017), publicly available tax records on happiness and life satisfaction in Norway (Perez-Truglia, 2020), perceived peer and manager salaries on effort and output (Cullen and Perez-Truglia, Forthcoming), pay inequality on attendance and output in India (Breza, Kaur, and Shamdasani, 2017), relative earnings on worker effort (Cohn et al., 2014) and on happiness and life-satisfaction (Brown et al., 2008; Clark, Kristensen, and Westergard-Nielsen, 2009; Clark and Oswald, 1996; Godechot and Senik, 2015; Luttmer, 2005).

The rest of this paper is structured as follows. In Section 1.2 we describe the pay transparency law in detail. Section 1.3 lays out a conceptual framework for transparency policies, Section 1.4 explains our data, sample selection, and our empirical strategy. We

[8] The Austrian and Danish transparency reforms share many institutional features; if anything the Austrian policy is more strict. Both reforms do not mandate to make the pay reports public, and nevertheless the Danish policy has led to a reduction in the gender wage gap. Further research is needed to better understand whether there are cross-country differences in how workers and firms engage with pay transparency policies, and whether transparency interacts with other labor market institutions.

present our results in Section 1.5, discuss the potential reasons behind the ineffectivness of the reform to affect the wage setting in Section 1.6, and the last section concludes.

## 1.2 Institutional Setting and the Pay Transparency Policy, 2011

In international comparisons, Austria has a relatively high gender pay gap. The unadjusted gender pay gap was 20 percent in 2017, being fifth highest in the European Union.[9] A commonly raised point in the public debate in Austria is that pay secrecy is a major obstacle to achieving equal pay because women might not know the degree of pay discrimination or have less precise information about pay schedules compared to their male colleagues.

In light of these debates, the Austrian government introduced a Pay Transparency law in 2011, serving two explicit goals: first, boosting female wages and second, thereby reducing the gender wage gap. To achieve these goals, firms have to produce and update internal gender pay gap reports every second year, disaggregated by occupation groups. These reports must include the number of employees within a gender-occupation cell and their average or median annual earnings, expressed in full-time equivalents. All components of pay must be included, but there is no obligation to separate them. It is important to note that employers have no discretion about the occupational groups, but they have to follow the pre-defined classifications in collective bargaining agreements.[10] Managerial positions are exempt from reporting requirements.

In principle, workers are almost universally covered by collective bargaining agreements. These define minimum wages at the industry level for different occupations, but firms and workers are free to bilaterally agree on wages above this floor. We are not aware of any precise evidence on the fraction of workers paid above required levels, but evidence on the wage structure suggests that they are not very binding. Differences in firm pay policies explain almost the same fraction of wage inequality in Austria as in the United States, suggesting that firms have a lot of flexibility in setting their pay

[9] Source: Eurostat (online data code sdg_05_20)

[10] The collective bargaining agreements are quite detailed in their occupational categories. For example, the wholesale and retail sector, which is the collective bargaining agreement with the highest number of employees in Austria, has 8 predefined occupational categories, 9 firm tenure groups, in addition to 2 regional categories.

policies, and are not much constrained by the collective bargaining agreements.[11]

In comparison to pay transparency legislation in other countries, the Austrian version is stricter and more detailed in various characteristics. First, to protect the anonymity of individuals, if less than 3 employees fall within a certain gender-occupation group, they are counted with the next larger occupational group. This is more comprehensive compared to Denmark and Germany, where firms have to aggregate cells with 10 and 7 employees respectively. The UK legislation is on an even more aggregated level, as it does not require a break down of income statistics by occupation. Second, reports must be made available to all employees via work councils where they can be accessed by any employee. In the absence of a works council, the report must be put on public display in a 'common (break) room'. Failure to compile these reports can lead to monetary fines and being directed by the courts to produce them. The wage reports are legislated to be company secret. Workers can discuss the contents of the report with their colleagues, union representatives, and legal advocates. However, communication of the contents to the outside are prohibited. Firms have no obligation to make these reports public, yet many public sector firms make theirs available online (see Appendix Table $1.A1$).

The implementation of the legislation was staggered over four years. Firms with more than 1000 workers came under the legislation in March 2011. Then in January of each subsequent year, firms with more than 500, 250, and finally 150 employees became subject to the reporting requirements in 2012, 2013, and 2014 respectively. Firms that grow and exceed the 150 employee threshold after 2014 have to produce a report in the first year they exceed the threshold. In 2011, about 30% of the Austrian workforce became subject to the legislation, which grew to 50% of workers by 2014 (see Appendix Figure $1.A1$). There are no other policy changes or legal requirements that specifically apply to these cutoffs and especially the 150-employee cutoff used in our baseline study.

Exploratory non-representative surveys conducted by the Austrian Chamber of Labor (*Arbeiterkammer*), the Austrian Trade Union Federation (OeGB), and the Austrian Federal Ministry for Education and Women's Affairs (AFMEW) in 2014 and 2015 study the level of compliance among firms and the dissemination of reports to employees. Evidence from these surveys (Arbeiterkammer (2014); Deloitte (2015)) show near universal compliance with the policy. Reports were shared with works councils promptly and information was distributed most frequently via intranet, announcements, employee

[11] See Gulyas and Pytka, 2020 for evidence on Austria, and Lamadon, Mogstad, and Setzler, 2020 for the US.

newsletters, etc. In more than half of the cases, council representatives reported close cooperation with their employers in preparing the reports and 80% reported that their employers were open to adopting measures addressing the gap.

We do not have precise information about what fraction of workers actively use the wage reports, but there is no reason to believe that pay reports are not widely known. The media regularly reports about the gender wage gap. In particular, this topic receives widespread attention on the so-called Equal Pay Days.[12] Around these dates, most newspapers and news stations discuss the existing gender pay gap in Austria, its roots and pathways to closing it. Pay reports are featured prominently in this debate, especially in the first four years after the reform.[13] We take this regular news coverage as evidence that the general public (and especially workers) are aware of the issue at hand and pay reports as way of addressing it. In addition, as mentioned above, the fact that many work councils are directly involved in the preparation of the wage reports suggests that this information should also percolate to workers.

## 1.3 Conceptual Framework

How should we expect pay transparency to affect the wage setting process? It has long been recognized that observationally similar workers are paid differently in the labor market. A recent literature emphasizes the role of firm pay in understanding wage differences across workers, which has been shown to explain around a third of the overall wage variation.[14] In models with frictional labor markets, more productive firms are willing to pay higher wages, as their opportunity cost of a vacancy is higher (Cahuc, Postel-Vinay, and Robin, 2006; Postel-Vinay and Robin, 2002). Since search is a time and resource intensive process in such frameworks, both parties would be willing to accept a range of wages. These range from reservation wages holding the worker to their outside option, up to wages where the worker appropriates the firm's maximum

[12]There are two Equal Pay Days in Austria: The first is in spring and marks the day until which women "work for free" in a given year based on the gender pay gap. The second is in fall and marks the date by which men would have earned the same annual income as women in full year (so to speak, from that day on, women work for free relative to men for the rest of the year).

[13]See for example https://www.tt.com/artikel/3502362/online-gehaltsrechner-soll-fuer-transparenz-sorgen (accessed Feb. 16th, 2021) or https://www.kleinezeitung.at/politik/innenpolitik/5298933/Equal-Pay-Day_Frauen-verdienen-in-ihrem-Leben-435000-Euro-weniger (accessed Feb. 15th, 2021)

[14]see e.g. Abowd, Kramarz, and Margolis, 1999, Card, Cardoso, and Kline, 2016, Song et al., 2018, among many others

willingness to pay. Typically, a particular bargaining protocol is assumed, where wages are pinned down by the bargaining power of workers. In such settings, wage differences within firms could arise due to differences between workers' bargaining power and outside options.

A less researched aspect in search frameworks is that asymmetric information between employers and workers and informational differences across workers about firms' willingness to pay can lead to differential wage outcomes. Therefore, pay transparency can alleviate these informational frictions and in turn affect wages and other labor market outcomes. If workers have different information about firms' output and willingness to pay, they would achieve different bargaining outcomes.[15] In particular, women might have less information than their male colleagues, possibly because of smaller workplace networks.[16] These information gaps could generate pay disparities both within and across gender lines. Transparency by design reveals more information about firms' willingness to pay and unequal pay schedules. If wage reports are company secret, this information empowers only current workers to challenge gender pay gaps and pay disparity in general. Instead, if wage reports are public, workers and especially women can direct their search towards more equitable and higher paying firms.

Beyond wages, transparency can affect job turnover through changes in job satisfaction. If workers perceive that they are underpaid and have little bargaining power to demand higher wages, we would expect them to have lower job satisfaction and higher quit rates (Card et al., 2012; Dube, Giuliano, and Leonard, 2019; Rege and Solli, 2015). In contrast, job satisfaction and retention might increase if within-firm transparency alleviates previously held concerns about unfair compensation.

On the firm side, transparency can induce firms to reduce wage dispersion out of equity concerns when large differences within the company become salient and information in wage reports begin to serve as reference points in negotiations. In addition, if wage reports are public information, wage and gender pay gap differences across firms would invite public scrutiny and criticisms that might pressure firms to correct their wage policies.

[15] See for example the framework in Cullen and Pakzad-Hurson, 2019

[16] Previous research shows that women are less informed about their market value than men (Babcock and Laschever, 2003), more private about their pay than men (Goldfarb and Tucker, 2011), and communicate about pay with their peers less often than men (Cullen and Pakzad-Hurson, 2019). According to a (Glassdoor, 2016) survey, globally 59% of men versus 51% of women believe they have a good understanding of how pay is determined at their company.

To summarize, internal wage reports can in theory be an effective policy tool to address wage differences within companies. But the above discussion makes it clear that transparency will only affect the wage setting under certain conditions. The Austrian transparency legislation only requires firms to compile wage reports, but does not mandate them to act upon pay gaps. Therefore, it becomes the workers' responsibility to challenge pay disparities. First, assuming that wage re-negotiations entail some costs on part of the worker, the revealed wage differences must be perceived as unjustified and large enough to warrant acting upon them. Second, workers must have the bargaining power to use this new information and demand higher wages. And finally, transparency as enacted in Austria only addresses information frictions in the wage setting due to differences in knowledge about firms' willingness to pay. If workers already had good information about how much their coworkers earn on average and therefore their employers' willingness to pay, it is likely that within-firm transparency would have no effects on the wage setting process.

In conclusion, it is a priori not clear whether internal wage reports will affect the gender wage gap and wage setting in general. Therefore, the empirical evaluation of the Austrian pay transparency policy not only estimates the efficacy of transparency legislation, but also the importance of informational differences in wage setting. Before we delve into these results, we describe our data and empirical strategy in the next section.

## 1.4 Data and Empirical Strategy

We use administrative employment records from the Austrian social security administration from 1997-2018 in our analysis. This data comprises of day-to-day information on the universe of employment spells subject to social security (Zweimüller et al., 2009). The data contains information on the yearly income at the person-establishment level, broken down by regular wages and bonus payments. It further contains basic socio-demographic information of workers such as age, gender, and citizenship. Except a flag for blue collar jobs, the dataset does not contain information on workers' occupation. Each establishment has a unique identifier, and we merge with this data information on its geographic location, 4-digit NACE industry classification, as well as (from 2007 onward) the firm size of the establishment's parent company. The information about overall firm size is crucial, since the law applies to firm size, and not establishment size.

Table 1.1: Sample Restriction and Composition

*The table below shows the composition of the sample under different sample restriction criteria. Column (3) is our main sample used in the baseline specification. Columns (4) and (5) show the sample means respectively for the treated and control group of establishments in pre-treatment years (2007-2013). The adjusted gender wage gap was computed by controlling for Austrian citizenship, a quartic age polynomial, work experience, establishment and year fixed effects.*

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Fraction Female | 0.469 | 0.417 | 0.435 | 0.442 | 0.426 |
| Fraction Austrian | 0.758 | 0.744 | 0.735 | 0.761 | 0.750 |
| Fraction Manufacturing | 0.174 | 0.244 | 0.242 | 0.279 | 0.235 |
| Fraction Blue-Collar | 0.427 | 0.474 | 0.507 | 0.512 | 0.514 |
| Age (yrs) | 38.9 | 38.9 | 38.4 | 38.2 | 38.0 |
| Establishment-Tenure (yrs) | 6.3 | 6.4 | 6.1 | 6.1 | 5.8 |
| ln(Daily Wage) | 4.389 | 4.459 | 4.411 | 4.407 | 4.401 |
| Gender Wage Gap | 0.363 | 0.369 | 0.339 | 0.358 | 0.329 |
| Adj. Gender Gap | 0.237 | 0.239 | 0.222 | 0.222 | 0.222 |
| Separation Rate | 0.128 | 0.117 | 0.121 | 0.122 | 0.128 |
| Fraction Topcoded | 0.057 | 0.067 | 0 | 0 | 0 |
| N | 41,429,703 | 5,269,153 | 4,914,038 | 1,039,328 | 1,651,146 |
| # Workers | 5,784,925 | 1,242,885 | 1,204,251 | 328,134 | 529,099 |
| # Establishments | 539,254 | 14,495 | 14,303 | 4,949 | 9,265 |
| Dominant Employers | ✓ | ✓ | ✓ | ✓ | ✓ |
| 75≤ Firm Size ≤225 | | ✓ | ✓ | ✓ | ✓ |
| Top-coded Removed | | | ✓ | ✓ | ✓ |
| Treated Establishments (≥150) | | | | ✓ | |
| Control Establishments (<150) | | | | | ✓ |
| Year <2014 | | | | ✓ | ✓ |

We select all employment spells from 2007-2018. For each worker-year pair, we select the dominant employer based on yearly income. This yields over 41 million person-year observations. Table 2.A7 presents descriptive statistics about the overall employment population as well as our estimation sample. The adjusted gender wage gap is above 20 percent in our dataset, although the true gender pay gap conditional on observables is likely much smaller. The social security dataset contains only few worker characteristics, but studies using survey data with a larger set of controls find the adjusted gender wage gap to be 7.2 percent in 2013 (Böheim, Fink, and Zulehner, 2020). For each worker-year observation, we compute the daily wage as yearly earnings from the dominant employer divided by the number of days employed at that establishment deflated to 2017 prices. One caveat of the administrative data is that it does not contain information on hours worked. Thus, we are only able to analyze the response of total daily wages, and not the hourly wage response. To make our control group as similar as possible to treated establishments, we focus our analysis on establishments that became subject to the law in 2014, i.e. establishments whose firm size were within a window around the 150 size threshold. Large firms are likely very different from the small firms in the control group, both along observed and unobserved dimensions of worker and firm characteristics, and so we drop them for our baseline estimation. In our main sample, we select all establishments with firm size between 75-225, but we consider robustness checks with other firm size windows as well as estimating the effect of the reform including establishments from all larger firms.

Since the social security administration only records income up to the maximum contribution limit, wage information is top-coded, which applies to 6 percent of our sample.[17] As we cannot observe any change in wages for this group, we drop top-coded spells in our baseline sample. Table 2.A7 shows that this selection does not change the worker composition much. In additional checks we explore the robustness of our results to either including top coded individuals, or excluding workers that were ever top coded during our study period.

These sample restrictions leave us with close to 4.9 million worker-year observations, generated by 1,204,251 workers employed across 14,303 distinct establishments. The worker and establishment characteristics of our baseline sample are overall quite similar to the whole population. The only significant difference is that manufacturing jobs are somewhat overrepresented in the baseline sample. They comprise 24 percent of all jobs, whereas the manufacturing share in the overall population is only 17 percent.

In our baseline sample, we assign treatment status based on the firm size in 2013, just before firms with 150-250 employees became subject to the policy in 2014. The last two columns in Table 2.A7 show that the treatment and control establishments had similar worker and establishment characteristics in the years before the policy was rolled out.

To estimate the causal effect of pay transparency on the gender wage gap as well as

[17] In 2016, the maximum monthly earnings used to calculate contributions was €4,860. There were no substantial changes in the maximum contribution threshold in Austria during our study period. It was essentially only valorized each year by the inflation rate.

on male and female wages we apply the following event-study model:

$$
\begin{aligned}
y_{ij(i,t)t} = &\sum_{k=2007}^{2018} \beta_1^k \mathbf{1}[t=k] * Male_i * Treat_{j(i,2013)} + \sum_{k=2007}^{2018} \beta_2^k \mathbf{1}[t=k] * Treat_{j(i,2013)} \\
&+ \beta_3 Male_i * Treat_{j(i,2013)} + \sum_{k=2007}^{2018} \gamma_k \mathbf{1}[t=k] * Male_i + \lambda_i + \lambda_j + \lambda_t + \varphi X_{it} + \epsilon_{ij(i,t)t},
\end{aligned}
\tag{1.1}
$$

where $i$ denotes a worker employed in establishment $j(i,t)$ in calendar year $t$. $\mathbf{1}[t=k]$ is a year dummy that takes the value one if $k$ equals $t$ and zero otherwise. $Male_i$ denotes the gender dummy that takes the value one if individual $i$ is male. $Treat_{j(i,2013)}$ denotes the treatment indicator which equals one if an establishment belongs to a firm which has 150 to 225 employees in 2013 and zero otherwise.[18] $X_{it}$ is a vector of individual, time-varying controls: It contains a quartic polynomial in age and its interaction with gender. $\lambda_i$ denotes the individual worker fixed effect. $\lambda_j$ and $\lambda_t$ respectively denote the establishment and calendar year fixed effects. Our outcome variable of interest is the log of daily wages at the worker-establishment-year level. We drop 2013 from the summation terms (i.e. the event-study coefficients $\beta_1^k$, $\beta_2^k$, $\gamma_k$, and $\lambda_t$.) Thus, the event-study coefficients $\beta_1^k$ on the triple interaction term measure the percentage points change in the gender wage gap in treated establishments relative to the control group and the base year 2013. If the pay transparency reform is effective in reducing the gender pay gap, the coefficient $\beta_1^k$ will be negative for $k > 2013$, i.e. the post-treatment years. Conversely, a positive coefficient implies that the gender pay gap has opened up. In addition, we are interested in the effects of pay transparency on male and female wages separately. The gender specific effects are measured with the coefficients $\beta_2^k$ for females and $\beta_1^k + \beta_2^k$ for males. Standard errors in all our analyses are clustered at the establishment level.

Our two-way fixed effects strategy implies that our effects are identified within- establishment and within-worker, i.e. the additional effect of this pay transparency reform after controlling for unobserved but time-constant worker and establishment characteristics. Workers who stay with their employers before and after the policy contribute to

[18] Assigning the treatment status based on the 2013 firm size is equivalent to estimating an intent-to-treat effect. To account for initial-treatment status violators in post-reform years, we consider a robustness exercise by estimating equation (1.1) for only those establishments that comply with their initial treatment assignment, thus not exceeding (dropping below) the 150 employee cutoff post 2013. We refer to this sample as the "Complier Sample". Complying firms account for 76 percent of worker-year observations in our baseline sample.

these effects only if their wages change as a result of the policy. This is also true for workers who move across establishments. Consequently, our results are not driven by sorting of higher individual fixed effect workers to higher paying establishments, which could be different across genders.[19]

We estimate equation (1.1) for our baseline sample, i.e. establishments whose firm size is around the lowest size cutoff of the reform to ensure their comparability with respect to (un)observables. Under the assumption that establishments of larger firms exhibit the same parallel trends, we can analyze the full staggered roll-out of the reform. To this end we are applying a staggered difference-in-difference design for all treated establishments, again accounting for response heterogeneity over time. We modify equation (1.1) as follows:

$$\begin{aligned} y_{ij(i,t)t} = &\sum_{k=-4}^{4} \beta_1^k \mathbf{1}[YST = k] * Male_i * Treat_{j(i,2010)} + \sum_{k=-4}^{4} \beta_2^k \mathbf{1}[YST = k] * Treat_{j(i,2010)} \\ &+ \beta_3 Male_i * Treat_{j(i,2010)} + \sum_{k=2007}^{2018} \gamma_k \mathbf{1}[t = k] * Male_i + \lambda_i + \lambda_j + \lambda_t + \varphi X_{it} + \epsilon_{ij(i,t)t}, \end{aligned} \tag{1.2}$$

where all variables have the same definition as above except that we now define the treatment status based on the firm size in 2010 and replace the year dummy $\mathbf{1}[t = k]$ with a "years-since-treatment" (YST) dummy $\mathbf{1}[YST = k]$. We choose 2010 as the base year for defining the treatment status as this is the last pre-treatment year for the largest firm size group (more than 1000 employees). Moreover, we replace the year dummy by the years-since-treatment dummy because the different firm size groups are treated at different points in time. Hence, we recenter the actual treatment for each establishment at YST equal to zero, which corresponds to different calendar years for each treatment group, e.g. 2011 for the largest firm size group (more than 1000 employees) and 2014 for the smallest firm size group (150 - 249 employees). We include four pre- and post-treatment years in our analysis, which corresponds to the number of pre-/post-treatment years we can observe for all treated firm-size groups. The $\beta_1^k$ coefficients inform us about the evolution of the gender pay gap in treated establishments relative to their specific treatment date and relative to never-treated establishments after con-

[19] In an alternative specification we include establishment-worker match fixed effects directly controlling for potential "assortative" matching. Both point estimates and confidence intervals are not sensitive to this alternative specification.

trolling for year, worker, and establishment specific heterogeneity (fixed effects), again omitting period $k = -1$.

Before presenting our results, we briefly discuss two key identifying assumptions for unbiased estimates in our context. First, we have to impose the parallel trend assumption: The gap between male and female wages in the control (75-149 employees) and treatment group (150-225 employees) exhibits the same trends absent any policy change. If this holds, we can attribute any post-transparency deviations between the groups to the policy. While not directly testable, the estimated coefficients $\beta_1^k$ for pre-treatment years show that the difference in the gender wage gap between treated and control groups is not significantly different from zero (see Figure 1.2). Note that this also precludes anticipation effects: If treated establishments respond to the reform prior to the actual reform date, for example by eliminating unfair pay practices, then this would also show up as a deviation from the parallel trend assumption.

A second concern is that firms use the time between the implementation of the reform in 2011 and its effective date in 2014 to downsize and locate themselves right below the 150 employee cutoff, thus avoiding treatment in 2014. If the worst offenders (largest gender pay gap) among this sample move below the cutoff, then our estimates will be biased towards zero. To show that this does not pose a threat to identification, we show in Figure 1.1 that the firm size distributions are almost identical in 2010 and 2014, and there is no evidence of bunching around the threshold. In Appendix Figure 1.A2 we check for violations of intended treatment rule by establishments after the policy was implemented and in Appendix Figure 1.A3 we plot the year-on-year transitions of establishments in treated and control groups across the size cutoff. These figures additionally show that even though there were some violations of the intended treatment rule, the proportions are in line with pre-policy firm size dynamics, thus further ruling out strategic bunching.

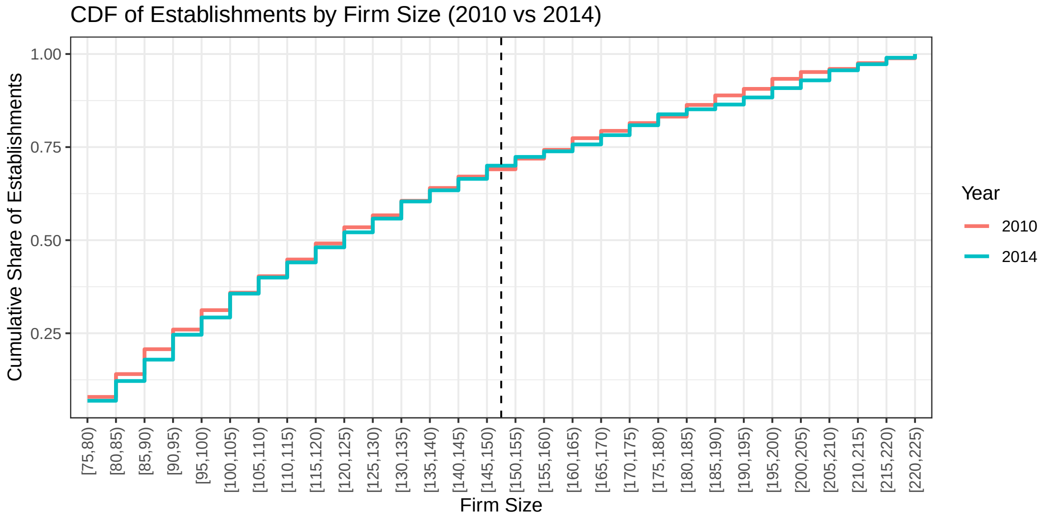
## 1.5 The Effects of Pay Transparency

### 1.5.1 Effects on Gender Wage Gap and Wages

In line with the primary goal of the Austrian Pay Transparency law, we begin by examining its effect on the gender gap in daily wages. Panel (a) in Figure 1.2 shows the estimated coefficients $\beta_1^k$ from equation (1.1), which measure the evolution of the gen-

Figure 1.1: Cumulative Firm Size Distribution of Establishments in Baseline Sample

*The figure below shows the cumulative distribution function of the firm size distribution for our baseline sample in 2010 (before the policy was announced) and 2014 (one year after the policy was fully implemented for all firms with more than 150 employees). The figure shows that there is virtually no change in the size distribution between these two years.*

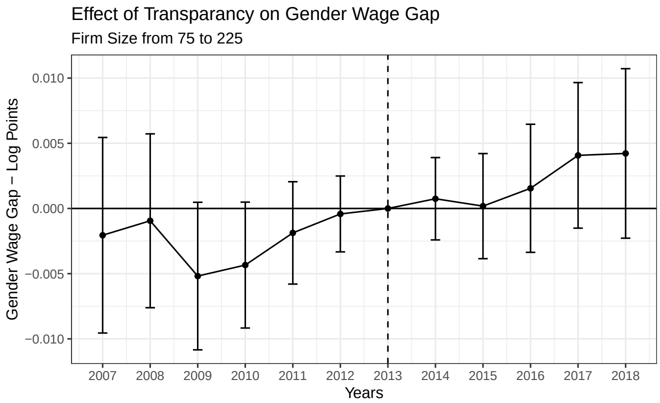



der wage gap (male wage premium) in treated establishments relative to those in the control group. First, we check that the parallel trend assumption is satisfied. Studying the coefficients in pre-treatment years, we find little evidence for any statistically and economically distinct evolution of the gender wage gap in treated versus control establishments. There is a noticeable, but statistically insignificant dip in the gender wage gap around the great recession. In Appendix Figure 1.A4 we show that this dip occurs in both treated and untreated establishments and is only somewhat (by about 0.5 percentage points) more pronounced in treated establishments. By the time the policy is implemented in 2014, the gender wage gap in both groups had recovered to their pre-recession levels.

Post-treatment, we also find little evidence for any significant and economically meaningful effects of the reform on the gender wage gap. The gender wage gap between treated and control group started opening up only in 2015, and we can rule out at the 95% confidence level that during our study period the policy narrowed the gender wage gap by more than 0.4 p.p.. In Panel (b) and (c) we plot the effects on male ($\beta_1^k + \beta_2^k$) and female ($\beta_2^k$) wages respectively. Female wages are virtually unchanged after 2013, whereas male workers in treated establishments have seen a modest increase of 0.25 p.p. compared to the control group. Both effects are statistically insignificant,

Figure 1.2: Effects of Pay Transparency on Gender Wage Gap and Daily Wages

*The figure below plots the evolution of the gender gap in daily wage (panel a), male (panel b) and female wages (panel c), in treated establishments relative to the control group in log points (Eq. (1.1)). The sample is restricted to establishments of firms with 75-225 employees. Treatment is assigned to establishments of firms which had more than 150 employees in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI. All regression results can be found in Appendix Table 1.A2.*



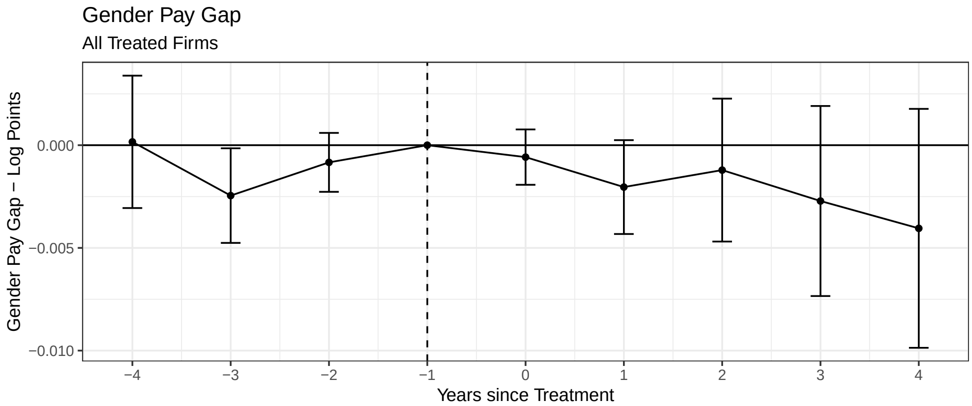

(a) Gender Wage Gap (Male-Female)


(b) Male $log$(Daily Wage)


(c) Female $log$(Daily Wage)

Figure 1.3: Effects of Pay Transparency on Gender Wage Gap

*The figure below plots the evolution of the gender gap in daily wages in treated establishments relative to the control group in log points based on the staggered difference-in-difference model in equation (1.2). The sample is restricted to establishments of firms above 75 employees. Treatment is assigned based on the 2010 firm-size and the treatment time is re-centered around 0, which is the first treatment year. We drop years outside our event window. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*

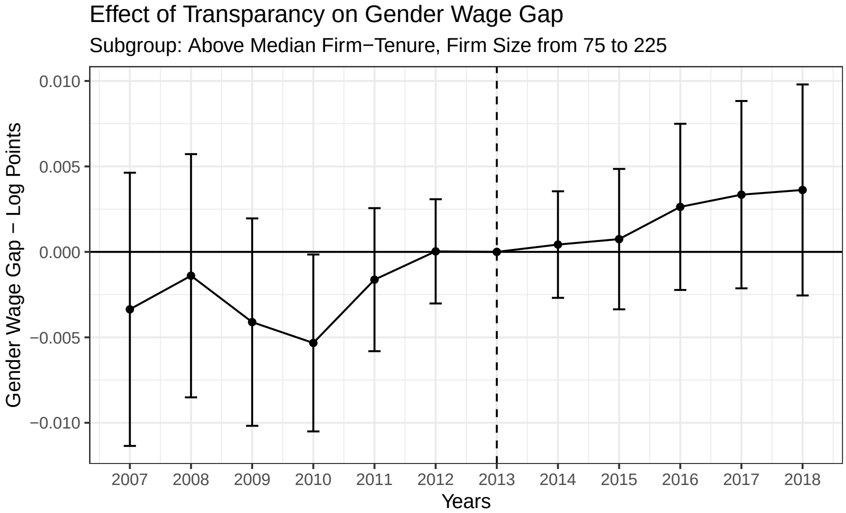



although they are precisely estimated. At the 95% confidence level, we can rule out that the reform affected wages by more than 0.5 percentage points in the years immediately after the roll-out and by more than 0.8 p.p. towards the end of our study period. Overall, there is little evidence to suggest that transparency has any economically significant effects on female workers. While our baseline sample focuses on firms around the threshold to make our control and treatment group as comparable as possible, we next investigate whether transparency had an effect in larger firms by studying the full roll-out over all firm size groups. Figure 1.3 presents the estimation results for $\beta_1^k$ from the staggered difference-in-difference model detailed in equation (1.2). Again, these coefficients inform us about the evolution of the gender pay gap (male wage premium) in treated establishments relative to those in the control group. Including all firm size groups eventually treated does not change the results found in the baseline sample. There are no discernible pre-trends and post treatment there is little evidence for any significant and economically meaningful effects on the gender wage gap. As above, these effects are precisely estimated and we can rule out any effect greater than 0.5 p.p. in the three years following the policy introduction.[20] Since pay reports are only avail-

[20]In independent work, Böheim and Gust (2021) confirm our main findings using a regression discontinuity design.

Figure 1.4: Effects of Transparency on Gender Wage Gap (GWG)

*The figure below plots the effects of pay transparency on the gender wage gap for workers with above median tenure where we compute firm-tenure in 2013, the year before treatment. The sample is restricted to establishments of firms with 75-225 employees in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




able to current employees, the reports might have a limited impact on wages of newly hired employees. Even after joining a company with a pay report, it might take some time until the employee is able to act upon the information provided in the wage reports and renegotiate their contract. Therefore, it is possible that transparency has significant effects only for those who have been with their current employer for a while. To investigate whether this group drives our zero results, we re-estimate equation (1.1) on the sample of workers with above 3.5 years of establishment tenure, which is the median value in our baseline sample. The results displayed in Figure 1.4 below show that there is no discernible effects of transparency for high tenure workers.

### 1.5.2 Robustness Checks

In the Appendix we show that the results of our baseline specification hold under multiple robustness checks with different sample and treatment definitions. In Appendix Figure 1.A5 and 1.A6 we restrict our sample to establishments with firm size between 100-200 employees and 125-175 employees in 2013 respectively. In contrast to our main analysis sample, we include all topcoded workers in Appendix Figure 1.A7 and drop all ever-topcoded workers in Appendix Figure 1.A8. For Appendix Figure 1.A9 we include only those establishments which do not change their intended treatment assignment based on their firm size in 2013. We also change the definition of treatment in the following two ways. In Appendix Figure 1.A10 we define establishment treatment status based on their firm size in 2010, instead of 2013. For Appendix Figure 1.A11 we assign treatment status to workers (instead of establishments) depending on whether they worked in an establishment with a firm size greater than 150 employees in 2013. In Appendix Table 1.A2, we re-estimate the gender wage gap results for our main sample with match fixed effects instead of worker and establishment fixed effects. Finally, in Appendix Figure 1.A12 we re-estimate the effects of transparency at the establishment-year level and thus on the establishment level gender wage gap.[21] All these specifications confirm our main results: pay transparency had no economic or statistically significant effects on the gender wage gap and individual wages.

### 1.5.3 Pay Transparency and Wage Dispersion

What explains the lack of any discernible effects of transparency on male and female wages? Perhaps the policy only led to wage compression, leaving the average wage unaffected. Wage increases for workers earning below average might have been compensated by wage reductions for highly paid individuals. To check whether this was indeed the case, we estimate the effects of the pay transparency on the establishment-level variance in male and female wages separately by estimating the following model in our baseline sample:

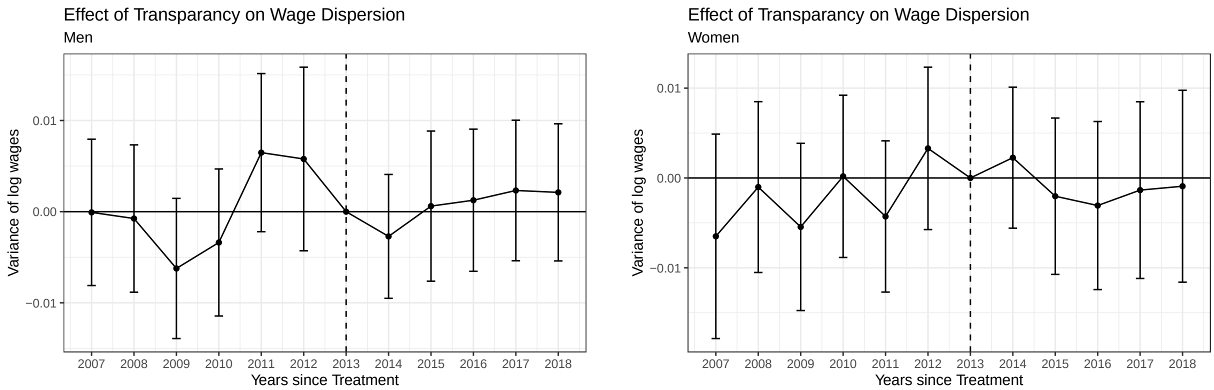
$$wvar_{jt} = \sum_{k=2007}^{2018} \beta^k \mathbf{1}[t = k] * Treat_{j(2013)} + \lambda_j + \lambda_t + \epsilon_{jt}, \tag{1.3}$$

where $wvar_{jt}$ is the gender-specific variance in log daily wages in establishment $j$ in year $t$, $Treat_{j(2013)}$ is a dummy which takes the value one for any establishment $j$ whose parent firm is larger than 150 employees in 2013, and the other variables have the same interpretation as in (1.1). A negative $\beta_k$ coefficient implies that the variance narrowed in post-treatment years relative to the control group, which implies wage compression. The results are displayed in Figure 1.5. There are no discernible pre-trends in wage variances for either men or women. We do not find statistically significant effects of transparency on the establishment-level variance in log wages for either men or women.

[21] The appendix section 1.A.5 describes the regression specification in detail.

Figure 1.5: Effects of Pay Transparency on Establishment-level Wage Variance

*The figure below plots the effects of transparency on the establishment-level variance in daily wages for male and female workers separately (Eq. (1.3)). The sample is restricted to establishments of firms with 75-225 employees in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% confidence intervals.*




(a) Male (b) Female

The mean of the establishment-level variance of log wages is 0.097 for males and 0.176 for females in the year before treatment. Thus, at the 95% confidence level, we can rule out that transparency narrowed establishment level wage dispersion for men or women by more than 5 percent relative to be baseline mean.

An alternative way to study the effects of the policy on wage compression is to estimate the impact separately on workers earning below and above their respective gender-specific median establishment wage. In the appendix, Figure $1.A13$ shows that the policy had little effect on the wages of any subgroup. All in all, we do not find any compelling evidence for wage compression within establishments.[22]

## 1.6 Why Was the Reform not Effective?

Why did the Austrian pay transparency law fail to narrow the gender wage gap? As we have already discussed in Section 1.2, the Austrian policy is in many aspects stricter than comparable laws in Europe and there was near-universal compliance with the policy. According to a survey of works councils (Arbeiterkammer, 2014), in 54% of cases employers cooperated with works councils in generating pay reports. 71% of respondents reported that the reports are informative and 63% claimed that they are useful for

[22]Including establishment-year level aggregates in (1.3) does not change our results.

work councils. Therefore, incomplete implementation and workers' unawareness are unlikely to explain the lack of policy effects.[23]

As argued in the the conceptual framework, transparency can only be effective if the within-occupation and within-firm gender wage gaps are large enough. The absence of detailed occupation information in the social security data does not allow us to quantify this gap. However, previous work (Böheim, Fink, and Zulehner, 2020) that controls for occupation information has found the adjusted wage gap in Austria to be 7.2 percent in 2013. Thus, within-firm and within-occupation gender differences are likely to be even smaller.[24]

Even if the transparency reform revealed large gender differences in firm pay policies, transparency itself might not remedy these differences. The Austrian pay transparency legislation does not require firms to act on revealed pay differences. Instead, it is the workers' responsibility to use the information provided to bargain for higher wages. Thus, the policy's ineffectiveness could also be grounded in low bargaining power of workers. If the reports show evidence of pay discrimination or unfair wage differences, but workers lack the bargaining power to renegotiate wages, we would expect job satisfaction to decline. In contrast, we would expect job satisfaction to increase if transparency leads workers to revise downwards their priors about unfair compensation. The social security data does not have a direct measure of job satisfaction, but we can use turnover rates as a proxy.[25] Past research has shown that workers who feel unfairly compensated have lower job satisfaction and higher quit rates (Card et al., 2012; Dube, Giuliano, and Leonard, 2019; Rege and Solli, 2015).

To study this channel, we estimate the effect of the policy on overall job separation rates by dropping the additional gender interaction from equation (1.1):

$$sepa_{ijt} = \sum_{k=2007}^{2018} \beta^k \mathbf{1}[t=k] * Treat_{j(i,2013)} + \sum_{k=2007}^{2018} \gamma_k \mathbf{1}[t=k] * Male_i + \lambda_j + \lambda_i + \lambda_t + \varphi X_{it} + \epsilon_{ij(i,t)t}, \tag{1.4}$$

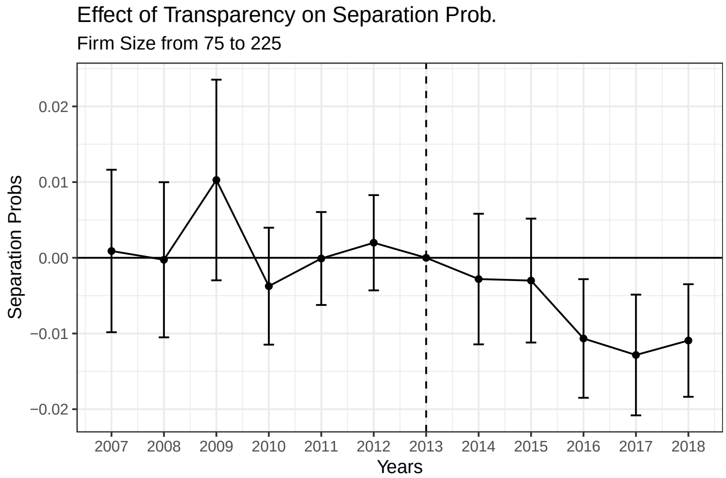
[23] As mentioned in section 1.2, gender pay gap in Austria is prominently discussed in media twice a year on "Equal Pay Days", once in spring and then again in fall. Pay reports are also often discussed in this context.

[24] This is also consistent with findings in Card, Cardoso, and Kline (2016), who show that within-firm gender pay gaps in Portugal is close to zero, and sorting explains the overwhelming majority of gender wage differences.

[25] Since pay reports are internal, we would not expect workers' outside options to change and therefore to confound effects on quit rates.

Figure 1.6: Effects of Transparency on Job Separation Rate

*The figure below plots the effects of pay transparency on the year-on-year job separation rate (Eq. 1.4). The sample is restricted to establishments of firms with 75-225 employees in 2013, and we pool male and female workers. Standard errors are clustered at the establishment level. The standard error spikes represent 95% confidence intervals.*




where $sepa_{ij(i,t)t}$ is one if individual $i$ separated in period $t$ from establishment $j$ and the rest of the variables follow the same definitions as in the baseline equation (1.1). As before, the year 2013 is omitted from the estimation of $\beta^k$, $\gamma_k$, and $\lambda_t$. Figure 1.6 shows that the transparency policy reduced the annual separation rate significantly in treated firms relative to the control group by over 1.1 p.p., which is a 9 percent reduction compared to pre-treatment levels.[26] In Appendix Figure $1.A14$ we show that these effects are similar for men and women.[27] The reduced turnover rate is perhaps indicative that transparency alleviated previously held concerns about unfair pay schedules among workers in general, as well as unfair gender pay gaps.

In addition, the Austrian transparency policy by design does not target an important determinant of the gender pay gap - the fact that men sort into better paying firms compared to women. Table 1.A3 in the appendix shows that in Austria, gender differences in sorting explain around ten percentage points of the unadjusted gender wage gap. But since wage reports are legislated to be company secret and hence not publicly available, they cannot directly affect the sorting component. Therefore, transparency legislation

[26]The separation rate is 0.122 in treated firms before the reform, see Table 1.1

[27]We estimate the gender specific effects of transparency on job separation using the specification of equation (1.1).

that requires firms to publicly disclose pay statistics, such as in the UK, could be more effective in closing the gender gap in firm pay (Duchini, Simion, and Turrell, 2020).[28] An additional advantage of the public nature is that the reported wage gaps can be discussed in the media which makes the policy more salient and can also put additional pressure on firms to equalize earnings (Blundell, 2020).[29]

Independent of the specific reasons why the Austrian transparency reform was not successful in narrowing the gender wage gap, requiring firms to act upon revealed wage differences or mandating wage reports to be public might lead to a more effective transparency policy.

## 1.7 Conclusion

Pay transparency is often prescribed as an instrument to close the gender pay gap, and reduce wage inequality. In this paper we study the causal effects of the 2011 Austrian pay transparency law which requires firms above a certain size threshold to publish reports on gender pay gap.

Using an event-study design and administrative data from social security records, we show that the transparency policy neither affected male and female wages nor did it narrow the gender wage gap. These effects are precisely estimated, and we can rule out at a 95% confidence level that the policy narrowed the gap by more than 0.4 p.p. by the end of our study period. We further show that this zero effect is not driven by wage compression, where wage increases below the median are compensated with wage cuts above the median.

In addition we find that pay transparency leads to a reduction in separation rates in treated firms. Past research has shown that workers who feel unfairly compensated have lower job satisfaction and a higher quit rate (Card et al., 2012, Rege and Solli, 2015, Dube, Giuliano, and Leonard, 2019). Therefore the lower separation rate might point towards higher job satisfaction and is perhaps indicative that transparency alleviated

[28] Another example is Canada, where public access to information about the salaries of university faculties led to a reduction in the gender wage gap (Baker et al., 2019).

[29] The Independent, a newspaper in the UK, regularly publishes the worst offenders in terms of gender pay gap based on the UK transparency reform. `https://www.independent.co.uk/news/business/news/gender-pay-gap-worst-offenders-each-sector-revealed-reporting-deadline-passes-a8290566.html`

previously held concerns about unfair pay schedules among workers.

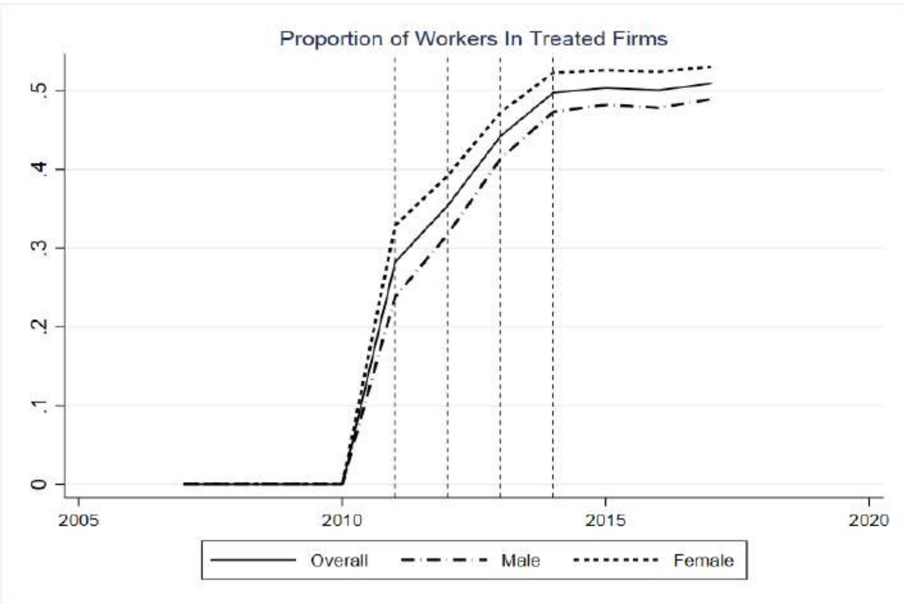
Our data does not allow us to definitively pin down the reasons behind the lack of policy effects on the gender wage gap. However, policies which require firms to act upon revealed wage differences or mandate wage reports to be public might be more effective in narrowing the gender wage gap.

# Appendix

## 1.A Appendix

### 1.A.1 Other Summary Figures

Figure 1.A1: Proportion of Workers Employed in Treated Establishments



## 1.A.2 Sample Income Report from the Public Sector

Table 1.A1: Income Report for 2016: All Federal Services

*The following table is from "Einkommensbericht 2017" of the Austrian Federal Government, Public Administration. It is publicly available at Einkommensbericht, 2017. The table illustrates how an income report can be written. The first column depicts the occupational groups/task groups as defined by collective bargaining agreements. The rows printed in bold summarize the statistics averaged for each occupation.task group. The same is repeated for employees in training and those who previously worked for the government, but are now employed in a (semi-) private company, e.g. postal services or telecommunications. All these tables are accompanied by brief discussion on why there are wage differences and measures taken to reduce differences that stem from factors not related to the seniority structure or composition within task groups (for example: office clerks and technicians are in the same group but technicians are paid more. The former group is mostly female, while the latter is mostly male, which explains some of the differences in remuneration schedules by group.*

| | Number of Workers | | Median Gross Income/Yr | | Mean Age | | Gender Pay Gap | Age Diff |
|---|---|---|---|---|---|---|---|---|
| Occupation Clusters | Men | Women | Men | Women | Men | Women | % | (Men-Women) |
| **Central Administration** | 23872 | 27002 | 45637 | 35799 | 49.2 | 46.1 | 21.6% | 3.1 |
| A1, v1 | 4157 | 3211 | 75141 | 61482 | 48.6 | 44.0 | 18.2% | 4.6 |
| A2, v2 | 7598 | 6454 | 57201 | 47898 | 49.7 | 45.9 | 16.3% | 3.8 |
| A3, v3, h1 | 6401 | 10721 | 38151 | 34285 | 49.8 | 46.7 | 10.1% | 3.1 |
| A4-7, v4-5, h2-5 | 4421 | 5962 | 28336 | 25749 | 46.5 | 45.1 | 9.1% | 1.5 |
| Service Rank: Central Administration | 756 | 553 | 78994 | 65742 | 57.3 | 56.0 | 16.8% | 1.4 |
| Data Services and Management | 539 | 101 | 60305 | 56189 | 46.7 | 48.5 | 6.8% | -1.8 |
| **Police and Law Enforcement (Executive)** | 27484 | 5230 | 51504 | 40776 | 44.8 | 34.2 | 20.8% | 10.5 |
| E1 | 649 | 42 | 81756 | 64668 | 52.3 | 44.4 | 20.9% | 7.9 |
| E2a | 9742 | 975 | 58561 | 46584 | 50.3 | 39.7 | 20.5% | 10.6 |
| E2b, Lowest Rank Officer | 15344 | 3519 | 48284 | 40797 | 43.0 | 34.5 | 15.5% | 8.5 |
| E2c, Aspirant | 1705 | 694 | 17442 | 17442 | 26.3 | 24.5 | 0.0% | 1.8 |
| Service Rank, Executive Office | 44 | 0 | 54334 | - | 54.8 | - | - | - |
| **Judges, District Attorneys (Judiciary)** | 1491 | 1746 | 91417 | 80341 | 48.4 | 43.9 | 12.1% | 4.5 |
| R3, III | 96 | 37 | 144402 | 123945 | 55.9 | 51.5 | 14.2% | 4.4 |
| R2, II | 106 | 85 | 111366 | 106649 | 54.0 | 52.3 | 4.2% | 1.7 |
| R1a, R1b, I | 739 | 1011 | 88651 | 80341 | 48.4 | 44.7 | 9.4% | 3.7 |
| Federal Court Judges | 225 | 195 | 96489 | 99331 | 52.4 | 50.9 | -3.0% | 1.4 |
| Judge Aspirants | 71 | 136 | 34192 | 34192 | 29.8 | 28.6 | 0.0% | 1.2 |
| Procurator General's Office | 12 | 6 | 128815 | 125434 | 52.7 | 49.5 | 2.6% | 3.2 |
| St2, STII | 55 | 30 | 90827 | 84100 | 46.3 | 45.1 | 7.4% | 1.2 |
| St1, STI | 187 | 246 | 81175 | 70271 | 43.9 | 39.3 | 13.4% | 4.6 |
| **Military Service** | 15661 | 421 | 41589 | 28777 | 41.6 | 31.1 | 30.8% | 10.4 |
| MBO1, MZO1 | 735 | 45 | 91956 | 78806 | 48.7 | 45.2 | 14.3% | 3.4 |
| MBO2, MZO2 | 2160 | 23 | 56766 | 43759 | 45.3 | 33.5 | 22.9% | 11.8 |
| MBUO1, MZUO1 | 6673 | 63 | 44411 | 34442 | 49.6 | 37.3 | 22.5% | 12.3 |
| MBUO2, MZUO2, MZO3 | 2477 | 92 | 34108 | 29580 | 33.1 | 31.6 | 13.3% | 1.5 |
| MZ Charge | 1684 | 171 | 27910 | 22792 | 24.1 | 25.3 | 18.3% | -1.3 |
| Service Rank: Military Service | 557 | 0 | 42654 | - | 55.1 | - | - | - |
| International Strike Force | 1375 | 27 | 29231 | 27493 | 24.1 | 26.2 | 5.9% | -2.1 |
| **Teachers** | 19339 | 30109 | 60584 | 52635 | 48.2 | 45.4 | 13.1% | 2.8 |
| L1, I1 | 14837 | 23628 | 64858 | 55453 | 49.0 | 46.1 | 14.5% | 3.0 |
| L2, I2 | 4156 | 5750 | 48396 | 43609 | 46.7 | 44.9 | 9.9% | 1.8 |
| L3, I3 | 123 | 118 | 24360 | 24599 | 45.9 | 47.0 | -1.0% | -1.2 |
| Foreign Exchange Teachers | 223 | 523 | 17154 | 17293 | 25.5 | 24.7 | -0.8% | 0.8 |
| **Lecturers (University)** | 679 | 852 | 69591 | 65002 | 52.4 | 50.9 | 6.6% | 1.5 |
| **Educational Board** | 171 | 143 | 85325 | 83103 | 56.6 | 56.0 | 2.6% | 0.6 |
| **Nursing and Health Services** | 91 | 175 | 44317 | 39369 | 48.1 | 47.8 | 11.2% | 0.4 |
| K2, k2 | 25 | 28 | 49982 | 43525 | 48.7 | 44.7 | 12.9% | 4.0 |
| K3, k3 | 7 | 11 | 56430 | 55410 | 55.2 | 55.8 | 1.8% | -0.7 |
| K4, k4 | 43 | 95 | 42875 | 40192 | 47.6 | 46.4 | 6.3% | 1.2 |
| K5, k5 | 8 | - | 40734 | - | 49.1 | - | - | - |
| K6, k6 | 15 | 34 | 32272 | 33825 | 46.6 | 50.7 | -4.8% | -4.1 |
| **Others** | 184 | 452 | 106960 | 106960 | 53.5 | 51.3 | 0.0% | 2.2 |
| Medical professionals | 168 | 449 | 106960 | 106960 | 55.4 | 51.4 | 0.0% | 4.0 |
| Others | 16 | 3 | 25269 | 27723 | 33.7 | 34.0 | -9.7% | -0.3 |

### 1.A.3 Bunching of Establishments

Figure 1.A2: Establishments Violating Intended Treatment Status based on Size Rule

*The figure below shows the establishment-size weighted fraction of establishments that violate intended treatment rule based on their firm sizes in 2010 and 2013, separately. Establishments would violate their intended treatment rule if they enter treatment either before the intended start year because of an increase in firm size, or they manage to delay treatment beyond their intended year by reducing firm size.*

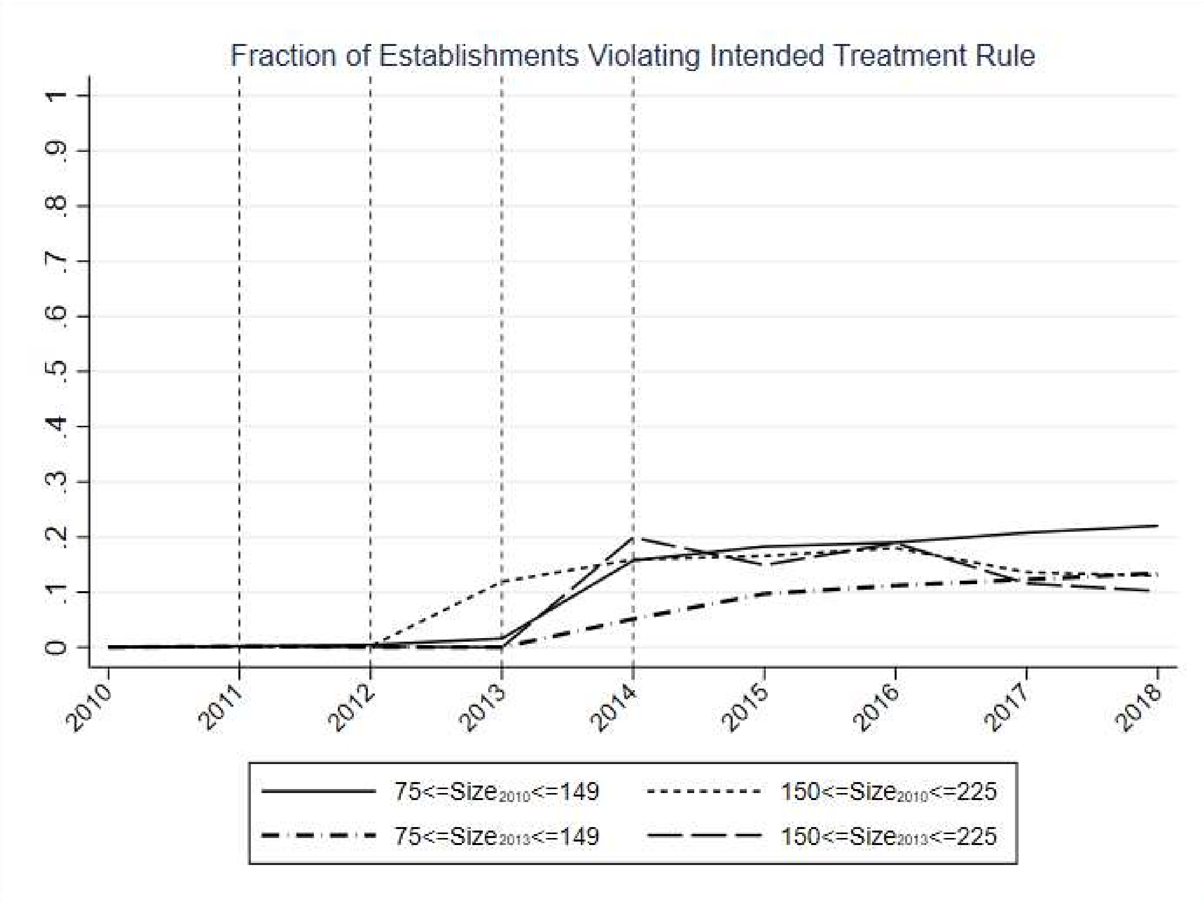
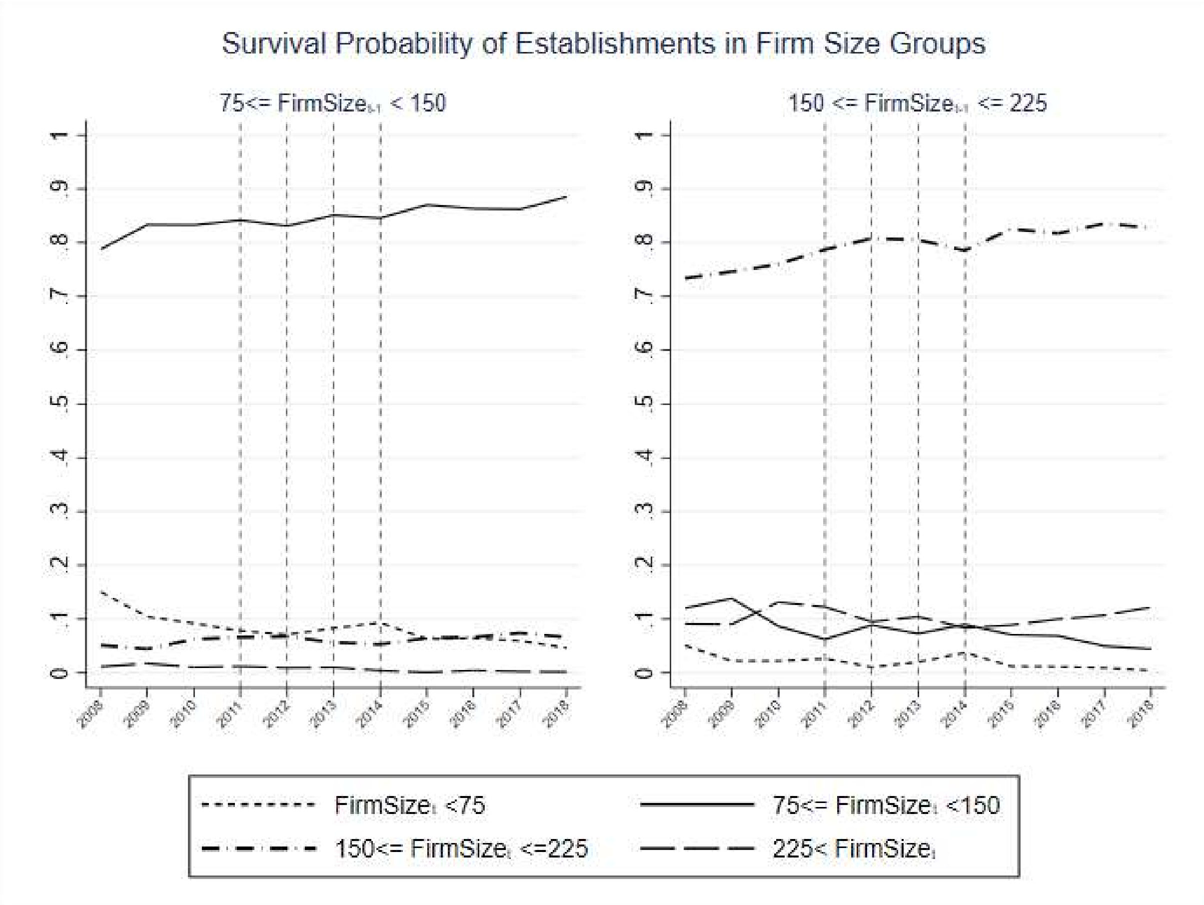


Figure 1.A3: Transitions of Establishments Across Firm Size Groups

*The figure below plots the fraction of establishments, weighted by establishment size, that survive in the same firm size group or transition to other firm size groups, relative to the number of establishments in each size group for the previous year. We do this exercise for the treated and control groups of establishments which represent those just above and below the 150 firm size-cutoff respectively.*



### 1.A.4 Robustness Checks

Figure 1.A4: Effects of Pay Transparency on Adjusted Gender Wage Gap (By Treatment Status)

*The figure below shows the evolution of the gender wage gap, separately for the treated and control group of establishments. The sample includes only establishments of firms which had between 75 and 225 employees in 2013, the year before treatment. Establishments of firms which had more than 150 employees in 2013, were assigned to treatment status, and others to the control group.*

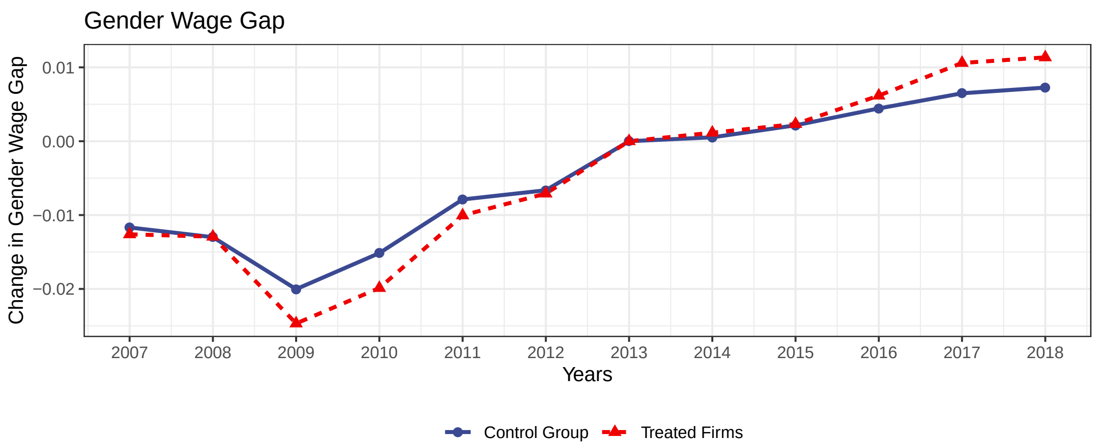
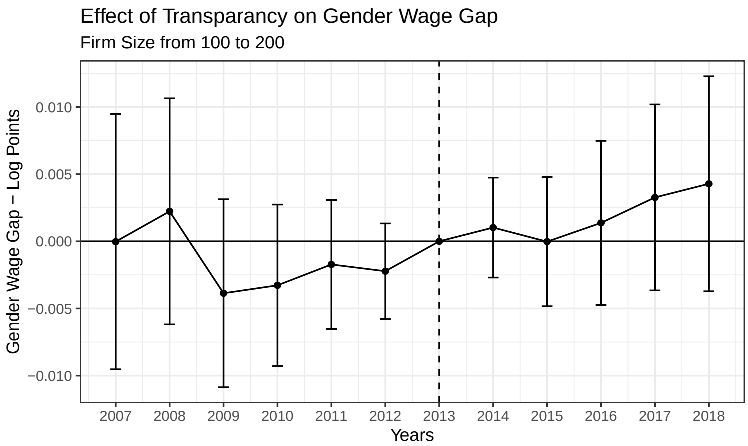


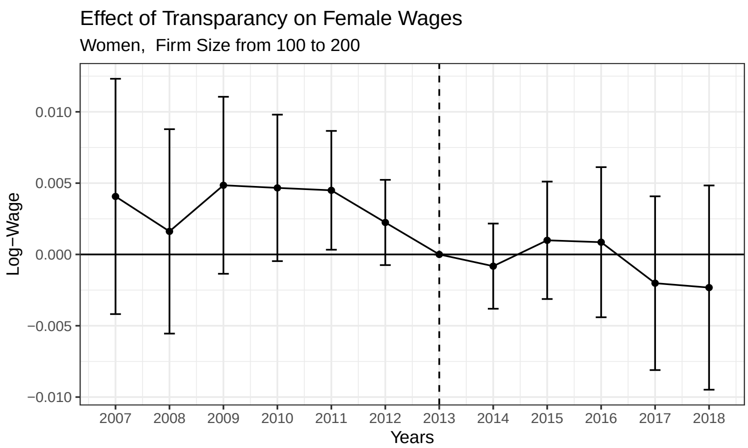
Figure 1.A5: Effects of Transparency on GWG and Daily Wage (100<=Firm Size<=200)

*The figure below plots the effects of transparency on gender wage gap (Panel (a)), and daily wages for male (Panel (b)) and female (Panel (c)) workers separately, in establishments of firms which had between 100-200 employees in 2013 (Eq. 1.1). Treatment is assigned to establishments of firms which had more than 150 workers in 2013. Standard errors are clustered at establishment level. The standard error spikes represent 95% CI.*




(a) Gender Wage Gap

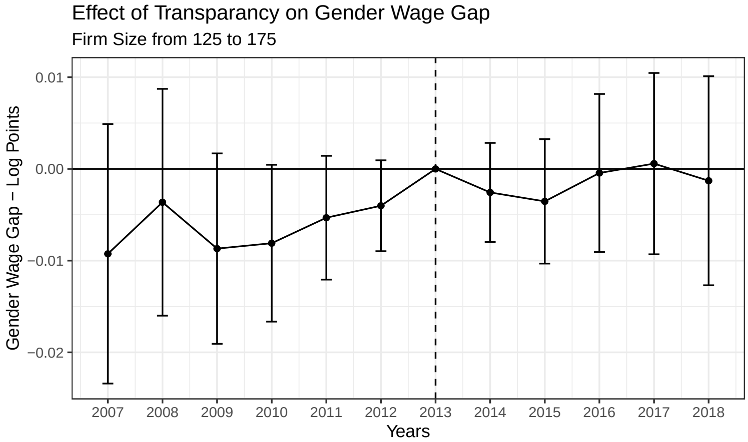

(b) Male Daily Wage




(c) Female Daily Wage

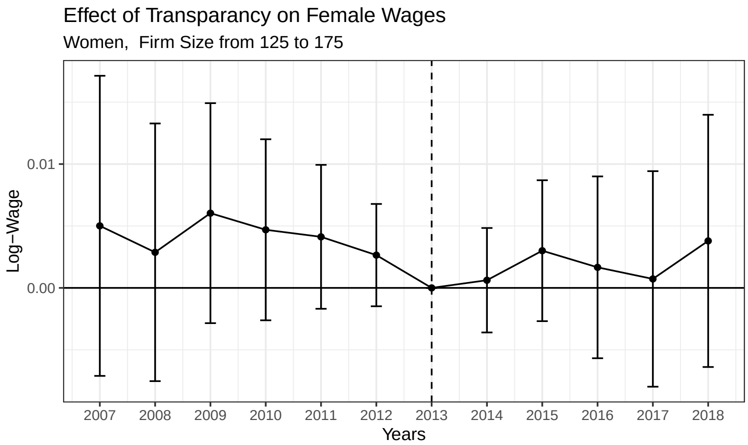
Figure 1.A6: Effects of Transparency on GWG and Daily Wage (125<=Firm Size<=175)

*The figure below plots the effects of transparency on the gender wage gap (Panel (a)), and on daily wages for male (Panel (b)) and female (Panel (c)) workers separately, in establishments of firms which had between 125-175 employees in 2013 (Eq. 1.1). Treatment is assigned to establishments of firms which had more than 150 workers in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




(a) Gender Wage Gap

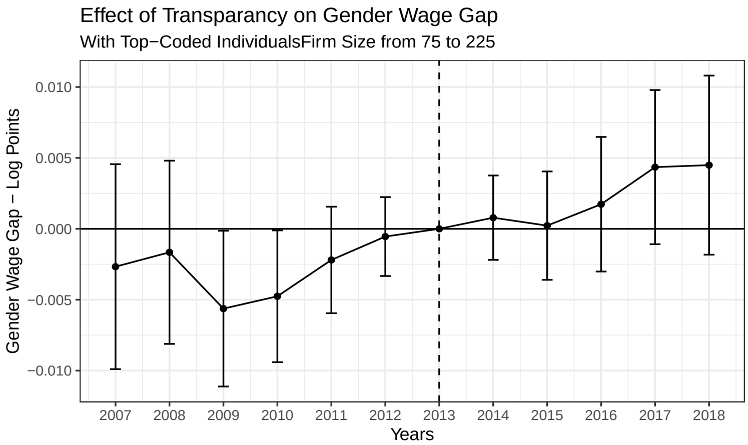

(b) Male Daily Wage




(c) Female Daily Wage

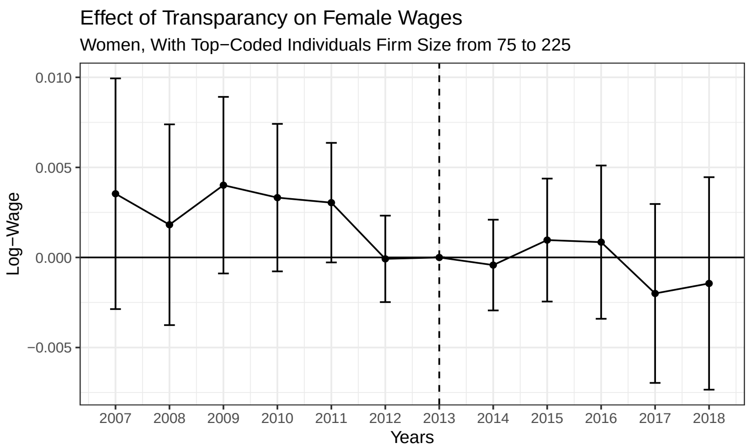
Figure 1.A7: Effects of Transparency on GWG and Daily Wage (With Top-Coded)

*The figure below plots the effects of transparency on the gender wage gap (Panel (a)), and on daily wages for male (Panel (b)) and female (Panel (c)) workers separately (Eq. 1.1). The sample is restricted to establishments of firms with 75-225 employees in 2013. All workers with top-coded daily wages are included in the sample, with their daily wage set to the year-specific top-coding. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




(a) Gender Wage Gap

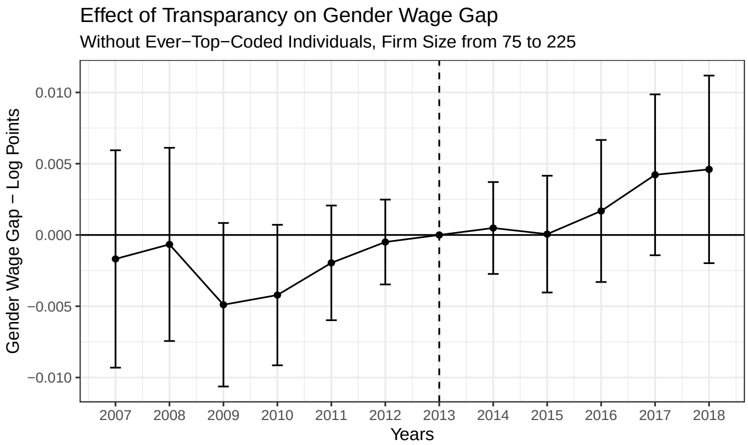

(b) Male Daily Wage




(c) Female Daily Wage

Figure 1.A8: Effects of Transparency on GWG and Daily Wage (Without Ever-Top-Coded)

*The figure below plots the effects of transparency on the gender wage gap (Panel (a)), and on daily wages for male (Panel (b)) and female (Panel (c)) workers separately (Eq. 1.1). The sample is restricted to establishments of firms with 75-225 employees in 2013. All workers who were ever top-coded in the sample period are dropped. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*



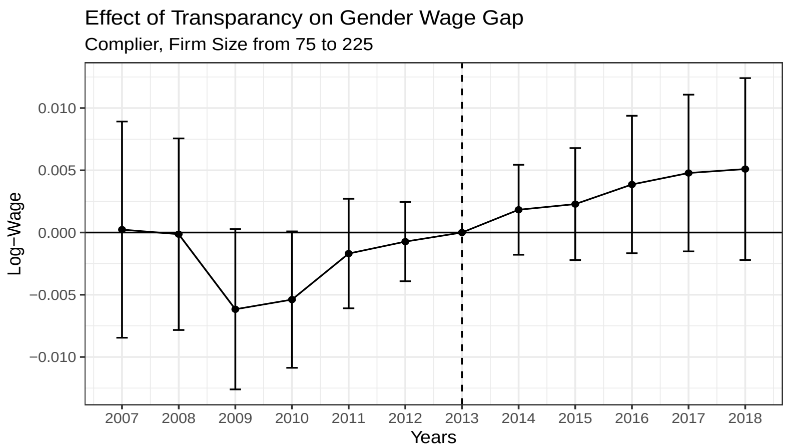

(a) Gender Wage Gap


(b) Male Daily Wage


(c) Female Daily Wage

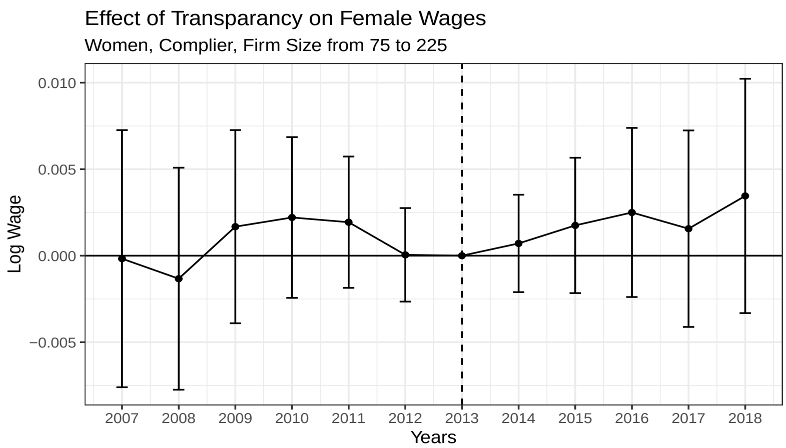
Figure 1.A9: Effects of Transparency on GWG and Daily Wage (Complier Sample)

*The figure below plots the effects of transparency on the gender wage gap (panel (a)), and on male (panel (b)) and female (panel (c)) workers separately, for those firms which do not change their treatment assignment after 2013. The sample includes only establishments of firms with 75-225 employees in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




(a) Gender Wage Gap

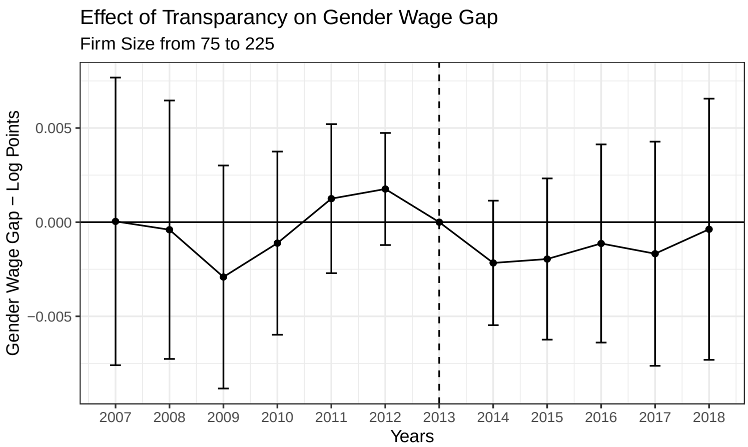

(b) Male Daily Wage




(c) Female Daily Wage

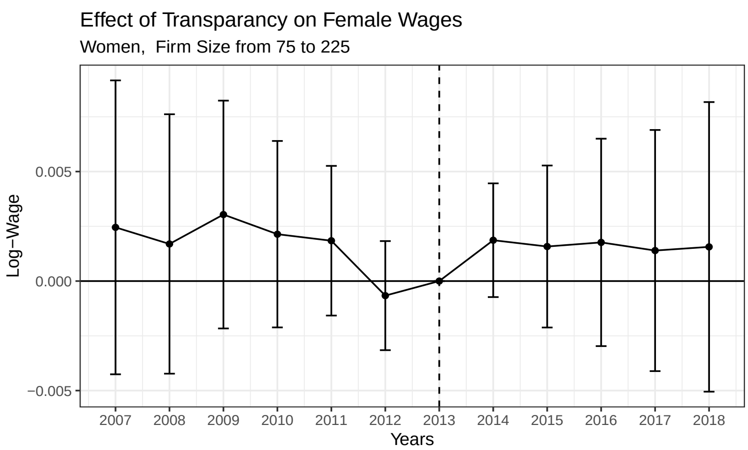
Figure 1.A10: Effects of Transparency on GWG and Daily Wage (Treatment Defined as of 2010)

*The figure below plots the effects of the transparency on gender wage gap (Panel (a)), and on daily wages for male (Panel (b)) and female (Panel (c)) workers separately. Treatment is assigned based on firm size in 2010, one year before the reform was announced. The rest is as specified in equation (1.1). Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




(a) Gender Wage Gap

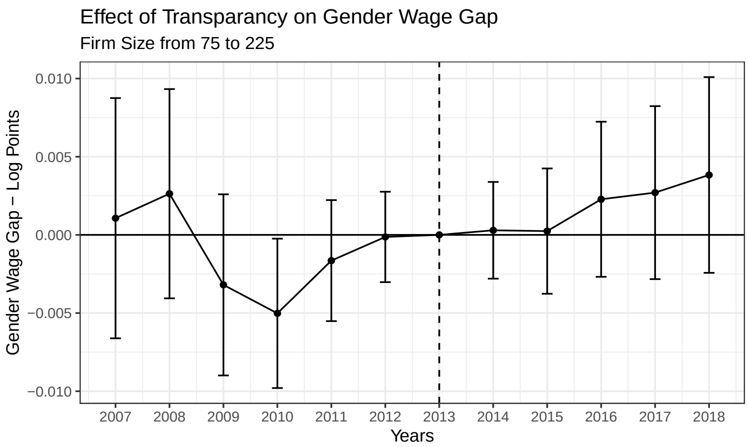

(b) Male Daily Wage




(c) Female Daily Wage

Figure 1.A11: Effects of Transparency on GWG and Daily Wage (Worker-level Treatment)

*The figure below plots the effects of transparency on the gender wage gap (Panel (a)), and on daily wages for male (Panel (b)) and female (Panel (c)) workers separately. Individuals are assigned to treatment status if they worked in an establishment whose firm size exceeded 150 employees in 2013, and to the control group otherwise. The rest is as specified in equation (1.1). Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




(a) Gender Wage Gap


(b) Male Daily Wage


(c) Female Daily Wage

Table 1.A2: Effects of Pay Transparency on Gender Wage Gap

| | *Dependent variable: ln(Daily Wage)* | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Male | 0.24*** | 0.32*** | | |
| | (0.003) | (0.004) | | |
| Male*Treat | 0.01 | 0.003 | -0.01* | |
| | (0.01) | (0.01) | (0.003) | |
| Male*Treat***1**[t=2007] | -0.01 | -0.01 | -0.002 | -0.001 |
| | (0.01) | (0.01) | (0.004) | (0.004) |
| Male*Treat***1**[t=2008] | -0.01 | -0.01 | -0.001 | 0.001 |
| | (0.01) | (0.01) | (0.003) | (0.004) |
| Male*Treat***1**[t=2009] | -0.01** | -0.01** | -0.01* | -0.01* |
| | (0.005) | (0.005) | (0.003) | (0.003) |
| Male*Treat***1**[t=2010] | -0.005 | -0.01 | -0.004* | -0.01** |
| | (0.004) | (0.004) | (0.002) | (0.002) |
| Male*Treat***1**[t=2011] | -0.004 | -0.005 | -0.002 | -0.003 |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| Male*Treat***1**[t=2012] | -0.002 | -0.002 | -0.0004 | -0.001 |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Male*Treat***1**[t=2013] | 0.00 | 0.00 | 0.00 | 0.00 |
| | - | - | - | - |
| Male*Treat***1**[t=2014] | -0.01** | -0.01** | 0.001 | 0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Male*Treat***1**[t=2015] | -0.01** | -0.01*** | 0.0002 | 0.001 |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| Male*Treat***1**[t=2016] | -0.01 | -0.01* | 0.002 | 0.002 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Male*Treat***1**[t=2017] | -0.001 | -0.002 | 0.004 | 0.003 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Male*Treat***1**[t=2018] | 0.002 | 0.001 | 0.004 | 0.003 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Treat∗**1**[t=2007] | 0.005 | 0.01 | 0.004 | 0.003 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Treat∗**1**[t=2008] | 0.003 | 0.004 | 0.002 | 0.001 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Treat∗**1**[t=2009] | 0.004 | 0.005 | 0.004 | 0.004 |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Treat∗**1**[t=2010] | 0.003 | 0.003 | 0.003 | 0.004* |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| Treat∗**1**[t=2011] | 0.01** | 0.01** | 0.003* | 0.003* |
| $\lambda_j$ | ✓ | ✓ | ✓ | |
| f(Age)$*\mathbb{I}^m$ | | ✓ | ✓ | ✓ |
| $\lambda_i$ | | | ✓ | |
| $\lambda_{ij}$ | | | | ✓ |

**Continued on next page**

**Table 1.A2 – continued from previous page**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | (0.003) | (0.002) | (0.002) | (0.002) |
| Treat∗**1**[t=2012] | 0.002 | 0.002 | 0.0001 | 0.0002 |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Treat∗**1**[t=2014] | 0.004** | 0.005** | -0.0004 | -0.0003 |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Treat∗**1**[t=2015] | 0.01*** | 0.01*** | 0.001 | 0.001 |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| Treat∗**1**[t=2016] | 0.01** | 0.01** | 0.001 | 0.0002 |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| Treat∗**1**[t=2017] | 0.002 | 0.003 | -0.002 | -0.001 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Treat∗**1**[t=2018] | -0.0001 | 0.001 | -0.001 | -0.001 |
| | (0.004) | (0.004) | (0.003) | (0.003) |
| Male∗**1**[t=2007] | 0.01*** | 0.01** | -0.04*** | -0.04*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Male∗**1**[t=2008] | 0.01*** | 0.01** | -0.03*** | -0.04*** |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| Male∗**1**[t=2009] | 0.001 | -0.001 | -0.04*** | -0.04*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Male∗**1**[t=2010] | 0.001 | -0.0001 | -0.03*** | -0.03*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Male∗**1**[t=2011] | 0.003 | 0.002 | -0.02*** | -0.02*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Male∗**1**[t=2012] | -0.002 | -0.002 | -0.01*** | -0.01*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Male∗**1**[t=2014] | 0.003* | 0.003* | 0.01*** | 0.01*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Male∗**1**[t=2015] | 0.003 | 0.002 | 0.01*** | 0.01*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Male∗**1**[t=2016] | 0.001 | 0.0001 | 0.02*** | 0.02*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Male∗**1**[t=2017] | -0.002 | -0.003 | 0.02*** | 0.02*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Male∗**1**[t=2018] | -0.003 | -0.01** | 0.03*** | 0.03*** |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| **1**[t=2007] | -0.04*** | -0.03*** | -0.05*** | -0.06*** |
| | (0.003) | (0.002) | (0.002) | (0.003) |
| **1**[t=2008] | -0.02*** | -0.01*** | -0.03*** | -0.03*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| $\lambda_j$ | ✓ | ✓ | ✓ | |
| f(Age)∗$\mathbb{I}^m$ | | ✓ | ✓ | ✓ |
| $\lambda_i$ | | | ✓ | |
| $\lambda_{ij}$ | | | | ✓ |

**Continued on next page**

**Table 1.A2 – continued from previous page**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **1**[t=2009] | -0.001 | 0.004** | -0.01*** | -0.01*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| **1**[t=2010] | -0.003** | 0.001 | -0.01*** | -0.01*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| **1**[t=2011] | -0.01*** | -0.01*** | -0.02*** | -0.02*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| **1**[t=2012] | -0.01*** | -0.01*** | -0.01*** | -0.01*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| **1**[t=2014] | 0.01*** | 0.01*** | 0.02*** | 0.0.02*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| **1**[t=2015] | 0.02*** | 0.02*** | 0.03*** | 0.0.03*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| **1**[t=2016] | 0.03*** | 0.02*** | 0.04*** | 0.0.04*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| **1**[t=2017] | 0.03*** | 0.03*** | 0.05*** | 0.05*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| **1**[t=2018] | 0.04*** | 0.03*** | 0.07*** | 0.07*** |
| | (0.002) | (0.002) | (0.002) | (0.003) |
| Age | | -0.04*** | | |
| | | (0.01) | | |
| AgeSq | | 0.73*** | 0.92*** | 1.05*** |
| | | (0.03) | (0.03) | (0.03) |
| AgeCu | | 1.62*** | 1.35*** | 1.18*** |
| | | (0.05) | (0.05) | (0.05) |
| AgeQuart | | -4.37*** | -3.99*** | -3.95*** |
| | | (0.10) | (0.09) | (0.09) |
| Male*Age | | 0.29*** | | |
| | | (0.01) | | |
| Male*AgeSq | | -1.58*** | -1.65*** | -1.74*** |
| | | (0.03) | (0.03) | (0.03) |
| Male*AgeCu | | -0.76*** | -0.60*** | -0.42*** |
| | | (0.05) | (0.05) | (0.05) |
| Male*AgeQuart | | 4.39*** | 3.69*** | 3.55*** |
| | | (0.11) | (0.09) | (0.10) |
| Observations | 4914038 | 4914038 | 4914038 | 4914038 |
| $R^2$ | 0.46 | 0.49 | 0.92 | 0.94 |
| Adjusted $R^2$ | 0.46 | 0.49 | 0.90 | 0.91 |

### 1.A.5 Analysis at the Establishment Level

In our main specification we estimate the effect of the Austrian pay transparency reform on individual (daily) wages. Here we present an alternative specification of our baseline model, in which we regress the gender pay gap of establishment $j$ in year $t$ ($GPG_{jt}$) on the interaction of the year indicator $\mathbf{1}[t = k]$ and the treatment indicator $Treat_{j(2013)}$. Thereby, we focus again on establishments of firms with 75-225 employees in 2013 and assign establishments with a firm size equal to or greater than 150 employees in 2013 to the treatment group:

$$GPG_{jt} = \sum_{k=2007}^{2018} \beta^k \mathbf{1}[t = k] * Treat_{j(2013)} + \lambda_j + \lambda_t + \epsilon_{jt}, \tag{1.5}$$

As in the baseline specification in equation (1.1), $\lambda_j$ and $\lambda_t$ denote the establishment and year fixed effects respectively. $\epsilon_{jt}$ denotes the idiosyncratic error term. As in the baseline specification, we drop the year 2013 from our estimation for $\beta^k$ and $\lambda_t$ due to collinearity concerns.

Figure 1.A12 plots the $\beta^k$ coefficients from estimating equation (1.5) for the establishments in our baseline sample. Overall, this analysis corroborates our baseline results: The Austrian pay transparency legislation had no discernible economic or statistically significant effect on the gender pay gap in treated establishments. Only in 2011 and 2012 we observe a small significant pre-trend in the gender pay gap. However, the gender pay gap is actually increasing rather than decreasing, such that we can rule out anticipation effects.

Figure 1.A12: Effect of Transparency on Establishment Level Gender Wage Gap

*The figure below plots the effects of pay transparency on the establishment-level gender wage gap using equation* (1.5). *The sample is restricted to establishments of firms with 75-225 employees in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% confidence intervals.*

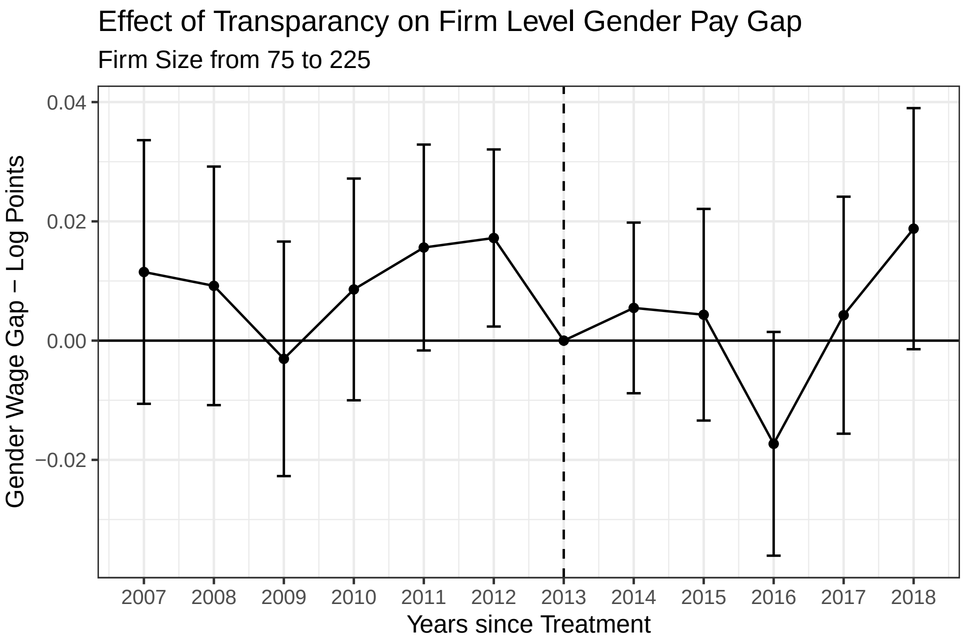
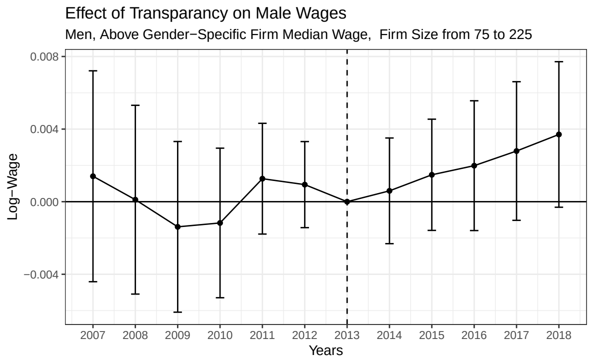


Figure 1.A13: Gender-Specific Effects of Transparency on Daily Wages

**[Above/Below Establishment-Level Gender-Specific Median Wage]**

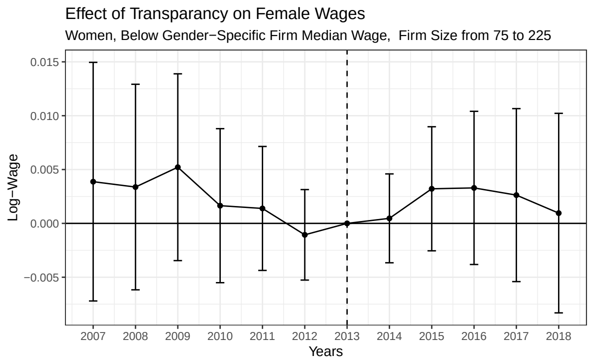
*The figure below plots the effects of transparency on male and female wages, for workers who earn above (top panels) and below (bottom panels) their gender-specific establishment-level median wage in 2013 (Eq. (1.1)), the year before treatment. Standard errors are clustered at the establishment level. The standard error spikes represent 95% CI.*




(a) Above Median Male


(b) Above Median Female

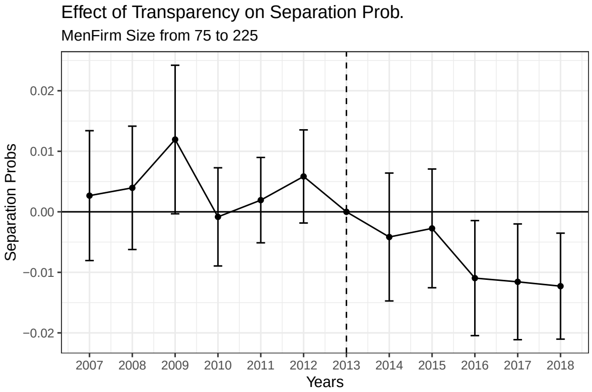

(c) Below Median Male



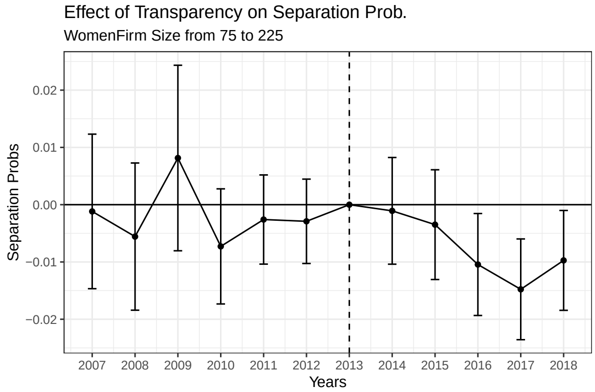

(d) Below Median Female

Figure 1.A14: Effects of Transparency on Job Separation Rate

*The figure below plots the effects of pay transparency on the year-on-year job separation rate for male and female workers (Eq. (1.4)). The sample is restricted to establishments of firms with 75-225 employees in 2013. Standard errors are clustered at the establishment level. The standard error spikes represent 95% confidence intervals.*




(a) Male




(b) Female

### 1.A.6 Gender Wage Gap Decomposition

We decompose the overall gender wage gap into a sorting component, which captures the fact that men and women work for different establishments, and a within establishment component that contains the gender wage gap originating from differences in pay policies towards men and women, as well as gender differences in other characteristics. Let's define the wage in a given year of worker $i$ with gender $g$ working at establishment $j(i)$ as $w^g_{i,j(i)}$. Subtracting and adding the respective female or male establishment average wage as shown in the following equation, allows us to decompose the gender wage gap into a sorting component and a within establishment component:

$$
\begin{aligned}
\frac{1}{N_M}\sum w^M_{i,j(i)} - \frac{1}{N_w}\sum w^W_{i,j(i)} =& \bar{w}^M - \frac{1}{N_W}\sum_{i=1}^{N_W}\left(\bar{w}^M_{j(i)} - (\bar{w}^M_{j(i)} - w^W_{i,j(i)})\right) \\
=& \underbrace{\bar{w}^M - \frac{1}{N_W}\sum_{i=1}^{N_W}\bar{w}^M_{j(i)}}_{\text{Sorting}} + \underbrace{\frac{1}{N_W}\sum_{i=1}^{N_W}(w^W_{i,j(i)} - \bar{w}^M_{j(i)})}_{\text{Within Establishment GPG}} \qquad (1.6) \\
=& \frac{1}{N_M}\sum_{i=1}^{N_M}\left(\bar{w}^W_{j(i)} + (w^M_{i,j(i)} - \bar{w}^W_{j(i)})\right) - \bar{w}^W \\
=& \underbrace{\frac{1}{N_M}\sum_{i=1}^{N_M}\bar{w}^W_{j(i)} - \bar{w}^W}_{\text{Sorting}} + \underbrace{\frac{1}{N_M}\sum_{i=1}^{N_W}(w^M_{i,j(i)} - \bar{w}^W_{j(i)})}_{\text{Within Establishment GPG}} \quad , \qquad (1.7)
\end{aligned}
$$

where $\bar{w}^W$ and $\bar{w}^M$ are average male and female wages, $w^W_{j(i)}$ and $w^M_{j(i)}$ is the average wage of females and male employees working at establishment $j(i)$. Table 1.A3 reports the findings of this decomposition for all treated firms pooled over all pre-treatment periods.

Table 1.A3: Decomposition Gender Wage Gap

*The sample is restricted to establishments of firms with 75-225 employees in 2013 and includes years before treatment (2007-2013).*

| | Gender Wage Gap | Sorting | Within Establishment |
|---|---|---|---|
| Decomposition (female dist. eq. (1.6)) | 0.358 | 0.108 | 0.250 |
| Decomposition (male dist. (1.7))) | 0.358 | 0.086 | 0.272 |

# Chapter 2

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

## 2.1 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, Gielen, and Luttmer, 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,

[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.

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 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 ap-

[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.

proximation 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 2.3.1): 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

[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.

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 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, Staubli, and Zweimüller (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, Bound, and Nichols, 2003) and sufficient statistic (Diamond and Sheshinski, 1995; Haller, Staubli, and Zweimüller, 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, Duggan, and Gruber (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, Seitz, and Siegloch (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, Moore, and Strand (2017), Meyer and Mok (2019), Mullen and Staubli (2016), and Ruh and Staubli (2019)), and the productivity of (rejected) claimants (e.g. Borghans, Gielen, and Luttmer (2014), Bound (1989), French and Song (2014), Kostol and Mogstad (2014), and Wachter, Song, and Manchester (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

## 2.2 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 2.E.4). The average replacement ratio of public DI amounts to 35% of past gross income (see Table 2.5.1), while the public DI schedule is progressive. The actual replacement ratios are greater/lower for very low/high in-

[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.

comes 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 2.5.1).

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 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

[7]Both of these features are included in the model in section 2.3.

[8]See Seibold, Seitz, and Siegloch (2021) for a discussion of priced (risk groups) vs. non-priced risk in private DI.

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, Duggan, and Gruber, 2014).

## 2.3 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.

### 2.3.1 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

[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.

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 ($l_{it} = 0$), 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:

$$\max_{\{c_k, A_{k+1}, l_k\}_{k=1}^{T}, pDI_0} V_{i0} = \sum_{t=0}^{T} \beta^t \mathbf{E}[U(X_{it}; S_{it})] \tag{2.1}$$

where $\beta$ 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 2.3.2), rejections from public insurance (section 2.3.3), and income risk (section 2.3.4). 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:

$$\begin{aligned} \frac{A_{it+1}}{1+r} + c_{it} + pDI_{it} * p_{it} &= A_{it} + y_{it} + y^s_{it} - SSC(y_{it}) - SSC(y^s_{it}) - TAX(y_{it} + y^s_{it}) \\ l_{it} &= L - hours_{it} - \theta \mathbf{1}[hours_{it} > 0] \\ A_{it} &\geq 0 \end{aligned} \tag{2.2}$$

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 2.E).

[10] As discussed in section 2.2, 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)).

Social security contributions are paid individually, while household income is taxed jointly. I describe the income process in section 2.3.4.

The second constraint in (2.2) 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 $\theta$ 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 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:

$$U(c_t, l_t; H_t) = \frac{(c_{it}^{\kappa} l_{it}^{1-\kappa} e^{-\varphi * \mathbf{1}[H_t = bad]})^{1-\gamma}}{1-\gamma} \tag{2.3}$$

where $\gamma$ denotes risk aversion, $\kappa$ the weight on consumption relative to leisure, and $\varphi$ expresses the (dis-)utility from disability (Low and Pistaferri, 2015). Intuitively, $\varphi$ informs us about how individuals would move consumption across health states if fully insured. A positive value of $\varphi$ 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 $\gamma$, $\kappa$, $\varphi$, and $\theta$ (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 2.A.

### 2.3.2 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 2.3.4 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 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:

$$disabled_{it} = \Phi(\zeta_0 + \zeta_1 * rg_{it}) \tag{2.4}$$

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:

$$\Pi_{it}(H_{t+1}; H_t, rg_i) = \pi(H_{t+1} = J | H_t = j) * \frac{dis\hat{a}bled(rg_{it})}{dis\hat{a}bled(rg = 3)} \tag{2.5}$$

for $J, j = good, disabled$.

Similar to Low and Pistaferri (2015), the process described in equation (2.5) 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 hetero-

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

[12] See section 2.2 for details on the risk group assignment.

geneity 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, Seitz, and Siegloch (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.

### 2.3.3 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 2.2. 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:

$$p_{it}(rg_i) = ppE(rg_i)RR^{private}Y_{it}(H_{it} = good). \tag{2.6}$$

[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.

[14]Recall from section 2.2 that prices are linear in the benefit level and thus can be expressed as a 'price-per-Euro'.

Private benefits are defined as a constant fraction $RR^{private}$ of full-time income in good health $Y_{it}(H_{it} = good)$. Subsection 2.4.2 and 2.5.2 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 2.7.3.

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 2.5.1):

$$benefits^{public} = RR^{public} * Y_{it}(H_{it} = good). \tag{2.7}$$

Equation (2.11) 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 average). As detailed in section 2.2, reasons for rejections are the failure to meet the minimum contribution period or the minimum health requirement, requiring a 62.5%

[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.

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.

### 2.3.4 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:

$$\log Y_{it} = \beta_0 + \sum_{k=1}^{4} \beta_k * age_{it}^k + \beta_5 \mathbf{1}[hours_{it} = 1] + \beta_6 * \mathbf{1}[rg_{it} = rg]) + \varepsilon_{it} + \epsilon_{it}. \tag{2.8}$$

$Y_{it}$ 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 $\mathbf{1}[rg_{it} = rg]$). $\varepsilon_{it}$ denotes a persistent shock process of income innovations following an AR(1) process (Guvenen, 2009; Low, Meghir, and Pistaferri, 2010):

$$\varepsilon_{it} = \rho \varepsilon_{it-1} + \eta_{it} \tag{2.9}$$

[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.

where $\eta_{it} \sim N(0, \sigma_\eta^2)$ and $\rho$ 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 $\epsilon_{it}$ 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 $\sigma_\epsilon^2$. The parameters governing the shock processes, $\{\rho, \sigma_\eta^2, \sigma_\epsilon^2\}$, are estimated directly from the data as described in section 2.5.2.

Controlling for the risk group in the income process is important to link income to risk and selection into private DI coverage (Seibold, Seitz, and Siegloch, 2021), which has relevant welfare effects as shown in section 2.7. 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 (2.8) 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 (2.8).[19]

Spousal income $y_{it}^s$ 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:

$$y_{it}^s = \beta_0^s + \sum_{k=1}^{4} \beta_k^s age_{it}^k + \beta_5^s log(Y_{it}) \tag{2.10}$$

The specification implies that spousal income is independent of their partner's health status for the reasons mentioned in section 2.3.1.

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 2.3.3, both benefits replace a given fraction of the

[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.

full-time labor income in good health, and can be simultaneously claimed:

$$B_{it} = Y_{it} * RR^{public} * \mathbf{1}[DIs_{it} = Admitted] + Y_{it} * RR^{private} * \mathbf{1}[pDI_0 = 1] \tag{2.11}$$

$B_{it}$ 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:

$$y_{it} = SSI \text{ if } \{0, B_{it}\} + y^s_{it} \leq SSI \;\; \& \;\; A_{it} \leq \bar{A} \tag{2.12}$$

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 2.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].

### 2.3.5 Why do not all people buy private DI?

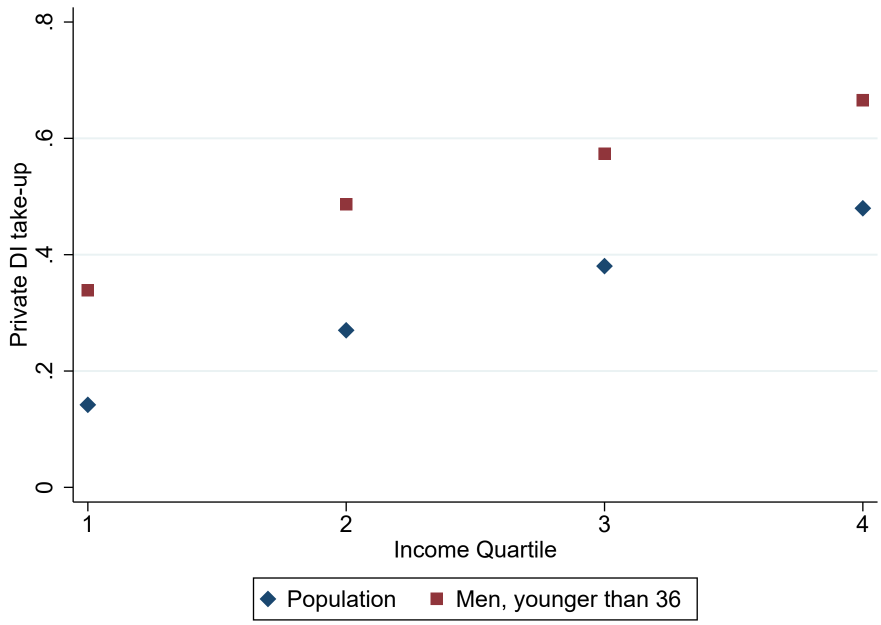
Basic economic theory predicts that risk averse individuals should always fully insure themselves if insurance is fair and no other frictions exist. However, as stated in the introduction, only about 34% (50%) of all people (men below 35) in Germany purchase private DI (EVS 2013). Since I estimate my 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 neg-

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

Figure 2.3.1: 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.*




ative correlation between income and risk implies that these high risk individuals have to pay a larger share of their income for insuring €1. Figure 2.3.1 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]

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

## 2.4 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.

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 2.B contains a detailed description of the data set construction.

### 2.4.1 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 2.A1 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, Seitz, and Siegloch (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, Dustmann, and Stevens (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 2.5 provides more information on the estimation procedure.

### 2.4.2 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 2.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 2.C. This allows me later to export the risk group - occupation mapping to the social security records (SIAB below). Moreover, it allows me to con-

[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, Seitz, and Siegloch (2021).

Table 2.4.1: 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 |

struct 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 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 2.B.2). As the premium is linear in benefit given the risk group (see Section 2.2), I recover the actual premium by multiplying the web-scrapped prices for insuring a Euro with the reported benefits from the data.

Table 2.4.1 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 2.A7 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.

### 2.4.3 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

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

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 (2.8), the labor market moments, the disability probability by risk group, and the population risk group distribution (see section 2.5).

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 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 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 2.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 individ-

[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.

[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.

uals 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 2.A9 presents the summary statistics and how the sample selection criteria affect the sample composition.

## 2.5 Estimation

I estimate the model described in section 2.3 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.

### 2.5.1 Values from the literature

Table 2.5.1 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 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 $\beta$ 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 2.E for details). Household income is assessed jointly based on the income tax schedule in Appendix 2.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.

[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.

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

Table 2.5.1: 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 | - |
| -$\beta$ | 0.987 | - |
| *Tax schedule and social security contributions* | | |
| -Income tax schedule | appendix 2.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 2.A11 | German Acturian Society |
| - Mortality risk | appendix table 2.A11 | 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 |

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 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

[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.

€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, Seitz, and Siegloch, 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 2.2).

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 2.A11 presents the respective probability and mortality probabilities.

The last panel of table 2.5.1 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.

### 2.5.2 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 (2.8) in a reduced-form fashion outside the model. The construction of the respective estimation samples is detailed in section 2.4. If not stated otherwise, the sample window always runs from 2001 to 2017. Table 2.5.2 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:

[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 2.5.2: 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.* (2.4) *(sample window: 2001-2017). Panel C reports the results from estimating the income equation* (2.8) *(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. (2.4) |
| $Prob(disabled(rg = 2))$ | $4.227 * 10^{-4}$ | |
| $Prob(disabled(rg = 3))$ | $6.476 * 10^{-4}$ | |
| $Prob(disabled(rg = 4))$ | $9.787 * 10^{-4}$ | |
| $Prob(disabled(rg = 5))$ | $14.592 * 10^{-4}$ | |
| *Num. Obs.* | 4,696,325 | |
| *Panel C: Income Process* | | |
| $\beta_0$ | 0.7730 | SIAB, eq. (2.8) |
| $\beta_1$ (age) | 0.0405 | |
| $\beta_2$ ($age^2$) | -0.0015 | |
| $\beta_3$ ($age^3$) | 2.46*$10^{-5}$ | |
| $\beta_4$ ($age^4$) | -1.91*$10^{-7}$ | |
| $\beta_5$ (full-time) | 0.7921 | |
| $\beta_6^k$ (risk group): | | |
| 2 | -0.2035 | |
| 3 | -0.5412 | |
| 4 | -0.7253 | |
| 5 | -0.7558 | |
| $\sigma_\eta^2$ | 0.0192 | |
| $\sigma_\zeta^2$ | 0.1265 | |
| $\sigma_\epsilon^2$ | 0.0404 | |
| $\rho$ | 0.9459 | |
| *Num. Obs.* | 5,143,326 | |
| Replacement ratio | 0.36 | contract data |

28.9% have a job falling into risk group 3, while the large majority (45.3%) work in high risk jobs. 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 (2.4). 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 (2.5) 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 (2.8) 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 2.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, Seitz, and Siegloch (2021) and Figure 2.3.1). 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 2.A14 presents the parameter estimates.

### 2.5.3 Method of Simulated Moments Approach

I estimate the four preference parameters of interest, risk aversion $\gamma$, consumption weight $\kappa$, (dis-)utility from disability $\varphi$, and the fixed cost of labor force participation $\theta$, 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 2.A.3 provides a more formal description of this method.

The fundamental model parameters are estimated based on the moments presented in table 2.5.3, which can be distinguished into three sets of moments: private DI shares, labor supply, and savings rates and assets. Appendix table 2.A13 presents the each data moment and its weight.

Table 2.5.3: 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 2.A13 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 |

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, Seitz, and Siegloch (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 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

[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.

mean and median assets are estimated on the pooled EVS estimation sample described in section 2.4.1. 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 $\gamma$ determines the consumption smoothing motive across time and states: A greater value of $\gamma$ 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 $\kappa$. Thus, the variation in leisure (full-time, part-time, no participation) helps identifying $\kappa$. The fixed cost of labor force participation $\theta$ 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, $\varphi$, the (dis-)utility from bad health, governs how people want to move consumption across health states (insurance motive). A greater value of $\varphi$ raises the value of an additional 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.

## 2.6 Results

This section presents the estimation results of the preference parameters from the model in section 2.3. 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.

### 2.6.1 Estimation Results

Table 2.6.1 presents the estimation results. The second column displays the parameter estimates derived from the method of simulated moments and the third column shows

Table 2.6.1: 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 2.5.3.*

| **Parameter** | **Value** | **Standard Error** | **Identification** |
|---|---|---|---|
| Risk aversion $\gamma$ | 6.232 | 0.453 | Mean and median assets |
| Consumption weight $\kappa$ | 0.495 | 0.003 | full-time and part-time shares, LF participation |
| Labor force participation cost $\theta$ | 0.161 | 0.01 | LF participation, full-time and part-time shares |
| Disutility from bad health $\varphi$ | 0.154 | 0.001 | mean private DI, mean and median assets |

the corresponding standard errors for each parameter.[35] The fourth column reports which moment identifies which parameter (see section 2.5.3). Overall, the parameter estimates are in line with the related literature and precisely estimated. The coefficient of relative risk aversion $\gamma$ 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 $\gamma$ lies at the upper end of this interval. The standard error in the third column of table 2.6.1 shows that $\gamma$ is precisely estimated.

The consumption weight $\kappa$ 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 $\kappa$ 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

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

error). Recall that a positive $\varphi$ implies that disability is a "bad" given the utility function in eq. (2.3), 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 2.A14 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 2.A15 reports the sensitivity of each parameter with respect to the different moments following Andrews, Gentzkow, and Shapiro (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.

### 2.6.2 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 leaves me confident about its performance.

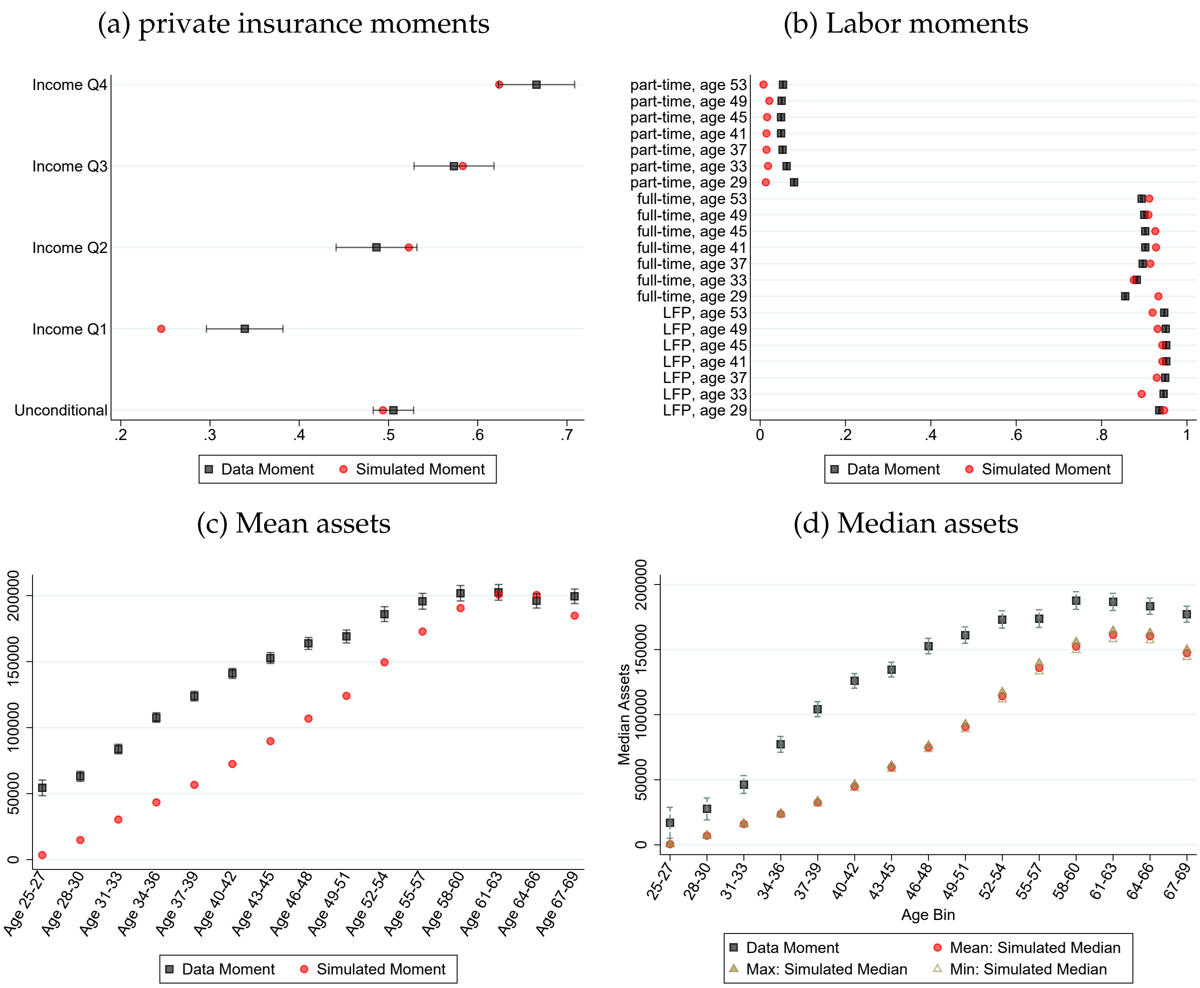
Figure 2.6.1 shows the fit between targeted (black) and simulated (red) moments.

[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 2.6.1: 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.*




(a) private insurance moments
(b) Labor moments
(c) Mean assets
(d) Median assets

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 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.

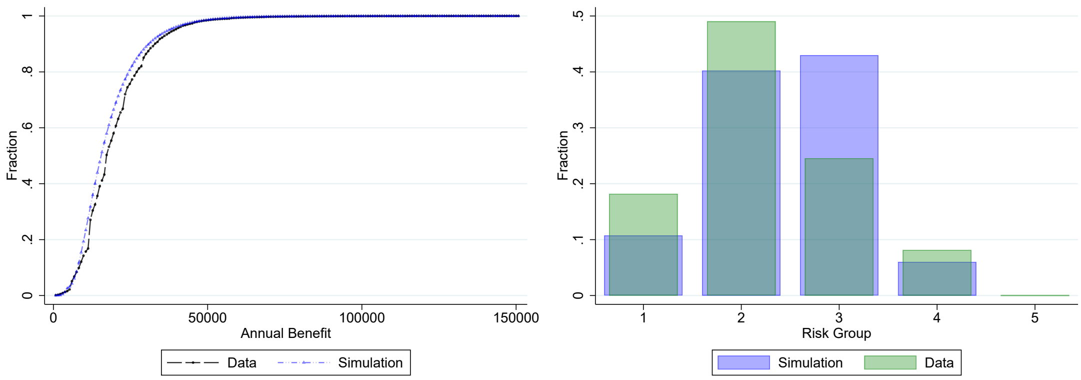
Besides closely fitting targeted moments, the model closely matches moments which were not explicitly targeted in the estimation. Figure 2.6.2 shows two private market moments conditional on private DI coverage. Panel (a) plots the cumulative distribution function (CDF) of private DI benefits. Despite only offering a single contract with the average replacement ratio of 36%, the model produces a benefit CDF (dotted line) which closely 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).

[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.

Figure 2.6.2: 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 2.A1 shows additional out-of-sample fit graphs.*




(a) CDF of Private Benefits (b) Risk group distribution

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, Seitz, and Siegloch (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 2.7), 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 2.A1 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.

## 2.7 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?

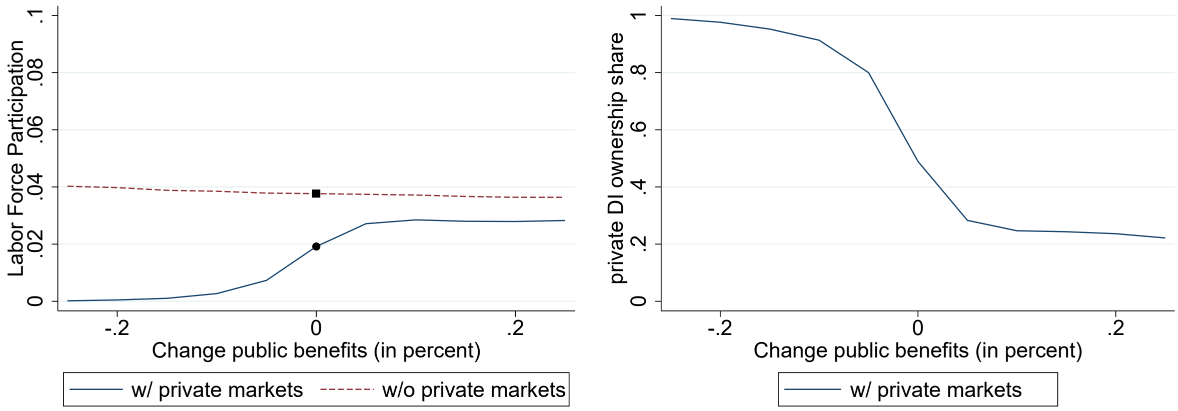
To answer these questions and evaluate welfare, I first quantify the behavioral responses to public and private DI coverage. The key question is to determine how private DI coverage distorts the labor supply of people eligible for DI benefits and how selection into private DI coverage varies with the 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 2.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.

Figure 2.7.1: 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.*




(a) Labor force participation (b) private DI take-up shares

### 2.7.1 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 2.7.2.

**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 2.7.1 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.

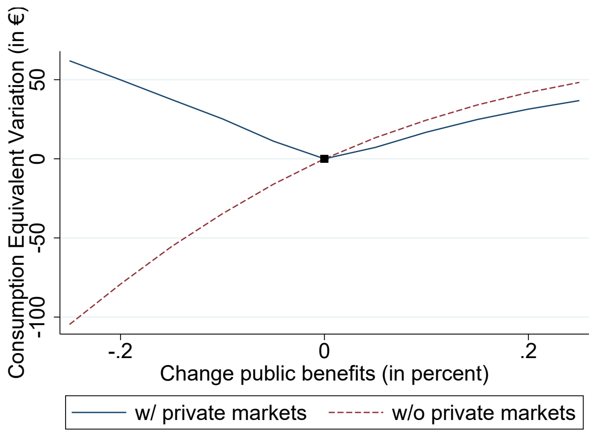
I find that private DI coverage reduces the LFP of the disabled across all considered benefit changes, which is captured by the gap between the solid and dashed line in Panel (a). For instance, at baseline (in black) the LFP of the disabled is reduced by 50% with private DI coverage. The additional moral 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 2.A2 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 2.A2) 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 2.7.2 plots the welfare gains under the alternative benefit generosity. Without a private DI market (dashed line), welfare unambiguously increases in benefit generosity, such that the increase in insurance value offsets the additional cost

Figure 2.7.2: 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.*

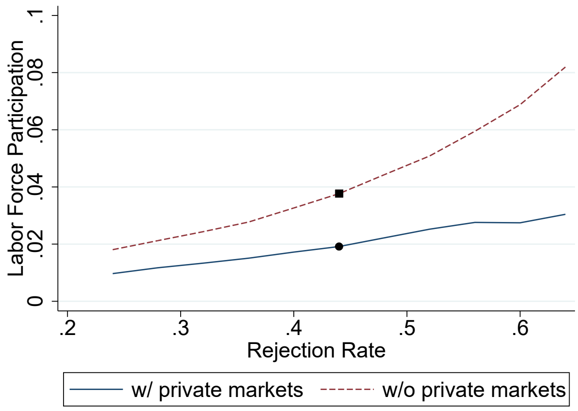



from more generous benefits. In contrast, benefit reductions lead to larger welfare gains when a private DI market exists (solid line). For lower benefit generosity, the selection into private DI on income weakens, so the moral hazard response of the marginal buyer is decreasing, as they are less likely to continue working even without private DI coverage (see Panel (e) in Appendix Figure 2.A2). 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.

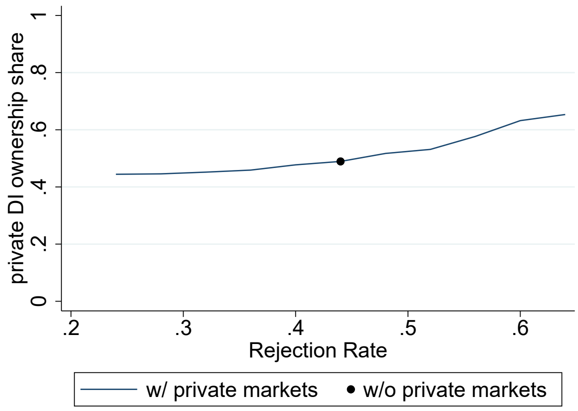
[39]The increase in welfare for higher benefit generosity is driven by the higher public insurance value. Yet, fewer people are covered by private DI, such that their total insurance value drops and the welfare gains in this group are smaller. Finally, the people still covered by private DI show the 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 2.7.3: 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.*




(a) Labor force participation




(b) private DI ownership shares

### Alternative Rejection Rates

Figure 2.7.3 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 2.A3 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.

The documented behavior has the following welfare implications summarized in Figure 2.7.4. 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.

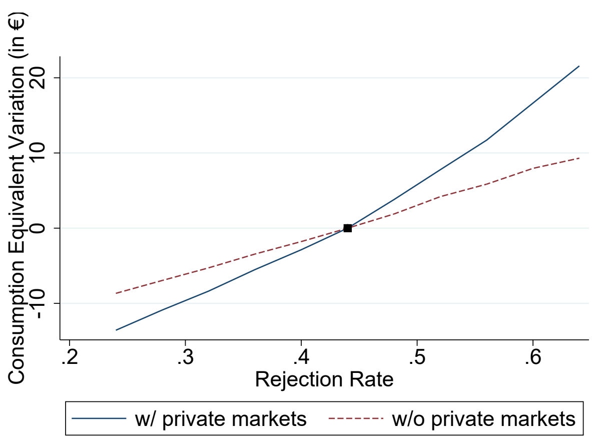
### 2.7.2 Welfare-Effects of private markets

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 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

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

Figure 2.7.4: 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.*




are computed under revenue-neutrality by the means of a lump-sum tax.

Figure 2.7.5 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.

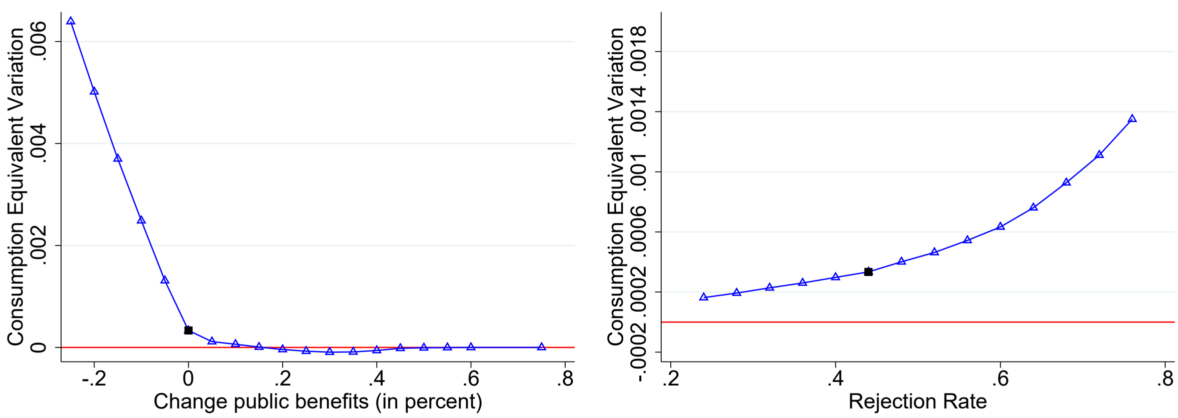
I find that under the current public schedule (in black), having a private DI market enhances welfare. As Panel (a) shows, having a private DI market is welfare-improving for less generous benefits, whereas it becomes welfare-reducing for more generous public DI benefits, before becoming zero at high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

[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.

[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.

Figure 2.7.5: 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.*




(a) Changes in benefit generosity

(b) Changes in rejection rates

viduals. 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.

[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.

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 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 2.7.5 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. in-

[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.

[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, Duggan, and Gruber (2014) and the BLS (Labor Statistics,

Table 2.7.1: 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) |

come distribution, disability risk, risk group distributions,...). I then compute the CEV for having a private market. The fourth column in Table 2.7.1 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) 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.

2020) for the USA and from Haller, Staubli, and Zweimüller (2020) and Kaniovski and Url (2019) for Austria.

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 2.7.5, 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.

### 2.7.3 Robustness Exercises

The results above are derived under the baseline specifications and assumptions. In Appendix 2.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 2.A14 presents the corresponding parameter estimates. Appendix Figures 2.A4 and 2.A5 show the welfare effects of changing the benefit generosity and the rejection rate respectively. Appendix Figures 2.A6 (benefit generosity) and 2.A7 (rejection rates) show the corresponding behavioral responses. Finally, Appendix Figure 2.A8 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.

## 2.8 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).

# Appendix

## 2.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.

### 2.A.1 Solution

The model needs to be solved numerically as no analytical solution to the problem described in section 2.3 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. (2.1) in $T = 60$ simplifies to the following problem because death occurs with certainty in the next period leaving the individual with zero utility:

$$V(S_T) = \max_{c_T, A_{T+1}} U(c_T, M) \tag{B1}$$

where $S_T$ is the set of state variables at time $T$. Since the per-period utility function $U(\circ)$ is given (eq. 2.3), 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 $T-1$:

$$V(S_{T-1}) = \max_{c_{T-1}, A_T} U(c_{T-1}, M) + s_{T-1} * \beta * \hat{V}(S_T | S_{T-1}) + (1 - s_{T-1}) * 0 \tag{B2}$$

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 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]

[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.

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

**(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.

### 2.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, Heckman, and Mosso (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 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. (2.8)) 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).

### 2.A.3 Estimation of preference parameters

The preference parameters of interest risk aversion $\gamma$, consumption weight $\kappa$, (dis-) utility from bad health $\varphi$, and labor force participation cost $\theta$ 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 $\mathbf{G}$ denote the difference between the data moments and the simulated moments:

$$G(\gamma, \kappa, \varphi, \theta) = \Sigma^{data}(\gamma_0, \kappa_0, \varphi_0, \theta_0) - \Sigma^{sim}(\gamma, \kappa, \varphi, \theta) \tag{B3}$$

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):

$$\begin{aligned} &M_t - \mathbf{E}[\hat{M}_t(\gamma, \kappa, \varphi, \theta)] = 0 \\ &0.5 - \mathbf{E}[\mathbf{1}[A_{ia} \leq median(\hat{A_{ia}}(\gamma, \kappa, \varphi, \theta))] = 0 \end{aligned} \tag{B4}$$

$A_{it}$ 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:

$$\min_{\gamma,\kappa,\varphi,\theta} G(\gamma, \kappa, \varphi, \theta)' W G(\gamma, \kappa, \varphi, \theta) \tag{B5}$$

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, Hendren, and Luttmer (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.

## 2.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 2.4).

### 2.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

[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.

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 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 2.A1 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 2.A1 shows, private DI coverage increased greatly among the cohorts who lost their coverage in 2001. Seibold, Seitz, and Siegloch (2021) study the effects of the reform on the private DI market in a related paper.

Table 2.A1: 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 |

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 top and bottom 1% of the household net income and asset distribution following Adda, Dustmann, and Stevens (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]).

### 2.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

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

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 2.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 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 2.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 2.A2 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),

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

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 webscrape 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 2.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 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-

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

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, Seitz, and Siegloch (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

### 2.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 (2.8), the labor market moments, the disability probability by risk group, and the population risk group distribution (see section 2.5).

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 2.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.

[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.

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. Appendix table 2.A9 presents the summary statistics and how the sample selection criteria affect the sample composition.

## 2.C Appendix: Occupation Code Assignment

As explained in appendix 2.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 2.A2 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 2.A2 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

[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 2.A2: 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.*

2.A7 in section 2.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.

## 2.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

[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".

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 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

[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.

Table 2.A3: 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.*

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 2.A3 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 2.A4 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

Table 2.A4: 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 2.A3.*

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 matched based on the "Line-by-Line" matching in section 2.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.

### 2.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 2.A7).

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 2.A5: 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 2.A5 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 2.A6 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 2.A5, 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

Table 2.A6: 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.*

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]

Finally, table 2.A7 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]

[59] For example, "Steuerassistent/Steuerfachgehilfe" are one occupation group in the insurance data but 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 2.A7: 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 2.A5.*

### 2.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, Seitz, and Siegloch (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 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

[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.

- 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
- 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 2.A8. 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

Table 2.A8: 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 |

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 2.A8 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.

## 2.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.

### 2.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 2.A9 presents some summary statistics of interest for the whole sample and after the imposing the sample selection criteria discussed in this section.

Table 2.A9: 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 |

### 2.D.2 Merging Risk Groups to Public Data

As discussed in appendix 2.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

[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 2.8

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 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.

## 2.E Appendix: German Institutional Setting

### 2.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 2.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.

### 2.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.

### 2.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 2.A10 depicts these fraction of annuity income that is subject to income taxation for different remaining terms.

Table 2.A10: Taxable Fraction of Annuity income

*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.*

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

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).

### 2.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]

$$hypPP_i = (\frac{1}{(T^k - T^0) * 12} * \sum_{j=0}^{T^k - T^0} actPP_{ij}) * (62 * 12 - T^k * 12) \tag{B6}$$

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:

$$Disc_{it} = 1 - \min\{0.108, (63 * 12 + 7 - T^k * 12)\} \geq 0.892 \ \forall t \tag{B7}$$

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.

$$DIb_{it} = (\sum_{j=0}^{T^k - T^0} actPP_{ij} + hypPP_i) * Disc_{it} * PensVal_{it} * HM_{it} \tag{B8}$$

**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 (B8).

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.

## 2.F Appendix: Health transition probabilities and mortality risk

Table 2.A11 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 2.A11: Mortality Risk for men in Germany, observation period 2011-2013

*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*

| 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 | | |

*continued*

*Table 2.A11 continued*

| Age | Mortality Risk | DAV1997 disability | DAV1997 recovery | DAV2018 disability | DAV2018 recovery |
|---|---|---|---|---|---|
| 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 | | | | |
| 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 | | | | |

## 2.G Appendix: Estimation of stochastic earnings components

The earnings process described in equation (2.8) is governed by two i.i.d. stochastic processes: An AR(1) persistent shock ($\varepsilon_{it}$) and a transitory income shock ($\epsilon_{it}$). The AR(1) process depends on the persistence term $\rho$ and the innovation variance $\sigma_\eta^2$ and the transitory shock only on its innovation variance $\sigma_\epsilon^2$.

I estimate these terms using the methods detailed in Guvenen (2009) and Low, Meghir, and Pistaferri (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 (2.8) on the SIAB. Their theoretical counterparts are the variance-covariance matrix ($\Sigma$) 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 $\varepsilon_{it}$ and $\epsilon_{it}$ are i.i.d., the variance and covariance of $u_{it}$ is then defined as (dropping the $i$ index for clarity of presentation):

$$var(u_t) = var(\varepsilon_t) + \sigma_\epsilon^2 \tag{B9}$$

$$\begin{aligned} cov(u_t, u_{t+j}) &= cov(\varepsilon_t + \epsilon_t, \varepsilon_{t+j} + \epsilon_{t+j}) \\ &= cov(\varepsilon_t, \varepsilon_{t+j}) + cov(\epsilon_t, \epsilon_{t+j}) \end{aligned} \tag{B10}$$

Given the transitory nature of $\epsilon_t$, $cov(\epsilon_t, \epsilon_{t+j}) = 0, \forall j > 0$ 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 $\rho$ (I define them recursively later):

$$var(\varepsilon_t) = \rho^2 var(\varepsilon_{t-1}) + \sigma_\eta^2 \tag{B11}$$

$$cov(\varepsilon_t, \varepsilon_{t+j}) = \rho cov(\varepsilon_t, \varepsilon_{t+j-1}) \tag{B12}$$

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 (B9) and (B10) 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 $\Sigma$ 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:

$$\mathbf{G}(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2) = \mathbf{vec}(\Sigma)(\rho, \sigma_\zeta^2, \sigma_\varepsilon^2, \sigma_\epsilon^2) - \mathbf{vec}(\hat{\Sigma}) \tag{B13}$$

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

$$\min_{\rho,\sigma_\zeta^2,\sigma_\varepsilon^2,\sigma_\epsilon^2} \mathbf{G}(\rho,\sigma_\zeta^2,\sigma_\varepsilon^2,\sigma_\epsilon^2)'\mathbf{W}\mathbf{G}(\rho,\sigma_\zeta^2,\sigma_\varepsilon^2,\sigma_\epsilon^2) \tag{B14}$$

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 2.5.2.

## 2.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:

$$LS = \frac{\hat{R} - R_0}{N_s} * \frac{(1+r)^{T_{retire}} * r}{(1+r)^{T_{retire}} - 1} \tag{B15}$$

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 ($\alpha$) 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 ($\hat{V}$):

$$\hat{V}((1-\alpha)c, l) = V_0(c, l) \tag{B16}$$

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

$$\alpha = 1 - (\frac{V_0(c, l)}{\hat{V}(c, l)})^{\frac{1}{\kappa * (1-\gamma)}} \tag{B17}$$

Since my estimation for $\gamma > 1$, I use this formula to compute $\alpha$. Thereby, a value for $\alpha$ 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,

[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.

negative values of $\alpha$ 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. (B17) this means $V_0$ ($\hat{V}$) corresponds to the expected utility without (with) a private DI market. Thus, $\alpha$ 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 2.7*

# 2.I Appendix: Additional Tables and Graphs

Table 2.A12: 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 2.5.1) per € insured |

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 2.A13: 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 |

*continued*

*Table 2.A13 continued*

| Moment | Data | Variance | Simulation | Standard Error |
|---|---|---|---|---|
| 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 |
| 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 |

*continued*

*Table 2.A13 continued*

| **Moment** | **Data** | **Variance** | **Simulation** | Standard Error |
|---|---|---|---|---|
| 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 | - |
| 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 2.A14: 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 $\gamma$ | 6.232 | 6.020 | 4.997 | 6.334 |
| Consumption weight $\kappa$ | 0.495 | 0.511 | 0.552 | 0.498 |
| Labor force participation cost $\theta$ | 0.161 | 0.230 | 0.372 | 0.202 |
| Disutility from bad health $\varphi$ | 0.154 | 0.160 | 0.131 | 0.159 |

Table 2.A15: 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, Gentzkow, and Shapiro (2017). The values shown document how mis-measuring a given moment would alter the parameter estimate by $\delta = value \times$ measurement error, such that the correct value would be $parameter + \delta$. 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** | $\gamma$ | $\kappa$ | $pc$ | $hc$ |
|---|---|---|---|---|
| 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 |

*continued*

*Table 2.A15 continued*

| **Moment** | $\gamma$ | $\kappa$ | $pc$ | $hc$ |
|---|---|---|---|---|
| PT2 | -1.97828 | -.0163166 | -.0419482 | .0080747 |
| LFP3 | -.577453 | .0419904 | -.0285609 | -.0177709 |
| FT3 | 10.7244 | .0515493 | .258681 | -.0223642 |
| 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 |

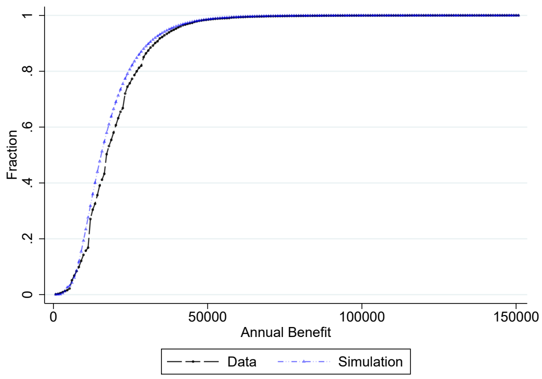
*continued*

*Table 2.A15 continued*

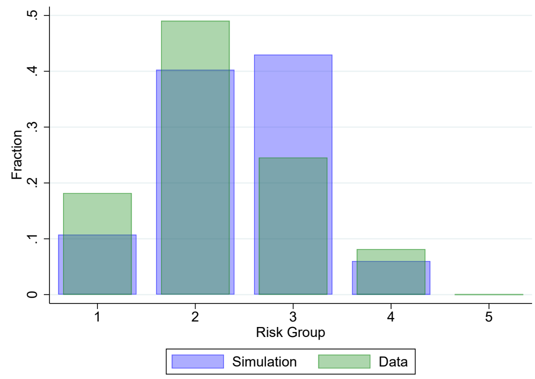
| **Moment** | $\gamma$ | $\kappa$ | $pc$ | $hc$ |
|---|---|---|---|---|
| Median4 | .42126 | .0011197 | .0103141 | -.0005393 |
| Median5 | .576927 | .0014818 | .0141623 | -.0007113 |
| Median6 | .408869 | .0010198 | .0100964 | -.0004764 |
| Median7 | .536233 | .001315 | .0132402 | -.0006175 |
| Median8 | .209686 | .0005916 | .0051118 | -.0002867 |
| Median9 | .223725 | .000638 | .0054306 | -.0003152 |
| Median10 | .209561 | .0005898 | .0051282 | -.0002799 |
| Median11 | .354117 | .0009338 | .0087256 | -.0004342 |
| Median12 | .339492 | .0009241 | .0083474 | -.0004305 |
| Median13 | .412423 | .0011592 | .0101075 | -.0005438 |
| Median14 | .299978 | .0007982 | .0073854 | -.00037 |
| Median15 | .610452 | .0016758 | .0149768 | -.0007854 |

Figure 2.A1: Out-of-sample fit of model

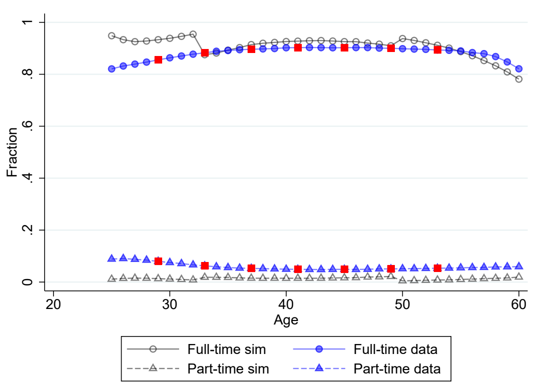
*The figure below presents the out-of-sample fit of simulated and data moments not targeted in the estimation. The data moments are estimated on the sample of employed men who are at least 25 years of age. Panel (a) shows the cumulative distribution of private DI benefits in the model(blue) and the data (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).*



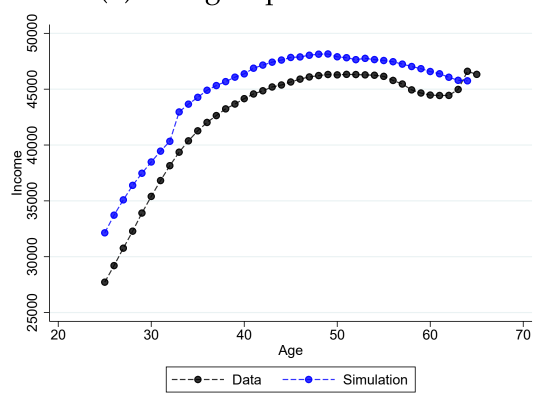

(a) CDF of Private Benefits




(b) Risk group distribution



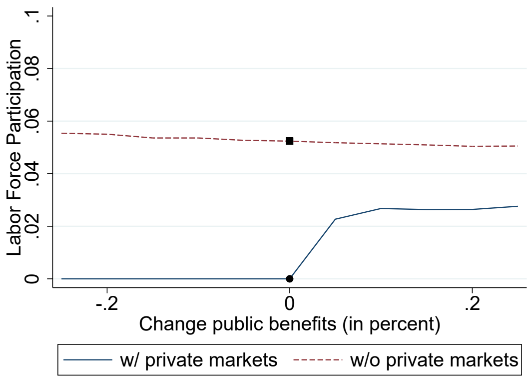

(c) Labor moments



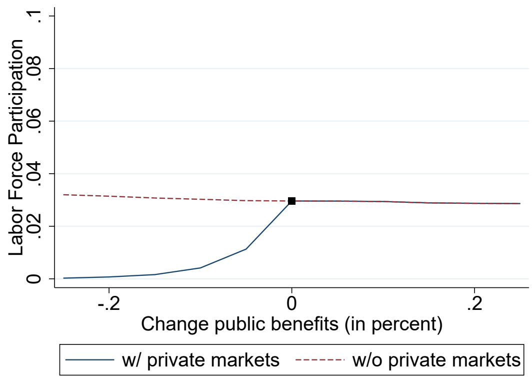

(d) Mean Income

Figure 2.A2: Labor Supply by private DI coverage at baseline - Benefit Generosity Changes

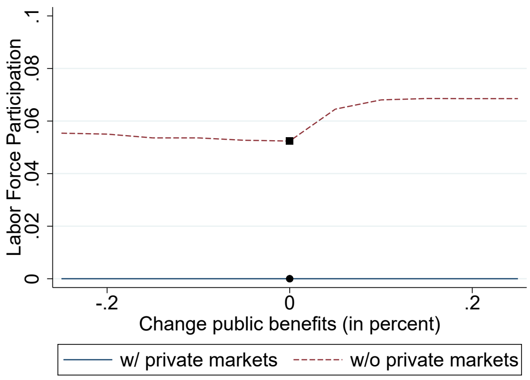
*The figure below presents the labor supply response for disabled individuals under alternative benefit generosity levels conditional on their private DI coverage at baseline. Panel (c) to (f) further condition on whether people continue to buy private DI or stop buying. The results are computed for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



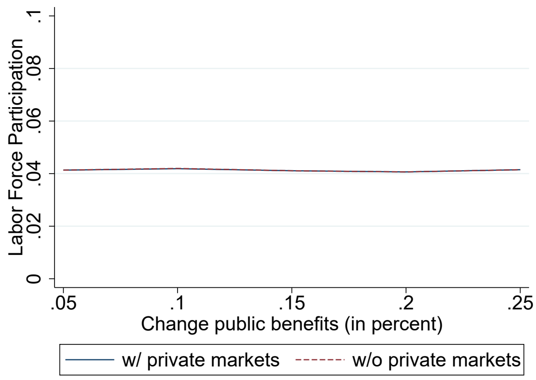

(a) Baseline Buyers



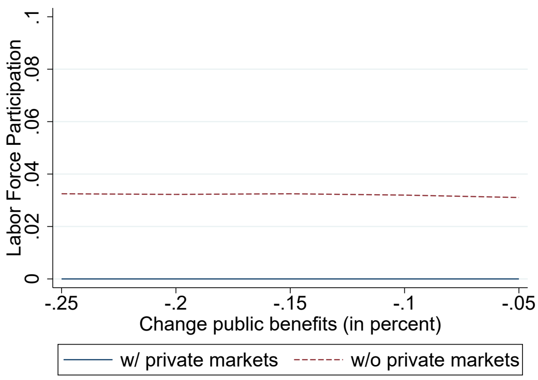

(b) Baseline Non-Buyers



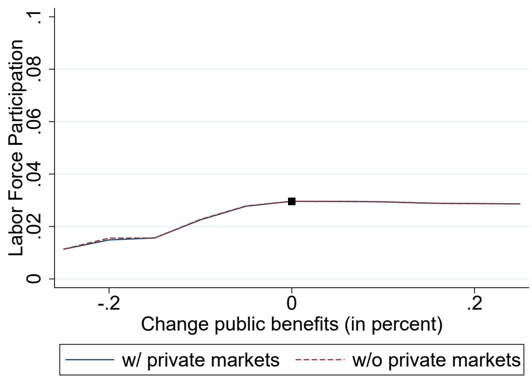

(c) Baseline Buyers (continue buying)




(d) Baseline Buyers (stop buying)



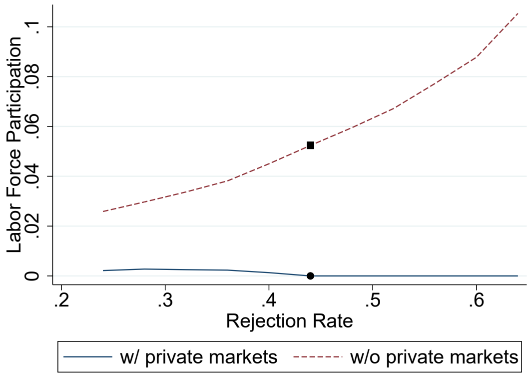

(e) Baseline Non-Buyers (start buying)



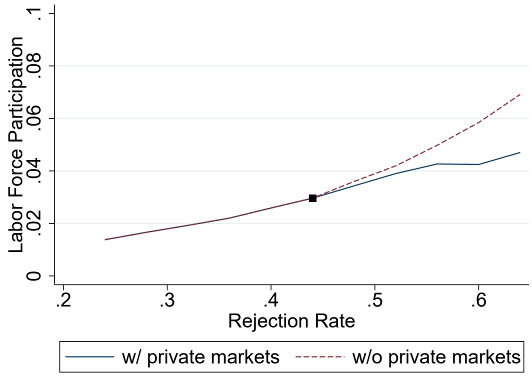

(f) Baseline Non-Buyers (still do not buy)

Figure 2.A3: Labor Supply by private DI coverage at baseline - Rejection Rate Changes

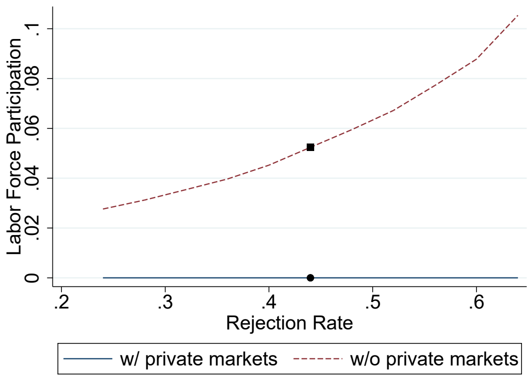
*The figure below presents the labor supply response for disabled individuals under alternative rejection rates conditional on their private DI coverage at baseline. Panel (c) to (f) further condition on whether people continue to buy private DI or stop buying. The results are computed for a population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



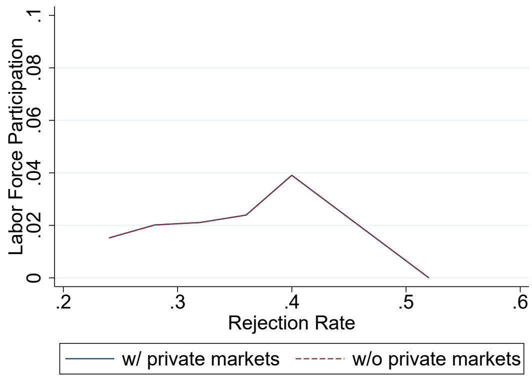

(a) Baseline Buyers



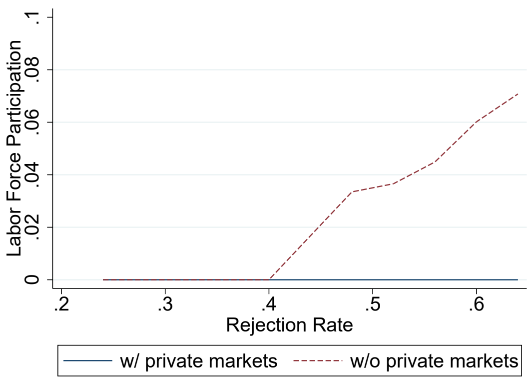

(b) Baseline Non-Buyers



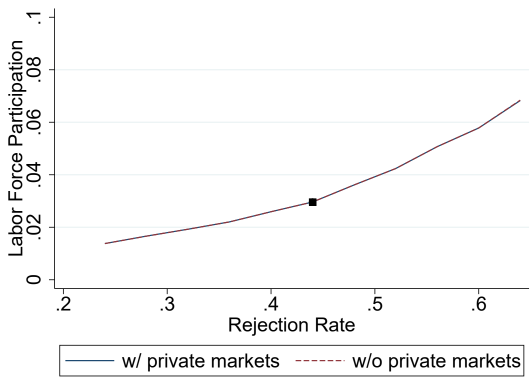

(c) Baseline Buyers (continue buying)




(d) Baseline Buyers (stop buying)



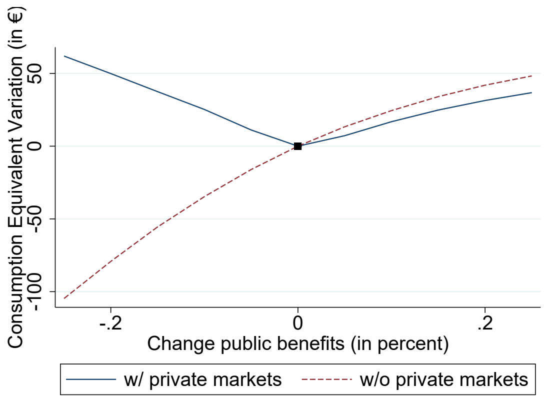

(e) Baseline Non-Buyers (start buying)



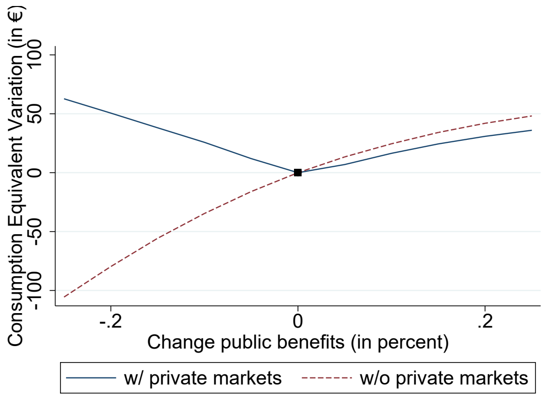

(f) Baseline Non-Buyers (still do not buy)

Figure 2.A4: Consumption - equivalent variation for changes in benefit generosity

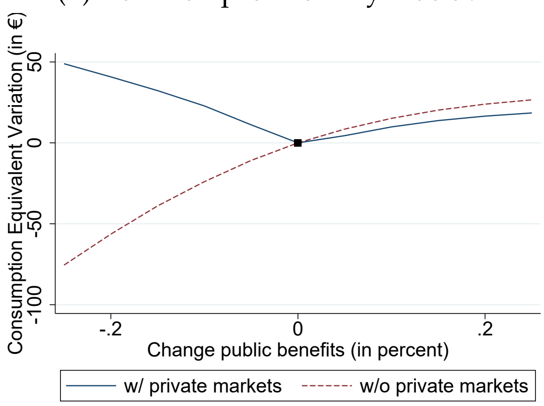
*The figure below presents the consumption-equivalent variation (CEV) for changes in the benefit generosity for the baseline specification (a), for a smaller retained productivity (b), and when an intensive margin is added to the problem (c). The CEV measures the change in expected life-time utility 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



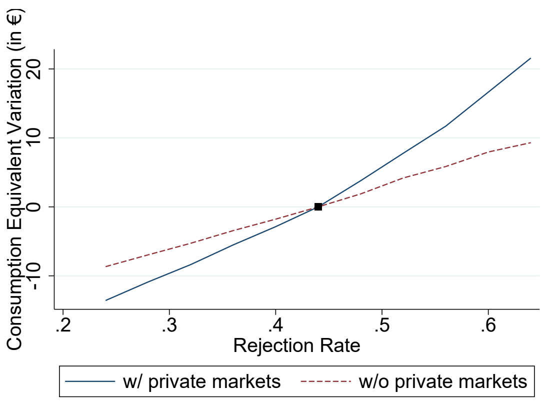

(b) Retained productivity = 38.5%



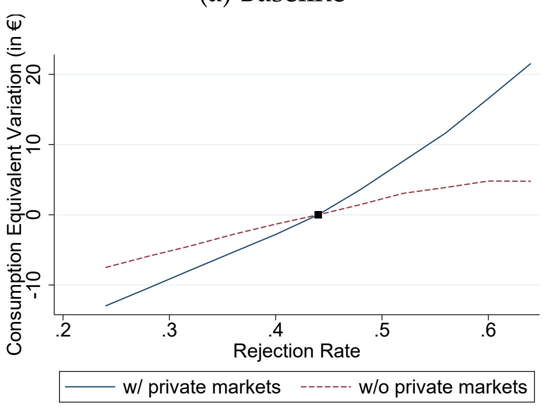

(c) Added intensive margin

Figure 2.A5: Consumption - equivalent variation for changes in rejection rates

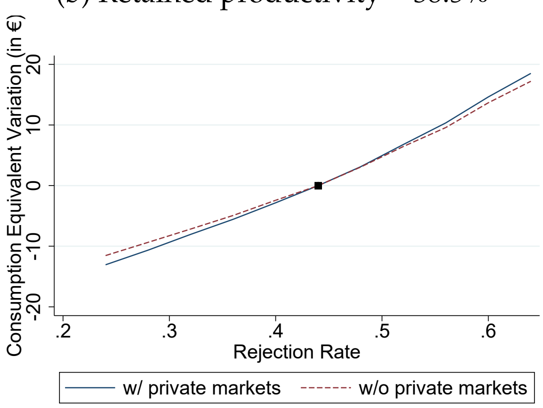
*The figure below presents the consumption-equivalent variation (CEV) for changes in the rejection rate for the baseline specification (a), for a smaller retained productivity (b), and when an intensive margin is added to the problem (c). The CEV measures the change in expected life-time utility relative 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



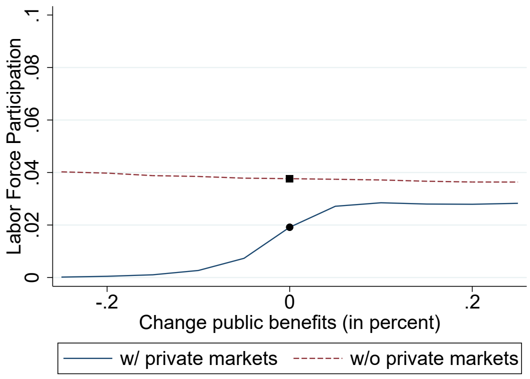

(b) Retained productivity = 38.5%



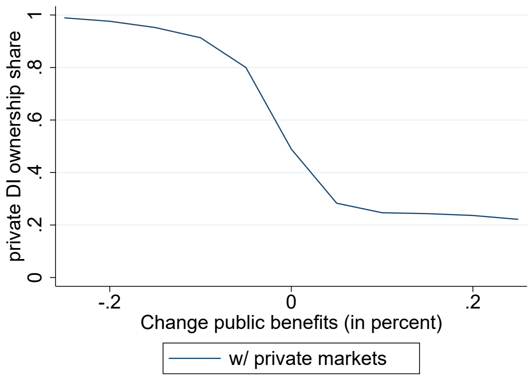

(c) Added intensive margin

Figure 2.A6: Labor force participation and mean private DI shares for changes in benefit generosity

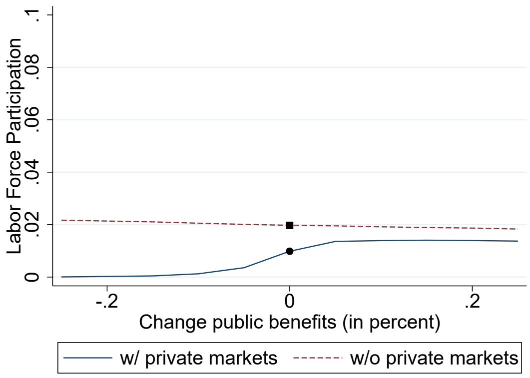
*The figure below presents the mean labor force participation of disabled individuals and the mean private DI ownership shares for alternative public DI benefit generosity. Panel (a) and (b) show the baseline results from the main text for the mean LFP and mean private DI shares respectively. Panel (c) 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.*



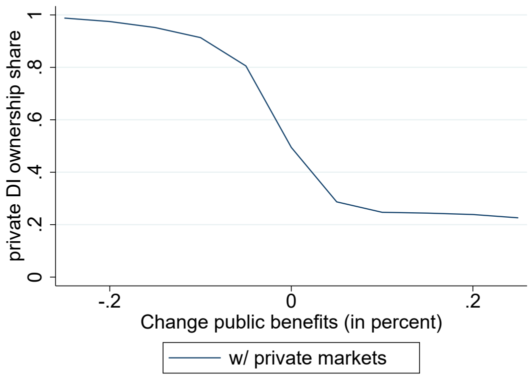

(a) Baseline, LFP



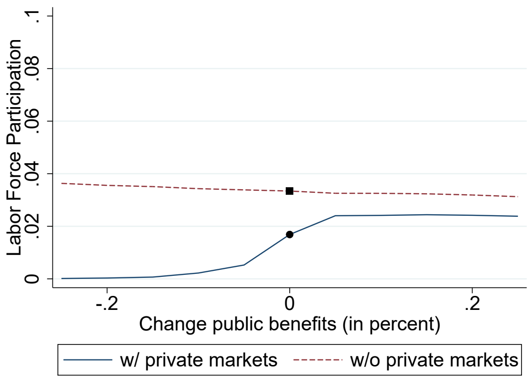

(b) Baseline, private DI shares



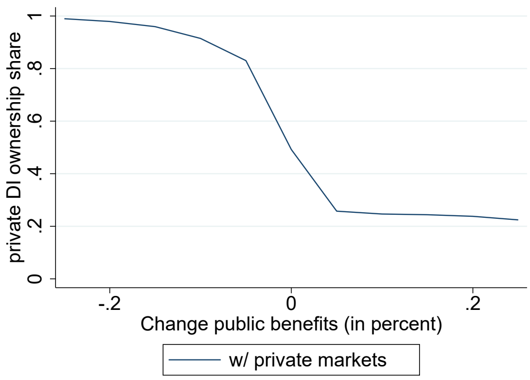

(c) Retained productivity = 38.5%, LFP




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



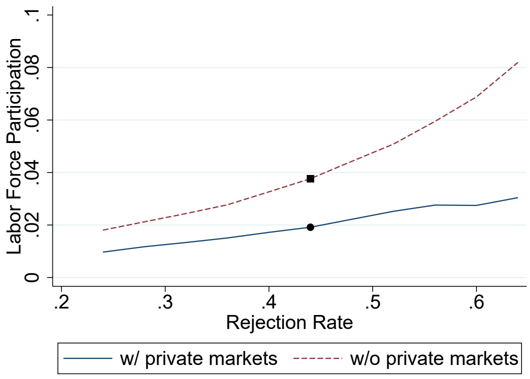

(e) Added intensive margin, LFP



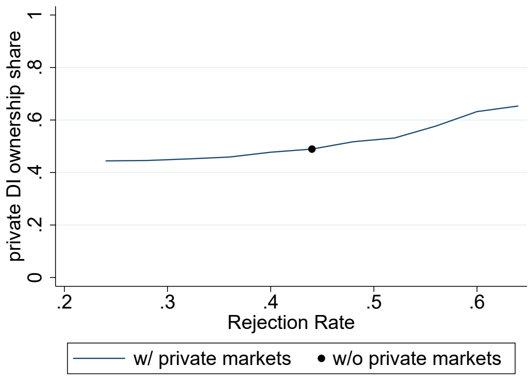

(f) Added intensive margin, private DI shares

Figure 2.A7: Labor force participation and mean private DI shares for changes in the rejection rate

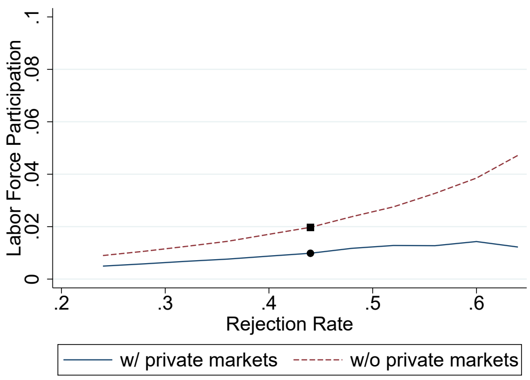
*The figure below presents the mean labor force participation of disabled individuals and the mean private DI ownership shares for alternative public DI rejection rates. Panel (a) and (b) show the baseline results from the main text for the mean LFP and mean private DI shares respectively. Panel (c) 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.*



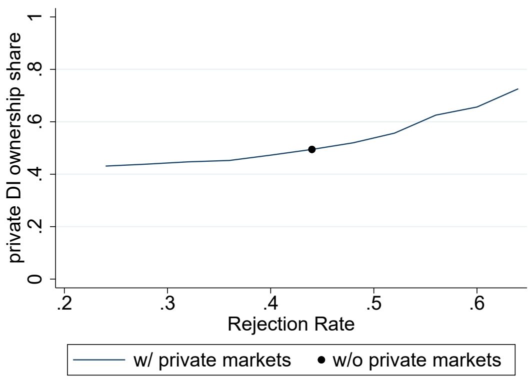

(a) Baseline, LFP



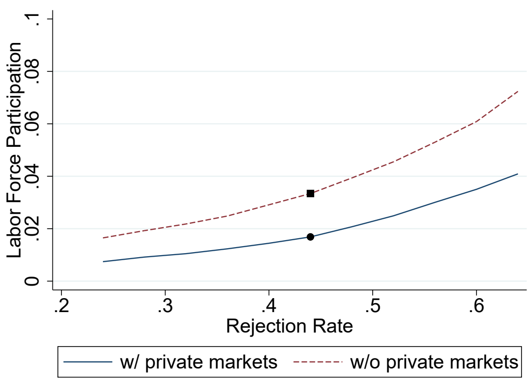

(b) Baseline, private DI shares



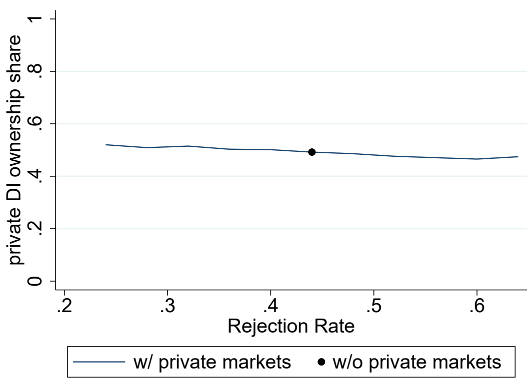

(c) Retained productivity = 38.5%, LFP




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



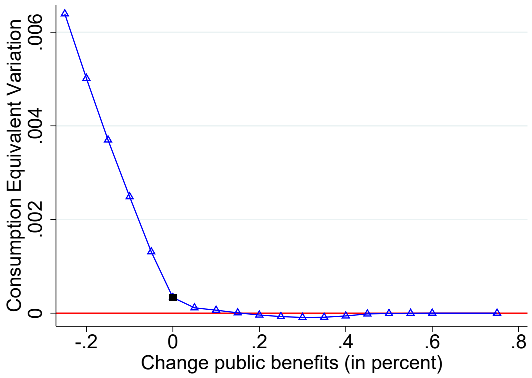

(e) Added intensive margin, LFP



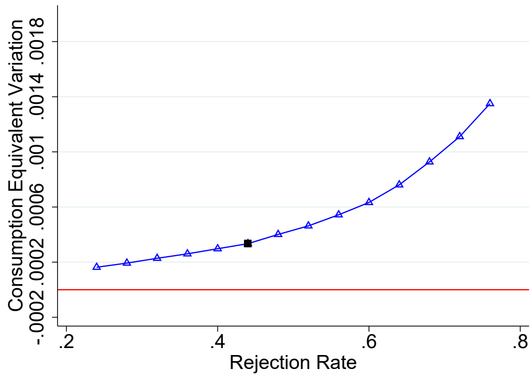

(f) Added intensive margin, private DI shares

Figure 2.A8: Welfare effects of private DI markets

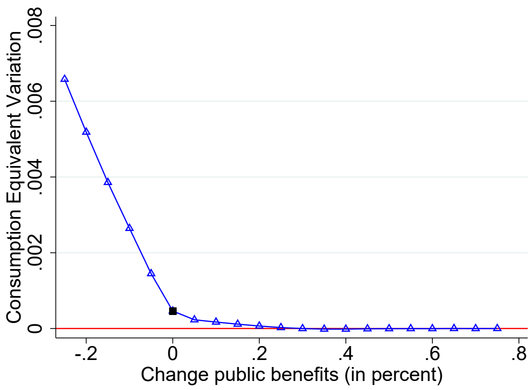
*The figure below presents the consumption-equivalent variation (CEV) for allowing for private DI markets under alternative policy schedules. The first (second) column depicts the CEV for changes in public benefit generosity (rejection rates). The first row shows the results derived under the baseline 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.*



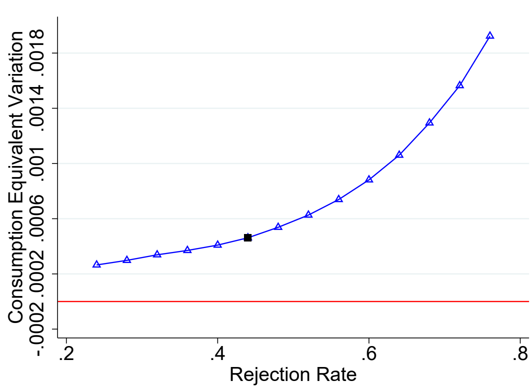

(a) Changes in benefit generosity, baseline



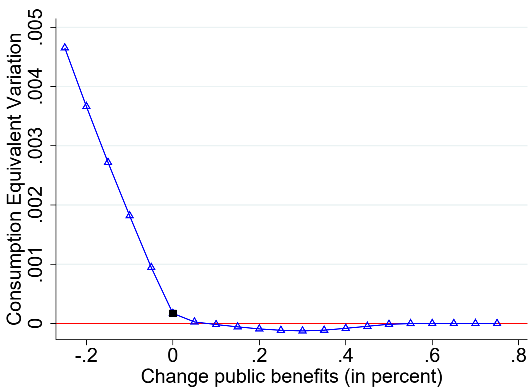

(b) Changes in rejection rates, baseline



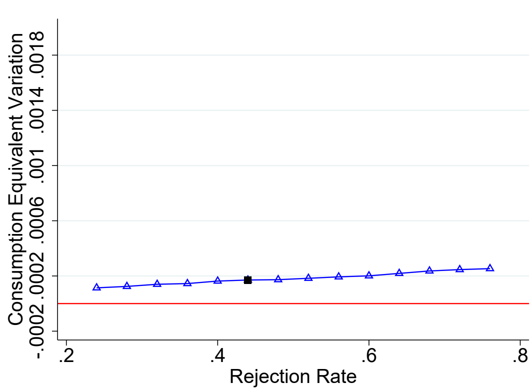

(c) Changes in benefit generosity, retained productivity = 38.5%




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




(e) Changes in benefit generosity, intensive margin




(f) Changes in rejection rates, intensive margin

# Chapter 3

# Privatizing Disability Insurance

Joint with Arthur Seibold and Sebastian Siegloch.

## 3.1 Introduction

Over the past decades, the number of individuals receiving public disability insurance (DI) benefits has risen rapidly in many countries. This growth in benefit receipt has made DI one of the largest social insurance programs in most OECD countries, which spend an average of 2% of GDP on public DI (OECD 2019). Due to the increasing fiscal burden of public DI, many governments face pressure to enact reforms reducing the generosity of these programs. While such reforms help improve fiscal sustainability, they usually come at the cost of providing lower benefits to individuals suffering from disability.

At the same time, private disability insurance markets exist in many countries.[1] If individuals value extra insurance, private DI could naturally help compensate for public DI cuts. Thus, a larger role of private DI is part of some policy proposals aimed at reducing public program costs (e.g. Autor and Duggan 2010). Opponents of this idea point out several potential problems in private DI markets. First, their efficiency may be impeded by adverse selection, such that adequate private insurance is not offered to everyone. Second, they could raise equity concerns, as premiums charged in private DI markets may not be equally affordable to all workers. Third, individuals may not make optimal private insurance choices. In fact, such concerns are key part of the rationale for public DI provision to begin with (Liebman 2015). However, there is remarkably little empirical evidence on these issues and the welfare impact of private DI markets.

In this paper, we provide novel empirical evidence on the functioning of private DI markets, and we investigate the welfare consequences of a larger role of private DI. Exploiting a reform that abolished one part of public DI for younger workers in Germany, we document significant crowding-in of private DI. Overall private DI take-up remains modest at around one quarter, but we do not find any evidence that this is driven by adverse selection on unpriced risk. On the contrary, private DI is concentrated among individuals with high income, high education and in low-risk occupations charged lower insurance premiums. In the second part of the paper, we turn to welfare implications. A key ingredient for welfare analysis are individual valuations of DI coverage, which we estimate in a revealed preferences approach. We find that in the absence of behavioral frictions, partly privatizing DI as done by the reform can be welfare improving. How-

[1] For instance, in the U.S., 33% of workers have private long-term DI as of 2019 (U.S. Bureau of Labor Statistics 2019). In Germany, 26% of workers have private DI as of 2015 (TNS Infratest 2015).

ever, a full public DI mandate could be justified by equity concern, where the planner values insuring low-income and high-risk individuals. Finally, we provide suggestive evidence that individuals may undervalue DI due to risk misperceptions, which can provide further grounds for a mandate.

There are two main challenges in studying private DI markets, which are presumably part of the reason why there is little existing evidence on these questions. First, in order to investigate to what extent private DI could compensate for public DI cuts, suitable variation in public DI coverage is needed. In the German setting, we can exploit a unique reform which sharply reduced the scope of public DI. The reform of 2001 abolished one part of public DI for younger workers, namely *own-occupation* insurance. Receiving own-occupation DI benefits requires not being able to work in one's previous occupation. In contrast, *general* DI benefits are based on more stringent eligibility criteria, requiring an individual to be unable to work in any occupation. Before the reform of 2001, both own-occupation and general DI were part of the social insurance system, but the reform abolished public own-occupation DI for cohorts born in 1961 and later. Importantly, the German private DI market offers contracts including own-occupation DI coverage, such that workers affected by the reform who wish to compensate for the loss of public DI coverage can do so by taking up private insurance.

A second challenge is the difficulty of obtaining comprehensive data on private DI take-up. To address this challenge, we combine a number of different data sources. First, we use microdata on all private DI contracts within a large insurance company, which is one of the top-10 in the German private DI market. Second, we obtained aggregate data on the overall private DI market from a leading rating agency, which compiles data from all insurers active in the private DI market. Third, we use administrative data on the universe of public DI claims between 1992 and 2014 provided by the German State Pension Fund. Finally, we use representative household survey data from the Income and Consumption Survey (EVS), which allows us to perform a number of checks to validate results from the insurer microdata. We find similar patterns in the survey and in available market-level data, suggesting that the insurer microdata is representative of the market along key dimensions.

In the first part of the paper, we provide empirical evidence on the functioning of the private DI market. We begin by studying *crowding-in* effects of the reform, that is the impact of public DI cuts on private DI take-up. On aggregate, we find substantial growth of the private DI market around the time of the reform. In order to identify a causal effect of the reform on private DI take-up, we use a difference-in-difference strategy around the cohort cutoff of the reform. We find that treated individuals born in the two years after the cutoff increase private insurance purchases by around two thirds compared to control cohorts born prior to the cutoff. We argue that this estimate is likely conservative, since we observe larger increases in take-up among younger workers born further away from the cutoff. Yet, even 15 years after the reform, overall take-up remains modest, as only 26% of workers hold private DI.

We find strong heterogeneity in private DI take-up by observable characteristics. In particular, individuals with high income and high education are much more likely to purchase private DI. For instance, take-up is 65% in the top income quintile, but only 7% to 11% in the bottom three quintiles. Heterogeneity by education is even more pronounced, with 80% take-up in the top education quintile, and only 5% to 8% in the bottom three quintiles. Moreover, there is important heterogeneity in take-up by priced *risk groups*, which insurers assign to workers based on occupations and which determine private DI premiums. Individuals in low risk groups who are charged low premiums are much more likely to take-up insurance than those in high risk groups where premiums are high. This result has two implications. First, individuals with the highest disability risk tend not to be covered by the private DI market. Second, since the relative premiums across risk groups are not far from actuarially fair, the large differences in take-up indicate strong responses of insurance demand to prices.

Next, we investigate risk-based selection into private DI in more detail. We implement a "positive correlation test", regressing post-reform private DI take-up within a three-digit occupation on disability risk among this occupation. Two features of our selection test are worth emphasizing. First, the relevant risk-based selection from an efficiency point of view is selection on *unpriced* risk. Thus, we condition on the priced component of risk by testing for selection within risk groups facing the same insurance prices. Second, an important issue with a correlation test is that it may confound selection and moral hazard. Our solution to this problem is to measure disability risk in an occupation only among workers in cohorts 1960 and older, who are still fully covered by public own-occupation DI. Since these workers are all observed under the same level of DI coverage, differences in observed claims should reflect differences in ex-ante risk rather than ex-post moral hazard responses.

We find no significant correlation between private DI take-up and unpriced risk. The point estimate from the correlation test is negative but close to zero. Importantly, this suggests that the private DI market is not impeded by adverse selection, which is often considered to be the main rationale for a public DI mandate. At first glance, the lack of adverse selection may seem surprising, as insurance should in principle be more valuable to higher-risk individuals. We present suggestive evidence that this could be explained by some individual characteristics driving advantageous selection. In particular, once we condition on education, the correlation of private DI take-up and risk becomes positive. This is consistent with higher-educated workers having stronger preferences for insurance, while they tend to work in lower-risk occupations. In other words, advantageous selection on unpriced characteristics could counterbalance potential adverse selection on unpriced risk, implying no overall adverse selection. Another potential explanation for the lack of adverse selection is that individuals may not correctly perceive their disability risk, which we discuss later in the paper.

In the second part of the paper, we turn to the welfare implications of (partly) privatizing disability insurance. The analysis builds on Einav, Finkelstein, and Cullen (2010), who show that insurance demand and cost curves can be used as sufficient statistics to assess welfare in insurance markets. In particular, our post-reform setting with insurance choice provides a unique opportunity to directly estimate individuals' willingness to pay for the DI coverage offered by the private market. Thus, we implement a revealed preferences approach, where observed choices reveal insurance valuations, absent behavioral frictions. To estimate demand elasticities, we exploit the price variation between risk groups. Intuitively, the insurer assigns occupations to a discrete number of risk groups based on underlying disability risk, such that there are occupations with similar risk facing different insurance prices around the risk group boundaries. We find sizeable jumps in insurance take-up in response to these quasi-discontinuities in prices, and the resulting average demand elasticity is -1.16. The second key statistic, namely the cost of providing DI, can be directly estimated based on realized DI claims in each risk group.

Our baseline welfare measure is the *net value* of DI, which expresses the willingness to pay for additional insurance relative to its cost, analogously to the marginal value of public funds (Finkelstein and Hendren 2020, Hendren and Sprung-Keyser 2020). Our main counterfactual of interest compares the post-reform status quo where DI is partly provided via the private market to a full public DI mandate including this extra coverage. We find an overall net value of a mandate of 0.76, implying that the revealed insurance valuation among individuals additionally covered by the mandate is only 76% of the cost of insuring them. This result reflects an efficiency advantage of the private insurance market. Since there is no significant adverse selection, the market covers the majority of individuals with sufficiently high willingness to pay, and a mandate would predominantly lead to additional coverage of those with valuations below the cost of insurance.

A first important caveat with this baseline result is that distributional concerns are not taken into account. In particular, the private DI market tends to leave low-income and high-risk individuals uninsured, which may be undesirable to a social planner with equity concern. To account for this, we extend the analysis and calculate the social net value of DI, applying social welfare weights based on expected lifetime income in each risk group. We find that a full public DI mandate has a social value exceeding its costs even under moderate equity concern given by a Utilitarian social welfare function and low risk aversion. Importantly, we note that the redistributive effects of a mandate hinge on the design of social insurance. A private insurance mandate does not achieve an increase in social net value, since the benefits of insurance to high-risk groups are counteracted by the high risk-based premiums charged to these workers. A public insurance mandate with income-based contributions, on the other hand, effectively redistributes to low-income, high-risk individuals.

A second caveat is that our revealed preferences approach assumes that individuals make optimal insurance purchase decisions, which has been called into question in recent literature (e.g. Chandra, Handel, and Schwartzstein 2019). Thus, in a second extension, we account for such behavioral frictions in a series of calibration exercises. We proceed in three steps. First, we calibrate risk preferences implied by observed private DI purchases in a simple model of insurance choice under a range of assumptions about the consumption drop upon disability. We find that relative risk aversion would have to very low for many individuals in order to rationalize low observed private DI take-up. Second, we argue that risk misperceptions could provide an alternative rationale for low revealed insurance valuations. In further calibrations, we find that individuals in higher-risk groups would have to underestimate disability risk by roughly one 30% to 60% to explain observed take-up. In the third step, we calculate the wedge between observed willingness to pay and normative valuations implied by calibrated risk misperceptions. The results suggest that willingness to pay of marginal buyers would be about 50% to 150% higher if they correctly perceived their disability risk. Finally, we find that implied normative valuations tend to exceed the cost of insurance, suggesting that risk misperceptions can provide an additional rationale for a mandate.

This paper contributes to a large and growing literature on disability insurance (see Low and Pistaferri 2020 for a recent review). Much of this literature focuses on the effect of public DI on labor supply and claiming decisions (Bound 1989, Gruber 2000, Autor and Duggan 2003, 2006, 2007, Autor, Duggan, and Lyle 2011, Staubli 2011, Wachter, Song, and Manchester 2011, Marie and Castello 2012, Maestas, Mullen, and Strand 2013, French and Song 2014, Kostol and Mogstad 2014, Borghans, Gielen, and Luttmer 2014, Koning and Lindeboom 2015, Liebman 2015, Autor et al. 2016, Burkhauser, Daly, and Ziebarth 2016, Deshpande 2016a,b, Mullen and Staubli 2016, Gelber, Moore, and Strand 2017, Autor et al. 2019, Ruh and Staubli 2019). There is little existing work on private DI markets, on the other hand. Exceptions include Autor, Duggan, and Gruber (2014), Stepner (2019) and Seitz (2021), who analyze moral hazard effects of private DI.

We make three main contributions to this literature. First, exploiting the unique German setting where a part the public DI mandate is removed, we provide novel empirical evidence on crowding-out and selection in private DI markets. To our knowledge, our findings constitute the first direct empirical evidence on these issues, which are key in assessing the welfare impact of policies expanding the role of private markets and choice in DI. Second, we further exploit our setting with insurance choice in order to estimate individual valuations for DI in a revealed preferences approach. Our approach is closely related to Cabral and Cullen (2019) who estimate a lower bound on the value of public DI using supplemental private DI purchases within a U.S. employer. Third, we assess the welfare consequences of the private DI market offering some coverage vs. a full public mandate. This complements and extends recent work analyzing welfare and the insurance-incentive trade-off within public DI (Low and Pistaferri 2015, Meyer and Mok 2019, Haller, Staubli, and Zweimüller 2020).

The remainder of this paper is organized as follows. Section 3.2 outlines context and data, Section 3.3 presents evidence on the crowding-in of private DI, Section 3.4 shows results on selection into private DI, Section 3.5 describes the validation exercises, Section 3.6 presents the demand and cost curve estimation, Section 3.7 discusses the welfare effects of private vs. public DI, and finally Section 3.8 concludes.

## 3.2 Context and Data

### 3.2.1 Institutional Context

**Public Disability Insurance.** In Germany, public disability insurance (DI) is administered by the State Pension Fund and shares many of its characteristics with DI programs in other countries. Enrolment in public DI is mandatory for all employed individuals, while most self-employed workers and civil servants are exempt. Contributions are levied as payroll taxes along with pension contributions. Enrolled workers become eligible for DI benefits in the event of a permanent disability. Moreover, eligibility requires having contributed for at least five years in total, and at least three out of the five years before the onset of disability. Upon application, a medical and work capacity assessment is carried out by the Pension Fund. Benefit calculation is based on a worker's contributions so far, assuming that they would have kept contributing according to their average pre-disability earnings until age 63. DI benefits are paid until the individual recovers from disability; otherwise, benefits are paid until the Normal Retirement Age, when they are converted into an old-age pension. Throughout their lifetime, 25.1% of workers claim public DI and the average gross replacement rate is 39% (own calculation based on public pension data).

Crucially for our purposes, the public DI system consists of two branches, *general DI* and *own-occupation DI*. The first branch pays benefits to workers suffering from a general disability (*Erwerbsunfähigkeit*), such that they are unable to work in any occupation for more than three hours per day. Common conditions leading to general disability include degenerative disc disease or severe burn-out/depression. The second branch, on the other hand, requires a so-called own-occupation disability (*Berufsunfähigkeit*) defined as being unable to work in their previous occupation. For instance, a bus driver suffering from severe vision impairment is unable to work in their occupation, but may be able to work in other occupations. Such own-occupation DI cases make up 13.2% of all public DI claims. Besides differences in work capacity assessment, the two DI branches also require separate applications and entail different benefit rules. Workers on own-occupation DI receive two thirds of general DI benefits, but face a somewhat less stringent earnings test.[2]

[2]General DI benefits are reduced for monthly earnings above EUR 400, whereas workers on own-occupation DI are allowed to earn at least EUR 700, depending on their prior earnings. Note that these

**The Reform of 2001.** Before 2001, all workers were covered both by general and own-occupation DI as part of the public DI mandate. However, rising expenditure on DI benefits stoked concerns about the fiscal sustainability of the program in the 1990s. This motivated a major reform in 2001 aimed at reducing public DI spending. Most importantly, the reform featured a sharp, cohort-based change in the scope of public DI: own-occupation DI coverage was abolished for birth cohorts 1961 and younger from 2001 onward. Besides this main element, the reform featured further changes equally affecting all cohorts, including gradually phased-in changes to benefit calculation.[3]

The timing of the reform was noteworthy. Initially, the reform was announced in December 1997 to take effect in January 1999. Importantly, the initial reform intended to abolish own-occupation DI for all workers and not only for younger cohorts. After a change of federal government and in the face of public opposition, the reform was retracted in late 1998. However, in December 2000, the reform was re-announced in its final form featuring the cohort cutoff, and it took effect in January 2001.

**Private Disability Insurance.** The market for private DI has existed since at least the 1920s in Germany. Around 70 insurance companies currently offer private DI contracts. Crucially, private DI always includes coverage of own-occupation disability risk, closely mirroring the pre-reform public DI system. Thus, workers affected by the reform can choose to purchase private DI to compensate for the loss in public own-occupation DI coverage. Private DI payouts are independent of the public DI system, such that they can also serve as a top-up in case a worker is eligible for public DI benefits.

An important difference to the public DI system is that private DI premiums are risk-based. In practice, the primary determinant of an individual's private DI premium is their occupation, whereby insurers map occupations into a discrete number of risk groups. The insurer from which our microdata originates uses five risk groups, and other insurers use similar numbers of groups during the period we study. Appendix

earnings test thresholds are adjusted every few years. The aforementioned figures apply between 2008 and 2017.

[3]More precisely, the reform altered two elements of benefit calculation. First, an adjustment factor was gradually introduced, featuring negative benefit adjustments similar to penalties for claiming old-age pensions early. Second, the hypothetical contribution period used for benefit calculation was gradually extended, somewhat counteracting the new penalties. In addition, the reform introduced the possibility of claiming partial DI benefits for individuals who are able to work between three and six hours per day. Finally, DI benefits are meant to be generally granted on a temporary basis after the reform, but in practice most beneficiaries still receive benefits permanently.

Table 3.A1 shows examples of frequent occupations in each risk group. Furthermore, insurance premiums can be adjusted for pre-existing medical conditions and risky private activities such as extreme sports, but this is relatively rare.[4] Finally, monthly premiums are actuarially adjusted to the individual's contract start and end date. The level of insured benefits can be specified individually. On average, monthly private DI payouts are EUR 836, a similar magnitude to the average benefits of EUR 711 in the public DI system (Allianz 2018). The majority of 85% of private DI contracts are purchased individually, and the remainder are bought via employers (FAZ 2012). Finally, private DI can be purchased either as a stand-alone product or bundled with other types of insurance, most commonly life insurance.

### 3.2.2 Data

An important challenge in studying private DI is that comprehensive, high-quality data on private insurance contracts and take-up is not readily available. We tackle this challenge by combining a number of data sources. First, we use microdata on all DI contracts in a large private insurance company. The insurer is among the top-10 in the private DI market, with a market share between 3% and 6%.[5] We observe private DI contracts existing in any of the years between 2012 and 2017, irrespective of the start date of the contract. The data contains information on contract start and end dates, insured benefits, risk groups as well as some sociodemographics including age and gender. Unfortunately, individual income and education are not included in the microdata. We thus match it with information on average income by occupation, age and gender measured in administrative labor market data[6] Similarly, we add education at the occupation level. Panel A of Table 3.2.1 shows summary statistics of the insurer microdata. Our main sample, which excludes contracts held by self-employed and civil servants, contains a high six-digit number of contracts. 61% of contract holders are male, the average purchase age is around 30 and the end age is around 63. Average monthly premiums are EUR 78 and insured monthly benefits are EUR 1383, and 55% of contracts were sold as a stand-alone product.

[4]In only 4% of private DI contracts, premiums are adjusted beyond risk-group specific prices. Moreover, only 4% of individuals are rejected at the contracting stage in the private DI market (GDV 2016). To our knowledge, this includes a few extremely risky occupations such as circus artists and explosives workers, as well as rejections due to pre-existing conditions or risky activities.

[5]For confidentiality reasons, we are unable to name the insurer or specify its market share more precisely.

[6]See Seitz (2021) for a detailed description of the insurer microdata and the occupation matching procedure.

As a second source of information on private DI, we have obtained aggregate data on the entire private DI market from a leading rating agency. This data, on which we draw mainly for the aggregate patterns shown in Sections 3.3 and 3.5, contains time-series information on the total number of private DI contracts, the shares of different types of contracts, as well as some information on the shares of contracts held by risk groups and age groups.

Third, we use administrative data on the universe of public DI claims between 1992 and 2014 provided by the German State Pension Fund.[7] This data contains information on the timing and type of DI claims, benefits, as well as information on individual earnings histories necessary to compute benefit eligibility and some sociodemographics including age, marital status and gender. Panel B of Table 3.2.1 shows summary statistics of the administrative data. In Column (1), 59% of all DI claimants are male, and the average claiming age is around 52. Monthly DI benefits are on average EUR 1078, and claimants' average earnings were EUR 2305 over all periods, and EUR 1307 in the period before the DI claim. Column (2) shows that compared to all DI claims, own-occupation DI claimants are more likely to be male and married, and their age and income tend to be slightly higher.

Finally, we use data from the Income and Consumption Survey (EVS), a representative household survey conducted by the German Federal Statistical Office. We focus on the 2013 wave of the survey, which contains information on households' private DI take-up. We use this data for complementary analyses, in particular for some of the validation exercises presented in Section 3.5. Appendix Table 3.A2 shows summary statistics of the survey data. 31% of households hold private disability insurance in 2013. Households' average labor earnings are around EUR 2185 per month, the average age of the household head is 44, 59% are male and the average household size is just above two.

**Representativeness of the Insurer Microdata.** An important question is how representative the insurer providing our microdata is for the private DI market. We argue that the insurer reflects the overall market well in key dimensions. First, the main features of private DI contracts described in Section 3.2.1, including the definition of dis-

[7] The data on public DI claims is a subset of administrative data on all public pension claims first used by Seibold (2021). We also use the full dataset on all pension claims to calculate some aggregate statistics, such as the distribution of occupations, risk groups, income and education.

ability, benefit levels, and contract durations offered, are similar across providers.[8] Second, the pricing of private DI contracts follows similar rules across insurers, assigning individuals to risk groups primarily based on occupations. As we show in Section 3.5, this results in similar relative prices across risk groups charged by different providers. Third, our insurer offers private DI to individuals across all occupations and industries. Thus, we observe private DI contracts of individuals belonging to 322 out of 334 3-digit occupations in the microdata. Fourth, the insurer has a countrywide presence and does not appear to specialize in particular geographic areas. This is illustrated in Appendix Figure 3.A1, showing the geographic distribution of its local insurance agencies. The insurer has agencies across all states and in all major cities, as well as in a large number of rural locations across the country. In addition, in Section 3.5, we present a number of validation checks of our main results using independent, representative data sources, which yield similar empirical patterns to the insurer microdata.

## 3.3 Crowding-In of Private Disability Insurance

The reform of 2001 abolishes public own-occupation DI for younger birth cohorts, which these individuals could compensate by purchasing private DI covering this risk. In this section, we study the effect of the reform on overall private DI take-up. We refer to the response of private insurance take-up to public DI cuts as a crowding-in effect, analogously to crowding-out effects following social insurance expansions studied in the literature (e.g. **BrownFinkelstein2008**, Chetty et al. 2014).

### 3.3.1 Overall Private DI Take-Up

We begin by showing aggregate patterns in public DI claims and private DI take-up in Figure 3.3.1. Panel (a) depicts the total number of public own-occupation DI claims by calendar month. Precisely at the time of the reform, there is a sharp drop in benefit claims, as the younger cohorts affected by the reform lose access to public own-occupation DI. Moreover, the figure indicates a continuing downward trend in claims over the years after the reform, as the share of workers in the older cohorts who are

[8] According to consumer advice, differences across private DI providers are more fine-grained, such as the precise definition of equivalent occupations, the minimum qualifying period of disability, whether benefits can be paid retroactively and whether coverage can be altered throughout the contract (see e.g. **BUVT2019**).

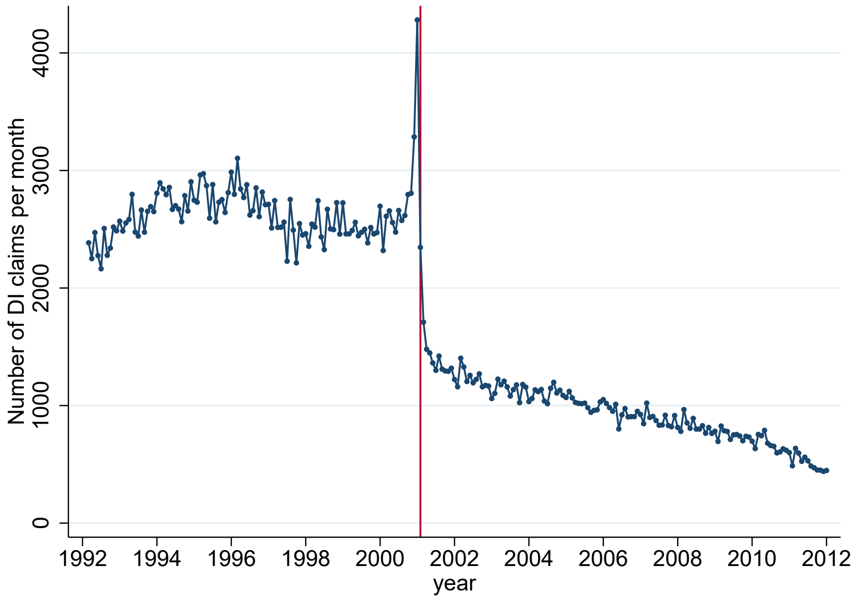
still eligible for own-occupation DI keeps declining. There also appears to be some re-timing of claims in the months just before the reform. Even though the spike just before January 2001 is sharp, the magnitude of these excess claims is small relative to the permanent reduction in claims after the reform. Panel (b) of Figure 3.3.1 shows overall private DI take-up over time. We calculate the take-up rate $Q_t = C_t/N_t$, where $C_t$ is the total number of private DI contracts and $N_t$ is the size of the relevant population. We obtain $C_t$ based on the rating agency data on all contracts in the market in each year, and we take $N_t$ as the total number of individuals contributing to social insurance from social insurance statistics. The figure shows a clear jump in private DI coverage around the time of the reform. By 2015, private DI take-up has increased to 26%, compared to around 10% in the years before the reform was first announced in 1997. This growth of the private DI market provides first suggestive evidence of a crowding-in effect. Yet, overall private DI take-up of around one quarter can be viewed as relatively modest, given that the reform fully removes public own-occupation DI coverage.

Appendix Figure 3.A2 shows some additional descriptive evidence on the private DI market. Panels (a) and (b) show alternative definitions of private DI take-up rates. First, Panel (a) shows take-up only the younger cohorts who are affected by the reform. We obtain this take-up rate by imputing the total number of contracts in the market held by cohorts 1961 and younger based on the corresponding share in the insurer microdata, and divide by the number of individuals contributing to social insurance in this age group. By 2015, the implied take-up rate among treated cohorts is 29%. The rate is very similar to overall take-up, as the share of individuals in these cohorts out of the total labor force is close to 90% by that time. Panel (b) shows overall take-up under an alternative definition of the relevant population, namely only currently employed individuals. Assuming private DI is only held by these individuals would result in a somewhat higher take-up rate of 38% in 2015. Next, Panel (c) shows that while most private DI contracts are bundled with other types of insurance, stand-alone DI grows fast in the years after the reform. Finally, Panel (d) presents a comparison of the number of contracts in the market vs. in the insurance company providing our microdata. The insurer microdata reflects the long-run trend in the market quite well, but we note some differences in the patterns over time. First, the insurer experiences higher overall growth than the market between 1995 and 2015. Second, the insurer grows somewhat less immediately after the reform, but instead exhibits more sustained growth over subsequent years.

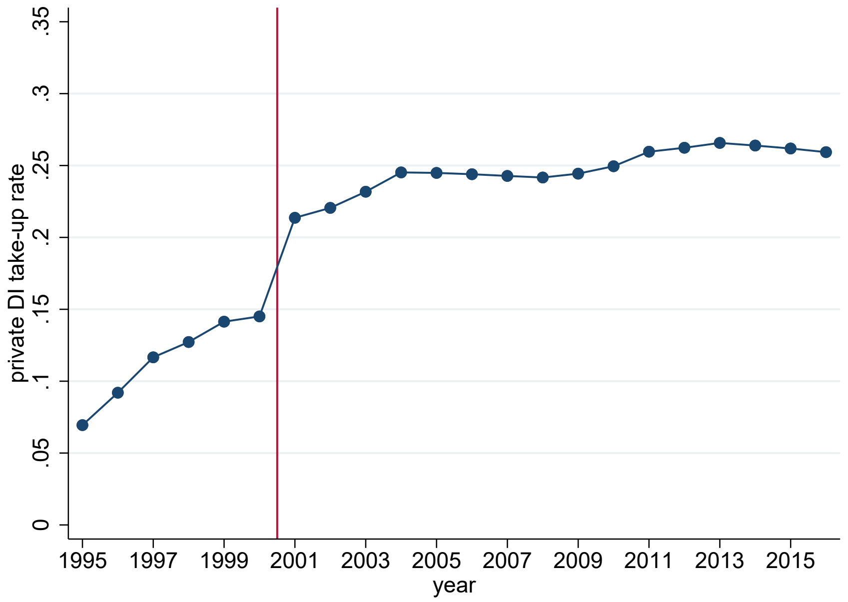
Figure 3.3.1: Crowding-In: Descriptive Evidence

*The figure shows the total number of public own-occupation DI claims (Panel a) and the overall private DI take-up rate (Panel b). In both panels, the vertical line denotes the time the reform of 2001 takes effect (January 2001).*

(a) Public Own-Occupation DI Claims




(b) Private DI Take-Up



### 3.3.2 Difference-in-Difference Estimation

The evidence above is suggestive of a crowding-in effect of the reform of 2001, but overall growth in the private DI market could be driven by a number of factors. In order to isolate a causal effect, we exploit the cohort cutoff of the reform to estimate a difference-in-difference specification. We run regressions of the following form:

$$Y_{ct} = \beta_0 + \beta_1 treat_c + \beta_2 treat_c \cdot post_t + \delta_t + \epsilon_{ct} \tag{3.1}$$

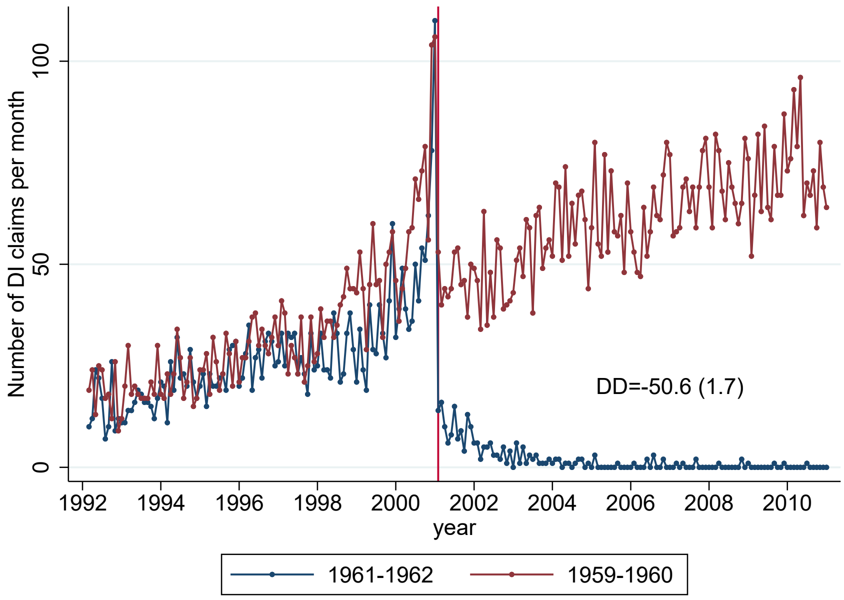
where $Y_{ct}$ denotes an outcome of cohort $c$ in calendar month $t$, $treat_c$ is an indicator for treated cohorts 1961 and younger, $post_t$ is an indicator for post-reform periods January 2001 and later, $\delta_t$ is a calendar month fixed effect, and $\epsilon_{ct}$ is an error term. The coefficient $\beta_2$ yields the difference-in-difference effect of interest. In the baseline specification, we focus on a narrow cohort window of +/– two years around the reform cutoff, comparing treated cohorts 1961-1962 to control cohorts 1959-1960. First, we investigate the effect of the reform on public own-occupation DI claims. Panel (a) of Figure 3.3.2 shows the number of claims by cohorts 1961-1962 vs. 1959-1960 over time. Before 2001, claims by both treated and control cohorts follow a similar increasing trend. Precisely in 2001, there is a sharp drop in claims by treated cohorts virtually to zero, while claims by the control group continue increasing similarly to before the reform.[9] Column (1) of Table 3.3.2 shows a highly significant difference-in-difference coefficient of -50.6, corresponding roughly to the number of monthly claims by treated cohorts just before the reform. Thus, the estimation confirms that the "first-stage" induced by the reform of 2001 is given by the virtually immediate and complete removal of public own-occupation DI coverage for younger workers. In addition, Column (2) of the table shows that the reform does not lead to spillovers into the other branch of public DI. The estimated effect on *any* type of public DI claims is, if anything, larger in magnitude than the effect on own-occupation DI claims, suggesting no benefit substitution towards general DI claims.

[9]The fact that claims by the treated cohorts do not drop precisely to zero in 2001 is likely due to delays in processing claims made before the reform.

Figure 3.3.2: Crowding-In: Difference-in-Differences

*The figure shows the number of public own-occupation DI claims (Panel a) and private DI purchases (Panel b) of individuals born in 1961-1962 (treated cohorts) vs. 1959-1960 (control cohorts). In both panels, the solid vertical line denotes the time the reform of 2001 takes effect (January 2001). In Panel (b), the dashed vertical line additionally demarcates the time the reform is first announced (December 1997). DD denotes the difference-in-difference coefficient estimated for the respective outcome with standard errors in parentheses (see Table 3.3.2 for details).*

(a) Public Own-Occupation DI Claims




(b) Private DI Purchases

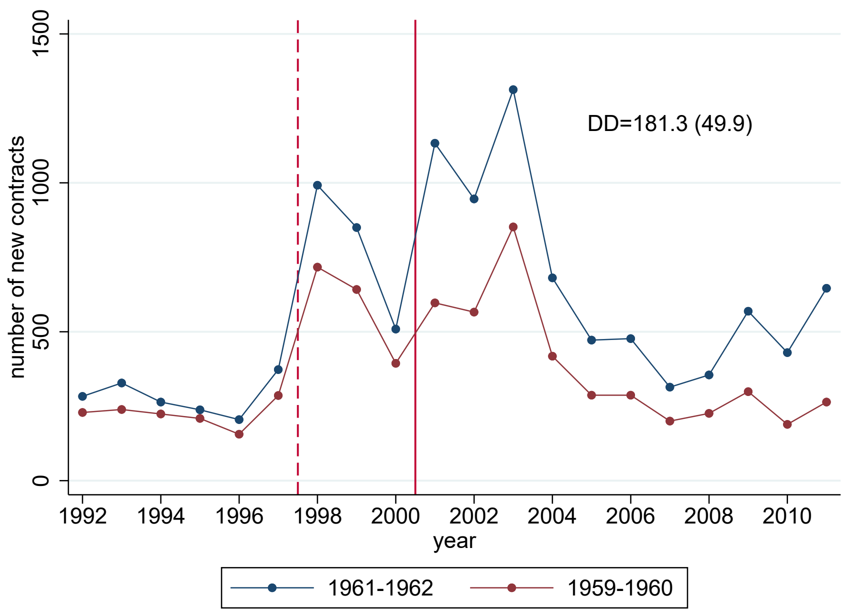


Table 3.2.1: Summary Statistics

*The table presents summary statistics of the insurer microdata (Panel A) and the adminstrative data on public DI claims (Panel B). In Panel A, "risk group" denotes risk groups assigned by the insurer to individuals based on their occupation. "Stand-Alone DI contract" denotes whether a contract was purchases on its own or in a bundle with other insurance products. Number of observations refers to number of private DI contracts, which we cannot show for the full sample due to confidentiality reasons. In Panel B, number of observations refers to number of DI claims.*

**Panel A: Insurer Microdata on Private DI Contracts**

| | (1) Full Sample | (2) Cohorts 1959-1962 |
|---|---|---|
| Male | 0.61 | 0.71 |
| | (0.49) | (0.45) |
| Income (monthly) | 4132.05 | 4422.10 |
| | (1385.57) | (1364.61) |
| Education (years) | 12.42 | 12.22 |
| | (1.97) | (2.03) |
| Risk Group | 1.96 | 2.55 |
| | (1.13) | (0.92) |
| Age at Purchase | 29.79 | 40.79 |
| | (7.81) | (4.95) |
| Age at Contract End | 62.53 | 60.18 |
| | (3.75) | (2.77) |
| Insured Benefits (monthly) | 1377.72 | 1553.75 |
| | (913.28) | (1242.95) |
| Insurance premium (monthly) | 77.82 | 106.67 |
| | (51.86) | (77.50) |
| Stand-Alone DI contract | 0.55 | 0.57 |
| | (0.50) | (0.50) |
| Observations | *confidential* | 18,659 |

**Panel B: Public DI Administrative Data**

| | (1) All DI Claims | (2) Own-Occupation DI Claims | (3) Cohorts 1959-1962 |
|---|---|---|---|
| Male | 0.59 | 0.82 | 0.53 |
| | (0.49) | (0.39) | (0.50) |
| Married | 0.66 | 0.77 | 0.51 |
| | (0.47) | (0.42) | (0.50) |
| Benefit claiming age | 51.80 | 53.84 | 43.34 |
| | (7.66) | (6.32) | (5.52) |
| Monthly benefit (Euros) | 1,077.85 | 867.57 | 856.94 |
| | (606.83) | (500.49) | (433.90) |
| Average monthly earnings before claim | 2,304.71 | 2,737.17 | 2,164.25 |
| | (1,109.40) | (1,010.85) | (1,230.59) |
| Monthly earnings in year before claim | 1,306.87 | 1,536.96 | 1,217.28 |
| | (1,026.46) | (1,101.52) | (1,005.15) |
| Education (years) | 10.39 | 10.35 | 10.64 |
| | (1.19) | (1.11) | (1.48) |
| Observations | 4,138,105 | 411,141 | 304,095 |

Table 3.3.2: Crowding-In: Difference-in-Differences

*The table shows results from the difference-in-difference regressions as described by equation* (3.1). *Regressions are run at the level of cohort* $\times$ *calendar month cells. Robust standard errors in parantheses. *** p<0.01, ** p<0.05, * p<0.1.*

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Public DI Claims | | Private DI Contracts | | |
| | Own-Occupation DI Claims | All Public DI Claims | Number of Purchases: All Contracts | Number of Purchases: Stand-Alone | Insured Benefits (All Contracts) |
| Treated×post | -50.57*** | -110.2*** | 15.11*** | 13.22*** | -462.2 |
| | (1.710) | (6.355) | (2.739) | (1.676) | (384.1) |
| Observations | 480 | 480 | 480 | 480 | 480 |
| R-squared | 0.935 | 0.990 | 0.939 | 0.939 | 0.926 |
| Mean (pre-reform) | 26.70 | 410.4 | 23.49 | 6.640 | 10,236 |
| Calendar month FE | yes | yes | yes | yes | yes |

Next, the main outcome of interest is the number of private DI purchases. To analyze these, we turn to the insurer microdata where we can observe individual characteristics. Panel (b) of Figure 3.3.2 depicts the number of private DI purchases by cohorts 1961-1962 vs. 1959-1960 over time.[10] Before the first announcement of the reform demarcated by the dashed vertical line, purchases by treated and control cohorts follow a very similar trend. After the first announcement, there is a clear increase in private DI purchases by both groups. This is consistent with the initial reform proposal affecting all cohorts. However, a clear differential increase in purchases by the treated cohorts occurs when the reform is implemented in 2001. Moreover, the differential effect on new contract purchases of the treatment group seems to persist in subsequent years. Column (3) of Table 3.3.2 presents the estimated effect on monthly private DI purchases. The coefficient of 15.1 is highly significant and corresponds to a 64% increase over pre-reform average monthly purchases of 23.5. In addition, Column (4) shows that the effect is mostly driven by newly purchased stand-alone DI contracts, where the estimated coefficient is 13.2. This suggests that individuals specifically buy additional DI contracts after the reform, rather than bundling DI with other insurance types. Finally, Column (5) shows the estimated effect on the amount of benefits insured in private DI contracts.

[10]The figure shows the annual number of private DI purchases, since the monthly contract data exhibits strong seasonality. Table 3.3.2 shows all effects estimated at the monthly level.

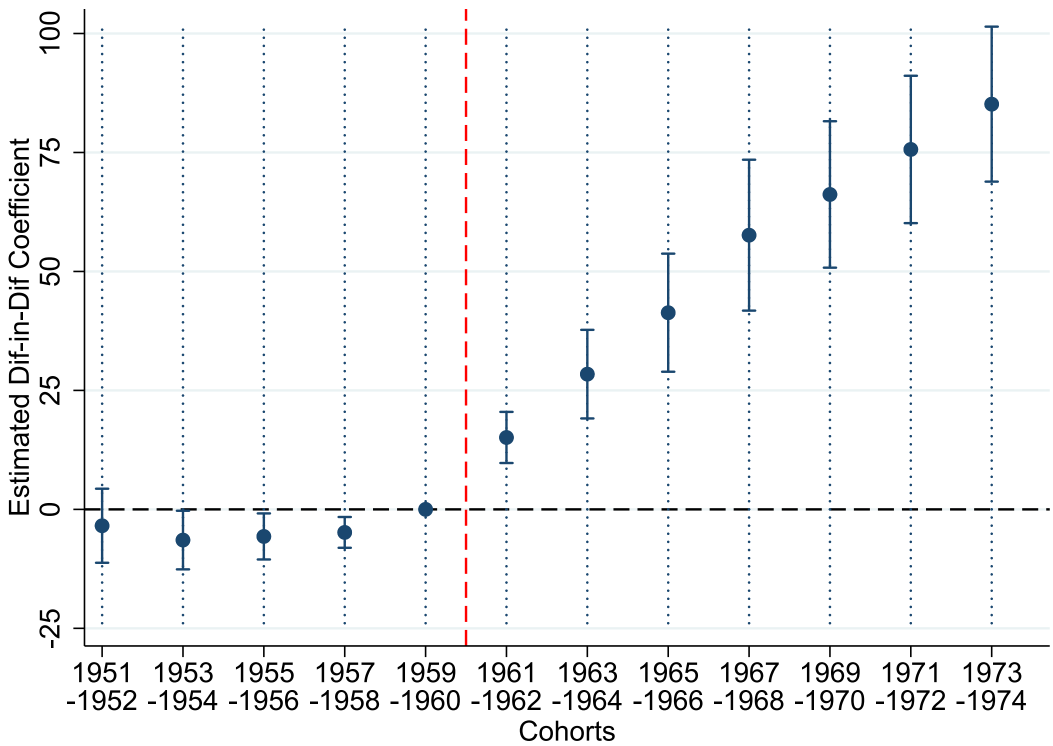
We find no significant effect along this "intensive margin" of private DI.[11] This motivates our focus on the extensive margin given by private DI take-up throughout this paper.

Appendix Table 3.A3 shows that these difference-in-difference results are robust to various alternative specifications. First, even though the treated and control cohorts in the baseline estimation are quite close in terms of age, there could be age-specific trends in private DI purchases. Panel A shows results from regressions based on equation (3.1) including cohort-specific linear trends. The estimated effect remains similar, and if anything the point estimates become slightly larger. Second, as explained in Section 3.2.1, the reform was first announced to take effect in 1998, but then retracted and re-announced for 2001. In the baseline estimation, the post-reform period is defined as January 2001 and later. This may understate the reform impact, as the initial announcement may already have an effect on private DI purchases. Panel B of Table 3.A3 shows difference-in-difference coefficients under different timing assumptions, including controlling for the period 1998 to 2000 with a separate indicator, omitting the years 1998 to 2000 or defining post-1998 as the post-reform period. Again, the estimated coefficients are slightly larger than the baseline effects, corresponding to increases between 72% and 81% relative to pre-reform purchases. Our baseline difference-in-difference estimation focuses on a narrow cohort window around the reform cutoff. This has the advantage of comparing relatively similar treated and control cohorts over time. However, this strategy is likely to lead to conservative estimates due to the age composition of the treatment group. Cohorts 1961-1962 are 39 to 40 years old at the time of the reform, while most individuals tend to purchase private DI at younger ages. In the full sample, the average purchase age is below 30 (see Table 3.2.1). In order to assess how the reform affects younger workers, we repeat the difference-in-difference estimation for a broader set of cohorts. Figure 3.3.3 shows estimated coefficients by cohort, where we replace the treated group in equation (3.1) by the respective cohorts denoted on the horizontal axis. Two main results emerge from the figure. First, the reform effect appears to be strongly increasing among younger cohorts. For instance, workers aged 29 to 30 at the time of the reform (cohorts 1971 to 1972) exhibit a roughly five times larger increase in the number of private DI purchases than the baseline treatment group. Second, the figure shows very small differences in private DI purchases between different cohorts born before the reform cutoff. Only our baseline control group exhibits a very small increase relative to

[11] See Appendix Figure 3.A3 for graphical results corresponding to Columns (2), (4) and (5) of Table 3.3.2.

Figure 3.3.3: Difference-in-Difference Effects by Cohort

*The figure shows difference-in-difference coefficients for a range of cohorts. The estimates correspond to coefficient* $\beta_2$ *from equation* (3.1)*, where the treatment group is given by the cohorts reported on the horizontal axis. Point estimates are shown along with 95% confidence intervals. The vertical line denotes the cohort cutoff of the reform of 2001, where all cohorts to the right are affected by the reform.*



older cohorts, but there are no differential trends in insurance purchases between cohorts further below the cutoff.

Finally, the difference-in-difference estimates are not directly comparable to overall take-up rates shown in Section 3.3.1, but a back-of-the-envelope calculation can illustrate such a comparison. For instance, we can calculate the predicted number of contracts held by cohorts 1961-1962 in 2015 based on pre-reform mean purchases, and add the estimated differential increase in purchases in the post-reform years. This would imply a 26% increase in the stock of private DI contracts held by the baseline treatment group who were treated at ages 39 to 40. Performing a similar calculation among the full set of treated cohorts from Figure 3.3.3 suggests a substantially larger rise in average private DI take-up by 193%. This magnitude is similar to the overall increase in private DI take-up from Figure 3.3.1, indicating that much of this growth can be attributed to an effect the reform.

## 3.4 Selection into Private Disability Insurance

### 3.4.1 Calculating Take-Up of Subgroups

In this section, we study which individuals select into private DI. The main challenge in doing so is that comprehensive microdata on the overall private DI market is not available. This challenge is faced by much of the literature investigating private insurance markets, which typically uses data from a specific insurer or employer (e.g. Einav, Finkelstein, and Cullen 2010, Autor, Duggan, and Gruber 2014, Cabral and Cullen 2019). We follow a similar approach and resort to the insurer microdata. Specifically, our goal is to use this data to calculate private DI take-up rates of subgroups:

$$Q_{g,t} = \frac{C_{g,t}}{N_{g,t}}$$

where $C_{g,t}$ denotes the number of private DI contracts held by subgroup $g$ at time $t$ and $N_{g,t}$ is the size of the respective subgroup. The denominator $N_{g,t}$ is relatively straightforward to obtain. We calculate sub-population sizes by cohort and gender from social insurance statistics. For the distribution of income, education and risk groups, we use the administrative public pension data, where income and education is observed and risk groups can be assigned based on occupations.

The key difficulty in calculating $Q_{g,t}$ lies in the numerator, as market-level data on the total number of contracts held by subgroups is not available. Using the insurer microdata, we calculate the number of contracts held by subgroup $g$ as

$$C_{g,t} = \sum_j \frac{c^j_{g,t}}{marketshare^j_t} \tag{3.2}$$

where $c^j_{g,t}$ is the number of contracts of type $j \in \{\text{stand-alone}, \text{bundled}\}$ within the insurer and $marketshare^j_t$ is the insurer's market share in the respective type of contract in year $t$. The approach requires the following assumption: Within type of contract and year, the market share of the insurer is constant across subgroups, i.e. $marketshare^j_{g,t} = marketshare^j_t \ \forall g$.

This assumption is certainly not innocuous, and its validity hinges on how representative the insurer is for the overall market. In Section 3.5, we present comprehensive validation checks of the resulting take-up rates. We find similar take-up patterns using representative household survey data and other independent data sources, confirming that the selection results we find in this section are present in the overall private DI market.

### 3.4.2 Selection on Observable Characteristics

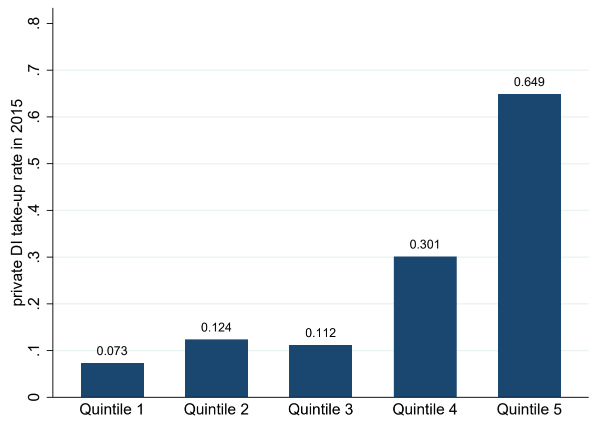
Figure 3.4.4 shows private DI take-up rates by observable characteristics, specifically by income, education, gender and risk group. All take-up rates are calculated in 2015, 15 years after the reform. To begin with, Panel (a) shows take-up rates by income quintile. The figure shows a striking positive correlation between private DI take-up and income. In the top income quintile, almost two thirds (65%) of individuals hold private DI. Private DI take-up in the fourth quintile is 30%, in the second and third quintiles take-up is 11% to 12%, and only 7% of individuals in the bottom quintile are covered by private DI.[12] Panel (b) shows an even stronger correlation of private DI take-up and education. 80% of individuals in the highest education quintile hold private DI, while take-up is 26% in the fourth quintile. In the bottom three quintiles, only 5% to 8% take up insurance. Panel (c) shows corresponding results by gender, suggesting that take-up is among men (30%) is somewhat higher than among women (20%). Next, we investi-

[12] Autor, Duggan, and Gruber (2014) similarly find that high-income individuals are more likely to take up private DI in the U.S.

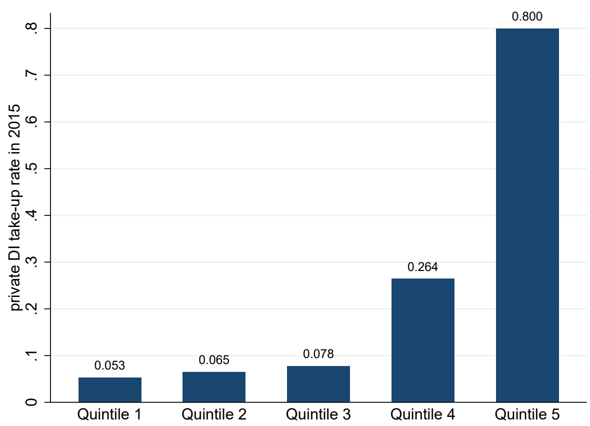
Figure 3.4.4: Private DI Take-Up by Observable Characteristics

*The figure shows private DI take-up rates in 2015 by income quintile (Panel a), education quintile (Panel b), gender (Panel c) and risk group (Panel d). In Panel (b), education is defined as years of schooling. All take-up rates are calculated as shown in equation* (3.2).

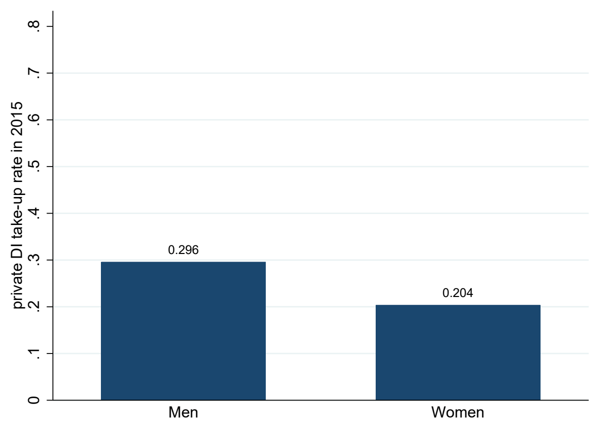
(a) By Income



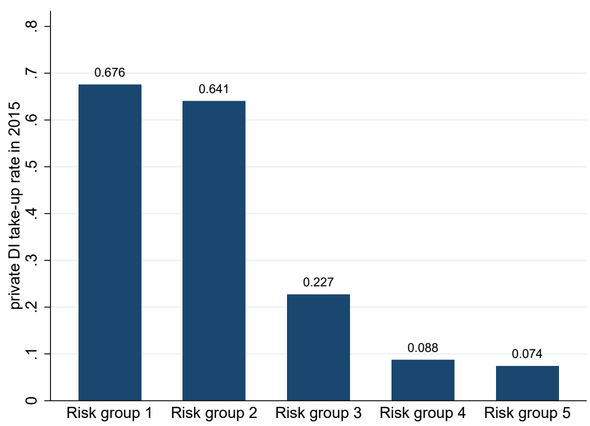

(b) By Education




(c) By Gender




(d) By Risk Group



gate private DI take-up by priced risk group. Recall that the insurer assigns individuals to one of five risk groups based on occupations, and these risk groups are the primary determinant of private DI premiums. Appendix Table 3.A4 summarizes risk groups. As expected, risk groups differ markedly in terms of lifetime disability risk, which we measure as the fraction of individuals claiming DI in the administrative data. Disability risk of individuals in risk group 1 is less than 5%, while it is 15% in risk group 2, 24% in risk group 3, 31% in risk group 4, and 40% in risk group 5. Moreover, the share of own-occupation DI claims increases with risk groups. For instance, only 8% to 11% of all DI claims in risk groups 1 and 2 are due to own-occupation disability, while the fraction is 32% in risk group 5. Accordingly, individuals are charged strongly varying insurance premiums depending on the risk groups they are assigned to. To insure EUR 1000 of monthly benefits at the age of 25, a worker in risk group 1 has to pay a monthly premium of EUR 32, compared to EUR 42 in risk group 2, EUR 68 in risk group 3, EUR 101 in risk group 4 and EUR 155 in risk group 5. Thus, premiums increase with risk groups roughly in line with disability risk, but there are some differences in pricing relative to risk which we revisit in Section 3.6.2. It is also worth noting that the population shares of risk groups differ substantially. 10% of the labor force work in an occupation in risk group 1, 17% in risk group 2, 35% in risk group 3, 38% in risk group 4, and only 0.6% in risk group 5. Finally, Panel (d) of Figure 3.4.4 shows a striking negative relationship between private DI take-up and risk groups. 68% and 64% of individuals in risk groups 1 and 2 hold private DI, respectively. Among risk group 3, private DI take-up is 23%, and only 9% and 7%, respectively, of individuals in risk groups 4 and 5 are covered by private DI.

These selection results have two key implications. First, they suggest that modest overall private DI take-up is driven by low take-up among individuals with low income, low education and high disability risk. On the other hand, there are groups with high insurance take-up of up to 80%, in particular the top income and education quintiles and the lowest risk group 1. These observations provide a first indication of potential equity issues in the private DI market, as vulnerable groups are much more likely to be without coverage. Second, low observed take-up among high-risk individuals is somewhat puzzling. Premiums are increasing with risk groups in a fashion not far from actuarially fair, and if individuals are well-informed about their risk, willingness to pay for insurance should increase with risk group. One potential explanation for the strong decline of take-up with risk groups is that individuals misperceive their risk, where high-risk individuals may under-estimate risk in particular. We return to this issue in Section 3.7.3.

As a complementary piece of evidence on heterogeneity in private DI take-up, we repeat the difference-in-difference analysis for each subgroup. Appendix Table 3.A5 shows results from estimating equation (3.1) separately by income, education, gender and risk group. The table reveals heterogeneity in crowding-in effects similar to simple differences in take-up. The estimated effect of the reform of 2001 on private DI purchases increases strongly with income and education, both in terms of absolute coefficient magnitudes and relative to pre-reform purchases. The effect on purchases by men is slightly larger than by women. Finally, the effects by risk groups have to be interpreted in relation to the size of each group. While raw coefficients are largest for risk groups 2 and 3, the increase in private DI purchases relative to group size are largest among risk groups 1 and 2. Strikingly, the reform seems to have lead only to a negligible number of additional purchases by individuals in the highest risk groups 4 and 5.

### 3.4.3 Risk-Based Selection

A crucial question for the efficient functioning of private DI markets is whether individuals select into purchasing insurance based on their risk. The classic theory of adverse selection predicts that high-risk individuals are more likely to purchase insurance, which leads to underprovision of insurance or even complete market unravelling (Akerlof 1970, Rothschild and Stiglitz 1976). To investigate this question, we implement a *positive correlation test* (**Chiappori2000**, Einav, Finkelstein, and Cullen 2010, Landais et al. 2021). The goal is to test whether there is a correlation between private DI take-up and unpriced risk, where a positive correlation would indicate adverse selection. Specifically, we run the following regression at the occupation level:

$$Q_j = \beta_0 + \beta_1 \pi_j + \beta_2 riskgroup_j + \epsilon_j \tag{3.3}$$

where $Q_j$ denotes private DI take-up of individuals in three-digit occupation $j$ in 2015, $\pi_j$ is a measure of disability risk in the occupation, and $riskgroup_j$ is the risk group assigned to the occupation by the insurer.[13]

[13] Note that risk groups are not necessarily the same for all individuals within a three-digit occupation for two reasons. First, the insurer sometimes changes the risk group assigned to an occupation over time.

Two features of this specification are worth emphasizing. First, we found a strong negative correlation of private DI take-up and risk groups in the previous section. Risk groups reflect an observed component of risk based on which the insurer prices contracts. However, in assessing whether there is adverse selection, it is key to estimate the correlation of private DI take-up and *unpriced* risk. Thus, the idea behind equation (3.3) is that $\beta_1$ captures selection on unpriced risk, after controlling for priced risk given by risk groups. Second, a potential pitfall of the correlation test is that ex-post measures of risk based on observed insurance claims may confound selection on ex-ante risk and moral hazard responses (see e.g. Landais et al. 2021). A correlation of DI take-up and claiming probabilities may be driven by certain risk types selecting into insurance (selection) or those with more insurance coverage becoming more likely to claim (moral hazard). In order to address this challenge and isolate risk-based selection, we calculate take-up among treated cohorts 1961 and younger, but we measure disability risk $\pi_j$ as the fraction claiming DI only among control cohorts 1960 and older. This risk measure should not be confounded by differential moral hazard, since all individuals in the control cohorts are still fully covered by public own-occupation DI, i.e. they are observed under the same insurance coverage. Figure 3.4.5 depicts the estimation results in binned scatterplots. First, Panel (a) shows the unconditional correlation of occupation-level private DI take-up and disability risk. This corresponds to estimating equation (3.3) without controlling for risk groups. There is a highly significant negative relationship between DI take-up and risk, with a slope coefficient of -1.38. This overall correlation is driven by a mixture of the negative relationship of DI take-up and risk groups documented in Figure 3.4.4, and any correlation of take-up and unpriced risk. Next, panel (b) of the figure shows the correlation of private DI take-up and unpriced risk, after controlling for priced risk. The relationship is remarkably flat, and the estimated slope coefficient corresponding to $\beta_1$ in equation (3.3) is small and statistically insignificant. In other words, we do not find any evidence of adverse selection from the point of view of the insurer: within priced risk groups, individuals with higher true disability risk are no more likely to select into purchasing insurance.

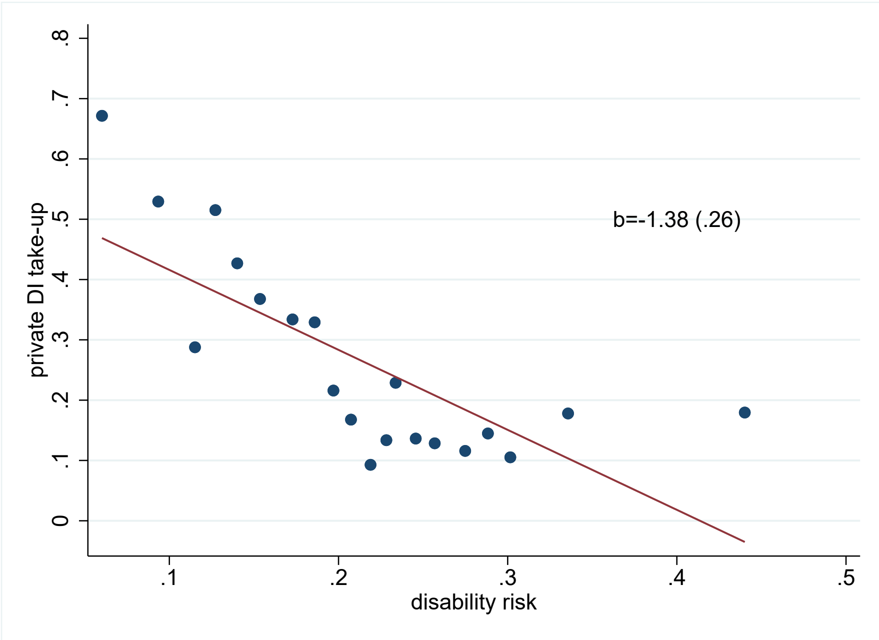
Appendix Table 3.A6 presents regression results based on equation (3.3). Column (1) shows the specification without controlling for risk groups, corresponding to Fig-

Second, occupation titles considered by the insurer may feature finer-grained distinctions not captured by the occupation classification, such as whether the individual mostly works inside an office. For the results shown here, we assign the average risk group to each occupation. Results remain very similar when considering the modal risk group within occupation.

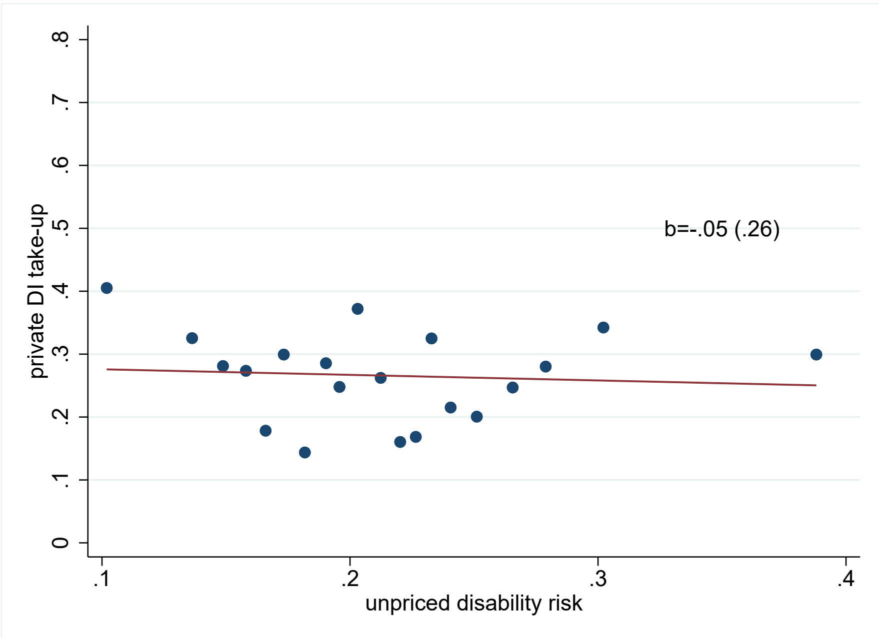
Figure 3.4.5: Risk-Based Selection

*The figure shows binned scatterplots of the correlation between private DI take-up in 2015 and disability risk at the three-digit occupation level. Panel (a) shows the unconditional correlation between take-up and risk, corresponding to estimating equation* (3.3) *without controlling for risk groups. Panel (b) shows the correlation between take-up and unpriced risk, controlling for risk groups.*

(a) Unconditional Correlation of Take-Up and Risk




(b) Selection on Unpriced Risk



ure 3.4.5, Panel (a). Column (2) shows a strong negative correlation between private DI take-up and risk groups, as expected from Figure 3.4.4. Column (3) shows the specification corresponding to Figure 3.4.5, Panel (b), where the relationship between private DI take-up and risk becomes insignificant once risk groups are controlled for. If anything, the point estimate on risk is slightly negative, which would imply advantageous selection into private DI. We also note that neither the coefficient on risk groups nor the explanatory power of the regression change much between Columns (2) and (3), which suggests that the overall negative correlation between risk and private DI take-up is fully driven by differences in take-up across risk groups.

In Columns (4) to (7) of Table 3.A6, we then add observable characteristics to the regression. This yields two additional insights. First, we can explore how risk-based selection changes conditional on different sets of observables. In Column (4), controlling for income hardly changes the coefficient on risk. However, Column (5) suggests that education may be a driver of advantageous selection. Once we control for education, the coefficient on risk turns sizeable and positive, albeit still insignificant due to a sizeable standard error. This indicates that the insurer may face adverse selection if pricing was conditional on education. In practice, not conditioning on education induces some advantageous selection, where individuals with higher education (who are less risky on average) are more likely to buy insurance, such that there is no overall adverse selection. In Column (6), controlling for gender does not alter selection much. Interestingly, adding further observables including economic training, marital status and an indicator for East Germany in Column (7) again turns the effect of risk close to zero and negative, suggesting that these characteristics may drive some adverse selection.

Second, Table 3.A6 is informative of which characteristics themselves predict private DI take-up. In Section 3.4.2, we show that income, education and risk groups exhibit a strong univariate correlation with take-up, but one may ask which of these remain "deep" predictors conditional on risk and other observables. Columns (5) to (7) suggests that income itself is not a significant driver of private DI take-up, once education and risk groups are controlled for. On the contrary, education remains highly positively correlated with take-up in all specifications. Similarly, although the effect of risk group somewhat shrinks when adding socioeconomic controls, it remains a significant negative predictor of take-up. Interestingly, working in an economically trained occupation has a positive impact on take-up beyond the influence of education. Column

(7) additionally indicates that private DI take-up is lower among females and married individuals.

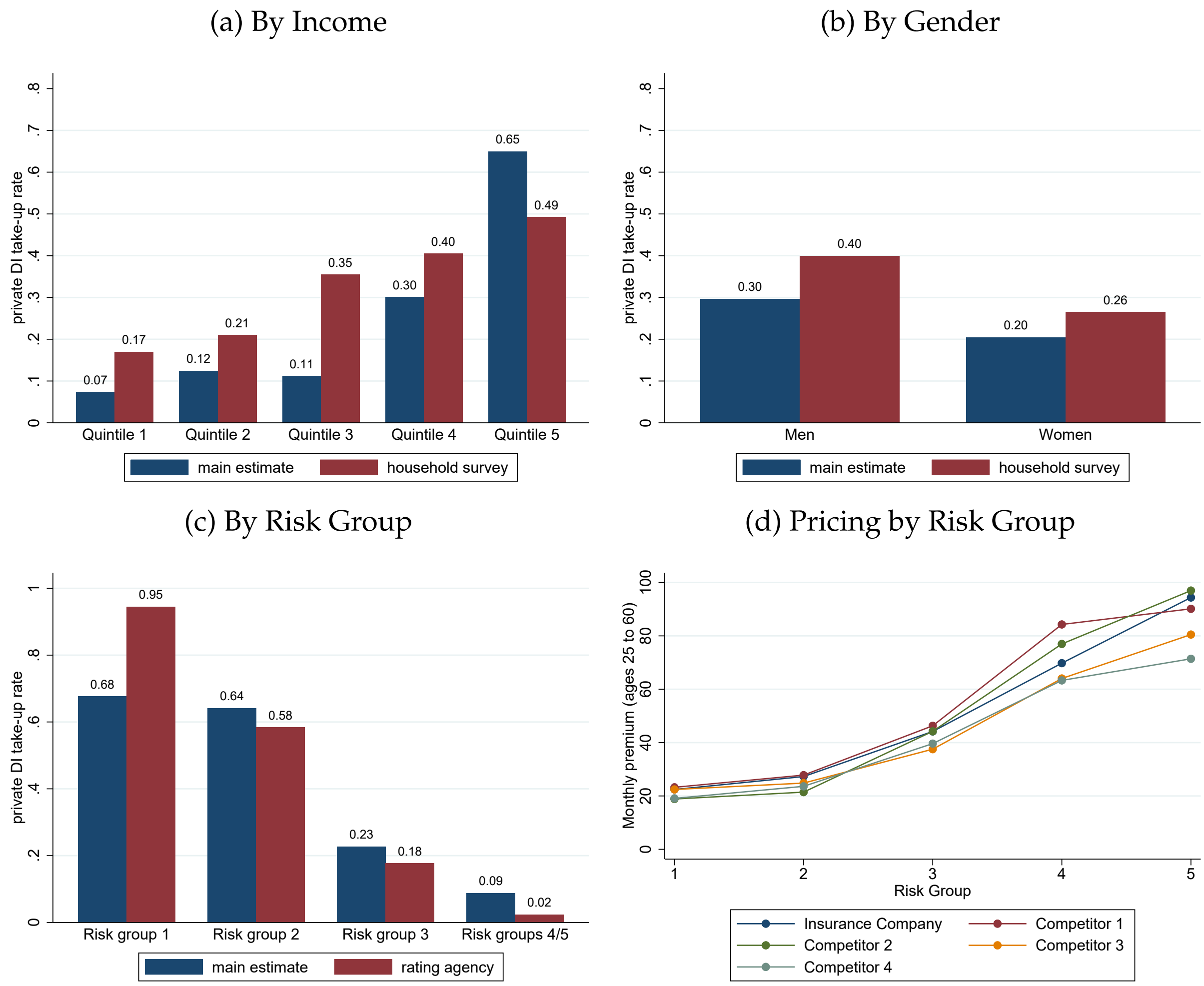
## 3.5 Validation of Empirical Results

Our empirical results on selection into private DI rely on the insurer microdata, as individual-level data on the entire market is not available. As discussed in Section 3.4.1, the validity of these findings depends on how representative the insurer is for the overall market. In this section, we present a number of validation checks using additional, independent data sources. Overall, we find similar patterns based on these alternative sources, confirming the validity of our main results.

To begin with, the overall private DI take-up we find is very similar to available independent survey results. For instance, a survey conducted by TNS Infratest (2015), a private survey company, found that 26% of working adults hold private DI, corresponding precisely to our main take-up rate estimate for the same year from Section 3.3.1. Moreover, we use data from the Income and Consumption Survey (EVS), a representative household survey conducted by the German Federal Statistical Office. According to the survey, overall private DI take-up by German households is 31% in 2013. The household-level figure is naturally somewhat larger than our individual-level estimate, since the average household has around two members (see Appendix Table 3.A2) any of whom may have individual private DI contracts. Nevertheless, take-up according to the household survey is remarkably close to our main estimate. Next, we turn to private DI take-up by subgroups. Panel (a) of Figure 3.5.6 shows take-up rates by income quintile in the household survey in comparison to our main estimates. The survey data confirms a clear positive relationship between private DI take-up and income. However, the gradient is somewhat less pronounced in the survey than in the main results. For instance, private DI take-up in the bottom quintile is 17% in the survey and 7% according to our main result, and in the top quintile it is 49% in the survey and 65% according to our main results. These differences could occur for two main reasons. First, we have to calculate both private DI take-up and income quintiles at the household level rather than at the individual level in the survey data. These household-level characteristics may mask additional variation across individuals. Second, the information in the survey is self-reported, and thus measurement error may be present to a larger extent than in the insurer microdata. This may further attenuate the relationship between private DI

Figure 3.5.6: Validating Take-Up Rates

*The figure collects various pieces of evidence supporting the validity of our main empirical results. Panels (a) and (b) show a comparison of the take-up rates we find based on the insurer microdata (blue bars) to take-up rates based on representative household survey data (red bars), by income quintile (Panel a) and gender (Panel b). Panel (c) compares take-up rates by risk group based on the insurer microdata (blue bars) to take-up rates based on the rating agency data (red bars). The rating agency data uses four harmonized risk groups, and we assign risk groups 4 and 5 from the insurer microdata to the fourth harmonized risk group. Panel (d) shows average monthly insurance premiums charged to the ten most frequent occupations in each risk group by the insurer providing our microdata and four large competitors.*



take-up and income.[14] Panel (b) of the figure shows a comparison of private DI take-up by gender. In the survey, 40% of households with a male household head have private DI, and the corresponding fraction is 26% with a female household head. This relative difference is similar to the private DI take-up of 30% and and 20% we find based on the insurer microdata.

In order to validate private DI take-up rates by risk groups, we use the rating agency data, which includes the shares of contracts by "harmonized" risk groups for the entire market. This information is based on insurers reporting the number of contracts in four risk groups defined by the rating agency. To our knowledge, these harmonized risk groups correspond largely to the risk groups used by the insurer providing our microdata, but the insurer additionally differentiates the fourth harmonized group into high (risk group 4) and very high risk (group 5). Panel (c) of Figure 3.5.6 shows implied take-up rates by harmonized risk group in comparison to our main estimates. Overall, we find a stark decreasing pattern of private DI take-up with harmonized risk groups. For the largest, medium-risk groups 2 and 3, take-up rates implied by the rating agency data and our main estimates are virtually the same. For the low-risk group 1 and the high-risk groups 4 and 5, the rating agency data displays even stronger heterogeneity in take-up than our main results. In harmonized risk group 1, the rating agency data implies a take-up rate close to 1 (95%), and in harmonized risk group 4, take-up is only 2%, compared to 8% in the insurer microdata. Of course, a caveat with using harmonized risk groups is that we do not know exactly how other insurers assign occupations to those, and thus the precise take-up rates should be taken with a grain of salt. Yet, the rating agency data confirms our result that private DI is predominantly taken up by low-risk groups.

Finally, as an additional piece of evidence, Panel (e) of Figure 3.5.6 shows a comparison of private DI pricing by different insurers. For this exercise, we web-scraped data on prices charged to the ten most frequent occupations in each risk group for those of the top-10 insurers offering online price calculators. The figure plots the average monthly premium by risk group for the insurer providing our microdata and four large com-

[14] We find similar differences in households' private DI take-up across education levels of the household head in the survey data. For instance, take-up is 47% in the highest category (graduate degree) and 18% in the lowest category (no vocational education). We do not include these results in Figure 3.5.6 since the self-reported education categories are not directly comparable to the quintile measure we use for the main results. Nevertheless, we note that the household survey confirms the positive relationship between private DI take-up and income.

petitors. In general, relative prices charged to different occupations are similar across insurers. There seem to be some differences in the level of insurance premiums, but all insurers levy similar relative risk surcharges on higher-risk occupations. This suggests that individuals in certain risk groups should have little reason to select specifically into the insurer providing the microdata, as its insurance pricing is representative of the overall market.

## 3.6 Value and Cost of Disability Insurance

### 3.6.1 Basic Conceptual Framework

Next, our aim is to quantify the value and cost of DI coverage offered by the private market, which are key inputs into assessing the welfare consequences of policy interventions in the provision of DI. Based on these two components, we can calculate the *net value* of DI, which we define as the value to recipients relative to the cost of insurance (see Section 3.7.1). Our analysis builds on Einav, Finkelstein, and Cullen (2010), who show that in order to evaluate welfare in insurance markets, the key sufficient statistics are given by insurance demand and cost curves. Similar frameworks have recently been used in related social insurance contexts, including DI and unemployment insurance (Cabral and Cullen 2019, Landais et al. 2021, Hendren, Landais, and Spinnewijn 2020).

Following this literature, we consider a population of heterogeneous individuals indexed by $\theta_i$, and $F(\theta_i)$ denotes the distribution of the population. Heterogeneity is unrestricted, and may include variation both in preferences for DI, such as varying risk aversion, and variation in individual disability risks. The first key component for welfare analysis is demand, or willingness to pay, for DI. Denote by $v(\theta_i)$ the utility of consumer $i$ from buying disability insurance, and by $p_k$ the insurance premium charged to individuals in risk group $k$. In a private market with insurance choice, the individual purchases DI if $v(\theta_i) \geq p_k$. Aggregate demand for private DI in group $k$ can be written as

$$D_k(p_k) = \int \mathbb{1}\left(v(\theta) \geq p_k\right) dF_k(\theta) = \Pr_k\left(v(\theta_i) \geq p_k\right)$$

In words, insurance demand corresponds to the share of individuals whose willingness to pay is above the premium within a given risk group.

The second component we require for welfare analysis is the cost of providing DI. We denote by $c(\theta_i)$ the expected cost associated with the potentially insured risk of individual $i$. Average cost at price $p_k$ is

$$AC_k(p_k) = \frac{1}{D_k(p_k)} \int c(\theta)\mathbb{1}\left(v(\theta) \geq p_k\right) dF_k(\theta) = \mathbb{E}_k\left(c(\theta_i)|v(\theta_i) \geq p_k\right)$$

Thus, the average cost curve is determined by the cost of providing insurance to those individuals who choose to buy insurance at a given price $p_k$. In addition, we can write marginal cost as $MC_k(p_k) = \mathbb{E}_k\left(c(\theta_i)|v(\theta_i) = p_k\right)$. The marginal cost curve captures the cost of providing insurance to the marginal individuals who purchase insurance exactly at price $p_k$.

Before we proceed to the empirical implementation, three aspects are worth noting. First, we assume that individuals make a discrete choice of whether to buy insurance or not (if such choice is permitted), and we abstract from the choice of insured benefit amounts in private DI contracts. This assumption is motivated by our results from Section 3.3.2, which suggest that individuals mainly respond along this extensive margin of insurance choice, whereas no significant responses occur along the intensive margin of insured benefits. Second, we follow the literature regarding the cost of providing DI and abstract from any other cost incurred by insurers, such as administrative cost. Third, since insurance prices depend on risk groups to which the insurer assigns individuals based on observable characteristics (occupations), we conduct the analysis separately for each risk group. In other words, the insurance demand and cost curves described above apply within risk groups where individuals vary only in unpriced characteristics.

### 3.6.2 Estimating Demand and Cost Curves

The first ingredient for welfare analysis is demand, or willingness to pay for DI. Our post-reform setting with insurance choice provides a unique opportunity to implement a revealed preference approach and directly estimate individual valuations of the DI coverage offered by the private market. Such an opportunity is rarely available, as public DI is fully mandated in most countries, leaving little choice for workers.[15] In particular, we use two empirical moments to estimate demand for DI. First, the observed post-reform take-up rate at given prices identifies one point on the demand curve of

[15] Cabral and Cullen (2019) follow a closely related but distinct approach, estimating a lower bound on the willingness to pay for public DI using supplemental private DI purchases of workers at a U.S. employer.

each risk group, anchoring its level. For this purpose, we can directly use the observed take-up rates shown in Panel (c) of Figure 3.4.4. Second, to estimate the slope, i.e. the responsiveness of demand to prices, we exploit the discontinuous price variation between risk groups. Assuming a constant elasticity of demand then allows us to construct demand curves of each risk group.

The slope of the demand curve captures the responsiveness of private DI take-up to insurance prices. To estimate such price responses, we run the following regression at the occupation level:

$$Q_j = \beta_0 + \beta_1 \pi_j + \sum_{k=2}^{5} \delta^k \mathbb{1}(riskgroup_j = k) + Z_j' \gamma + \epsilon_j \tag{3.4}$$

where $Q_j$ denotes private DI take-up by three-digit occupation $j$, $\pi_j$ is a measure of disability risk, $\mathbb{1}(riskgroup_j = k)$ is an indicator for occupation $j$ being assigned to risk group $k$ by the insurer and $Z_j$ is a vector of control variables. Again, we measure take-up among treated cohorts in 2015 and disability risk only among control cohorts. Equation (3.4) captures the idea that a discrete number of risk groups are assigned to occupations based on a continuous running variable, namely occupation-level disability risk $\pi_j$. Thus, at the boundaries between risk groups, similar occupations with very similar or even the same disability risk are assigned to different risk groups and thus face different prices. The coefficients $\delta^k$ capture the jump in private DI take-up between risk groups $k$ and $k-1$ conditional on underlying risk, which we interpret as a response to the local, discrete difference in insurance premiums between the two groups.

This specification is similar to equation (3.3), but there are two important differences. First, we include indicators for risk groups in order to separately estimate the jump in private DI take-up for each adjacent pair of risk groups. In order to better capture the discrete variation between risk groups, we additionally define risk groups as the modal risk group within each occupation. Second, our preferred specification includes control variables $Z_j$, such as income, gender and education. We do not include these characteristics in the main correlation test based on equation (3.3), since they are not priced by the insurer. However, it can be important to add these controls in equation (3.4) if occupations in different risk groups differ in terms of observable characteristics in a way correlated with private DI take-up.

Based on the estimated regression coefficients, we can then calculate the demand elasticity at the boundary between risk groups $k$ and $k-1$ as

$$\hat{\varepsilon}^k = \frac{(\hat{\delta}^k - \hat{\delta}^{k-1})/\overline{Q_j}^{k,k-1}}{\Delta p^{k,k-1}/\overline{p_j}^{k,k-1}} \tag{3.5}$$

where $\overline{Q_j}^{k,k-1}$ and $\overline{p_j}^{k,k-1}$ are average private DI take-up and average premiums among occupations belonging to risk group $k$ and $k-1$, respectively, and $\Delta p^{k,k-1}$ is the difference in premiums between groups $k$ and $k-1$.[16] Figure 3.6.7 illustrates the estimation graphically. In Panel (a), we rank occupations by disability risk within risk group in order to depict the variation in prices and DI take-up in a stylized way. The blue line shows the sizeable jumps in premiums between risk groups. The black dashed line shows a linear fit of private DI take-up within risk group, revealing large jumps in take-up at the risk group boundaries. The elasticity calculation in equation (3.5) relates these jumps in demand to the price variation between the respective groups. Next, Panel (b) shows binned scatterplots of private DI take-up by actual disability risk, corresponding directly to the estimation from equation (3.4). Similarly to Panel (b) of Figure 3.4.5, the relationship between DI take-up and underlying disability risk is slightly downward-sloping within risk group. There appears to be sizeable overlap in underlying risk across risk groups. On the one hand, this is perhaps surprising as one may expect the insurer to assign risk groups in a less "fuzzy" way.[17] On the other hand, the large overlap implies that there are many instances of occupations with the same disability risk facing different premiums, providing us with sufficient statistical power to estimate price responses. Indeed, the figure indicates clear, large jumps in private DI take-up conditional on underlying risk across all adjacent risk group pairs, suggesting sizeable demand responses of demand to insurance premiums.

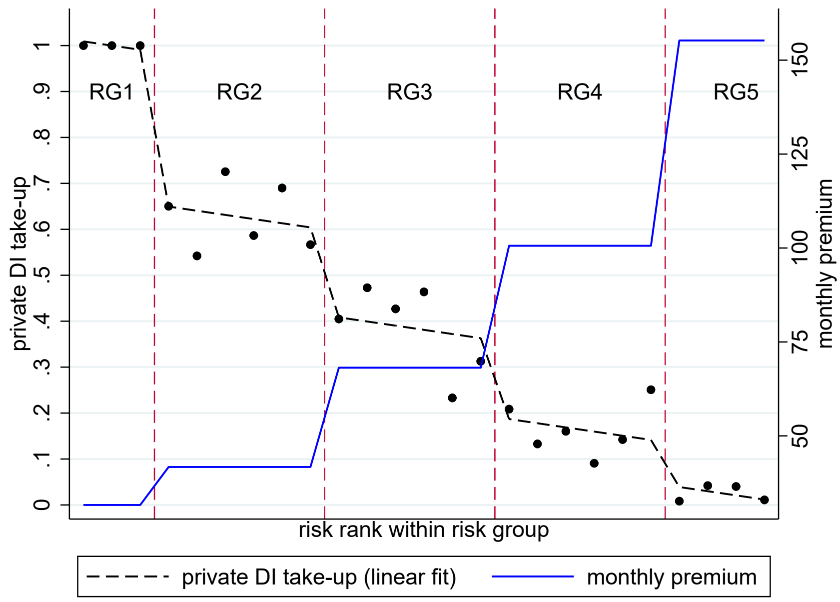
[16]In contrast to the expected price calculation described in equation (3.7), we calculate $\Delta p^{k,k-1}$ and $\overline{p_j}^{k,k-1}$ directly based on monthly insurance premiums charged to the respective risk groups. We do this because the relevant jump in prices at the risk group boundaries is the percentage change in premiums conditional on risk, which is directly given by the percentage change in monthly premiums.

[17]One potential reason for the fuzziness in risk group assignment is that the insurer may not have had sufficiently comprehensive data on lifetime DI claiming probabilities by occupation at the time. This argument is consistent with the fact that the insurer carried out a major overhaul of risk groups for new private DI contracts after the end of our sample period.

Figure 3.6.7: Demand Responses to Insurance Prices

*The figure presents evidence of demand responses to insurance premiums. In Panel (a), we rank three-digit occupations by disability risk within risk for a stylized depiction of jumps in premiums and take-up rates between risk groups. The blue line shows monthly private DI premiums, which increase discontinuously at the risk group boundaries. The black dots denote average private DI take-up in risk bins, and the dashed black line shows a linear fit within risk group. Panel (b) shows binned scatterplots of private DI take-up by disability risk at the three-digit occupation level, corresponding to the regression shown in equation* (3.4).

(a) Take-Up vs. Price by Risk Ranks




(b) Take-Up by Actual Risk

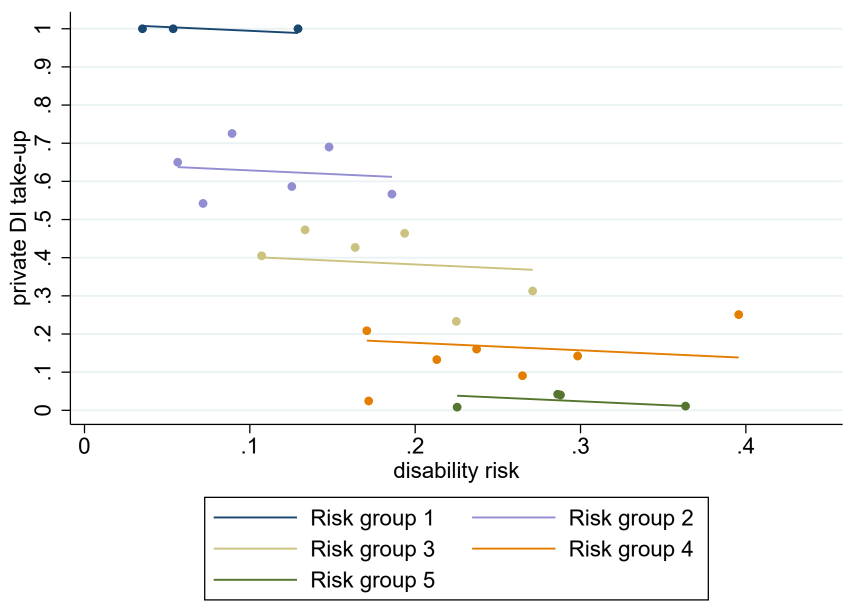


Table 3.6.3: Demand Elasticity Estimation

*The table shows results from the demand elasticity estimation. The first row shows the percentage change in price between adjacent risk group pairs. The next two rows show estimates of the corresponding percentage change in private DI take-up. The estimates are based on regression results shown in Appendix Table 3.A7. "Without controls" indicates that the respective figure is obtained from a regression without controls. "With controls" indicates that income, education, gender, marital status, economic training and residence in East Germany are included as controls. The bottom two rows show elasticity estimates, relating the respective percentage change in take-up to the percentage change in price as shown in equation (3.5). For each outcome, Column (1) shows the weighted average of the estimates among the different risk group pairs from Columns (2) to (5). Bootstrapped standard errors are shown in parantheses.*

| | (1) Average | (2) Groups 1-2 | (3) Groups 2-3 | (4) Groups 3-4 | (5) Groups 4-5 |
|---|---|---|---|---|---|
| dp/p | 0.398 | 0.246 | 0.439 | 0.370 | 0.536 |
| dQ/Q | | | | | |
| Without controls | -0.675 | -0.563 | -0.495 | -0.814 | -0.829 |
| | (0.051) | (0.133) | (0.152) | (0.153) | (0.088) |
| With controls | -0.468 | -0.274 | -0.141 | -0.571 | -0.885 |
| | (0.111) | (0.210) | (0.170) | (0.152) | (0.321) |
| Elasticity | | | | | |
| Without controls | -1.791 | -2.285 | -1.129 | -2.201 | -1.548 |
| | (0.146) | (0.541) | (0.347) | (0.415) | (0.165) |
| With controls | -1.155 | -1.110 | -0.322 | -1.542 | -1.646 |
| | (0.293) | (0.852) | (0.388) | (0.412) | (0.597) |

Table 3.6.3 shows results from the demand elasticity estimation.[18] The average price difference between adjacent risk groups is 40%, and the average unconditional jump in private DI take-up at the risk group boundaries corresponds to a 68% reduction in demand for insurance. Including controls (income, gender, education, economic training, marital status and residence in East Germany) yields a response of 47%. The demand elasticity estimation then relates the demand response to the jump in price for each pair of adjacent risk groups. In on our preferred specifications including controls, we find an average demand elasticity across all risk groups of -1.16. Without controlling for observables, the average elasticity is -1.79. Elasticity estimates among the different risk groups are close to the average, except the estimate between risk groups 2 and 3 where we find a smaller elasticity of -0.32. Overall, there is no clear increasing or decreasing pattern of elasticities with risk groups. This motivates our assumption of a constant

[18] In addition, we show regression results directly corresponding to equation (3.4) in Appendix Table 3.A7.

elasticity along the demand curve.[19]

The second ingredient required for welfare analysis is the cost of providing disability insurance. We calculate the expected cost of insuring individual $i$ belonging to risk group $k$ as

$$c_{i,k} = \sum_{t=0}^{T_i} \Pi_{k,t} b_i \delta_t \tag{3.6}$$

where $T_i$ is the contract end date relative to a contract start date normalized to zero, $\Pi_{k,t}$ is the cumulative disability risk among risk group $k$ in period $t$, $b_i$ is the level of insured benefits, and $\delta_t = \frac{1}{(1+r)^t}$ is a discount factor. We use a discount rate of $r = 3\%$ and as before, we measure disability risk as the ex-post realized risk of claiming DI benefits in the administrative data. Appendix Figure 3.A4 shows empirical risk paths for each risk group. As expected, lifetime disability risk increases strongly with risk groups (see also Appendix Table 3.A4). Risk paths by age evolves quite similarly across groups, with most disability claims occurring between between ages 45 and 60. We calculate $c_{i,k}$ for each individual in the insurer microdata, and then take the average expected cost within risk group. To construct average cost curves, it is crucial that we do not find evidence of adverse or advantageous selection in Section 3.4.3. Since there is no significant correlation between private DI take-up and disability risk within risk group, average costs are constant with respect to the level of demand, resulting in flat cost curves. Moreover, as average cost is constant, average cost and marginal cost curves coincide. Finally, two important features of cost curves are worth noting. First, the cost estimates can be interpreted as inclusive of a fiscal externality due to moral hazard responses to DI coverage, since our risk measure is based on ex-post observed claims. Second, we assume that the cost of providing insurance is the same across private and public DI systems.[20]

Throughout the subsequent analysis, we consider prices in terms of expected insur-

[19] Alternatively, the literature often assumes a linear demand curve (e.g. Einav, Finkelstein, and Cullen 2010, Landais et al. 2021). In our case, the magnitude of demand responses estimated at different risk group cutoffs suggest a constant elasticity may be a better approximation than a linear curve.

[20] Unfortunately, the insurer microdata does not provide information on claims over a sufficiently long period to directly compare private and public DI claims. However, some aggregate calculations on private DI claiming risk are provided by the German Actuarial Society (**DAV2018**). Panel (f) of Appendix Figure 3.A4 shows private DI claiming risk from this source, calculated for a representative individual. There are some differences in the timing of claims, but overall disability risk is remarkably similar to observed in public DI claims, providing suggestive evidence that our assumption of equal cost is likely a good approximation.

ance premiums paid by individuals and received by the insurer:

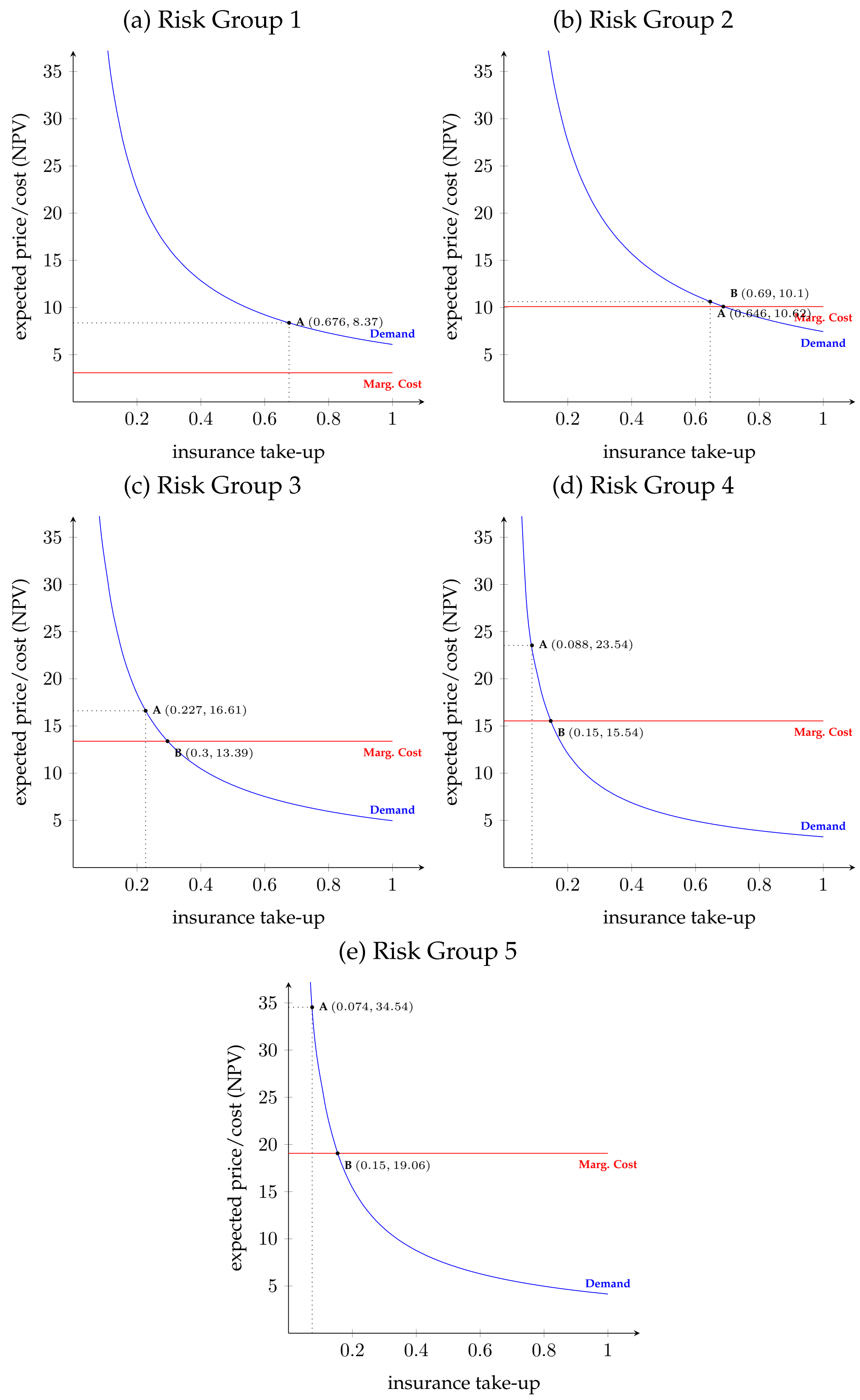
$$p_{i,k} = \sum_{t=0}^{T_i}(1 - \Pi_{k,t})\tilde{p}_k\delta_t \tag{3.7}$$

where $\tilde{p}_k$ is the per-period premium charged to risk group $k$. Again, we calculate $p_{i,k}$ for each individual in the insurer microdata and take average expected premiums by risk group $p_k = \mathbb{E}_k(p_{i,k})$. Thus, willingness to pay for insurance and the welfare measures described below are expressed in terms of certainty equivalents. Figure 3.6.8 plots the estimated demand and cost curves by risk group. In each panel, the horizontal axis denotes the fraction of the respective risk group covered by private DI, ranging from zero to one. Demand curves rank individuals from high to low willingness to pay on the horizontal axis and show the fraction of individuals whose willingness to pay is at least equal to a given price. Cost curves show the marginal/average cost associated with insuring the set of individuals willing to purchase insurance at this price. In Panel (a), the expected cost of insuring individuals in risk group 1 is low as this group faces the lowest disability risk. The estimated willingness to pay is above the cost of providing insurance at any level of take-up. Panel (b) shows corresponding results for risk group 2, for whom the cost of insurance is already substantially higher. The demand curve also indicates somewhat higher willingness to pay for DI among risk group 2, but demand and cost curves intersect at an insurance take-up rate of 69%. Thus, willingness to pay is below the cost of insurance for 31% of individuals. In Panel (c), the cost of insuring risk group 3 is higher again, while the demand curve is lower than that of risk group 2. In fact, willingness to pay is above cost for only 30% of individuals in risk group 3. Similarly, in Panels (d) and (e), risk groups 4 and 5 are even costlier to insure, but willingness to pay revealed by observed demand and price responses are low. Thus, the cost of insurance is above willingness to pay for 85% of individuals in the two highest-risk groups.

In addition, Figure 3.6.8 is informative of the difference between premiums charged in the private DI market and the expected cost of insuring each risk group. There are notable differences in implied profit markups across risk groups. Premiums are substantially above expected costs for risk group 1, indicating sizable profits from insuring the lowest-risk individuals. For risk group 2, on the other hand, premiums are very

Figure 3.6.8: Demand and Cost Curves

*The figure presents own-occupation DI demand and cost curves estimated as described in Section 3.6.2. The horizontal axes display private DI take-up rates between zero and one, and the vertical axes show expected prices and cost as defined in equations (3.6) and (3.7). Each panel shows the demand curve (blue line) and the marginal/average cost curve (red line) for the risk group indicated in the panel title. Points A denotes the private market equilibrium in each risk group, with associated insurance take-up and price in parantheses. Points B denote the points of intersection of demand and marginal cost curves, associated take-up and price in parantheses.*



close to actuarially fair.[21] Similarly, the markup is modest for risk group 3. For risk group 4 and especially risk group 5, markups appear to be larger again.

Appendix Table 3.A8 further quantifies the value and cost of insurance implied by the estimated curves. Willingness to pay and cost in the table are calculated for a private DI contract insuring a 30% income replacement rate and scaled relative to lifetime income. Across all groups, median willingness to pay is 0.93% of income and the expected cost of providing this coverage is 1.47% of income. In line with strongly varying disability risk across groups, we estimate an insurance cost of 0.33% of income in risk group 1, 1.09% in risk group 2, 1.47% in risk group 3, 1.72% in risk group 4, and 2.14% in risk group 5. On the contrary, median valuations do not appear to increase with risk. Our estimates suggest a willingness to pay for private DI of 1.13% of income in risk group, 1.42% in risk group 2, 0.96% in risk group 3, 0.63% in risk group 4, and 0.82% in risk group 5.

### 3.6.3 Decomposing Willingness to Pay

So far, we estimate willingness to pay and cost for the full coverage provided by private DI in the post-2001 setting. This includes coverage of own-occupation DI risk, but private DI can also serve as a top-up insurance if the worker qualifies for public DI in the case of a general disability. In this section, we propose a decomposition of DI valuations into these two components, exploiting differences in insurance take-up over time.

We begin by writing an individual's total disability risk as the sum of two-sub risks: $\pi = \pi_g + \pi_o$, where $\pi_g$ is the risk of a general disability, and $\pi_o$ is the risk of an own-occupation disability. In the post-reform setting, an observed willingness to pay for private DI captures the sum of valuations for own-occupation DI coverage and top-up insurance for general disability risk:

$$v^{post} = v_g(b_g, \Delta) + v_o(0, \Delta)$$

where $v_j(b_j, \Delta)$ denotes the valuation for an amount $\Delta$ of private insurance against risk $\pi_j$, $j \in g, o$, given public DI coverage $b_j$ against that risk.[22] In the pre-reform setting, on

[21] In fact, when the insurer carried out an overhaul of risk groups after the end of our sample period, one major goal was to introduce more fine-grained groups to replace the former risk group 2. This is consistent with the pricing of risk group 2 not being fully optimal from the point of view of the insurer.

[22] For simplicity, we drop type $\theta$ from the notation here.

the other hand, private DI is purely a top-up insurance, such that

$$v^{pre} = v_g(b_g, \Delta) + v_o(b_o, \Delta)$$

Thus, the difference in willingness to pay post-reform vs. pre-reform can be interpreted as a lower bound on the valuation for insurance against own-occupation disability risk:

$$v^{post} - v^{pre} = v_o(0, \Delta) - v_o(b_o, \Delta) \leq v_o(0, \Delta) \tag{3.8}$$

Furthermore, we can obtain an upper bound on the valuation for own-occupation DI. For this, we assume that the drop in consumption upon own-occupation disability is smaller or equal to the drop in consumption upon general disability. This is likely to hold, since individuals can still work in other occupations in the event of own-occupation disability, while general disability requires being unable to work in any occupation. The assumption implies $v_o(b_o, \Delta) \leq \frac{\pi_o}{\pi} v^{pre}$,[23] and in turn

$$v^{post} - \frac{\pi_g}{\pi} v^{pre} \geq v_o(0, \Delta) \tag{3.9}$$

Hence, the difference between post-reform willingness to pay and the share $\pi_g/\pi$ of pre-reform willingness to pay provides an upper bound on valuations for own-occupation DI. Finally, the corresponding fraction of pre-reform willingness to pay can be interpreted as a lower bound on the valuation for top-up insurance against general DI risk:

$$\frac{\pi_g}{\pi} v^{pre} \leq v_g(b_g, \Delta) \tag{3.10}$$

To empirically implement this decomposition, we construct pre-reform demand for private DI based on observed pre-reform take-up by risk group (see Appendix Figure 3.A5), using the elasticity estimates from Section 3.6.2. Results from the decomposition are shown in Panel A of Appendix Table 3.A8. We find a median valuation for own-occupation DI between 0.43% and 0.50% of lifetime income, and a lower bound on the valuation for top-up insurance against general DI risk of 0.43%. Thus, roughly half of the post-reform willingness to pay for private DI is attributed to insurance against own-occupation disability risk. Moreover, the estimates suggest that valuations for

[23]To see this, note that $v_o(b_o, \Delta) \approx \frac{\pi_o}{\pi} v^{pre}$ if the drop in consumption upon own-occupation and general disability was the same. If the drop in consumption upon own-occupation disability is smaller, insurance against this risk becomes less valuable, such that $v_o(b_o, \Delta) < \frac{\pi_o}{\pi} v^{pre}$.

own-occupation DI decrease with risk groups, whereas general DI valuations tend to increase with risk groups. Panel B of the table additionally shows a decomposition of the cost of private DI into own-occupation DI and the general DI top-up. We calculate these costs analogously to equation (3.6), using observed shares of claims of the two types of DI. Since own-occupation DI accounts for a modest share of all claims (see Appendix Table 3.A4), the expected cost of providing own-occupation DI is 0.19% of lifetime income, compared to 1.28% for general DI.

## 3.7 Welfare Effects of Privatizing Disability Insurance

### 3.7.1 Baseline Welfare Calculations

Based on these estimated demand and cost curves, we can assess welfare in the private DI market. As our main welfare measure, we define the *net value* as the value of DI to the insured relative to the cost to the insurer. In the private market where individuals have the choice whether to purchase DI coverage, the net value is given by

$$NV^{priv} = \frac{\sum_k n_k \Big[ \int v(\theta) \mathbb{1}(v(\theta) \geq p_k) dF_k(\theta) \Big]}{\sum_k n_k \Big[ \int c(\theta) \mathbb{1}(v(\theta) \geq p_k) dF_k(\theta) \Big]} \tag{3.11}$$

where $n_k$ denotes the size of risk group $k$. In the market, the net value is thus given by the value of DI to those choosing to take it up, i.e. for whom $v(\theta) \geq p_k$, divided by the cost of providing DI to them. Since we estimate private DI valuations in the presence of baseline public DI coverage, $NV^{priv}$ should be interpreted as the net value of extra coverage provided by the private market.

Our main counterfactual of interest is the introduction of an insurance mandate providing the level of coverage offered by the private DI market to all workers. Starting from the private market equilibrium, the net value of introducing the mandate is

$$\Delta NV^{mand} = \frac{\sum_k n_k \Big[ \int v(\theta) \mathbb{1}(v(\theta) < p_k) dF_k(\theta) \Big]}{\sum_k n_k \Big[ \int c(\theta) \mathbb{1}(v(\theta) < p_k) dF_k(\theta) \Big]} \tag{3.12}$$

A mandate ensures all individuals are covered, but it leads to some crowding out of existing private insurance. Individuals whose willingness to pay is above the market price already purchased private DI, and the mandate expands coverage to those individuals whose willingness to pay is below the market price.[24]

Our net value measures express the value of providing insurance per Euro of spending, analogously to the marginal value of public funds (Finkelstein and Hendren 2020, Hendren and Sprung-Keyser 2020). A reform can be deemed welfare-improving if its net value is greater than one, i.e. it generates value exceeding its costs.[25] For our counterfactual, $\Delta NV^{mand} > 1$ would imply that mandating the coverage offered by private DI (on top of the existing baseline public DI coverage) is welfare-improving, while $\Delta NV^{mand} < 1$ would imply that providing this extra coverage via the private market is preferable. These welfare effects can be graphically illustrated using the demand and cost curves estimated in Section 3.6.2. Panel (a) of Figure 3.7.9 depicts the net value provided by the private DI market for the case of risk group 3. The total area under the demand curve up to equilibrium take-up corresponds to the numerator in equation (3.11), and the area under the marginal cost curve corresponds to the denominator. In addition, the figure shows the standard decomposition of willingness to pay into consumer surplus (area $A$ between willingness to pay and the price), producer surplus (area $B$ between the price and marginal cost) and cost (area $C$ below the marginal cost curve). Thus, net value in the private DI market is the sum of areas $A$, $B$ and $C$ divided by total cost $C$. Appendix Figure 3.A6 shows analogous graphs for all risk groups. The private DI market generates a surplus, as those individuals with the highest willingness to pay choose to purchase private DI. Consumer surplus is particularly large in risk groups 1 and 2, where individuals exhibit the highest valuations of insurance. Producers receive the largest surplus from risk groups 1, 4 and 5, where markups are highest.

Panel (b) of Figure 3.7.9 illustrates the welfare effects of introducing a mandate starting from the private market, again for the case of risk group 3. Insuring all individuals entails additional costs given by the area under the cost curve between equilibrium take-up and complete take-up of 100%. This corresponds to the sum of areas $F$ and $G$. Expanding insurance to additional consumers yields value $D+G$, but they have to pay premiums equal to areas $D+E+F+G$, implying a net loss in consumer surplus of $-(E+F)$. Insurers, on the other hand, gain surplus equal to area $D+E$. Thus, the over-

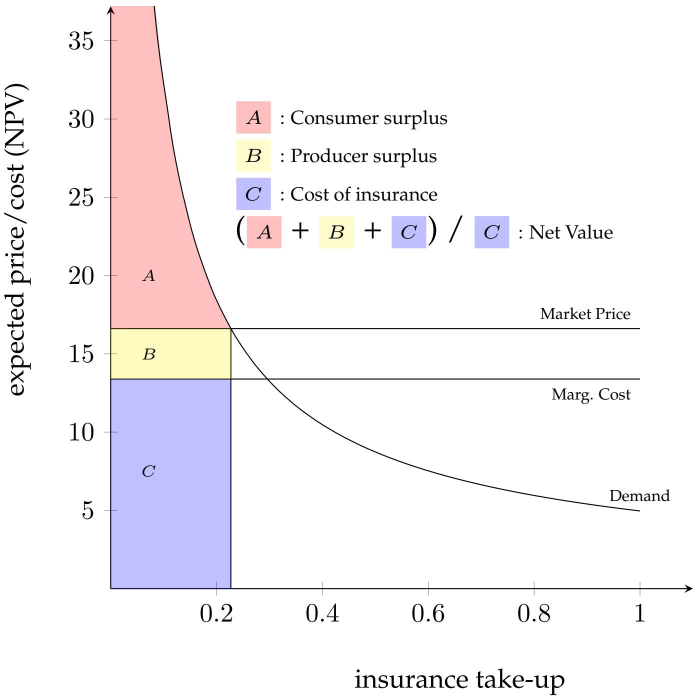
[24] In the absence of a private DI market, the net value of the mandate would be given by $NV^{mand} = \frac{\sum_k n_k \left[\int v(\theta) dF_k(\theta)\right]}{\sum_k n_k \left[\int c(\theta) dF_k(\theta)\right]}$.

[25] Instead of dividing the value of insurance by its cost, we could alternatively calculate the difference between the two. In this setting, we prefer to take the net value as the ratio of the two, since the resulting numbers are unit-free and easily interpretable.

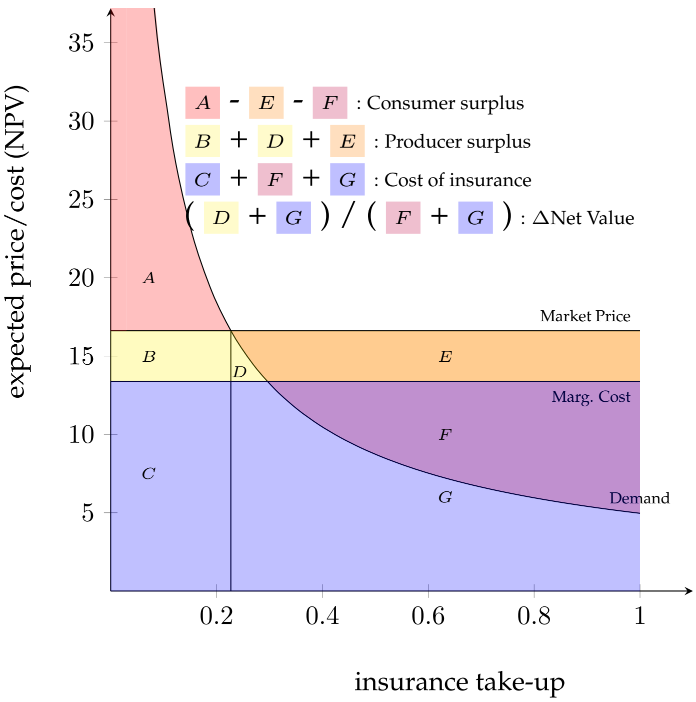
Figure 3.7.9: Welfare Calculations

*The figure illustrates our welfare calculations for the case of risk group 3. Panel (a) depicts welfare in the private DI market equilibrium, where the net value is given by the total area under the demand curve ($A+B+C$) divided by the area under the cost curve ($C$). Panel (b) illustrates the net value of introducing a DI mandate. The mandate increases DI take-up from the market equilibrium to 1. The net value of the reform is given by the additional area under the demand curve ($D+G$) divided by the additional cost ($F+G$). In both panels, net value can be further decomposed as explained in the respective legend. See Appendix Figures 3.A6 and 3.A7 for graphs for all risk groups.*

(a) Private DI Market




(b) Net Value of Introducing a DI Mandate



all net value of the mandate is given by $D + G$ relative to $F + G$, which is clearly below one. Appendix Figure 3.A7 shows corresponding graphs for all risk groups. The net value of a mandate is below one for all groups except risk group 1. Mandating private DI coverage would have sizeable negative welfare effects for higher risk groups in particular, since the observed willingness to pay is low relative to cost for most individuals in these groups.

Table 3.7.4: Welfare Effects of Insurance Mandate

*The table shows the net value of mandating the DI coverage offered by the private insurance market. Panel A shows the baseline net value calculated as shown in equation* (3.12)*. Panel B shows the social net value calculated as in equations* (3.13) *and* (3.14)*, under different social welfare functions indicated in the row titles. Panel C shows the net value under risk misperceptions, based on calibrated normative insurance valuations from Appendix Table 3.A10.*

| | (1) | (2) | (3) |
|---|---|---|---|
| | Private DI Mandate | Public DI Mandate | |
| | | Lump-Sum Contributions | Income-Based Contributions |
| **Panel A: Baseline Calculation** | | | |
| Net Value | 0.762 | 0.762 | 0.762 |
| **Panel B: Social Net Value** | | | |
| Utilitarian, $\sigma$=1 | 0.704 | 0.941 | 1.218 |
| Utilitarian, $\sigma$=3 | 0.612 | 1.142 | 1.743 |
| Utilitarian, $\sigma$=5 | 0.549 | 1.221 | 1.960 |
| Utilitarian, $\sigma$=8 | 0.488 | 1.255 | 2.064 |
| Rawlsian | 0.131 | 1.455 | 2.328 |
| **Panel C: Net Value under Risk Misperceptions** | | | |
| Hand-to-mouth ($\sigma$=0.44) | 1.506 | 1.506 | 1.506 |
| Hand-to-mouth + SA floor ($\sigma$=0.84) | 1.100 | 1.100 | 1.100 |
| High $\Delta C$ + SA floor ($\sigma$=1.16) | 1.469 | 1.469 | 1.469 |
| Low $\Delta C$ + SA floor ($\sigma$=3.03) | 1.418 | 1.418 | 1.418 |

Panel A of Table 3.7.4 shows results of our baseline net value calculation based on equation (3.12). We find a net value of introducing a private DI mandate of 0.76. In a way, this result is not too surprising given our empirical findings. First, we do not find adverse selection, which would lead to inefficiently low insurance take-up in the private market, and which is often considered a key rationale for a mandate. Second, insurance premiums are only somewhat above marginal costs for most risk groups. Accordingly, the private DI market seems to cover the majority of individuals whose willingness to pay is above the cost of insuring them. Third, the value of own-occupation DI revealed by insurance choices appears to be low for many individuals, especially in the higher risk groups. This is reflected both by the low general level of willingness to pay and by the sizeable demand elasticities, which imply that the valuation of insurance declines fast among the uninsured.

Overall, our baseline welfare calculations suggest that starting from a full public DI mandate, partly privatizing DI is welfare-improving. Conceptually, these results are closely related to the reform of 2001, which privatized insurance against own-occupation disability. However, it is important to note that the counterfactual should be interpreted as a broader reform, removing own-occupation risk coverage while also cutting benefit levels.[26] In the following sections, we consider two extensions that may justify a full mandate, namely equity concerns and risk misperceptions.

### 3.7.2 The Social Value of a DI Mandate

A first potential rationale for mandating additional DI coverage may be equity concerns. Recall that the private DI market disproportionately covers high-income and low-risk individuals. A mandate would extend coverage to more low-income and high-risk individuals, on whom a social planner concerned with equity may place particular weight. In order to account for such distributional issues, we write the social net value of introducing a mandate as:

$$\Delta SNV^{mand} = \frac{\sum_k n_k \left[ \lambda_k \overbrace{\int (v(\theta) - p_k)\, \mathbb{1}(v(\theta) < p_k) dF_k(\theta)}^{\text{Consumer surplus}} + \overbrace{\int p_k \mathbb{1}(v(\theta) < p_k) dF_k(\theta)}^{\text{Insurer revenue}} \right]}{\sum_k n_k \left[ \int c(\theta) \mathbb{1}(v(\theta) < p_k) dF_k(\theta) \right]} \tag{3.13}$$

The first term in the numerator captures the additional net utility individuals in risk group $k$ derive under a mandate, corresponding to their valuation minus the price. The total change in consumer surplus among risk group $k$ is multiplied by $\lambda_k$, the social welfare weight on individuals in this group. The second term in the numerator reflects additional revenue to the insurer, corresponding to the sum of producer surplus and cost in Figure 3.7.9. Like our baseline measure, the social net value then relates these two components to the change in the cost of providing insurance.[27]

Equation (3.13) considers a private insurance mandate where individuals are compelled to purchase private DI at market prices. However, in our setting, extra DI cover-

[26] The main reason why we focus on the counterfactual corresponding to a broader reform is that most of our empirical results apply to the DI coverage offered by the private market. This allows us to credibly calculate the welfare effects of different ways of providing this coverage, while analyzing the welfare effects of sub-components would require additional assumptions.

[27] Both insurer revenue and cost carry a weight of one, corresponding to the average social welfare weight in the population.

age was part of the social insurance system before the reform of 2001, where employed individuals are mandated to participate and pay social insurance contributions rather than risk-based premiums. In order to evaluate such a public insurance mandate, we have to take into account that contributions may differ from market prices $p_k$:

$$\Delta SNV^{pub} =$$

$$\frac{\sum_k n_k \left\{ \lambda_k \left[ \int (v(\theta) - p_k)\, \mathbb{1}(v(\theta) < p_k) dF_k(\theta) + \overbrace{\int (p_k - p_k^{pub}) dF_k(\theta)}^{\text{Pricing effect}} \right] + \int p_k^{pub} \mathbb{1}(v(\theta) < p_k) dF_k(\theta) \right\}}{\sum_k n_k \left[ \int c(\theta) \mathbb{1}(v(\theta) < p_k) dF_k(\theta) \right]} \tag{3.14}$$

where $p_k^{pub}$ denotes contributions paid by individuals in risk group $k$. Compared to equation (3.13), a public insurance mandate thus entails an additional pricing effect, where all individuals experience a change in surplus equal to the difference between private market premiums and social insurance contributions. In particular, we consider two scenarios of public mandates. On the one hand, the government may insure everyone in a public DI system with lump-sum contributions irrespective of risk and income. We calculate the required level of lump-sum contributions as the average cost of providing coverage equivalent to private DI across all risk groups. On the other hand, contributions could be income-based. This reflects the situation in typical real-world social insurance systems, where contributions are levied as a proportion of an individual's gross income. Again, we calculate the required contribution rate such that total contributions equal the cost of providing insurance to all individuals.

In order to obtain welfare weights, we require a social welfare function. As is common in the literature, we assume a Utilitarian social welfare function, such that welfare weights are given by the marginal utility from consumption in each group. Moreover, we assume constant relative risk aversion (CRRA) utility $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ with marginal utility $u'(c) = c^{-\sigma}$, where $\sigma$ is the coefficient of relative risk aversion. We then calculate social welfare weights for each risk group based on average expected lifetime income in that group, scaled such that the average weight in the population is equal to one. Appendix Table 3.A9 shows average income and resulting social welfare weights by risk group. Expected income decreases monotonically with risk groups. On average, individuals in risk group 1 earn more than double the income of those in risk group 5. We consider a range of values of risk aversion between 1 and 8, where higher $\sigma$ entails stronger higher relative welfare weights on higher-risk groups. In addition, we show results under a more extreme variant of equity concern given by Rawlsian social preferences, where the planner only places weight on the worst-off individuals in risk group 5.

Table 3.7.4 shows results from the social net value calculations. Column (1) suggests that a private DI mandate would lower welfare, regardless of the degree of equity concern. In fact, stronger equity concern decreases the social net value of a private DI mandate. This occurs because a private DI mandate is a regressive policy. As can be seen in Appendix Figure 3.A7, forcing all individual to purchase insurance at market prices entails larger reductions in consumer surplus among higher risk groups, since they have to pay higher prices relative to a low revealed willingness to pay. Column (2) shows welfare effects of a public DI mandate with lump-sum contributions. Note that for our baseline net value without social welfare weights, pricing in an insurance mandate does not affect welfare, as it does not affect total of surplus but only its distribution. However, with sufficient equity concern, a public DI mandate with lump-sum contributions can improve welfare relative to the private market. We find that for $\sigma$ between 2 and 3, the social net value of such a mandate becomes greater than one. Intuitively, lump-sum contributions imply redistribution towards higher-risk individuals on whom the planner places greater weight since they have lower expected lifetime income. In Column (3), this redistributive effect is exacerbated in the scenario with income-based contributions. Since lower-risk groups have higher average income, they now have to pay the highest contributions. Thus, the social insurance system with income-based contributions raises revenue from low-risk, high-income groups, and redistributes towards high-risk, low-income groups by providing them with additional insurance at premiums below risk-based market prices. This redistribution is highly valued by a social planner with equity concern. Even under low risk aversion given by $\sigma$=1, the social net value of the DI mandate with income-based contributions is above 1. For $\sigma$=3, the social net value is 1.74, and under a Rawlsian social welfare function the social net value is 2.33.

We conclude that equity concern can provide a rationale for including the DI coverage currently offered by the private market in the public DI mandate. For such a reform to improve social welfare, it is crucial to implement non-risk based contributions as is done in real-world social insurance systems. Instead enforcing a private insurance mandate would entail even greater welfare losses in the presence of equity concern than under pure efficiency considerations.

### 3.7.3 Risk Misperceptions

A second potential rationale for policy interventions in the DI market could be given by behavioral frictions. So far, our welfare analysis assumes that individuals make optimal insurance purchase decisions, such that we can interpret observed private DI demand as indicative of individuals' true valuations. However, a growing literature documents behavioral frictions in insurance choices (e.g. Ericson and Sydnor 2017, Chandra, Handel, and Schwartzstein 2019). In our setting, two observations point towards a role for such choice frictions. First, private DI take-up is positively correlated with education and economic training, conditional on income, risk and other observables. Thus, low take-up may be concentrated among individuals with low financial literacy who are less likely to make optimal insurance choices. Second, higher-risk groups who are charged higher insurance premiums are less likely to take up private DI. Accordingly, we find in Section 3.6.2 that willingness to pay for insurance does not increase with risk. Indeed, a number of surveys suggest that most German workers tend to underestimate disability risk (e.g. Swiss Life 2018, Forsa 2020), implying that they likely undervalue insurance.

The main empirical challenge is to disentangle such behavioral biases from variation in true risk preferences. Workers in higher risk groups may exhibit low willingness to pay for insurance because they misperceive their disability risk, or due to low risk aversion. In this section, we present calibration exercises approaching this challenge in three steps. First, we calibrate risk preferences implied by observed insurance purchase decisions in each risk group, and we argue that risk aversion appears to be implausibly low for many workers. Second, we calibrate a simple model of risk misperceptions which can rationalize low willingness to pay for insurance in higher risk groups. Third, we calculate the wedge between observed willingness to pay (with misperceptions) and normative willingness to pay (without misperceptions), and re-do welfare calculations based on normative valuations.

We begin by asking what level of risk aversion would be implied by observed insurance purchase decisions in each risk group. Individuals buy insurance if the discounted expected utility with insurance exceeds $V_1$ utility without insurance $V_0$. We can write an indifference condition for the marginal individual purchasing insurance as

$$\underbrace{\sum_{t=0}^{T} \delta^t \left[(1-\Pi_t)u(c_H^0) + \Pi_t u(c_L^0)\right]}_{V_0 \text{ (utility without DI)}} = \underbrace{\sum_{t=0}^{T} \delta^t \left[(1-\Pi_t)u(c_H^1) + \Pi_t u(c_L^1)\right]}_{V_1 \text{ (utility with DI)}} \tag{3.15}$$

where $T$ is the end date of the insurance contract relative to start date normalized to zero, $\Pi_t$ is cumulative disability risk in period $t$ and $\delta^t$ is a discount factor. $c_H^0$ and $c_L^0$ denote consumption levels when not disabled ($H$) and disabled ($L$), respectively, without insurance, and $c_H^1$ and $c_L^1$ denote the corresponding consumption levels with insurance.

For the calibration, we assume again CRRA preferences $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and we plug in average income, insured benefits, contract duration, and cumulative risk paths by risk group. Furthermore, differences in consumption levels across disabled and non-disabled states are a crucial input for the calibration. Incomes can be written as $y_H^0 = w$, $y_L^0 = \underline{y}_0$, $y_H^1 = w - p$ and $y_L^1 = \underline{y}_1 + b$, where $w$ is the individual's wage, $\underline{y}_0$ is an income floor for uninsured individuals, $p$ is the insurance premium, $b$ is the insured benefit, and $\underline{y}_1$ is the income floor for insured individuals (which may differ from $\underline{y}_0$, for instance due to a means test). We consider a range of consumption scenarios. To begin with, we consider hand-to-mouth consumers whose consumption equals income in each state, either with or without a consumption floor given by basic social assistance. In addition, we use estimates of the drop in consumption upon disability based on Meyer and Mok (2019).[28]

Under these assumptions, we can calibrate risk aversion $\sigma$ of the marginal buyer in each risk group. Results are shown in Appendix Table 3.A10. It is important to note that the marginal buyer whose risk aversion is calibrated is at very different percentiles of willingness to pay for DI across risk groups, as shown in Panel A. For instance, 68% of individuals in risk group 1 take up private DI and thus the marginal buyer is at the 32nd percentile of willingness to pay, whereas in risk group 5 take-up is only 7% such that the marginal buyer is at the 93rd percentile. In Panel B, depending on the assumption

[28] An important issue with the consumption drop estimates from Meyer and Mok (2019) is that these are reported for individuals covered by disability insurance. We are not aware of any estimates of the consumption drop upon disability in the absence of insurance. Thus, we choose two estimates from Meyer and Mok (2019) that may come closest to consumption drops without insurance. The first is their finding of a 77% drop in earnings before public transfers upon long-term disability. Second, Meyer and Mok (2019)) report a drop in income after public transfers of 28% and a corresponding drop in consumption of 25%, while the income drop before public transfers is 53%. A back-of-the-envelope calculation results in a hypothetical drop in consumption without public transfers of 53%·25%/28%=47%.

about consumption levels, we find implied risk aversion coefficients between 0.44 and 3.03 for the marginal individual in risk group 1. In the remaining risk groups, especially in groups 2 to 4, risk aversion implied by observed insurance take-up is considerably lower between 0.03 and 1.34. Interestingly, risk aversion does not appear to decrease monotonically with risk groups. Calibrated risk aversion in risk group 2 is particularly low, which is due to insurance premiums being close to actuarially fair for this group, such that even under modest risk aversion insurance take-up should be higher than the observed rate. We also note that risk aversion of the marginal buyer in group 5 is higher than in groups 2 to 4. In principle, insurance should be highly valuable to these high-risk individuals, but this is counteracted by two forces. First, insurance premiums for risk group 5 are high, even relative to their high disability risk, and second, basic social assistance provides sizeable insurance against inability to work in the absence of formal DI given their low average income. Finally, a direct comparison of the risk aversion estimates across groups is complicated by the fact that the marginal buyer in the high risk groups is at very high percentiles of willingness to pay. For instance, the risk aversion estimates of 0.26 to 1.87 in risk group 5 must be interpreted in the sense that 93% of individuals in this group have risk aversion of *at most* 0.26 to 1.87, whereas the risk aversion of 0.44 to 3.03 in risk group 1 applies to an individual closer to the median among this group. Overall, observed insurance choices would imply very low risk aversion for many individuals, especially in risk groups 2 to 4. The implied values for the coefficient of relative risk aversion are considerably lower than most estimates from the literature on insurance choices.[29]

In the second calibration step, our goal is to investigate whether risk misperceptions can rationalize low willingness to pay for DI exhibited by many individuals. We denote individuals' perceived disability risk by $\hat{\Pi}_t \neq \Pi_t$. In particular, we consider risk misperceptions of the form $\hat{\Pi}_t = \alpha\Pi_t$, where $\alpha$ denotes the degree of bias. The indifference condition governing insurance choice of the marginal buyer is

$$\underbrace{\sum_{t=0}^{T} \delta^t \left[(1-\hat{\Pi}_t)u(c_H^0) + \hat{\Pi}_t u(c_L^0)\right]}_{V_0(\hat{\Pi}_t)} = \underbrace{\sum_{t=0}^{T} \delta^t \left[(1-\hat{\Pi}_t)u(c_H^1) + \hat{\Pi}_t u(c_L^1)\right]}_{V_1(\hat{\Pi}_t)} \tag{3.16}$$

[29] Studies on insurance choices typically yield larger estimates of risk aversion ranging between 2 and 8 (e.g. French 2005, Lockwood 2018, Jacobs 2020, Landais et al. 2021) and some work implies much larger values (e.g. Cohen and Einav 2007; Sydnor 2010). Seitz (2021) estimates a coefficient of around 6 in the German setting, which is identified based on observed asset holdings.

Under the assumptions on the utility function and consumption levels described above, we can use equation (3.16) to calibrate $\alpha$ for the marginal buyer in each risk group. However, we additionally require a benchmark level of risk aversion. To obtain this, we assume that risk group 1 perceives disability risk correctly, and that other groups have the same true risk aversion as group 1 where we found values of $\sigma$ between 0.44 and 3.03. Panel C of Appendix Table 3.A10 shows resulting estimates of $\alpha$. Under virtually all specifications, we find that individuals in risk groups 2 to 5 substantially underestimate their disability risk. The proportional underestimation reflected by $\alpha$ is roughly between 30% and 60% in most specifications. Only under hand-to-mouth consumption and basic social assistance, risk groups 4 and 5 is found not to underestimate risk. We conclude that even under modest levels of true risk aversion, risk misperceptions can explain low observed valuations of DI.

In the third step, we calculate the wedge between observed willingness to pay and normative willingness to pay implied by these risk misperceptions. Observed willingness to pay is implied by the indifference condition $V_0(\hat{\Pi}_t) = V_1(\hat{\Pi}_t)$ ((3.16)), and corresponds to the empirical willingness to pay of the marginal buyer. Normative willingness to pay, on the other hand, is implied by $V_0(\Pi_t) = V_1(\Pi_t)$, that is the hypothetical indifference condition of the marginal buyer without any risk misperception. Panel D of Appendix Table 3.A10 shows estimated ratios between normative and observed willingness to pay. The results suggest that the true value of insurance to marginal buyers is up to 2.6 times higher than the valuation implied by observed choices. In line with the misperception results, we find that undervaluation tends to be most severe among risk groups 2 to 4.

Finally, we return to our welfare calculations. We can interpret the above results as an internality, where individuals do not internalize the full value of DI. In Panel C of Table 3.7.4, we show results from net value calculations based on equation (3.11), where we replace observed demand $v(\theta)$ in each risk group by normative valuations implied by the results from Panel D of Appendix Table 3.A10. We find a net value of mandating private DI coverage between 1.10 and 1.51. In other words, average normative valuations exceed the cost of providing insurance for individuals who choose not to buy private DI in the market. Hence, risk misperceptions can provide an additional rationale for mandating the coverage currently offered by the private DI market.

### 3.7.4 Extensions and Robustness

Our main welfare calculations compare the value to the direct cost of providing extra DI. However, there could be various types of indirect costs associated with increasing DI coverage via a mandate. In this section, we present extensions of the welfare analysis taking into account such indirect costs in the spirit of a more complete marginal value of public funds calculation. Overall, we find that allowing for indirect costs can have some quantitative impact on the welfare effects of a full DI mandate, but our main results remain largely unaffected.

To begin with, mandating extra DI coverage is likely to impose additional moral hazard costs onto the public baseline insurance, as it includes top-up insurance in case the worker also qualifies for public general DI benefits. To quantify this channel, we use the estimate of Seitz (2021) who finds that taking up private DI increases public DI claims by 4pp. (16%) in the German setting. As shown in Panel A of Appendix Table 3.A11, taking into account this additional moral hazard lowers the net value of a mandate. A second indirect cost may arise when a public DI mandate is financed by income-based contributions, as these contributions tax earnings. Thus, a standard fiscal externality from additional payroll taxes may arise. We calibrate this channel based on the Harberger triangle calculation of Feldstein (1999), where we assume an elasticity of taxable income of 0.3 and use marginal and average income tax rates simulated for the average individual in each risk group. In Panel B of the table, the distortion from raising contributions again lowers the net value of a public DI mandate. We note that this fiscal externality likely provides an upper bound, as some studies suggest that social insurance contributions induce much smaller fiscal externalities than income taxes (e.g. Lehmann, Marical, and Rioux 2013).

Moreover, providing additional DI could impose a positive fiscal externality on other social programs. In particular, covering all workers with own-occupation DI may reduce their propensity to claim basic social assistance in the case of a disability. We incorporate this externality in Panel C, which shows that the net value of mandating private DI increases. The change in net value is small, however, since social assistance is relatively low in the German setting and for many claims baseline public DI is still available. Finally, Panel D shows the combined effect of all these indirect effects. Qualitatively, results remain very similar to the baseline calculations. Quantitatively, the net value of a full public DI mandate becomes somewhat smaller, such that a higher degree of equity concern ($\sigma$ around 3) would be needed to justify the mandate.

As a further robustness exercise, we allow for some risk-based selection in the private DI market. We do not find significant selection in Section 3.4.3 and thus argue that cost curves are flat in the main welfare analysis. However, the estimation results shown in Figure 3.4.5 carry some statistical noise, such that we cannot exclude some degree of selection. To quantify the range of potential slopes of cost curves, we invert the specification from equation (3.3), regressing claiming probabilities on take-up within risk groups. We find a point estimate of -0.3pp., with a 95% confidence interval between -2.8pp. and +2.3pp. These results imply small degrees of selection. The point estimate corresponds to a -1.0% difference in claims between between individuals with and without private DI, and the confidence interval includes adverse selection with a 9.3% difference in claims up to advantageous selection with a -11.4% difference. In Panels E and F of Table 3.A11, we replicate the welfare analysis under these bounds on selection. Adverse selection somewhat increases the net value of a mandate and advantageous selection somewhat decreases it, but the results are qualitatively unaffected by the small degrees of selection we cannot exclude.

## 3.8 Conclusion

In this paper, we provide novel empirical evidence on the functioning of private DI markets. We show significant crowding-in of private DI when the scope of public DI is reduced, but overall take-up remains relatively modest. In particular, high-risk, low-income and low-education individuals are less likely to take-up private insurance. Yet, we do not find any evidence of adverse selection on unpriced risk. Our welfare analysis highlights the policy implications of these findings. If observed willingness to pay reflects individuals' true valuation of DI, providing extra DI coverage via a private DI market with choice is welfare-improving compared to a full mandate. Yet, equity concerns provide a potentially important rationale for a public DI mandate, as this would lead to additional coverage predominantly for low-income and high-risk individuals. In addition, we argue that risk misperceptions could explain low observed demand for DI of many workers, which may provide further grounds for policies increasing take-up such as a mandate.

To our knowledge, the German setting is unique in that one branch of the public DI mandate was fully removed. This allows us to provide first-time evidence on partly replacing public DI with a private insurance market. However, a key issue to bear in mind is that our empirical results are specific to to the type of coverage offered by private DI in this setting, combining insurance against own-occupation disability and more general top-up insurance. In principle, one could think of similar reforms privatizing insurance against other sub-risks of disability, such as insurance against short-term disability or against disability due to selected types of medical conditions. But of course our findings cannot simply be extrapolated to privatizing any part of DI coverage. Nevertheless, we believe the issues studied in this paper are likely to be relevant for other DI reforms aimed at an increased role of private insurance. Further research in this area will be highly valuable, as many governments are implementing reforms cutting public DI generosity.

# Appendix

## 3.A Appendix Figures and Tables

Figure 3.A1: Geographical Presence of Insurer

*The figure shows the geographical distribution of local insurance agencies of the insurer providing our microdata. A local agency is present in the counties marked in red, and no agency is present in those marked in gray.*

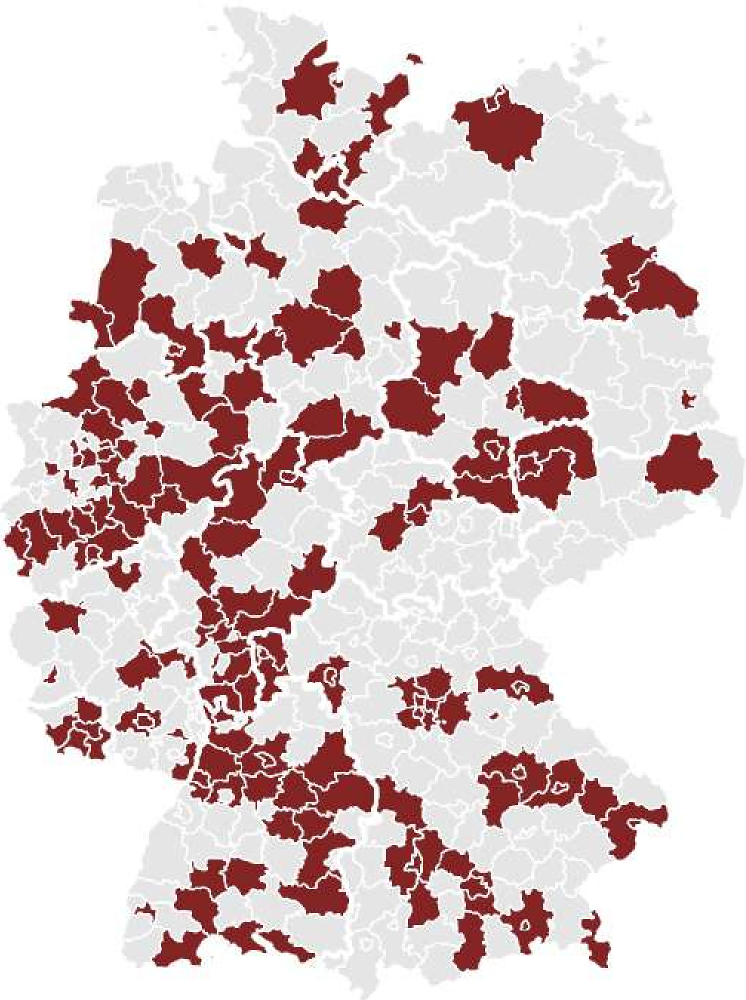
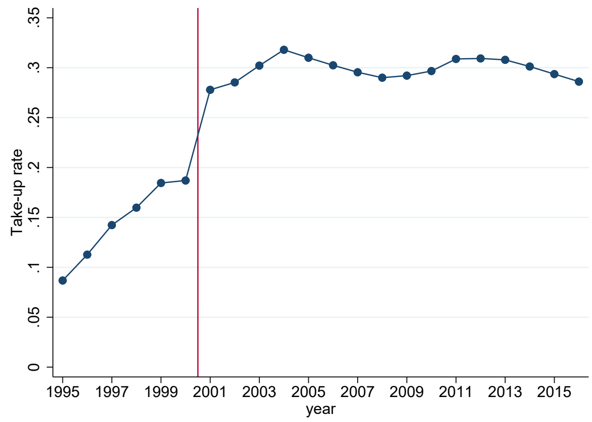


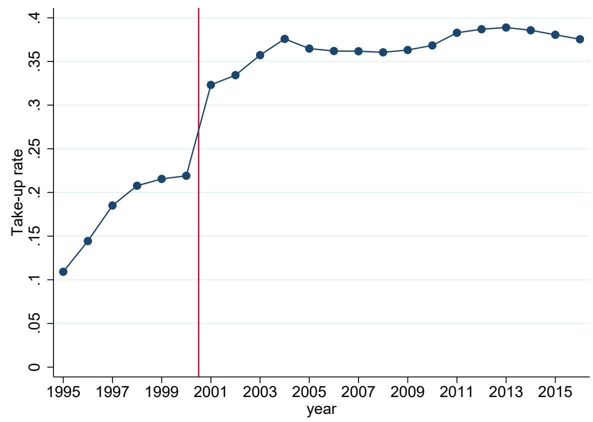
Figure 3.A2: Additional Descriptive Evidence on Private DI Take-Up

*The figure shows additional evidence on overall private DI take-up. Panel (a) shows take-up only among cohorts 1961 and younger who are affected by the reform. To obtain this take-up rate, we impute their number of contracts by applying the share of contracts held by these cohorts from the insurer microdata to the total number of contracts in the market. Panel (b) shows an alternative calculation of private DI take-up, relating the total number of contracts only to the number of currently employed individuals rather than all individuals contributing to social insurance. Panel (c) replicates Panel (b) of Figure 3.3.1 by type of private DI contract. Finally, Panel (d) shows a comparison of the total number of contracts in the market (based on the rating agency data) to the number of contracts in the insurer microdata. For confidentiality reasons, we are unable to specify the the scale of the insurer microdata series.*

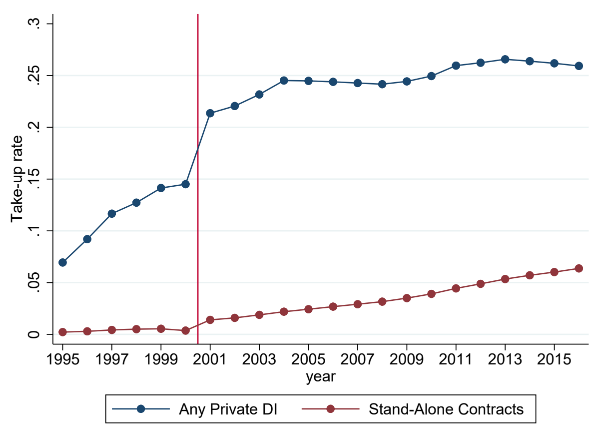
(a) Only Cohorts 1961 and Younger




(b) Only Currently Employed Individuals




(c) By Type of Contract




(d) Number of Contracts: Market vs. Company

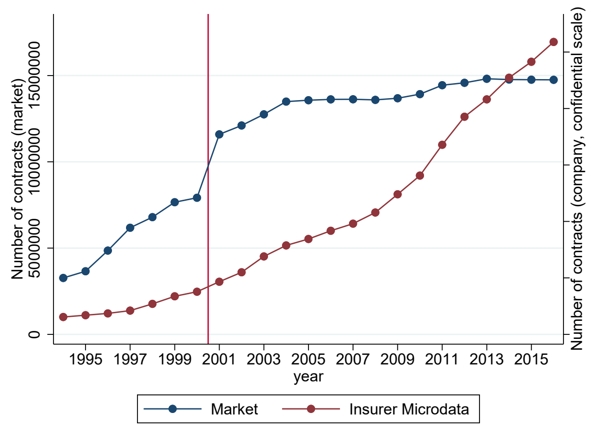
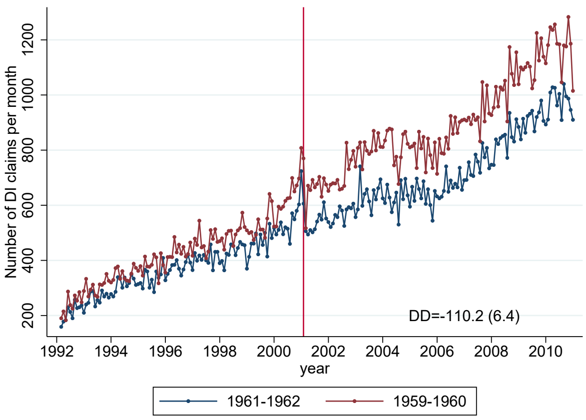


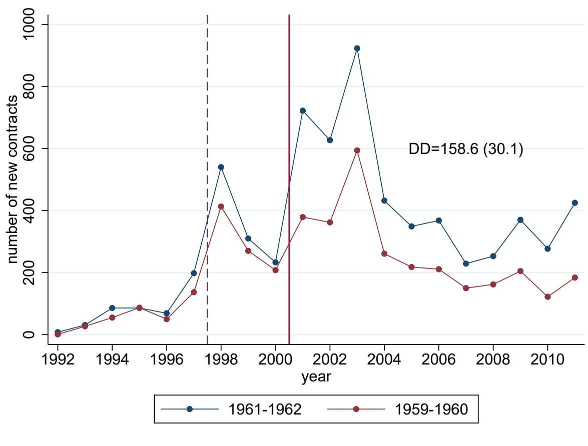
Figure 3.A3: Additional Difference-in-Difference Results

*The figure shows the number of public DI claims (Panel a), stand-alone private DI purchases (Panel b) and insured benefits in private DI contracts of individuals born in 1961-1962 (treated cohorts) vs. 1959-1960 (control cohorts). In all panels, the solid vertical line denotes the time the reform of 2001 takes effect (January 2001). In Panels (b) and (c), the dashed vertical line additionally demarcates the time the reform is first announced (December 1997). DD denotes the difference-in-difference coefficient estimated for the respective outcome with standard errors in parentheses (see Table 3.3.2 for details).*

(a) All Public DI Claims




(b) Private Stand-Alone DI Purchases




(c) Private DI Benefits

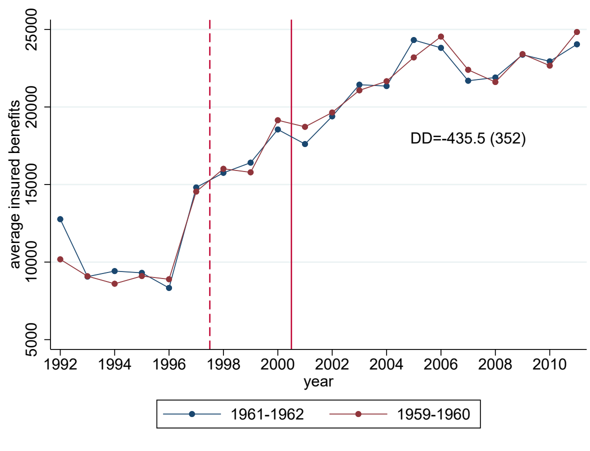


Figure 3.A4: Disability Risk Paths

*The figure shows the cumulative fraction of individuals claiming DI benefits. Panels (a) to (e) show the fraction claiming public DI benefits by age in each risk group. Panel (f) shows a comparison of public DI claims among all risk groups to private DI claiming risk calculated by the German Actuarial Association for a representative individual.*

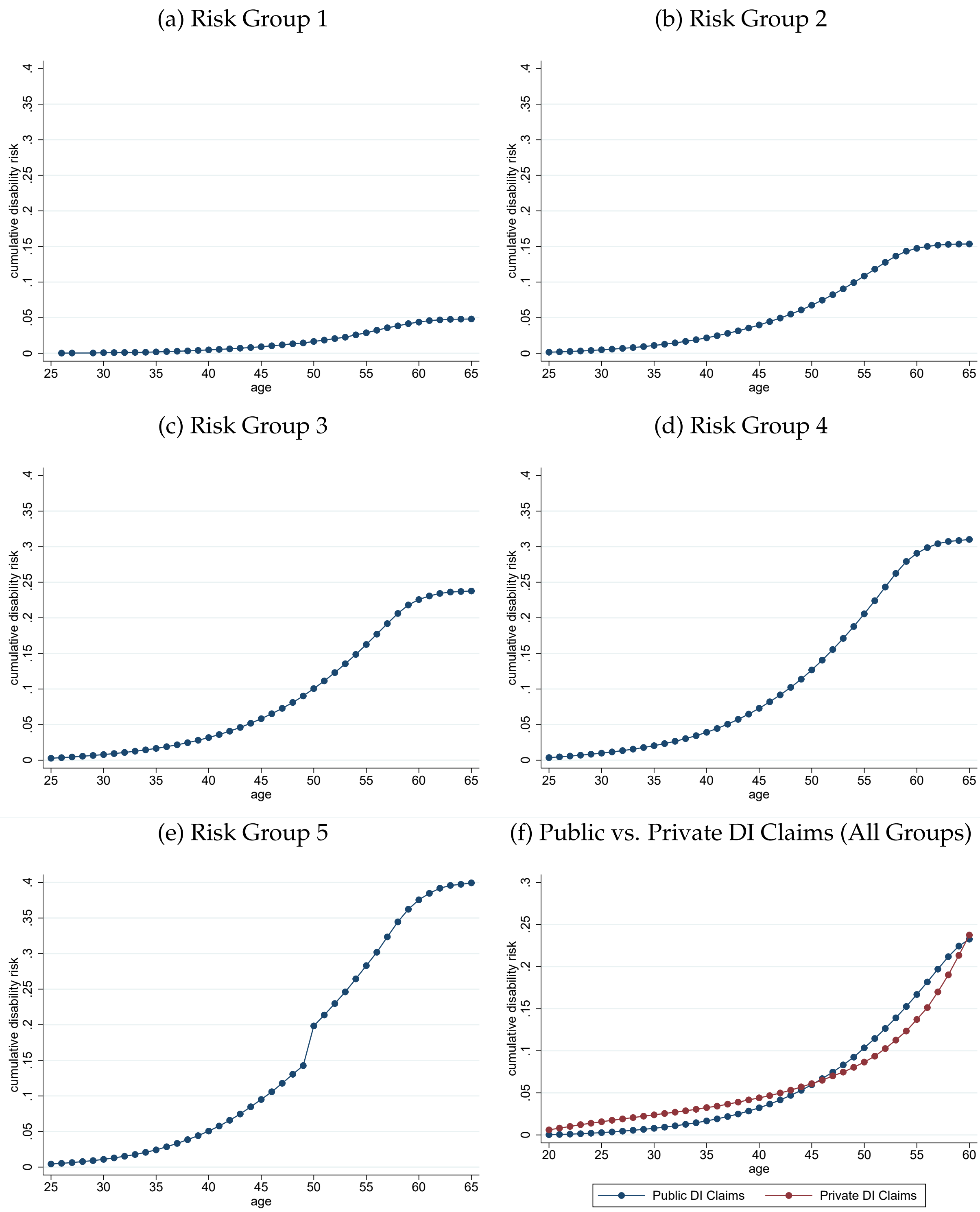


Figure 3.A5: Private DI Take-Up Pre- vs. Post-Reform

*The figure shows private DI take-up rates in 2015 by risk group in 2015 (Panel a) and in 1997, the year before the reform of 2001 was announced (Panel b). All take-up rates are calculated as shown in equation* (3.2).

(a) Post-Reform (b) Pre-Reform

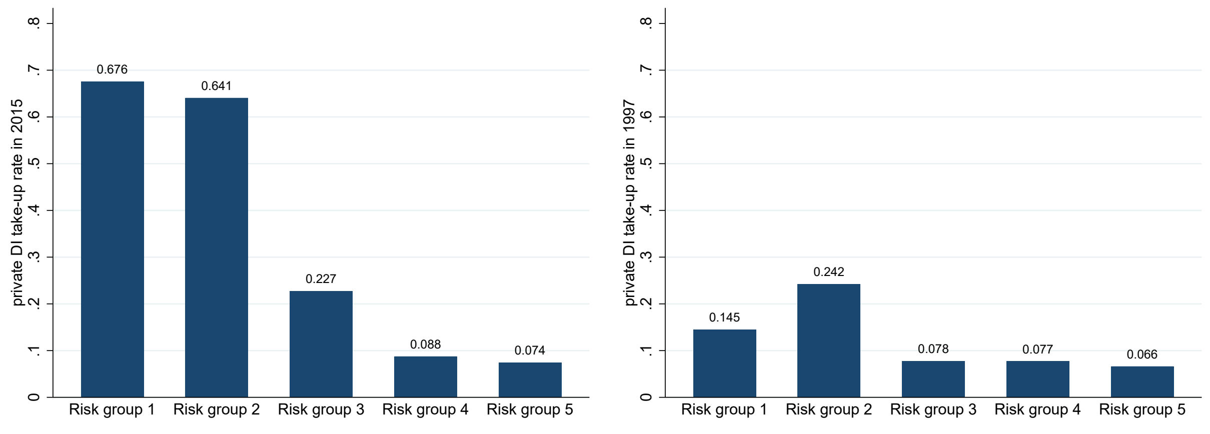


Figure 3.A6: Net Value in Private DI Market

*The figure depicts welfare in the private DI market equilibrium by risk group. In each panel, the net value is given by the total area under the demand curve ($A + B + C$) divided by the area under the cost curve ($C$). Net value can be further decomposed as explained in the figure legend.*

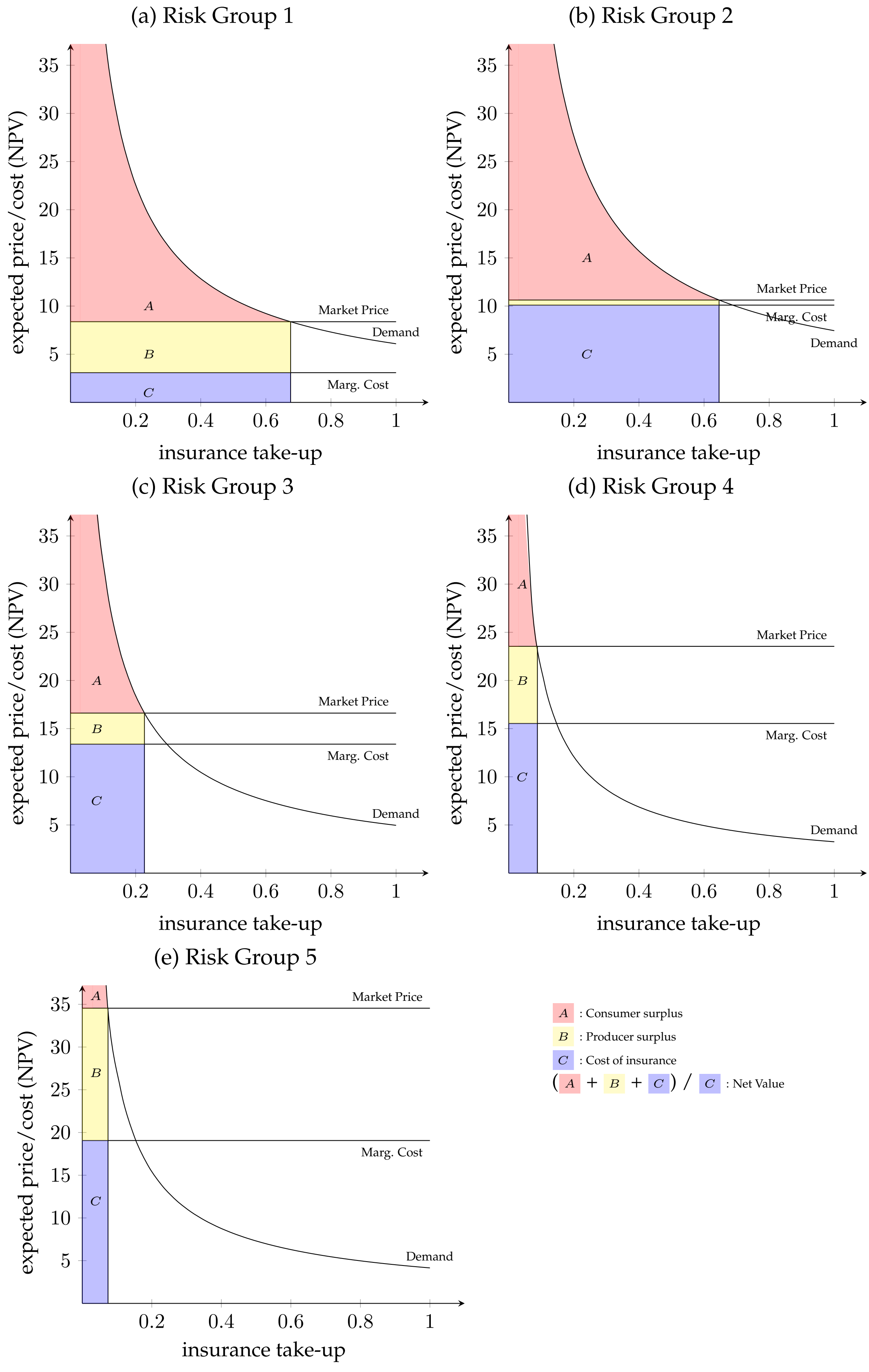


Figure 3.A7: Net Value of Introducing a DI Mandate

*The figure shows the net value of introducing a DI mandate by risk group. The mandate increases DI take-up from the market equilibrium to 1. The net value of the reform is given by the additional area under the demand curve (D + G) divided by the additional cost (F + G). Net value can be further decomposed as explained in the figure legend.*

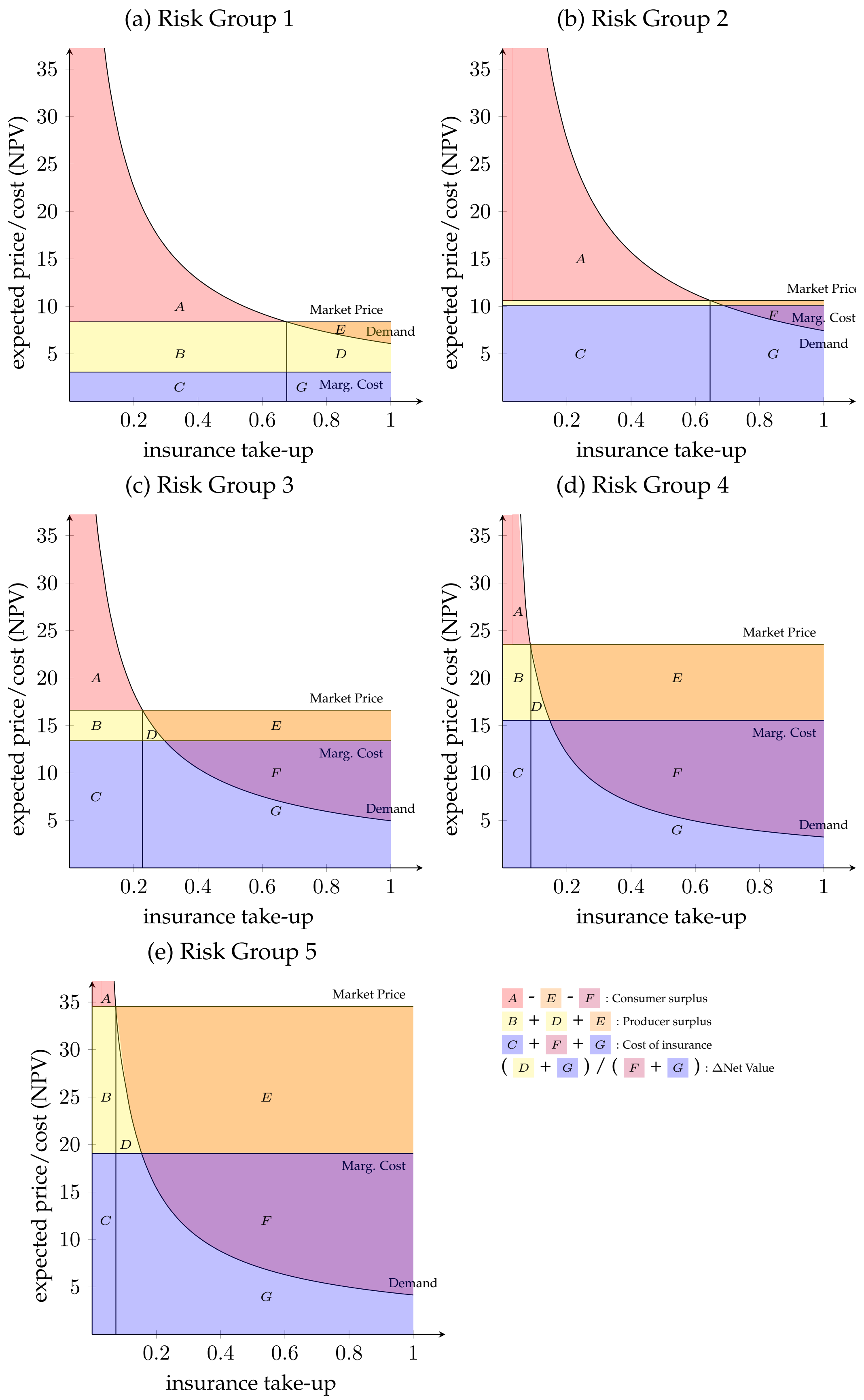


Table 3.A1: Occupations and Risk Groups

*The table shows examples among the most frequent occupation titles in each risk groups, based on the insurer microdata. * denotes occupations included in risk group 1 under the condition that the individual works mostly inside an office.*

| Risk group | Frequent occupation titles |
|---|---|
| RG 1 | Medical doctor (no surgeon), civil engineer*, business economist*, managing director*, business consultant*, tax consultant, pharmacist, computer scientist*, economist*, accountant* |
| RG 2 | Commercial clerk, surgeon, dentist, managing director, executive assistant, business consultant, construction engineer, IT technician, lawyer, bank clerk |
| RG 3 | Physiotherapist, high school teacher, sales clerk, educator, secretary, social worker, electrical engineer, hotel clerk, administrative clerk, beautician |
| RG 4 | Carpenter, nurse, metalworker, plumber, mason, hairdresser, painter, driver, roofer, car mechanic, electrician, toolmaker, tiler, gardener, waiter |
| RG 5 | Baker, dairy worker, firefighter, miner, road builder, pipe cleaner, steelworker, concrete worker, warehouse worker, excavation worker |

Table 3.A2: Summary Statistics: Household Survey Data

*The table shows summary statistics of the 2013 wave of the Income and Consumption Survey (EVS).*

| | (1)<br>All households | (2)<br>Employed households |
|---|---|---|
| Private DI owner | 0.31 | 0.35 |
| | (0.46) | (0.48) |
| Gross labor income (annual) | 26,218.6 | 35,103.9 |
| | (23,384.1) | (20,629.2) |
| Age | 44.09 | 43.39 |
| | (11.83) | (11.17) |
| Male | 0.59 | 0.61 |
| | (0.49) | (0.49) |
| Household size | 2.01 | 2.09 |
| | (1.14) | (1.15) |
| Observations | 31,452 | 21,037 |

Table 3.A3: Difference-in-Differences: Robustness

*Panel A shows results from difference-in-difference regressions as described by equation* (3.1). *Columns (1) and (3) replicate the baseline estimation and Columns (2) and (4) additionally control for a linear time trend interacted with an indicator for treated cohorts. Panel B shows difference-in-difference regressions with varying timing assumptions. Column (1) replicates the baseline estimation, Column (2) additionally controls for an indicator for the period 1998 to 2000 and its interaction with the indicator for treated cohorts, Column (3) omits the period 1998 to 2000 from the estimation, and Column (4) defines the post-reform indicator as post-1998 instead of post-2001. All regressions are run at the level of cohort × calendar month cells. Robust standard errors in parantheses. *** p<0.01, ** p<0.05, * p<0.1.*

**Panel A: Controlling for Cohort-Specific Trends**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Number of Private DI Purchases | | | |
| | All Contracts | | Stand-Alone | |
| Treated × post-2001 | 15.11*** | 17.38** | 13.22*** | 17.33*** |
| | (2.739) | (7.107) | (1.676) | (4.297) |
| Observations | 480 | 480 | 480 | 480 |
| R-squared | 0.939 | 0.939 | 0.939 | 0.940 |
| Calendar month FE | yes | yes | yes | yes |
| Group-specific trend | no | yes | no | yes |
| Mean (pre-reform) | 23.49 | 23.49 | 6.640 | 6.640 |

**Panel B: Robustness to Timing of Reform**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Number of Private DI Purchases (All Contracts) | | | |
| | baseline (post-2001) | control for 1998-2000 | omit 1998-2000 | post-1998 |
| Treated × post | 15.11*** | 19.04*** | 16.96*** | 17.48*** |
| | (2.739) | (2.539) | (2.456) | (2.202) |
| Observations | 480 | 480 | 384 | 480 |
| R-squared | 0.939 | 0.940 | 0.944 | 0.940 |
| Calendar month FE | yes | yes | yes | yes |
| Mean (pre-reform) | 23.49 | 23.49 | 23.49 | 23.49 |

Table 3.A4: Risk Groups and Disability Risk

*The table shows information on risk groups assigned by the insurer to individuals based on their occupations. Column (1) shows the share of each risk group out of the labor force based on occupations observed in the administrative public pension data. Column (2) shows the fraction of individuals in each risk group claiming public DI benefits at any age. Columns (3) shows the share of own-occupation DI claims out of all DI claims. Columns (4) to (6) show the monthly premium (in EUR) charged to an individual insuring EUR 1000 of private DI benefits by risk group and contract start age, for a fixed contract end age of 65.*

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Risk group | Share of labor force | Lifetime disability risk | Share of own-occupation DI claims | Monthly insurance premium for contract start at... | | |
| | | | | age 25 | age 35 | age 45 |
| All | 100% | 25.07% | 13.20% | 72.83 | 83.53 | 98.15 |
| RG 1 | 9.72% | 4.81% | 10.85% | 31.61 | 35.95 | 43.22 |
| RG 2 | 16.96% | 15.35% | 8.06% | 41.73 | 49.08 | 57.50 |
| RG 3 | 35.14% | 23.77% | 12.57% | 68.15 | 79.91 | 93.73 |
| RG 4 | 37.55% | 31.02% | 15.74% | 100.60 | 113.31 | 133.03 |
| RG 5 | 0.62% | 39.94% | 31.95% | 155.24 | 175.78 | 210.68 |

Table 3.A5: Difference-in-Difference Results by Subgroup

*The table shows results from difference-in-difference regressions as described by equation (3.1) for subgroups specified in the column titles. Regressions are run at the level of cohort × calendar month cells. Robust standard errors in parantheses. *** p<0.01, ** p<0.05, * p<0.1.*

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | **A: Private DI Contracts by Income** | | | | | **B: by gender** | |
| | Quintile 1 | Quintile 2 | Quintile 3 | Quintile 4 | Quintile 5 | Men | Women |
| Treated × post | 0.316 | 1.573*** | 1.240*** | 4.377*** | 7.602*** | 8.460*** | 6.652*** |
| | (0.454) | (0.512) | (0.458) | (0.820) | (1.330) | (1.870) | (1.108) |
| Observations | 480 | 480 | 480 | 480 | 480 | 480 | 480 |
| R-squared | 0.838 | 0.889 | 0.818 | 0.886 | 0.944 | 0.944 | 0.897 |
| Mean (pre-reform) | 3.180 | 2.720 | 2.710 | 6.350 | 7.940 | 17.24 | 6.250 |
| Calendar month FE | yes | yes | yes | yes | yes | yes | yes |

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **C: Private DI Contracts by Risk Group** | | | | |
| | Risk group 1 | Risk group 2 | Risk group 3 | Risk group 4 | Risk group 5 |
| Treated × post | 1.756*** | 6.833*** | 6.699*** | 0.0749 | 0.0909* |
| | (0.571) | (1.063) | (1.221) | (0.674) | (0.0470) |
| Observations | 480 | 480 | 480 | 480 | 480 |
| R-squared | 0.899 | 0.933 | 0.913 | 0.898 | 0.516 |
| Mean (pre-reform) | 2.720 | 7.640 | 6.190 | 6.030 | 0.0700 |
| Calendar month FE | yes | yes | yes | yes | yes |

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **D: Private DI Contracts by Education** | | | | |
| | Quintile 1 | Quintile 2 | Quintile 3 | Quintile 4 | Quintile 5 |
| Treated × post | 0.0867 | 0.355 | 0.920** | 3.861*** | 9.889*** |
| | (0.388) | (0.414) | (0.419) | (0.807) | (1.546) |
| Observations | 480 | 480 | 480 | 480 | 480 |
| R-squared | 0.786 | 0.866 | 0.856 | 0.887 | 0.947 |
| Mean (pre-reform) | 2.100 | 3.180 | 3.250 | 4.780 | 10.18 |
| Calendar month FE | yes | yes | yes | yes | yes |

Table 3.A6: Risk-Based Selection

*The table shows regression results on the correlation between private DI take-up in 2015 and disability risk at the three-digit occupation level. Column (1) corresponds to estimating equation* (3.3) *without controlling for risk groups, Column (2) shows a specification where actual disability risk is omitted, and Column (3) corresponds to the specification shown in equation* (3.3). *Columns (4) to (7) add varying set of control variables to the regression. Robust standard errors in parantheses. *** p<0.01, ** p<0.05, * p<0.1.*

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | Dependent Variable: Private DI Take-Up | | | | | | |
| Actual Disability Risk | -1.383*** | | -0.0513 | -0.0341 | 0.374 | 0.327 | -0.0343 |
| | (0.260) | | (0.263) | (0.263) | (0.288) | (0.289) | (0.243) |
| Risk Group | | -0.257*** | -0.252*** | -0.245*** | -0.147*** | -0.149*** | -0.137*** |
| | | (0.0238) | (0.0323) | (0.0324) | (0.0386) | (0.0384) | (0.0385) |
| Log income | | | | 0.0391** | 0.0204 | 0.0184 | 0.0209 |
| | | | | (0.0160) | (0.0139) | (0.0141) | (0.0161) |
| Education (years) | | | | | 0.114*** | 0.113*** | 0.104*** |
| | | | | | (0.0201) | (0.0203) | (0.0203) |
| Female | | | | | | -0.0631 | -0.279*** |
| | | | | | | (0.0590) | (0.0771) |
| Economic training | | | | | | | 0.161* |
| | | | | | | | (0.0946) |
| Married | | | | | | | -1.363*** |
| | | | | | | | (0.431) |
| East Germany | | | | | | | 0.254* |
| | | | | | | | (0.150) |
| Observations | 293 | 293 | 293 | 293 | 293 | 293 | 293 |
| R-squared | 0.126 | 0.270 | 0.270 | 0.277 | 0.359 | 0.361 | 0.398 |

Table 3.A7: Demand Elasticity Estimation Regressions

*The table shows results from the demand elasticity estimation regressions described by equation* (3.4). *Columns (1) and (2) estimate the average jump in private DI take-up between risk groups, and Columns (3) and (4) estimate the jump in take-up separately at each risk group boundary. Table 3.6.3 converts the estimates into implied demand elasticities.*

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Dependent Variable: Private DI Take-Up | | | |
| Actual disability risk | -0.0513 | -0.0343 | -0.226 | -0.0579 |
| | (0.263) | (0.243) | (0.248) | (0.240) |
| Risk group | -0.252*** | -0.137*** | | |
| | (0.0323) | (0.0385) | | |
| Risk group 2 | | | -0.345*** | -0.183 |
| | | | (0.0574) | (0.128) |
| Risk group 3 | | | -0.581*** | -0.249* |
| | | | (0.0486) | (0.139) |
| Risk group 4 | | | -0.784*** | -0.399*** |
| | | | (0.0494) | (0.147) |
| Risk group 5 | | | -0.926*** | -0.549*** |
| | | | (0.0501) | (0.155) |
| Log income | | 0.0209 | | 0.0201 |
| | | (0.0161) | | (0.0162) |
| Female | | -0.279*** | | -0.301*** |
| | | (0.0771) | | (0.0787) |
| Education (years) | | 0.104*** | | 0.112*** |
| | | (0.0203) | | (0.0204) |
| Economic training | | 0.161* | | 0.178* |
| | | (0.0946) | | (0.0968) |
| Married | | -1.363*** | | -1.312*** |
| | | (0.431) | | (0.440) |
| East Germany | | 0.254* | | 0.211 |
| | | (0.150) | | (0.148) |
| Observations | 293 | 293 | 293 | 293 |
| R-squared | 0.270 | 0.398 | 0.266 | 0.403 |

Table 3.A8: Value and Cost of Insurance

*The table shows estimated willingness to pay and cost of disability insurance. Column (1) shows median valuations and cost among all workers, and Columns (2) to (6) show median valuations and cost by risk group. Besides willingness to and cost for the full private DI coverage, the table also displays a decomposition into own-occupation and general DI as described by equations* (3.8), (3.9) *and* (3.10).

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | All | RG 1 | RG 2 | RG 3 | RG 4 | RG 5 |
| **Panel A: Median Willingness to Pay (in % of Income)** | | | | | | |
| Full private DI coverage | 0.930 | 1.134 | 1.418 | 0.958 | 0.633 | 0.818 |
| Own-occupation DI (lower bound) | 0.434 | 0.812 | 0.780 | 0.559 | 0.065 | 0.070 |
| Own-occupation DI (upper bound) | 0.499 | 0.847 | 0.832 | 0.609 | 0.154 | 0.309 |
| General DI top-up (lower bound) | 0.431 | 0.288 | 0.586 | 0.349 | 0.479 | 0.509 |
| **Panel B: Median Cost of Insurance (in % of Income)** | | | | | | |
| Full private DI coverage | 1.469 | 0.327 | 1.094 | 1.469 | 1.720 | 2.140 |
| Own-occupation DI | 0.185 | 0.035 | 0.088 | 0.185 | 0.271 | 0.684 |
| General DI top-up | 1.284 | 0.291 | 1.006 | 1.284 | 1.450 | 1.457 |

Table 3.A9: Social Welfare Weights

*The table shows average income and social welfare weights by risk group. "Income (NPV)" denotes the net present value of expected lifetime income calculated at age 25. Social welfare weights are calculated for the social welfare function specified in the row titles, and serve as an input into the social net value calculations shown in Panel B of Table 3.7.4.*

| | (1) RG 1 | (2) RG 2 | (3) RG 3 | (4) RG 4 | (5) RG 5 |
|---|---|---|---|---|---|
| Income (annual) | 64,605 | 54,998 | 40,648 | 35,202 | 31,546 |
| Income (NPV) | 1,524,574 | 1,269,566 | 926,151 | 794,701 | 702,268 |
| Social welfare weights | | | | | |
| Utilitarian, $\sigma$=1 | 0.629 | 0.752 | 1.021 | 1.183 | 1.332 |
| Utilitarian, $\sigma$=3 | 0.225 | 0.385 | 0.962 | 1.496 | 2.135 |
| Utilitarian, $\sigma$=5 | 0.074 | 0.180 | 0.831 | 1.734 | 3.137 |
| Utilitarian, $\sigma$=8 | 0.013 | 0.053 | 0.610 | 1.980 | 5.116 |
| Rawlsian | 0.000 | 0.000 | 0.000 | 0.000 | 161.290 |

Table 3.A10: Risk Misperception: Calibration Results

*The table shows results from the calibrations described in Section 3.7.3. Panel A shows the willingness-to-pay percentile of the marginal buyer among the risk group indicated by the column title. Panel B shows the calibrated coefficient of relative risk aversion $\sigma$ of the marginal buyer under the assumption about consumption levels indicated in the respective row title. Panel C shows calibrated risk underestimation $\alpha$, i.e. the ratio of perceived to actual disability risk, of the marginal buyer. Panel D shows the implied ratio of normative willingness to pay to observed willingness to pay.*

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | RG 1 | RG 2 | RG 3 | RG 4 | RG 5 |
| **Panel A: Location of marginal buyer** | | | | | |
| Percentile of willingness to pay in group | 0.324 | 0.359 | 0.773 | 0.912 | 0.926 |
| **Panel B: Risk aversion of marginal buyer** | | | | | |
| Hand-to-mouth | 0.437 | 0.034 | 0.115 | 0.194 | 0.264 |
| Hand-to-mouth + SA floor | 0.841 | 0.304 | 0.581 | 0.902 | 1.216 |
| High $\Delta C$ + SA floor | 1.158 | 0.085 | 0.287 | 0.492 | 0.769 |
| Low $\Delta C$ + SA floor | 3.030 | 0.241 | 0.811 | 1.339 | 1.866 |
| **Panel C: Risk underestimation of marginal buyer** | | | | | |
| Hand-to-mouth ($\sigma$=0.44) | 1.000 | 0.394 | 0.469 | 0.543 | 0.642 |
| Hand-to-mouth + SA floor ($\sigma$=0.84) | 1.000 | 0.448 | 0.705 | 1.086 | 1.586 |
| High $\Delta C$ + SA floor ($\sigma$=1.16) | 1.000 | 0.397 | 0.475 | 0.554 | 0.720 |
| Low $\Delta C$ + SA floor ($\sigma$=3.03) | 1.000 | 0.400 | 0.484 | 0.557 | 0.674 |
| **Panel D: Implied normative WTP/observed WTP** | | | | | |
| Hand-to-mouth ($\sigma$=0.44) | 1.000 | 2.596 | 2.187 | 1.880 | 1.594 |
| Hand-to-mouth + SA floor ($\sigma$=0.84) | 1.000 | 2.255 | 1.440 | 0.901 | 0.600 |
| High $\Delta C$ + SA floor ($\sigma$=1.16) | 1.000 | 2.532 | 2.133 | 1.817 | 1.400 |
| Low $\Delta C$ + SA floor ($\sigma$=3.03) | 1.000 | 2.447 | 2.027 | 1.743 | 1.455 |

Table 3.A11: Welfare Calculations: Extensions and Robustness

*The table shows the net value of mandating the DI coverage offered by the private insurance market under the various extensions of our welfare calculations described in Section 3.7.4.*

| | (1) | (2) | (3) |
|---|---|---|---|
| | Private DI Mandate | Public DI Mandate | |
| | | Lump-Sum Contributions | Income-Based Contributions |
| **Panel A: Moral Hazard Effect on Baseline Insurance** | | | |
| Net Value | 0.668 | 0.668 | 0.668 |
| Social Net Value, $\sigma$=1 | 0.616 | 0.824 | 1.067 |
| Social Net Value, $\sigma$=3 | 0.535 | 1.000 | 1.526 |
| Social Net Value, $\sigma$=5 | 0.480 | 1.069 | 1.715 |
| **Panel B: Fiscal Externality from Social Insurance Contributions** | | | |
| Net Value | 0.762 | 0.762 | 0.612 |
| Social Net Value, $\sigma$=1 | 0.704 | 0.941 | 0.942 |
| Social Net Value, $\sigma$=3 | 0.612 | 1.142 | 1.323 |
| Social Net Value, $\sigma$=5 | 0.549 | 1.221 | 1.480 |
| **Panel C: Reduction in Social Assistance Claims** | | | |
| Net Value | 0.776 | 0.776 | 0.776 |
| Social Net Value, $\sigma$=1 | 0.713 | 0.954 | 1.232 |
| Social Net Value, $\sigma$=3 | 0.614 | 1.154 | 1.757 |
| Social Net Value, $\sigma$=5 | 0.546 | 1.231 | 1.974 |
| **Panel D: Combining A to C** | | | |
| Net Value | 0.680 | 0.680 | 0.547 |
| Social Net Value, $\sigma$=1 | 0.625 | 0.835 | 0.836 |
| Social Net Value, $\sigma$=3 | 0.537 | 1.010 | 1.168 |
| Social Net Value, $\sigma$=5 | 0.477 | 1.078 | 1.306 |
| **Panel E: Some Adverse Selection** | | | |
| Net Value | 0.777 | 0.777 | 0.777 |
| Social Net Value, $\sigma$=1 | 0.719 | 0.962 | 1.248 |
| Social Net Value, $\sigma$=3 | 0.628 | 1.171 | 1.791 |
| Social Net Value, $\sigma$=5 | 0.565 | 1.253 | 2.014 |
| **Panel F: Some Advantageous Selection** | | | |
| Net Value | 0.745 | 0.745 | 0.745 |
| Social Net Value, $\sigma$=1 | 0.686 | 0.916 | 1.184 |
| Social Net Value, $\sigma$=3 | 0.594 | 1.109 | 1.689 |
| Social Net Value, $\sigma$=5 | 0.531 | 1.185 | 1.898 |

# Bibliography

Abowd, John M., Francis Kramarz, and David N. Margolis (1999). "High Wage Workers and High Wage Firms." In: *Econometrica* 67.2, pp. 251–333.

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

Akerlof, George A. (1970). "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." In: *Quarterly Journal of Economics* 84.3, 488–500.

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.

Allianz, Lebensversicherungs AG (2018). *Einkommensvorsorge 2018 - Faktisch überlegen: bei Antrag, während der Laufzeit und im Leistungsfall*. Presentation.

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.

Arbeiterkammer (2014). "Einkommensberichte - Erfahrung aus Sicht der BetriebsrätInnen." In: *Arbeitskammer Online Publications*.

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez (2011). "Top Incomes in the Long Run of History." In: *Journal of Economic Literature* 49.1, pp. 3–71.

Autor, David and Mark Duggan (2010). *Supporting Work: A Proposal for Modernizing the U.S. Disability Insurance System*. The Center for American Progress and The Hamilton Project.

Autor, David, Mark Duggan, Kyle Greenberg, and David S. Lyle (2016). "The Impact of Disability Benefits on Labor Supply: Evidence from the VA's Disability Compensation Program." In: *American Economic Journal: Applied Economics* 8.3, 31–68.

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 (2003). "The Rise in Disability Rolls and the Decline in Unemployment." In: *Quarterly Journal of Economics* 118.1, pp. 157–205.

— (2006). "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." In: *Journal of Economic Perspectives* 20.3, pp. 71–96.

— (2007). "Distinguishing Income from Substitution Effects in Disability Insurance." In: *AEA Papers & Proceedings* 97.2, pp. 119–124.

Autor, David H., Mark G. Duggan, and David S. Lyle (2011). "Battle Scars? The Puzzling Decline in Employment and Rise in Disability Receipt among Vietnam Era Veterans." In: *AEA Papers & Proceedings* 101.3, pp. 339–344.

Babcock, Linda and Sara Laschever (2003). "Women Don't Ask: Negotiation and The Gender Divide." In: *Princeton University Press*.

Baker, Michael, Yosh Halberstam, Kory Kroft, Alexandre Mas, and Derek Messacar (2019). "Pay Transparency and The Gender Pay Gap." In: *NBER Working Paper 25834*.

Bennedsen, Morten, Elena Simintzi, Margarita Tsoutsoura, and Daniel Wolfenzon (Forthcoming). "Do Firms Respond to Gender Pay Gap Transparency?" In: *The Journal of Finance*.

Berufsunfähigkeitsversicherungen Heute (BUVT) (2019). *Berufsunfähigkeitsversicherung – Klauseln und Regelungen*. Consumer advice website.

Blau, Francine D. and Lawrence M. Kahn (2017). "The Gender Wage Gap: Extent Trends and Explanations." In: *Journal of Economic Literature* 55.3, pp. 789–865.

Blundell, Jack (2020). "Wage responses to gender pay gap reporting requirements." In: *Working Paper*.

Böheim, René, Marian Fink, and Christine Zulehner (2020). "About time: The narrowing gender wage gap in Austria." In: *Empirica*, pp. 1–41.

Böheim, René and Sarah Gust (2021). "The Austrian pay transparency law and the gender wage gap." In: *CESifo Working Paper*.

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.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani (2017). "The Morale Effects of Pay Inequality." In: *The Quarterly Journal of Economics* 133.2, pp. 611–663.

Brown, Gordon D.A., Jonathan Gardner, J. Oswald Andrew, and Jing Qian (2008). "Does Wage Rank Affect Employees' Well-Being?" In: *Industrial Relations* 47.3, pp. 355–389.

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

Burkhauser, Richard V., Mary C. Daly, and Nicolas R. Ziebarth (2016). "Protecting working-age people with disabilities: experiences of four industrialized nations." In: *Journal of Labour Market Research* 49, pp. 367–386.

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.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin (2006). "Wage bargaining with on-the-job search: Theory and evidence." In: *Econometrica* 74.2, pp. 323–364.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016). "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." In: *The Quarterly Journal of Economics* 131.2, pp. 633–686.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez (2012). "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." In: *American Economic Review* 102.6, pp. 2981–3003.

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, Benjamin Handel, and Joshua Schwartzstein (2019). "Behavioral economics and health-care markets." In: *Handbook of Behavioral Economics – Foundations and Applications 2*. Ed. by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson. North-Holland, 459–502.

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

Chetty, Raj, John Friedman, Soren Leth-Petersen, Torben Nielsen, and Tore Olsen (2014). "Active vs. Passive Decisions and Crowd-out in Retirement Savings Accounts: Evidence from Denmark." In: *Quarterly Journal of Economics* 129.3, pp. 1141–1219.

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.

Clark, Andrew E., Nicolai Kristensen, and Niels Westergard-Nielsen (2009). "Job Satisfaction and Co-Worker Wages: Status or Signal?" In: *The Economic Journal* 119.536, pp. 430–447.

Clark, Andrew E. and Andrew J. Oswald (1996). "Satisfaction and Comparison Income." In: *Journal of Public Economics* 61.3, pp. 359–381.

Cohen, Alma and Liran Einav (2007). "Estimating Risk Preferences from Deductible Choice." In: *American Economic Review* 97.3, pp. 745–788.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frederic Schneifer (2014). "Social Comparison and Effort Provision: Evidence from a Field Experiment." In: *Journal of The European Economic Association* 12.4, pp. 877–898.

Cullen, Zoe B. and Bobak Pakzad-Hurson (2019). "Equilibrium Effects of Pay Transparency in a Simple Labor Market." In: *Harvard Business School Working Paper*.

Cullen, Zoe B. and Ricardo Perez-Truglia (Forthcoming). "How Much Does Your Boss Make? The Effects of Salary Comparisons." In: *Journal of Political Economy*.

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.

Deloitte (2015). "Einkommenstransparenz - Gleiches Entgelt für gleiche und gleichwertige Arbeit." In: *Publications, Bundesministerium für Bildung und Frauen* 1.1, pp. 1–35.

Deshpande, Manasi (2016a). "Does Welfare Inhibit Success? The Long- Term Effects of Removing Low- Income Youth from the Disability Rolls." In: *American Economic Review* 106.11, 3300–3330.

— (2016b). "The Effect of Disability Payments on Household Earnings and Income: Evidence from the SSI Children's Program." In: *Review of Economics and Statistics* 98.4, 638–654.

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

Dube, Arindrajit, Laura Giuliano, and Jonathan Leonard (2019). "Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior." In: *American Economic Review* 109.2, pp. 620–63.

Duchini, Emma, Stefania Simion, and Arthur Turrell (2020). "Pay Transparency and Cracks in the Glass Ceiling." In: *Working Paper*.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen (2010). "Estimating Welfare in Insurance Markets Using Variation in Prices." In: *Quarterly Journal of Economics* 125.3, pp. 877–920.

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.

Ericson, Keith M. and Justin Sydnor (2017). "The Questionable Value of Having a Choice of Levels of Health Insurance Coverage." In: *The Journal of Economic Perspectives* 31.4, 51–72.

FAZ (2012). *Schutz vom Chef*. Newspaper article, Frankfurter Allgemeine Zeitung.

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

Feldstein, Martin (1999). "Tax Avoidance and the Deadweight Loss of the Income Tax." In: *Review of Economics and Statistics* 81.4, pp. 674–680.

Finkelstein, Amy and Nathaniel Hendren (2020). "Welfare Analysis Meets Causal Inference." In: *Journal of Economic Perspectives* 34.4, 146–167.

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/-versicherungsunternehm 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.

Glassdoor (2016). "Global Salary Transparency Survey - Employee Perceptions of Talking Pay." In.

Godechot, Oliver and Claudia Senik (2015). "Wage Comparisons In and Out of Firm: Evidence from a Matched Employer-Employee French Database." In: *Journal of Economic Behavior and Organization* 117, pp. 395–410.

Goldfarb, Avi and Catherine Tucker (2011). "Technology, Age, and Shifting Privacy Concerns." In: *Mimeo*.

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.

Gruber, Jonathan (2000). "Disability Insurance Benefits and Labor Supply." In: *Journal of Political Economy* 108.6, pp. 1162–1183.

Gulyas, Andreas and Krzysztof Pytka (2020). *Understanding the Sources of Earnings Losses After Job Displacement: A Machine-Learning Approach*. CRC TR 224 Discussion Paper Series. University of Bonn and University of Mannheim, Germany. URL: https://EconPapers.repec.org/RePEc:bon:boncrc:crctr224_2020_131v2.

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.

Hendren, Nathaniel, Camille Landais, and Johannes Spinnewijn (2020). "Choice in Insurance Markets: A Pigouvian Approach to Social Insurance Design." In: *forthcoming, Annual Review of Economics*.

Hendren, Nathaniel and Ben Sprung-Keyser (2020). "A Unified Analysis of Government Policies." In: *Quarterly Journal of Economics* 135.3, pp. 1209–1318.

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.

Koning, Pierre and Maarten Lindeboom (2015). "The Rise and Fall of Disability Insurance Enrollment in the Netherlands." In: *The Journal of Economic Perspectives* 29.2, 151–172.

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`.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler (2020). "Income volatility, taxation and the functioning of the US labor market." In: *Working paper*.

Landais, Camille, Arash Nekoei, Peter Nilsson, David Seim, and Johannes Spinnewijn (2021). "Risk-Based Selection in Unemployment Insurance: Evidence and Implications." In: *American Economic Review* 111.4, pp. 1315–55.

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

Lehmann, Etienne, Francois Marical, and Laurence Rioux (2013). "Labor income responds differently to income-tax and payroll-tax reforms." In: *Journal of Public Economics* 99.3, pp. 66–84.

Liebman, Jeffrey B. (2015). "Understanding the Increase in Disability Insurance Benefit Receipt in the United States." In: *Journal of Economic Perspectives* 29.2, pp. 123–150.

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.

Low, Hamish and Luigi Pistaferri (2020). "Disability insurance: theoretical trade-offs and empirical evidence." In: *Fiscal Studies* 41.1, pp. 129–164.

Luttmer, Erzo F.P. (2005). "Neighbors as Negatives: Relative Earnings and Well-Being." In: *The Quarterly Journal of Economics* 120.3, pp. 963–1002.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand (2013). "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." In: *American Economic Review* 103.5, 1797–1829.

Marie, Olivier and Judit Vall Castello (2012). "Measuring the (income) effect of disability insurance generosity on labour market participation." In: *Journal of Public Economics* 96.1-2, pp. 198–210.

Mas, Alexandre (2016). "Does Disclosure Affect CEO Pay Setting? Evidence from the Passage of the 1934 Securities and Exchange Act." In: *Mimeo*.

— (2017). "Does Transparency Lead to Pay Compression?" In: *Journal of Political Economy* 125.5, pp. 1683–1721.

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.

Morchio, Iacopo and Christian Moser (2019). "The Gender Gap: Micro Sources and Macro Consequences." In.

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.

OECD (2010). *Sickness, Disability and Work - Breaking the Barriers*. OECD Press.

OECD (2019). *OECD Social Expenditure Database*. database.

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.

Perez-Truglia, Ricardo (2020). "The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment." In: *American Economic Review* 110.4, pp. 1019–54.

Postel-Vinay, F. and J.M. Robin (2002). "Equilibrium wage dispersion with worker and employer heterogeneity." In: *Econometrica* 70.6, pp. 2295–2350.

Rege, Mari and Ingeborg F. Solli (2015). "Lagging Behind The Joneses: The Impact of Relative Earnings on Job Separation." In: *Mimeo*.

Rothschild, Michael and Joseph Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." In: *Quarterly Journal of Economics* 90.4, pp. 29–649.

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.

Schmidt, Cornelius (2012). "Does Transparency Increase Executive Compensation?" In: *Mimeo*.

Seibold, Arthur (2021). "Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities." In: *American Economic Review* 111.4, pp. 1126–1165.

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

Seitz, Sebastian (2021). *The Moral Hazard Cost of Private Disability Insurance and its Welfare Consequences*. Working paper.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter (2018). "Firming Up Inequality." In: *The Quarterly Journal of Economics* 134.1, pp. 1–50.

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/`.

Staubli, Stefan (2011). "The impact of stricter criteria for disability insurance on labor force participation." In: *Journal of Public Economics* 95.9-10, pp. 1223–1235.

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

Sydnor, Justin (2010). "(Over)insuring Modest Risks." In: *American Economic Journal: Economic Policy* 2.4, pp. 177–199.

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

TNS Infratest (2015). *Finanzmarktdatenservice (FMDS) 2015*. Report.

U.S. Bureau of Labor Statistics (2019). *National Compensation Survey: Employee Benefits in the United States, March 2019*. Tech. rep.

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.

Zweimüller, Josef, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf, and Simon Büchi (2009). "Austrian Social Security Database." In: *IEW Working Paper No. 410, Institute for Empirical Research in Economics*.

# Curriculum vitae

| | |
|---|---|
| 2016–2022 | University of Mannheim (Germany)<br>*PhD in Economics* |
| 2017–2018 | Yale University (USA)<br>*Visiting Scholar* |
| 2014–2016 | University of Mannheim (Germany)<br>*MSc in Economics* |
| 2015 | Akita International University (Japan)<br>*Visiting Student* |
| 2011–2014 | University of Mannheim (Germany)<br>*BSc in Economics* |
| 2013 | University of North Carolina at Asheville (USA)<br>*Visiting Student* |