Adverse Selection and Moral Hazard in Social Insurance for Entrepreneurs*

Youssef Benzarti† Jarkko Harju‡ Tuomas Matikka§
Ella Mattinen¶ Alisa Tazhitdinova†

Abstract

This paper estimates the extent of adverse selection and moral hazard in social insurance for entrepreneurs using rich administrative data and exogenous variation from Finland. We use a reform that allows certain entrepreneurs to freely choose their level of insurance contributions. We use this reform along with a difference-in-differences approach to implement empirical tests of adverse selection and moral hazard. First, we find evidence of small moral hazard effects when considering whether entrepreneurs who opt into higher levels of social insurance tend to use it more. Second, by considering whether entrepreneurs with higher pre-reform risks select into higher levels of social insurance, we find no evidence of adverse selection. These findings are important for the design of social insurance for entrepreneurs.

Keywords: entrepreneurs; social insurance; moral hazard, adverse selection.

JEL: H55, J32, L26.

*We thank seminar participants at Tampere University, VATT, the SustAgeable Consortium Meeting, and the 2023 National Tax Association Meetings for insightful comments.

†University of California, Santa Barbara and NBER
‡Tampere University, FIT, VATT and CESifo
§VATT, FIT and CESifo
¶Tampere University, FIT and VATT

1 Introduction

Self-employment, independent contract work and small business ownership have been steadily rising in recent decades in OECD countries (Boeri et al., 2020). The majority of such small business owners are exposed to higher levels of income volatility than wage earners, and find social (e.g. unemployment and disability) insurance desirable (Audoly, 2022; Boeri et al., 2020). Yet, these individuals are rarely covered by the traditional social insurance schemes: most OECD countries provide some form of social insurance coverage to the self-employed, but it tends to be either of voluntary or income-invariant nature, and therefore only ensures against extreme poverty. Insuring small business owners is challenging as they are likely to be well informed about the riskiness of their businesses and may have a greater ability to take advantage of insurance benefits. As a result, social insurance provision is expected to be highly adversely selected, and costly because of the higher moral hazard responses. However, the actual extent of adverse selection and moral hazard among small business owners have not been measured, largely due to the lack of identifying variation.

In this paper, we take advantage of a unique setting in Finland, to provide the first evidence on the extent of adverse selection and moral hazard in the social insurance (SI) market for the small business owners. In Finland, small business owners have the ability to choose the extent of their SI coverage, thus determining the generosity of pension, unemployment, sick day pay, and parental leave benefits they receive. To measure adverse selection and moral hazard, we take advantage of an exogenous policy change: prior to 2011, only business owners whose ownership share exceeded 50% could choose the level of social insurance benefits. Meanwhile, those with smaller ownership shares were defaulted into a traditional SI scheme, similarly to the wage earners. Starting in 2011, the threshold was lowered to 30%. As a result of this policy change, individuals who owned between 30–50% of their company could now more freely determine the level of SI contributions, thus providing exogenous variation in insurance coverage. This policy change allows us to estimate moral hazard responses and measure the extent of adverse selection using linked administrative

data on individuals' SI contributions, SI benefit claims, as well as individual and business tax returns. We proceed as follows.

First, we use a difference-in-differences (DiD) approach to estimate moral hazard responses. We compare SI contribution choices and resulting claims of individuals who were affected by the policy change (i.e. those with 30–50% ownership share) to those who were unaffected (i.e. those with less than 30% or greater than 50% ownership shares), before and after the reform. We show that the policy change led to an immediate and lasting decrease in SI coverage: treated individuals decreased SI contributions by an average of 20%. Using benefit claims as an outcome variable, we are able to estimate average change in moral hazard responses as a result of this 20% decrease in SI coverage. To see this note that assignment to treatment/control group is fixed over time. Therefore, when we compare pre- and post- claims of treated/control individuals, we compare claims of the same group of individuals. This implies that the adverse selection channel is shutdown, and the simple difference recovers moral hazard responses as long as individuals' risk profiles are time-invariant. The DiD estimator allows us to relax this assumption and account for common changes in risk-profiles over time. Our DiD estimates suggest that sick pay, parental leave, unemployment and pension moral hazard responses are weak and economically insignificant, resulting in a 1–2% increase in claims in response to a 20% change in SI coverage.

Second, to recover the extent of adverse selection, we compare post-reform claims of individuals with different choices of social insurance in the post-reform period. It is well-known that the differences in claims are due to differences in individuals' risk profiles, i.e. adverse selection, and individuals' moral hazard responses. Therefore we can back out the extent of adverse selection by differencing out the moral hazard component using our DiD estimates discussed in the previous step. Using this approach, we find that the extent of adverse selection is limited in this market – we do not detect positive correlations between claiming behavior and social insurance coverage.

Our results suggest that in the social insurance market for entrepreneurs that we

study, and for the types of SI we consider (sick pay, parental leave, pension, and UI), moral hazard responses are weak, and adverse selection is not very prevalent. The latter suggests that a social insurance mandate is not necessary, and instead can be distortionary, as it can impose unnecessary liquidity constraints on young firms (Benzarti et al., 2020). On the other hand, the weak levels of moral hazard responses suggest that entrepreneurs do not abuse the SI system. Therefore, the government offer coverage that is of similar generosity is that for the wage earners.

To the best of our knowledge, we are the first study to estimate moral hazard and adverse selection responses of entrepreneurs. Previous literature has estimated the degree of adverse selection and the extent of moral hazard in other settings (e.g. Finkelstein and McGarry, 2006; Einav et al., 2010; Finkelstein and Poterba, 2004; Landais et al., 2021). Focusing on entrepreneurs is important because they face unique risks, making the provision of social insurance an important policy question for them, and have greater ability to respond to incentives, thus making moral hazard concerns more acute.

2 Institutional Setting and Data

2.1 The Finnish Social Insurance System

Overview. The Finnish social security system is financed by taxes and social insurance contributions (paid by employees and employers), and provides benefits for individuals who live in Finland on a permanent basis, or work temporarily. The system provides a wide range of benefits, including old age, disability, sickness, unemployment, health-care, childbirth, rehabilitation and other benefits. Most social security benefits are based on previously earned incomes and therefore are a function of contributions (e.g., earnings-related pension, unemployment allowance), but some minimum-level benefits are not (e.g., health-care, national pension, basic unemployment allowance). Annual insurance contributions are income-tax deductible.

TyEL vs YEL Contribution Schemes. Wage earners receive the aforementioned social insurance benefits under the TyEL insurance scheme ("Työnantajan eläkevakuutus" in Finnish). Social insurance contributions of TyEL-insurees are fixed: government-set rate applies to individuals' labor earnings. In 2016, the total contribution rate was 23.6% (i.e. the sum of employee plus employer rates).[1]

On the other hand, self-employed individuals and owners of partnerships and privately held corporations access social insurance under one of two schemes. The YEL insurance scheme ("Yrittäjän eläkelaki" in Finnish) applies to any entrepreneur (including self-employed and partnership owners) who owns, alone or together with family members, at least a *certain* percent of their firm, and to those who work in these firm.[2] In addition, business owners who hold a leading position in a privately held corporation (such as the CEO or chairman of the board) and own a certain percentage share of the company's shares are also considered to be YEL entrepreneurs. The above conditions are binding, and entrepreneurs cannot opt out of the YEL insurance scheme. However, if the above conditions are not met, entrepreneurs are automatically subject to the TyEL insurance scheme, i.e. the same insurance program as for the wage earners.

The main difference between the TyEL and YEL schemes is that while the TyEL owners have no discretion over their contribution levels, the YEL owners can freely set the contribution level within a certain range (in 2016, between €7,557 and €171,625), and can adjust it at any point in time. We refer to their chosen contribution level as the *social insurance (SI)* income, and to their overall earnings as the *labor income (LI)*. The contribution rate that applies to the SI income under the YEL scheme is similar to the contribution

[1] Contribution rates vary slightly over time and are set annually. The contribution rates also vary by age, and are higher for older individuals. In 2016, the TyEL rate for persons aged 53 or older was is 25.2 percent instead of 23.6%.

[2] In addition, to qualify for the YEL, entrepreneurs must meet the following conditions: they have to be 18 to 67 years old, their firm must be at least four months old, and the work contribution they derive from the firm has to have value of at least €7,557 per year (in 2016). The Self-employed Persons' Pension Act (HE 1272/2006) is available online here (in Finnish): https://www.finlex.fi/fi/laki/alkup/2006/20061272 (accessed May 24, 2019). More information in English can be found here: https://www.ilmarinen.fi/en/self-employed-person/self-employed-persons-pension-insurance/yel-contributions (accessed May 24, 2019).

rate under the TyEL scheme (23.7% instead of 23.6%).[3] There are no difference in income taxation between YEL and TyEL owners.

To summarize, YEL owners have significantly more discretion over their level of mandatory social insurance contributions than TyEL owners. Since future benefit entitlements are a function of contribution levels, when YEL owners change their contribution SI, they implicitly change the level of benefits they will be entitled to in the future. On the other hand, the contributions and benefit entitlements of TyEL owners are set automatically as a function of their labor income LI. For this reason, we refer to all entrepreneurs to whom YEL rules apply as *unconstrained* owners, and all entrepreneurs to whom TyEL rules apply as *constrained* owners.

2011 Reform. As stated earlier, to qualify for the YEL insurance scheme, the business owner must own, alone or together with family members, at least a *certain* percent of their firm. Before 2011, this threshold was set at 50%. In 2011, this threshold was decreased to the current level of 30%. This means that from 2011 onward, business owners with a 30–50% ownership share could more freely determine their level of social insurance contributions.[4] Note that, by law, entrepreneurs with ownership shares between 30% to 50% could opt out of switching to YEL scheme for 3 years. However, in our data, we see that all of the switches occurred immediately after the reform. Apart from this reform, there were no other notable changes to the social insurance program during the time period we study.

Benefit Entitlements. Generally speaking, earnings-related benefits are a function of individual's SI incomes in the preceding 12-month period, with the exception of pension benefits, which are a function of individual's career-long SI incomes. Most benefits feature

[3]YEL contribution rates and the minimum and maximum YEL income levels vary slightly over time and are determined annually. The contribution rates also vary by age, and are higher for older individuals. In 2016, the YEL contribution rate for persons aged 53 to 62 was 25.1% instead of 23.7%. In addition, the YEL contribution rates are lower for start up business-owners: 19.6% for persons aged 53 to 62, and 18.4% for others.

[4]More information about the reform (HE 135/2010) can be found here (in Finnish): https://www.finlex.fi/fi/esitykset/he/2010/20100135 (accessed May 24, 2019).

a floor, so that individuals with no or small income during the relevant period receive a minimum benefit. In this study we will focus on four earnings-related benefits: parental allowance, sick pay benefits, unemployment benefits and retirement benefits.

Parental allowance can be claimed by either parent or jointly up to a maximum of six months. A typical parental allowance is claimed when the child turns three months old (after the conclusion of the maternity/paternity allowances) and ends when the child turns nine months old. Until 2015, the benefit was equal to 70% of individual's SI in the past 12 months for SI incomes up to €32,892 with lower replacement rates for higher SI levels. For example, in 2010, an individual with $SI =$ €50,000 would receive a daily parental allowance of €99.6.[5]

Sick pay benefit can be claimed when individual faces a long-term health issue. During the first ten days of a sickness spell, an individual has the right to the full salary if her ability to work is hindered by her illness or injury. Only after these ten days pass, individual can claim sickness benefits up to a maximum of 300 working days. The replacement rate follows the same rules as the ones described for the parental allowance.

Earnings-related *unemployment benefits* are calculated based on the SI six months preceding the start of the unemployment spell. The replacement rate was approximately 58% for SI below approximately 42,000 euros in 2023, with lower replacement rates for higher income levels. For example, an individual with €48,000 of annual SI will receive monthly benefits of €2100 per month or €99 per workday, resulting to a 52,5% replacement rate. There is a waiting period of five days before one starts to receive unemployment benefits, and the benefits can be claimed up to a maximum of 500 days.

To qualify for earnings-related unemployment benefits, an individual must have been working for the past ten months preceding the unemployment spell and has been a member of an unemployment fund for the past six months (membership is voluntary and

[5]Once the parental allowance has been exhausted, either parent can claim a childcare subsidy until the child turns three years old. The childcare subsidy has lower replacement rates compared to the parental allowance, but the structure is similar.

imposes a small cost). If these requirements are not fulfilled, the unemployment benefits are set at the basic minimum level, approximately 37 euros per workday in 2023. The basic minimum level of unemployment benefits can be received indefinitely (including after conclusion of earnings-related benefits) as long as the claimant registers with an unemployment office as an unemployed job seeker, is in search of full-time work, and is fit to work.

Pension benefits are calculated based on all the earnings an individual has received during her working career, whether as an employee or as an entrepreneur. Yearly earnings are weighted equally (at 1.5% in 2023), so that the level of pension increases with the number of working years and one's SI contributions irrespective of when these contributions are made during one's career. If earnings-based pension is very low, it is supplemented with a guarantee pension to provide a minimum basic income for elderly individuals.

2.2 Data

We use several administrative data sets in this study: (1) contributions data from the two largest private Finnish pension companies managing the social insurance of the YEL entrepreneurs, (2) individual-level annual benefit claims data, and (3) individual and corporate tax return data covering all entrepreneurs in Finland, which include both firm and owner-level outcomes and other characteristics.

YEL Contributions Data. YEL contribution-level data is obtatined from the two largest Finnish pension companies, that together cover approximately 70% of all entrepreneurs under the YEL insurance scheme. This data is available from 2006 to 2014 and it contains the base on which the insurance is calculated, the amount of annual contributions, and the relevant contract dates.

Tax Return Data. We use tax return data covering the full population of business owners and their firms from 2006 to 2016, extracted from the Finnish Tax Administration database. The data includes information on the financial statements and tax records of all Finnish

businesses and their main owners. The data contains detailed information on individuals' total income and its composition. Importantly, the data allows us to observe SI contributions of the business owners under the TyEL scheme.

Benefit Claims Data. We observe all benefits claimed by the TyEL and YEL business owners. We focus on the following earnings-related benefit claims: sick pay, unemployment benefits, parental leaves, and retirements. The data provides information on the amount of benefits received, by combining this data with information on individuals' SI contributions, we are able to accurately back out the number of days claimed using statutory benefit formulas. We identify individuals' retirement dates from the pension benefit claims. Our analysis focuses on the following claims measures. For *sick pay*, we record the number of sick days claimed during the year for for TyEL owners, while for YEL we record the max(0, sick days - 10) to account for differences in sick pay claim rules. For *parental leave benefits,* we record the number of days taken for parental leave in a given year, both for TyEL and YEL owners. For *pensions* claims, we record whether or not individual retired in a given year. Finally, for the measurement of unemployment spells for entrepreneurs is difficult because no formal displacement notice is recorded and because it is not possible to observe whether a given business owner belongs to the unemployment fund or not.[6] for this reason, we assume that all unemployed owners receive flat rate minimum benefits and calculate the number of unemployed days accordingly, both for YEL and TyEL owners.

Survey Data. We use survey data from two separate surveys conducted by Statistics Finland regarding the entrepreneur's pension system. The first survey carried out in 2013 garnered 1573 responses, with 74% identifying as entrepreneurs. The second survey was conducted in 2017, with 1072 responses (13.4% response rate), out of whom 94.8% were under the YEL scheme. The surveys ask respondents about their businesses, factors influencing low

[6] FSA (2010) reports that approximately 27,000 entrepreneurs (out of 200,000+) were members of the unemployment fund in 2010.

insurance coverage, health and well-being, absences related to family or illness, awareness of, attitudes toward, and the perceived need for the social insurance system, as well as insights into their retirement plans. Specifically, we focus on questions pertaining to pensions, social insurance and subjective well-being. By combining this data with information on individual's SI contributions, we are able to discern correlations between SI contribution levels and subjective dispositions.

Sample Restrictions. Since we only focus on business owners of privately-held corporations, we exclude all other businesses from the sample. Summary statistics for our main sample are available in Table 1.

3 Conceptual Framework

Our reduced-form empirical approach takes advantage of the 2011 reform which allowed a subset of entrepreneurs to freely choose their level of SI contributions and therefore the magnitude of future SI benefits. In this section, we describe how we can use this reform to identify the extent of moral hazard and adverse selection in our setting.

In our conceptual framework, a business owner i's claim in year t can be represented by the sum of two components: his naturally-occurring claims $True(risk_i^t)$, which are determined by the owner's risk type $risk_i^t$, and his moral hazard response $MH(SI_i^t/LI_i^t)$. The latter is assumed to be driven by the generosity of his replacement rate SI_i^t/LI_i^t, the ratio of reported social insurance income relative to his labor income. Thus, $Claim_i^t = True(risk_i^t) + MH(SI_i^t/LI_i^t)$, and for simplicity, we assume that business owners only vary in their risk type, and there is no randomness in claims across years.

Prior to 2011, business owners whose business share was between 30% and 50% contributed according to their income level, so that a given business owner's social insurance income was the same as their income, i.e. $SI_i^{pre} = LI_i^{pre}$. Since all business owners faced the same replacement rate, their moral hazard components were the same, $MH(SI_i^{pre}/LI_i^{pre}) =$

$MH(SI_j^{pre}/LI_j^{pre})$.

After 2011, treated business owners were able to choose SI_i^{post} according to their risk type. Figure 1 illustrates the incentives of these individuals. Generally speaking, lower-risk individuals (e.g. individual A in Figure 1) prefer lower insurance coverage $SI_i^{post} < LI_i^{post}$, while higher-risk individuals (i.e. individual D) prefer higher coverage $SI_i^{post} > LI_i^{post}$. Importantly, note that since the reform does not directly affect business owners' risk types, $True(risk_i^{post}) = True(risk_i^{pre})$.

In this deterministic setting, a simple difference of each individual's claims before and after the 2011 reform recovers each individual's moral hazard response, i.e.:

$$Claim_i^{post} - Claim_i^{pre} = MH(SI_i^{post}/LI_i^{post}) - MH(SI_i^{pre}/LI_i^{pre}). \tag{1}$$

In practice, the identifying assumption $True(risk_i^{post}) = True(risk_i^{pre})$ may not hold for exogenous reasons. For example, individuals' risk types are unlikely to stay constant over time and claims may vary from year to year. We can account for such exogenous fluctuations using a difference-in-differences design, thus comparing treated individuals' claims before and after the reform to those of a control group. Two plausible control groups exist in our setting: the always constrained individuals and the always unconstrained individuals.

Next, in order to estimate the extent of adverse selection, we must separate moral hazard component from the adverse selection, since after the reform, differences in treated individuals' claims reflect the sum of adverse selection and moral hazard:

$$Claim_j^{post} - Claim_i^{post} = \underbrace{[True(risk_j^{post}) - True(risk_i^{post})]}_{AdverseSelection} + \underbrace{[MH(SI_j^{post}/LI_j^{post}) - MH(SI_i^{post}/LI_i^{post})}_{MoralHazard}. \tag{2}$$

In this deterministic setting, we can use (1) and the fact that $MH(SI_j^{pre}/LI_j^{pre}) = MH(SI_i^{pre}/LI_i^{pre})$

to estimate the moral hazard component,

$$MoralHazard = [MH(SI_j^{post}/LI_j^{post}) - MH(SI_j^{pre}/LI_j^{pre})] - [MH(SI_i^{post}/LI_i^{post}) - MH(SI_i^{pre}/LI_i^{pre})], \quad (3)$$

and therefore back out the extent of adverse selection from (2). In practice, the identifying assumptions $True(risk_i^{post}) = True(risk_i^{pre})$ (implicit in (1)) and $MH(SI_j^{pre}/LI_j^{pre}) = MH(SI_i^{pre}/LI_i^{pre})$ may not hold for exogenous reasons. In addition to facing fluctuating risk types and risk preferences, business owners may vary in the magnitude of moral hazard responses. We can account for these heterogeneity by controlling for individuals' demographic characteristics and using income controls. However, we will not be able to separate adverse selection from moral hazard if the two are correlated. For example, if high-risk individuals exhibit higher levels of moral hazard, even when holding the replacement rate constant (i.e. $SI_i^{pre}/LI_i^{pre} = SI_j^{pre}/LI_j^{pre}$), then our measure of adverse selection will be biased upward.

4 Moral Hazard Responses

4.1 Empirical Approach

We measure moral hazard responses using an event study approach around the 2011 reform. Our treatment group consists of individuals whose business ownership share was between 30% and 50% in 2010.[7] We use two comparison groups. Our first comparison group, *the always restricted*, consists of business owners who own between 10% and 30% of their company shares. These individuals remain in the T-status and therefore cannot choose SI contributions throughout the period of study. Our second control group, *the always unrestricted*, consists of business owners who own 50–70% of the company shares. These Y-owners are free to

[7] We show that business ownership shares remain stable throughout owner's business life. For this reason, more conservative definitions based on ownership share throughout the period of study or using all pre-reform years yield similar results. See Appendix Figure A.1 for details.

choose their SI contributions throughout the period of study. In the main text, we show the results that use first comparison group as a control. Arguably, this group remains untreated throughout the period of study and thus satisfies the assumptions of a canonical DiD model (Tazhitdinova and Vazquez-Bare, 2023). Nonetheless, the results are qualitatively equivalent when using the second control group instead, and are available in Appendix C.1. Note that our event study design is robust to the issues highlighted by de Chaisemartin and D'Haultfoeuille (2020); Sun and Abraham (2021); Callaway et al. (2021); Goodman-Bacon (2021) because the treatment occurs in the same year for all treated units.

Our event studies estimate how the reported SI income and the number of benefit days claimed (altogether and separately for sick pay leave, parental pay leave and unemployment pay) change for the treated business owners relative to the control:

$$Outcome_{it} = \sum_{\substack{\ell=2006 \\ \ell \neq 2010}}^{2014} \beta_\ell \; Treat_i \; 1_{t=\ell} + \eta_t + \delta X_{it} + \varepsilon_{it}. \tag{4}$$

Subscript i identifies business owners, and t year; δ_t are year fixed effects, and X_{it} are individual controls: age, gender, municipality, and industry. The coefficients of interest β_ℓ measure average moral hazard responses of an average treated individual who switched from an average pre-reform level of SI coverage to average post-reform level of coverage.

Following the discussion in Section 3, our identification relies on two assumptions: (1) individual risk types do not change as a result of the reform, and (2) the risk profiles of the treated and control individuals evolve similarly over time. In other words, we allow individuals' risk-based claims $True(risk_i^t)$ to fluctuate over time for each individual i, but assume that the nature of these fluctuations is exogenous to ownership status and is similar for treated and control individuals. The first assumption is plausibly satisfied due to the nature of the reform and the fact that the treatment/control status is determined by the pre-reform ownership share. This assumption ensures that the adverse selection channel is shut down, and our estimates recover moral hazard responses. The second assumption is the

standard parallel trends assumption and allows us to account for exogenous fluctuations in claims.

Since the 2011 reform allowed individuals to adjust their SI coverage according to their insurance needs and preferences, SI coverage could go up or down. For this reason, we also estimate (4) separately for 3 groups of individuals: individuals who chose to substantially decrease their SI coverage, those who chose to substantially increase it, or leave it approximately unchanged. To implement these specifications, we break business owners into three groups based on the average value of SI/LI in the post-reform years: those with $\frac{1}{4}\sum_{t=2011}^{2014} SI_i^t/LI_i^t \leq 0.8$, those with $0.8 \leq \frac{1}{4}\sum_{t=2011}^{2014} SI_i^t/LI_i^t \leq 1.1$, and those with $\frac{1}{4}\sum_{t=2011}^{2014} SI_i^t/LI_i^t > 1.1$. Approximately 52.3% of treated individuals fall into the first group, 33.3% fall into the second group, and 14.3% in the last group. The results shown in the main text compare each treated group to the same comparison group – the always restricted business owners. In Appendix C.1, we show the results where we compare each group to a *corresponding* group of always unrestricted business owners (i.e. those whose average SI/LI is less than 0.8, between 0.8 and 1.1, or over 1.1); the results are qualitatively similar.

4.2 First stage: Choice of SI Income

Figure 2 documents how SI income choices changed as a result of the reform. Figure 2(a) shows that prior to 2011, treated owners' SI income followed a path similar to that of the always restricted owners. Immediate after the 2011 reform, we see a dramatic reduction in treated individuals' SI income. Figure 2(b) shows the distribution of one-year changes among *treated* entrepreneurs immediately before (2009–2010) and after the reform (2010-2011). This figure shows that the average response conceals a fair amount of heterogeneity. Relative to the pre-reform distribution that a showed a large mass to the right of zero (reflecting small increase in SI income as labor income grew over time), the whole distribution shifted to the left. We see a large decrease in the number of individuals who did not change their SI income

and large increase in the number of individuals who decreased their SI income by over 25%.

Figures 2(c) and (d) present event study estimates of equation (4). Figure 2(c) shows that the reform immediately decreased the SI/LI replacement rate by 10 log points and by over 20 log points by 2014. Figure 2(d) shows that the response was indeed heterogeneous, with some business owners increasing their replacement rates while others decreasing it. However, for all 3 groups, we see a decreasing pattern over time, which may reflect learning about one's risk type and benefit use.

Additional evidence is available in Appendix A. Figure A.2(a) plots business owner's SI income against their labor income LI, separately for the unrestricted (Y owners) and restricted (T owners) business owners. Figure A.2(a) shows that with the exception of low-income individuals who are bound by the minimum SI provisions, unrestricted owners choose lower levels of SI coverage than restricted individuals. Figure A.2(b) shows the distribution of chosen SI incomes of Y owners. The mean SI income is €25,000 (mean labor income is €38,000), and we see that many business owners choose round numbers when picking the value of SI income.

4.3 Second Stage: Number of Benefit Days Claimed

In Figures 3 and 4 we present our main estimates of moral hazard responses. Figure 3 shows event study estimates of (4) for all treated individuals combined and using always restricted business owners as the control. Figure 4 repeats this analysis, but separately for individuals with SI/LI below 0.8, between 0.8 and 1.1 and above 1.1.

In Figure 3(a), the outcome variable is the total number of benefit days claimed, including sick leaves, parental leaves and unemployment benefit days. Panels Panels (b)–(d) estimate (4) separately for each benefit categories. In all figures with the exception of Figure 3(e), we see no evidence of diverging trends, suggesting that treatment and control group follow each other well.

Figure 3(a) shows that as a result of the reform, the number of benefit days

claimed increased by between 2-3 days per year. This result is surprising since the majority of business owners decreased their replacement rates and therefore one would expect a decrease in the number of benefit days claimed rather than an increase. Figures 3(b)-(e) show that this increase is driven by increased reliance on sick pay and parental pay. In contrast, the number of unemployment insurance days appear to remain largely unchanged, except for the first year after the reform, where we see a sharp 4-day increase in the number of UI days claimed.

Figures 3(d) and (e) show the effects on unemployment days. When using always restricted as a control group, we have some issues of different pre-trends that we need to take into account in interpreting these results. However, when using entrepreneurs which are always unrestricted as a control group, the pre-trends are significantly better (like when using restricted as a control group), and in this specification we observe a gradual decrease over time. This suggest a small moral hazard behavior, and this effect seems to also drive the results for total absent days in Panel (a).

Figure 3 shows that for many benefits, we see a small increase in claims, despite the fact that the majority of individuals reduced their SI benefits. One potential explanation for this result is that it is driven by a small minority of individuals who increased their SI benefits. We test this possibility next. Figure 4 examines how claims changed depending on individuals' choice of SI income. As discussed in Section 4.1, we break treated business owners into three groups depending on the average post-reform replacement rate: those with SI/LI below 0.8, between 0.8 and 1.1 and above 1.1 (recall Figure 2(d)). If individuals responses are in part driven by moral hazard considerations, we should see an increase in benefit claims for individuals who increased their SI income, and a decrease in claims for individuals who decreased their SI income.

Figure 4 provides mixed evidence in support of moral hazard presence. On the one hand, we see a larger increase in total days claimed among individuals who increased SI income above their labor income than among individuals who decreased their SI income.

In particular, these business owners are more likely to claim sick pay and basic UI days. However, this pattern does not hold for all benefit types – e.g. we do not see it for parental pay days and earnings-based UI days. Furthermore, both individuals who increased their SI income and individuals who decreased their SI appear to increase their claims on average. Similar to Figure 3, we see a small increase in the number of sick pay and parental pay claims.

4.4 Discussion

Discuss potential explanations: quality of comparison rules, differences in SI claim rules, awareness of benefits and contributions. In progress.

5 Adverse Selection Responses

The 2011 reform allowed entrepreneurs to choose social insurance levels more freely according to their risk type. For example, individuals who expect to claim sick pay would prefer to have higher levels of coverage than individuals who do not expect to be sick. Therefore, the ability to sort according to one's risk type should result in a positive relationship between one's risk type and one's chosen insurance level – i.e. adverse selection.

In this section we turn to estimating the extent of such adverse selection in our setting. In Section 5.1, we explore whether individuals who choose higher levels of SI coverage claim benefits more frequently, and in Section 5.2, whether their claiming rates change over time and to what extent they vary with business owners' insurance needs. In Section 5.3, we study anticipatory responses: whether individuals increase SI coverage level prior to a predictable claim event and decrease SI coverage thereafter.

5.1 Relationship Between SI Income Choice and Benefit Claims

A well-known approach to estimating a lower bound on the extent of adverse selection is to measure the relationship between pre-reform claims and post-reform insurance choices. This approach shuts down the moral hazard channel by sorting individuals on immutable pre-reform claims and therefore reveals to what extent plausibly higher-risk individuals choose higher-coverage contracts. Unfortunately, this test may not reveal the full extent of adverse selection because individuals' risk types may change over time, making pre-reform claims an imperfect proxy for individuals' current risk type.

We conduct such test in Panel A of Figure 5, which plots average pre-reform claims (in years 2007-2010) by bin of SI/LI, and separately for sick pay days, parental pay days, and unemployment pay. Panel A shows a weak relationship between one's pre-reform claims and post-reform SI replacement rate choice. The weak correlation suggests that either our setting does not exhibit adverse selection, or that the relationship is biased towards zero because individuals' risk types vary too much over time. Indeed, in our setting, we find that claims do not show high degree of correlation over time (see Appendix Figure XXX).

Panel B of Figure 5 shows equivalent relationship but for average post-reform claims (years 2011-2014). This analysis is not biased towards zero, but it conflates adverse selection responses with moral hazard responses, since both lead to a positive relationship between one's claims and one's SI coverage choice. Nonetheless, we again see a weak relationship between claims and SI coverage level. Figure 5 shows that we can confidently reject that there is a strong positive correlation between post-reform claims and post-reform social insurance contributions. In fact, some of the coefficients are negative suggesting the presence of advantageous selection.

To estimate the extent of actual adverse selection, we must differentiate between adverse selection and moral hazard. We do so by following the strategy described in Section 3: we use equation (2) to recover the extent of adverse selection by subtracting moral hazard responses estimated in Section 4.3 from the total differences in claims observed in Panel B

of Figure 5. Since we are not able to recover moral hazard response for each individual, we measure the extent of adverse selection for three groups of business owners: individuals who chose SI coverage with $SI/LI \leq 0.8$, $0.8 \leq SI/LI \leq 1.1$, and $SI/LI > 1.1$. These results are summarized in Table 2 (in Progress).

5.2 Learning and Relevance

Our results in Section 5.1 suggest that there is only a small amount of adverse selection in our setting. In this section, we explore why this might be the case by considering three explanations: learning, and the bundle constraint.

First, since our treated individuals are new to the choice of SI income, they might not be fully aware of the SI cost-benefit relationship, or of their own risk type. Both types of information frictions suggest that as individuals learn over time, we should see an increasingly positive relationship between claims and contributions. Indeed, Appendix Figure B.9 lends support to this possibility by showing that the always unrestricted individuals exhibit a more pronounced and positive relationship between SI contribution levels and contemporaneous claims than treated individuals. To investigate the effect of learning directly, we compare the claims-insurance relationship immediately after the reform (in 2011–2012) to 3-4 years after (in 2013–2014). Figure 6 shows the relationship for total claims, while Appendix Figures B.10-B.12 consider each claim type individually. Overall, we do not see a consistent increase in positive correlation for the treated individuals over time. Combined with evidence from Figure B.9, this suggests that the learning process is slow and takes longer than the four years we considered in Figure 6.

Generally speaking, it is difficult to distinguish which type of information friction individuals experience – whether they are not aware of their risk type or of the cost-benefit link. Since for parental leaves, individuals become aware about their risk type many months before the benefits are claimed, differences in claims-insurance relationship should reflect learning about the cost-benefit link. Figure 7 explores whether the claims-insurance rela-

tionship strengthens for second+ child compared to the first child. For both treated and always unrestricted individuals, we see a notable increase in positive association between parental pay claims and SI insurance levels for second and higher count children. This highlights the importance of learning about the benefit rules for insurance choice decisions.

The second reason why we might observe a weak relationship between SI income choice and claims is because the SI contributions provide entrepreneurs with a bundle of benefits – sick pay, parental pay, unemployment and pension benefits. As a result, an individual may not be able to sort correctly for each benefit choice, and may instead choose their insurance coverage based on one or two benefit types he is most likely to claim. If this is the case, we should observe a strong positive claims-insurance relationship for "important" benefits and no claims-insurance relationship for "unimportant" benefits. To test this hypothesis, for each individual, we identify the benefit type that they claim the most (in euros). We find that in a given year, one benefit type typically accounts for the majority of benefits claimed in that year. Then for each benefit type, we divide individuals into two groups – entrepreneurs for who this type of benefit accounts for the majority of their claims, and entrepreneurs who rarely claim this benefit type, and therefore should not make decisions based on its coverage schedule.

Figure 8 presents the results from this exercise.

5.3 Choice of SI Income in Anticipation of Future Claims

Individuals may have private information about the likelihood of future claims, allowing them to modify their contribution rates accordingly. For example, business owners have an incentive to increase SI income if they expect to suffer from a major illness, if they are expecting a child, or if they plan to retire, etc. Such anticipatory responses should manifest in an increase of SI income just prior to the benefit claim and a decrease after the claiming period concluded.

Let t_0 identify the year in which a major claim is made. Then we can study the

evolution of SI/LI_i^t around time t_0 and relative to period t_{-3}, i.e 3 years prior to the claim, using specification:

$$SI/LI_i^t = \alpha + \beta_{-2} \cdot t_{-2} + \beta_{-1} \cdot t_{-1} + \beta_0 \cdot t_0 + \beta_1 \cdot t_1 + \beta_2 \cdot t_2 + X_{it} + \varepsilon_{it}, \tag{5}$$

estimated on 6 years around the major claim. If SI income is *actively* chosen by the business owners, then statistically significant coefficients β_j, $j = -2, ..., 2$ reflect anticipatory responses. We carry out this approach for all of our outcomes: sick leave, parental leave, and unemployment receipt. We also estimate this specification for retirement decision, a major event that entrepreneurs are most likely to anticipate. For this outcome, we consider a longer pre-event time period (seven years). However, once an individual retires, they stop making SI contributions, therefore we do not include post-event indicators.

Figure 9 shows how entrepreneurs modify their contribution rates before and after experiencing a sickness, having a child, or experiencing unemployment, while Figure 10 shows how entrepreneurs modify their contribution rates before retirement. Note, that for all of these benefit types, the benefits received are calculated based on the SI contributions in the 12 month preceding the claim.

For parental leave and UI claims, we see an increase in SI income choice in the year of claiming. Intuitively, an entrepreneur has full control over their UI claiming and therefore

6 Policy Implications

7 Conclusion

References

Audoly, R. (2022): "Self-employment and Labor Market Risks," Mimeo.

Benzarti, Y., J. Harju, and T. Matikka (2020): "Does Mandating Social Insurance Affect Entrepreneurial Activity?" *American Economic Review: Insights*, 2, 255–268.

Boeri, T., G. Giupponi, A. B. Krueger, and S. Machin (2020): "Solo Self-Employment and Alternative Work Arrangements: A Cross-Country Perspective on the Changing Composition of Jobs," *Journal of Economic Perspectives*, 34, 170–195.

Callaway, B., A. Goodman-Bacon, and P. H. C. Sant'Anna (2021): "Difference-in-Differences with a Continuous Treatment," arXiv Working Paper 2107.02637.

de Chaisemartin, C. and X. D'Haultfoeuille (2020): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–2996.

Einav, L., A. Finkelstein, and M. R. Cullen (2010): "Estimating Welfare in Insurance Markets Using Variation in Prices," *The Quarterly Journal of Economics*, 125, 877–921.

Finkelstein, A. and K. McGarry (2006): "Multiple dimensions of private information: evidence from the long-term care insurance market," *American Economic Review*, 96, 938–958.

Finkelstein, A. and J. Poterba (2004): "Adverse selection in insurance markets: Policyholder evidence from the UK annuity market," *Journal of political economy*, 112, 183–208.

FSA (2010): "Työttömyyskassat 2010," Annual Report the of Financial Supervisory Authority.

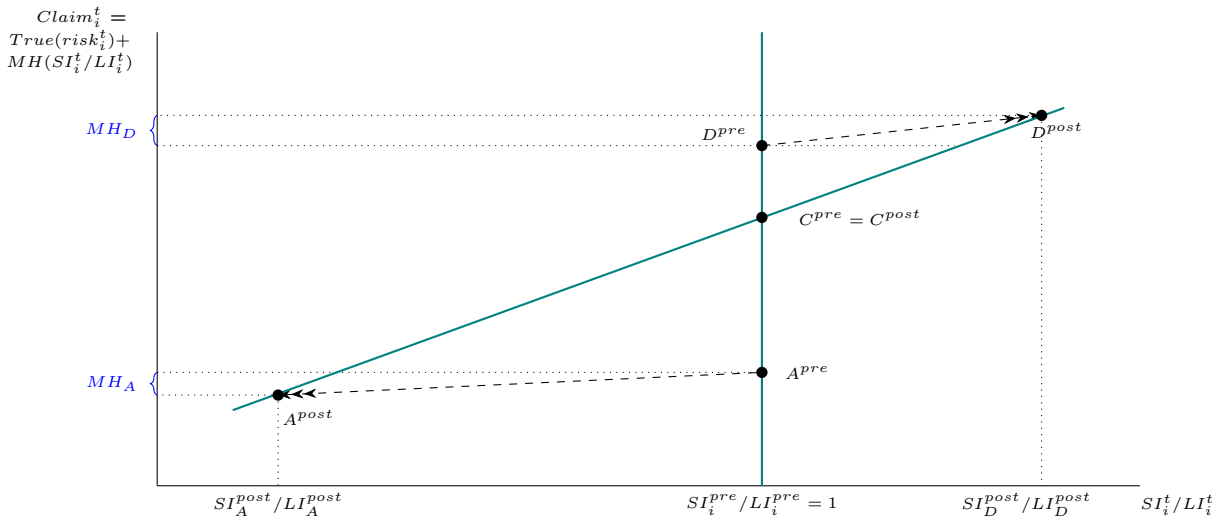
Goodman-Bacon, A. (2021): "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 225, 254–277.

Landais, C., A. Nekoei, P. Nilsson, D. Seim, and J. Spinnewijn (2021): "Risk-based selection in unemployment insurance: Evidence and implications," *American Economic Review*, 111, 1315–1355.

Sun, L. and S. Abraham (2021): "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 225, 175–199.

Tazhitdinova, A. and G. Vazquez-Bare (2023): "Difference-in-Differences with Unequal Baseline Treatment Status," NBER Working Paper 31063.

Figure 1: Illustration of the 2011 Reform Incentives and Identification Approach

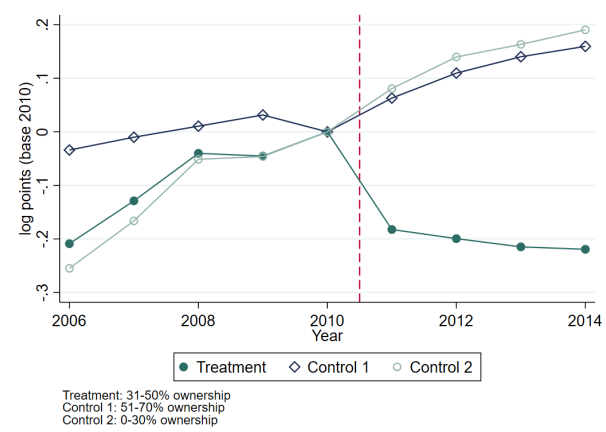


Notes: Prior to the reform, everybody contributed according to their income level, $SI_i^{pre} = LI_i^{pre}$ for all individuals. After the reform, individuals can choose SI income, and therefore sort according to their risk type: low risk types (individual A) are likely to choose lower-coverage insurance contracts with $SI_i^{post}/LI_i^{post} < 1$ while high-risk individuals (individual D) are likely to choose higher-coverage contracts with $SI_j^{post}/LI_j^{post} > 1$. Finally, some individuals may choose to report $SI_k^{post} = LI_k^{post}$ (individual C).

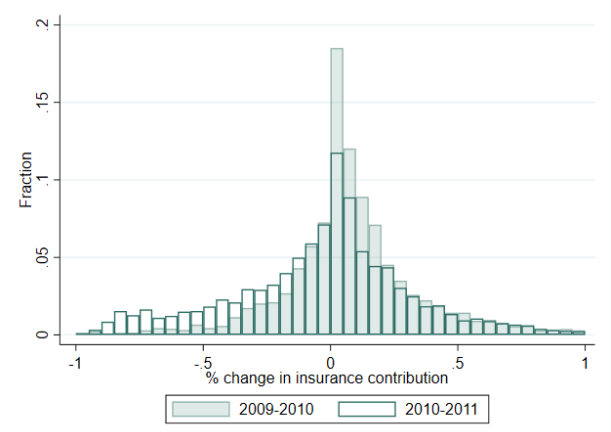
Differences in given a given individual's claims before and after the reform measure the extent of moral hazard: e.g., MH_A measures the moral hazard response from switching from contract with $SI_A^{pre}/LI_A^{pre} = 1$ to contract with $SI_A^{post}/LI_A^{post} < 1$. Differences in individuals' claims levels prior to the reform reveal the extent of risk type and preference heterogeneity present in the population. If individuals' risk types and claims are stable over time, then the correlation between post-reform insurance contracts SI_i^{post}/LI_i^{post} and pre-reform claims $Claim_i^{pre}$ reflect the extent of adverse selection. However, if risk types and claims are not stable over time, then one can recover the extent of adverse selection by subtracting the moral hazard component from the overall differences in claims between individuals with different insurance contracts. See discussion in Section 3.

Figure 2: Changes to Reported SI Income

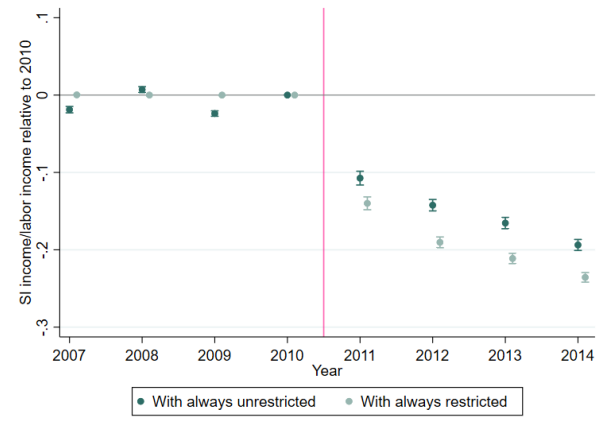
(a) SI Income: Time Series



(b) SI Changes for Treated Individuals



(c) SI/LI : Event Study Estimates

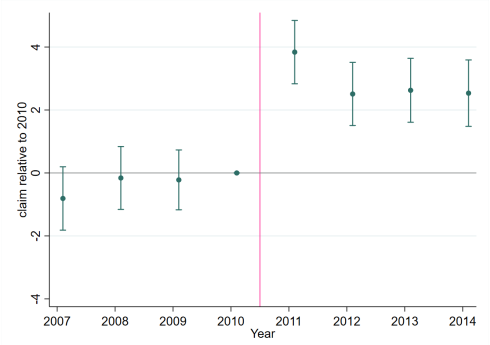


(d) SI/LI Heterogeneity: Event Study Estimates

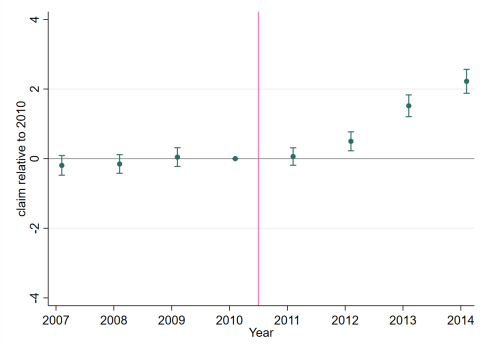
Notes: Figure (a) shows the evolution of social insurance income over time relative to 2010 for treated individuals (those with ownership share between 30-50%), the always restricted (those with ownership share between 10% and 30%) and the always unrestricted (those with ownership share between 50% and 70%). Figure (b) shows the distribution of changes in insurance contributions between 2009 and 2010 and between 2010 and 2011 for the treated individuals. Figure (c) shows coefficient estimates from an event study specification (4), which compares the ratio of SI income to labour income of treated individuals to that of the always restricted or always unrestricted business owners, relative to 2010. Figure (d) again shows coefficient estimates from an event study specification (4), but after separating treated individuals into three groups based on the average SI/LI in post-reform years (2011-2014): those with SI/LI below 0.8, between 0.8 and 1.1, and above 1.1. In figure (d), the control group consists of all always restricted individuals whose ownership share is between 10 and 30%. (See also Figure C.13.) Standard errors are clustered at the individual level, 95% confidence intervals shown.

Figure 3: Moral Hazard Responses – Number of Days Claimed

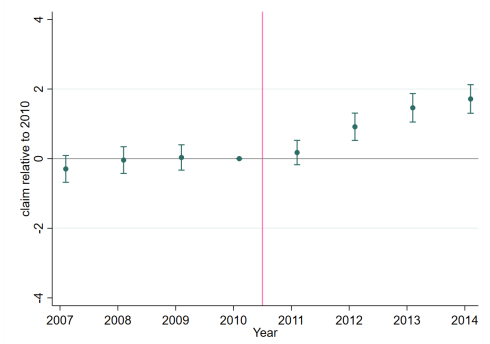
(a) Total Absent Days



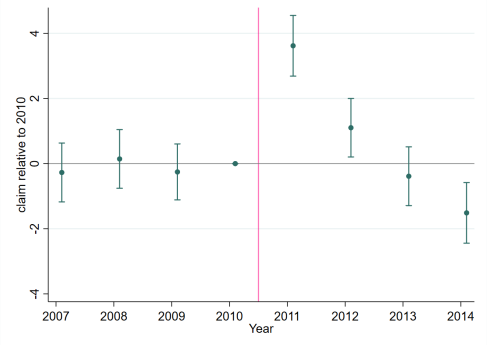
(b) Sick Pay Days



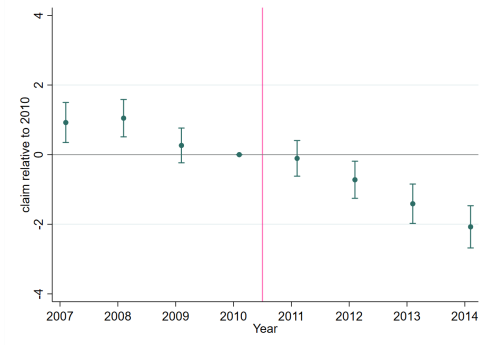
(c) Parental Pay Days



(d) Basic UI Pay Days



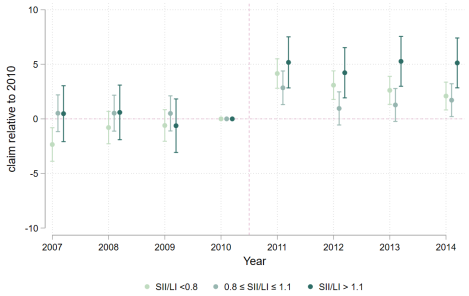
(e) Earnings-Based UI Pay Days



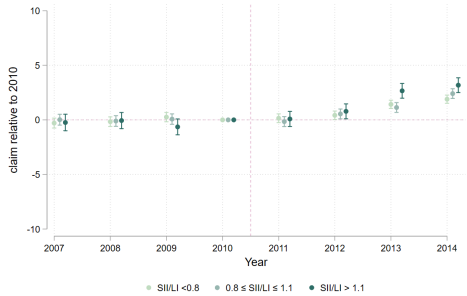
Notes: These figures show the results of estimating (4), where the outcome variable measures the number of days per year an individual claims absence (including zeros), and the control group consists of all always restricted individuals (those with ownership share between 10% and 30%). In Figure (a), the outcome is the total number of absent days, including absences due to sickness, parental leave, and unemployment. In Figures (b), (c), (d) and (e), the outcome is the number of absent days due to sickness, parental leave, basic unemployment, and earnings-based unemployment, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown. Equivalent results but using the always unrestricted individuals as a control group are available in Figure C.14. Extensive and intensive margin responses shown in Appendix B.1.

Figure 4: Moral Hazard Response Heterogeneity – Number of Days Claimed

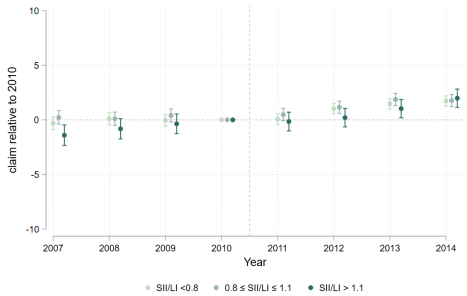
(a) Total Absent Days



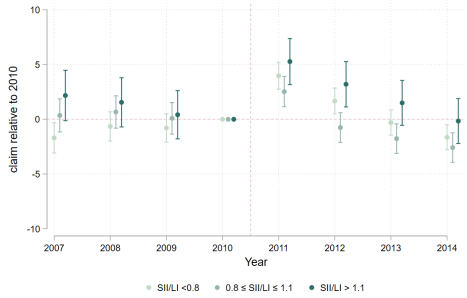
(b) Sick Pay Days



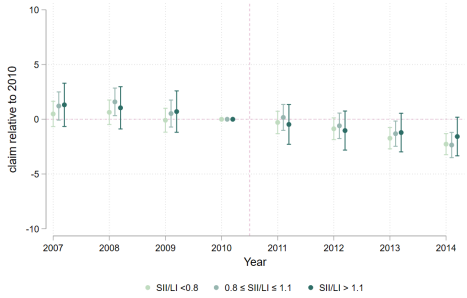
(c) Parental Pay Days



(d) Basic UI Pay Days



(e) Earnings-Based UI Pay Days

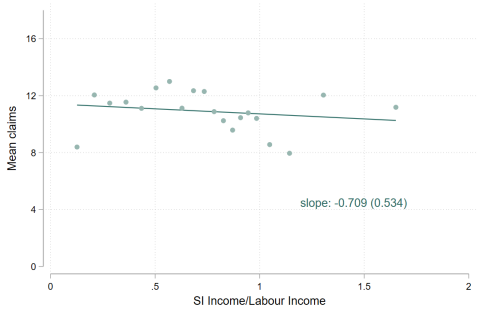


Notes: These figures show the results of estimating (4), where the outcome variable measures the number of days per year an individual claims absence. In each figure, the treated individuals are separated into three groups based on the average SI/LI in post-reform years (2011-2014): those with SI/LI below 0.8, between 0.8 and 1.1, and above 1.1. The control group consists of all always restricted individuals (whose ownership share is between 10 and 30%). In Figure (a), the outcome is the total number of absent days, including absences due to sickness, parental leave, and unemployment. In Figures (b), (c), (d) and (e), the outcome is the number of absent days due to sickness, parental leave, basic unemployment, and earnings-based unemployment, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown. Equivalent results but using the always unrestricted individuals as a control group are available in Figure C.15.

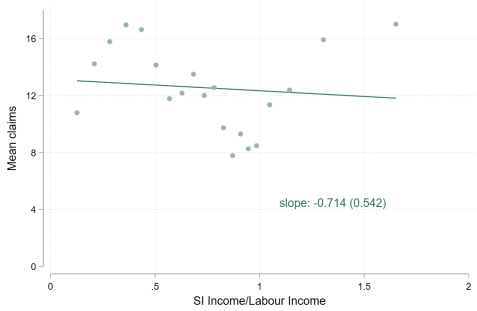
Figure 5: Adverse Selection – Number of Days Claimed

Panel A: Pre-reform Claims **Panel B: Post-reform Claims**

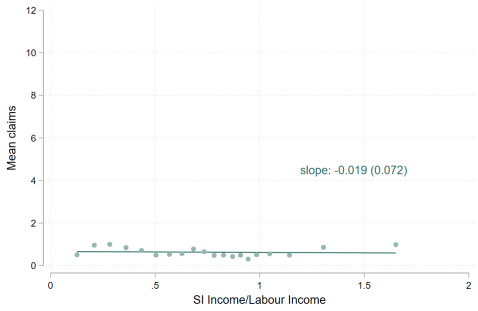
(a) All claims



(b) All claims

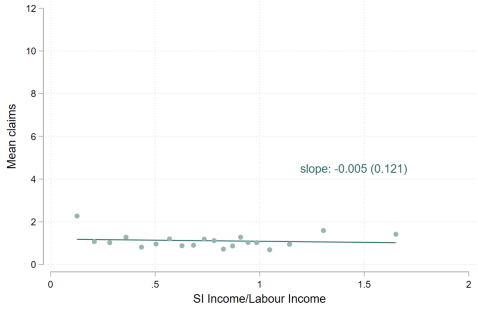


(c) Sick Pay



(d) Sick Pay

(e) Parental Pay



(f) Parental Pay

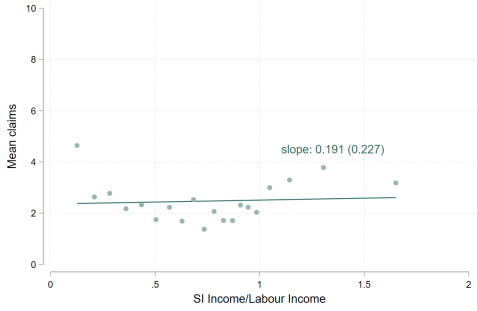
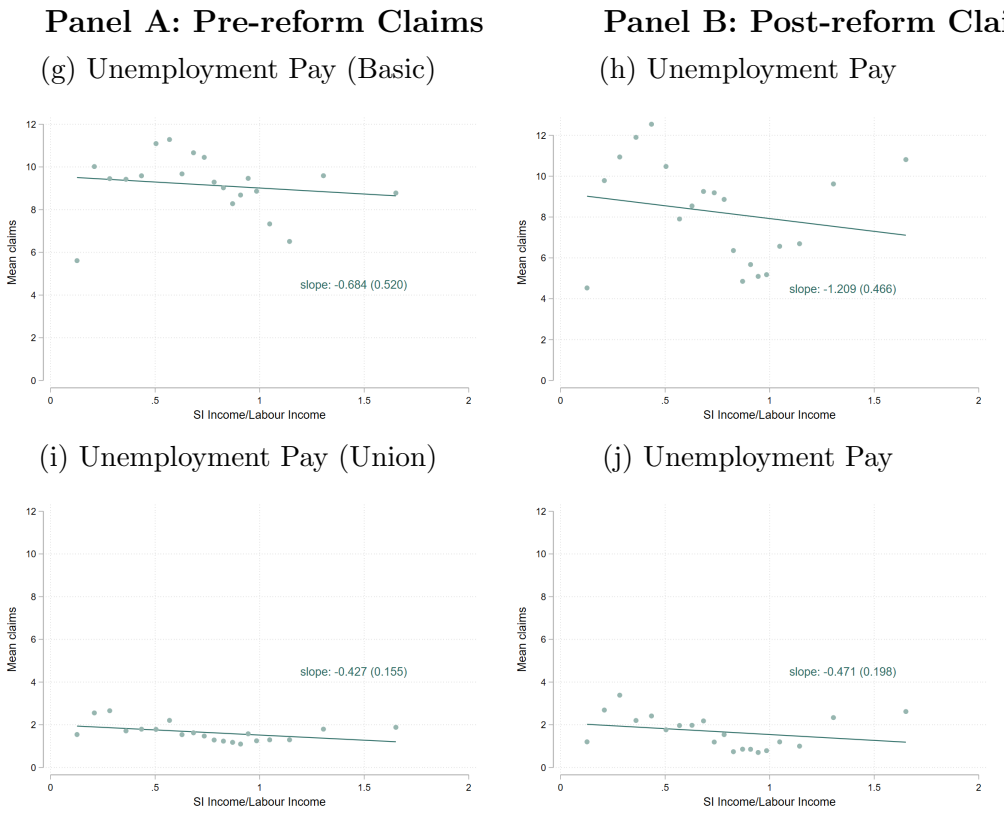
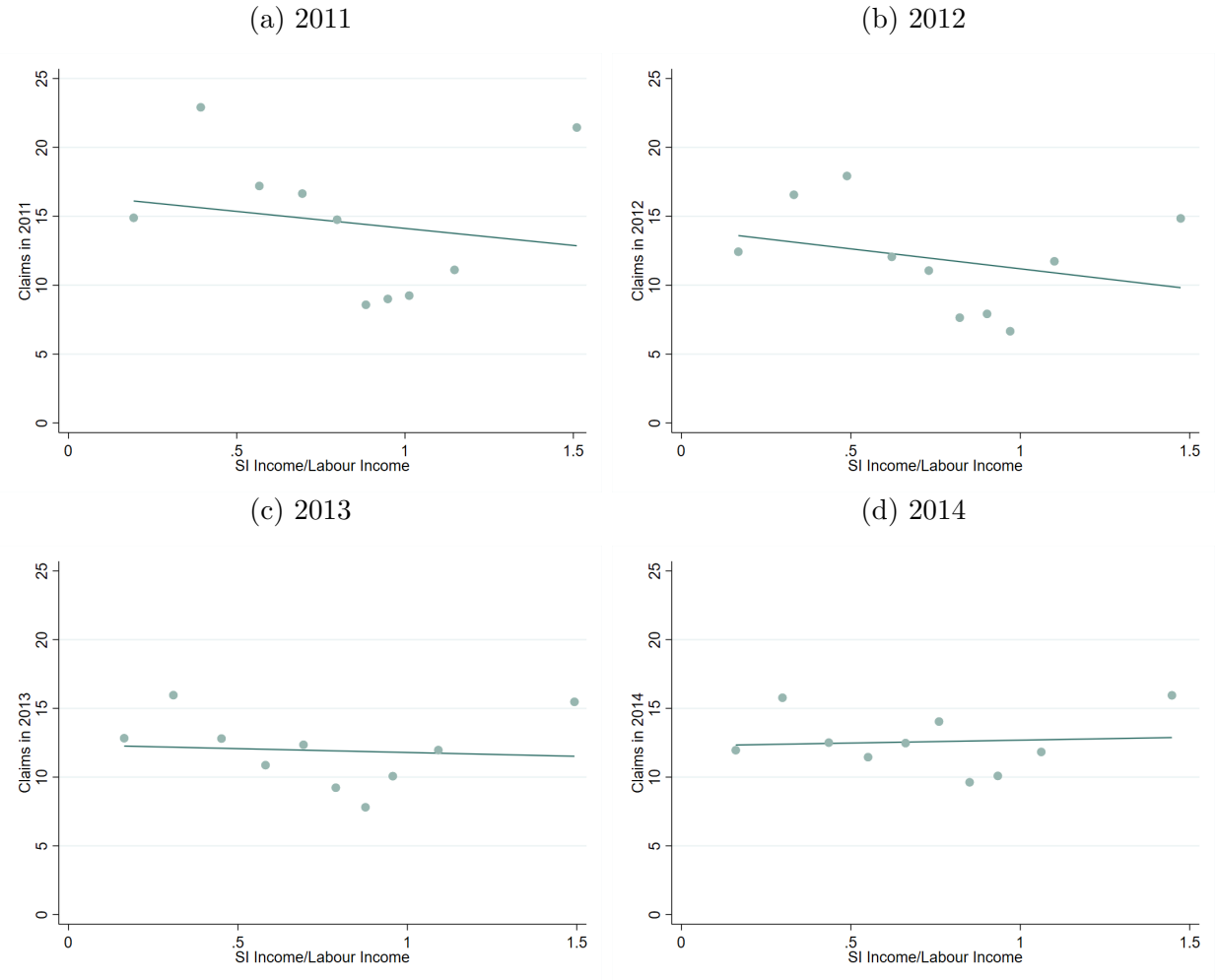


Figure 5: Number of Days Claimed



Notes: These figures show the average number of benefit days claimed by the treated individuals (those with ownership share between 30-50%) in 2006-2010 (Panel A) and in 2011-2014 (Panel B) by bins of SI/LI in 2011-2014. Each SI/LI bin contains 5% of 2011-2014 person-year observations. The y-axis shows the average number of days claimed (including zeros): in Figures (a) and (b) the outcome is the total number of absent days, in (c) and (d) – the number of sick pay days, in (e) and (f) – the number of parental pay days, in (g) and (h) – the number of basic unemployment pay days, and in (i) and (j) – the number of earnings-based unemployment days. The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry. Intensive margin figures shown in Appendix B.2.

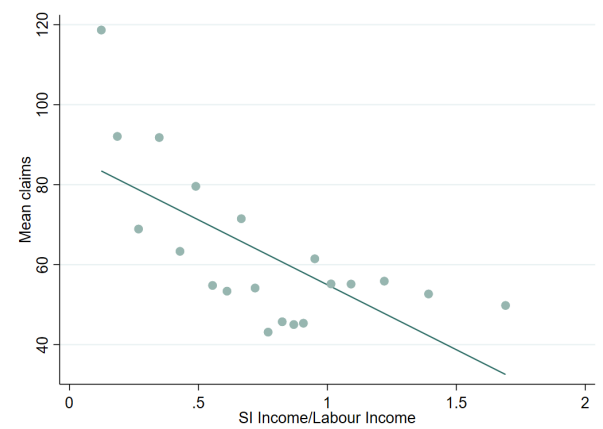
Figure 6: Learning Over Time – Total Number of Days Claims



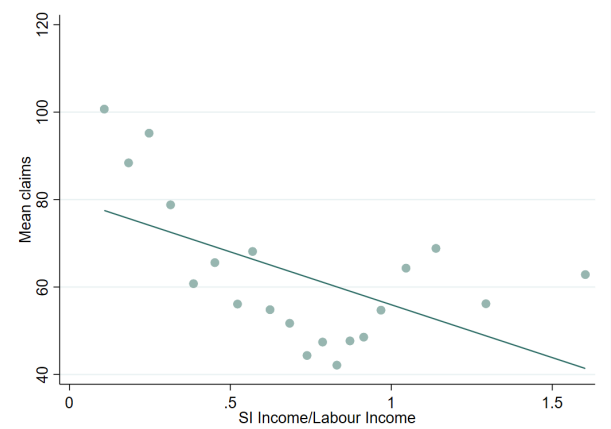
Notes: These figures show the average number of benefit days claimed (including zeros) by the treated individuals in 2011, 2012, 2013, and 2014, by bins of SI/LI in 2011-2014. Each SI/LI bin contains 10% of 2011-2014 person-year observations. The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry. See Figures B.10-B.12 for equivalent figures for each claim type separately.

Figure 7: Parental Leave Claim Heterogeneity

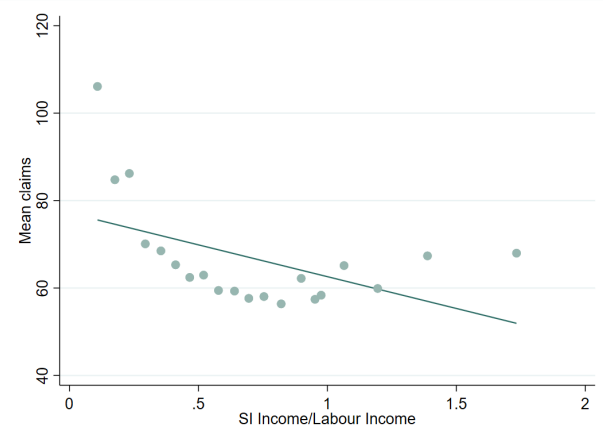
(a) Treated Individuals: First Child



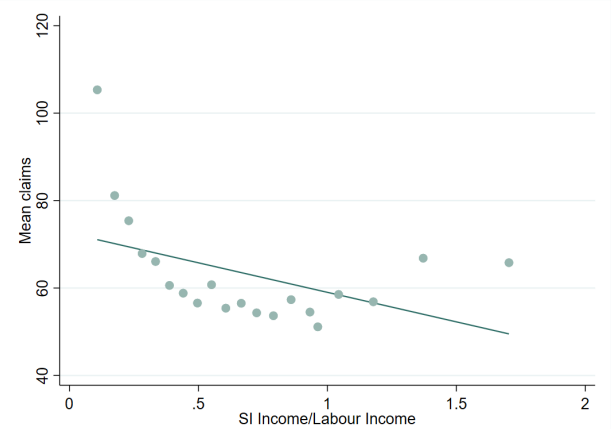
(b) Treated Individuals: Second+ Child



(c) Always Unrestricted: First Child



(d) Always Unrestricted: Second+ Child

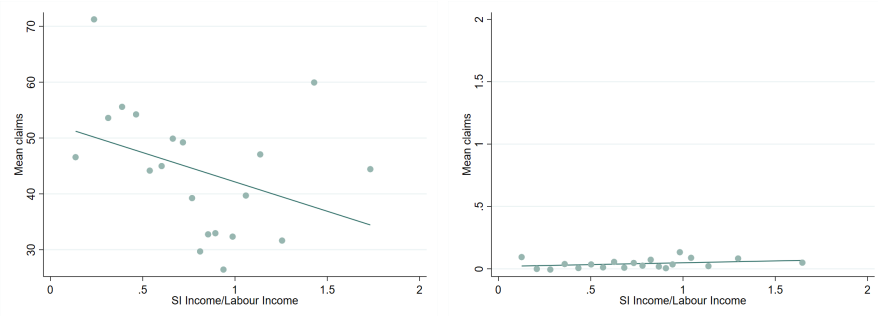


Notes: These figures show the average number of parental days claimed, conditional on claiming, by the treated individuals (those with ownership share between 30% and 50%) and always unrestricted individuals (those with ownership share between 50% and 70%), by bins of SI/LI in 2011-2014. Each SI/LI bin contains 5% of 2011-2014 person-year observations. The y-axis shows the average number of parental leave days claimed, conditional on claiming. Figures (a) and (c) show the number of days claimed for the first child, while figures (b) and (d) – for the second, third, etc child. The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

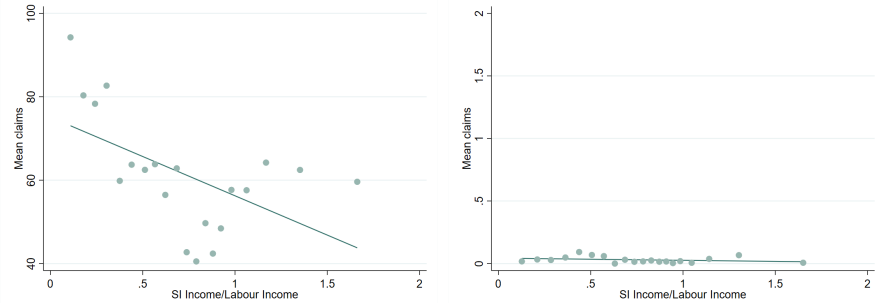
Figure 8: Adverse Selection – Number of Days Claimed

Panel A: Important Benefit **Panel B: Unimportant Benefit**

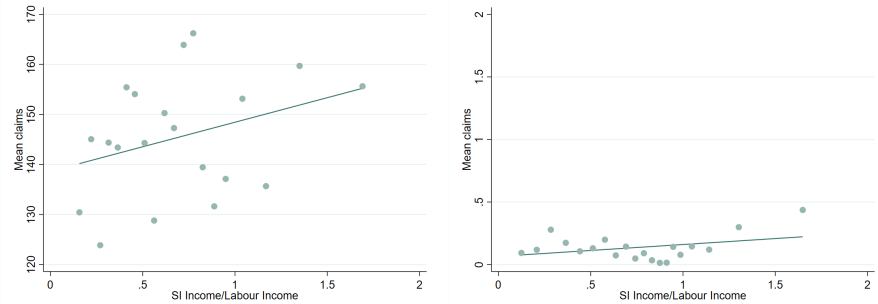
(a) Sick Pay (b) Sick Pay



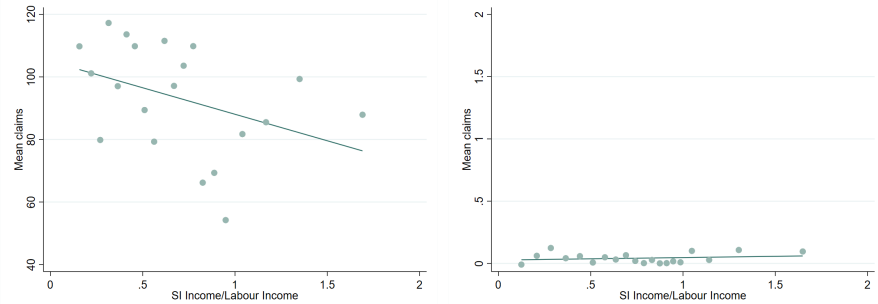
(c) Parental Pay (d) Parental Pay



(e) Unemployment Pay (Basic) (f) Unemployment Pay (Basic)



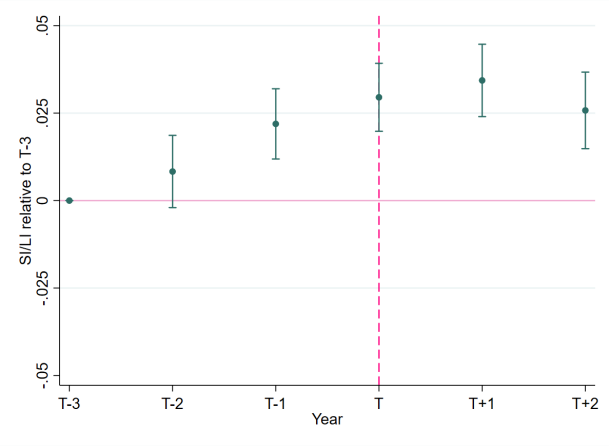
(g) Unemployment Pay (Union) (h) Unemployment Pay (Union)



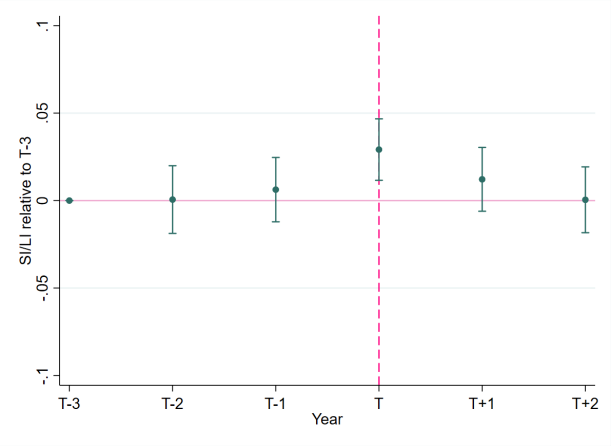
Notes: These figures show the average number of benefit days claimed by the treated individuals in 2011-2014, for whom that benefit is most important (Panel A) and for those, to whom the benefit is not (Panel B) by bins of SI/LI in 2011-2014. Each SI/LI bin contains 5% of 2011-2014 person-year observations. The y-axis shows the average number of days claimed (including zeros for panel B): in Figures (a) and (b) the outcome is the total number of sick pay days, in (c) and (d) – the number of parental pay days, in (e) and (f) – the number of basic unemployment pay days, in (g) and (h) – the number of earnings-based unemployment pay days. The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

Figure 9: Anticipatory Responses to Sick Leaves, Parental Leaves and UI Claims

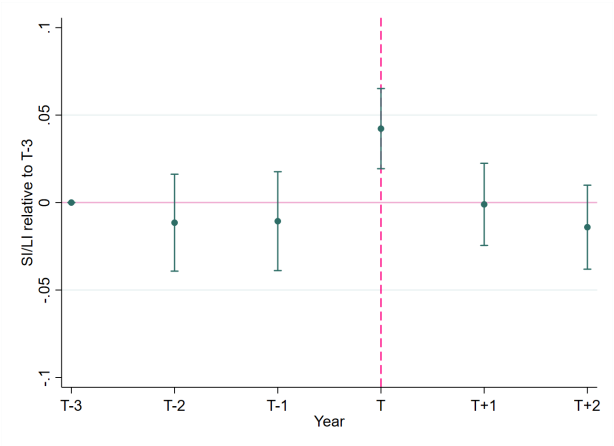
(a) Sick Pay Leave



(b) Parental Leave



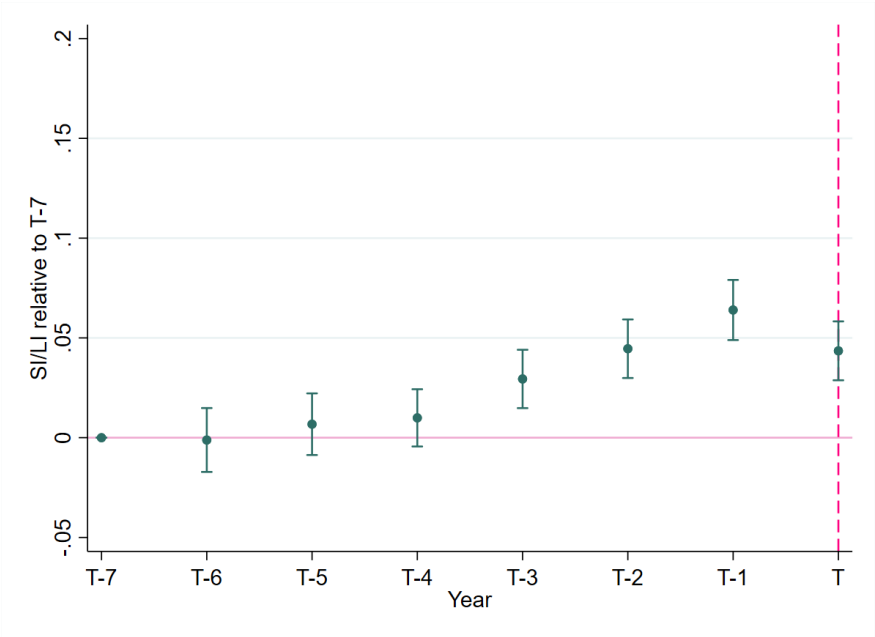
(c) UI Claim



Notes: Figures (a), (b) and (c) show the estimation results estimating equation (5). Time T on the horizontal axis represents the year when an individual gets ill, has a baby or gets unemployed, respectively. Figure (a) shows changes in insurance level through time for the unrestricted group of entrepreneurs with respect to the baseline T-3, three years prior to getting sick, controlling for age, gender, municipality and industry with year fixed effects. Similarly, Figure (b) shows changes in insurance level through time for the unrestricted group of entrepreneurs with respect to the baseline T-3, three years prior to having a baby. Figure (c) shows changes in insurance level through time for the unrestricted group of entrepreneurs with respect to the baseline T-3, three years prior to getting unemployed. The 95% CI are derived from standard errors that are clustered on individual level.

Figure 10: Anticipation of Retirement Decision

(a) Changes in SI/LI



Notes: Figure (a) shows the estimation results estimating equation (5). Time T on the horizontal axis represents the year an individual retires. Figure (a) shows changes in insurance level through time for the unrestricted with respect to the baseline T-7, seven years prior to retirement but controlling for age, gender, municipality and industry with year fixed effects. The 95% CI are derived from standard errors that are clustered on individual level.

Table 1: Summary Statistics (year 2010)

Overall Observations	194,975 shareholders								
	Unrestriced (YEL) (52.6%)			Restricted (TyEL) (38.7%)			Treatment (TyEL → YEL) (8.7%)		
	Ext.	Mean	sd	Ext.	Mean	sd	Ext.	Mean	sd
Female	24.1%			35.6%			19.7%		
Age		47.24	10.37		43.47	12.15		42.51	9.57
Labor Income		38059.15	46410.52		43713.69	47357.55		41794.72	31396.24
SI Income		25131.47	22530.75		42175.20	45689.11		40294.86	29505.24
Insurance Contributions		5023.14	5034.20		8152.58	9138.06		8058.07	5833.22
Sick Days (proxy)	3.9%	63.97	69.68	2.2%	57.35	68.81	1.3%	46.73	54.11
Parental leave (proxy)	2.0%	63.40	74.53	3.5%	73.19	78.38	2.5%	60.09	79.74
Unemployment (proxy)	2.1%	153.97	113.63	7.2%	162.67	112.68	2.4%	143.50	110.55
Pension	13.1%	18994.99	19205.32	11.7%	21593.98	24012.78	2.7%	12643.35	11512.20

Notes: This table presents the summary statistics for our main sample in the year 2010.

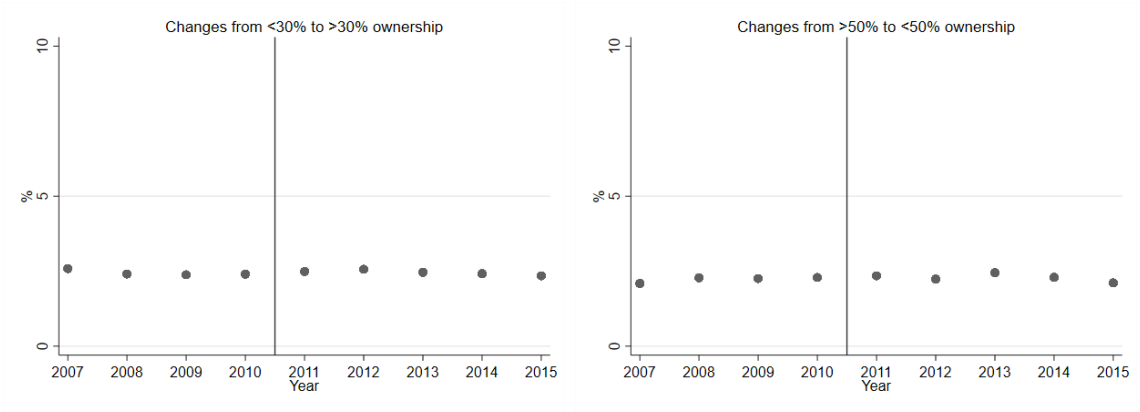
Table 2: Summary of Reduced Form Estimates

	Average	$SI/LI < 0.8$	$SI/LI \in [0.8, 1.1]$	$SI/LI > 1.1$
Moral Hazard Responses – Extensive Margin (DiD estimates):				
Average SI change	$-35.9\%^{***}$	$-70.5\%^{***}$	$-6.5\%^{***}$	$8.7\%^{***}$
Overall Benefits	0.3%	$-0.56\%^{**}$	$0.5\%^{**}$	$1.9\%^{***}$
Sick Days	$1.1\%^{***}$	$0.57\%^{***}$	$0.87\%^{***}$	$2.0\%^{***}$
Parental Leave	0.2%	0.02%	$0.58\%^{**}$	$0.75\%^{**}$
Unemployment	0.03%	-0.10%	-0.11%	$0.45\%^{*}$
Pension	0.10%	-0.29%	$-1.7\%^{***}$	$-1.4\%^{***}$
Moral Hazard Responses – Intensive Margin (DiD estimates):				
Sick Days				
absolute	0.17	0.27^{**}	0.16	0.94^{***}
arcsinh	0.78%	0.58%	$2.03\%^{**}$	$6.0\%^{***}$
log	$31.8\%^{***}$	$22.3\%^{***}$	19.1%	$38.2\%^{**}$
Parental Leave				
absolute	-0.32^{**}	-0.26^{*}	0.16	0.23
arcsinh	$-1.40\%^{**}$	$-1.8\%^{**}$	0.94%	1.8%
log	$25.4\%^{***}$	$34.8\%^{***}$	3.9%	$19.8\%^{**}$
Unemployment				
absolute	-0.72^{***}	-0.83^{***}	-0.81^{***}	-0.28
arcsinh	-0.35%	-0.98%	-1.0%	1.8%
log	$-31.9\%^{***}$	$-22.8\%^{***}$	$-17.0\%^{**}$	$-28.2\%^{***}$
Adverse Selection – Pre-Reform Claims:				
All Claims	-0.565 (0.327)			
Sick Days	0.013 (0.068)			
Parental Leave	0.044 (0.128)			
Unemployment	-0.639 (0.292)			

Notes: The table presents the reduced form estimates for moral hazard responses as well as the coefficients for the adverse selection tests. The control group used to form the estimates for the moral hazard responses in the always unrestricted. The heterogeneity results are also shown in the table. Here the control group is the always unrestricted group as well. The results are pooled difference in differences estimates for the pre and post years. The adverse selection coefficients represent the slope of the correlation line with robust standard errors in parenthesis.

A Additional Descriptive Evidence

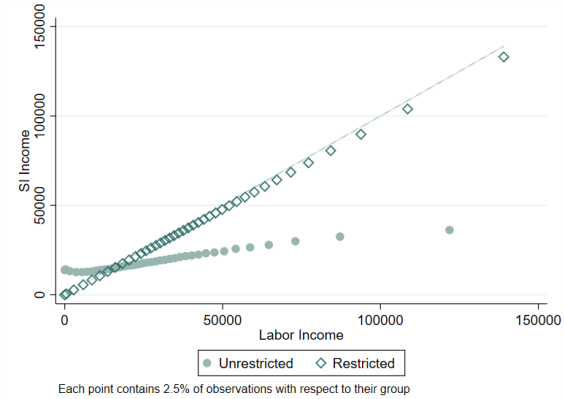
Figure A.1: Ownership changes to the treatment group



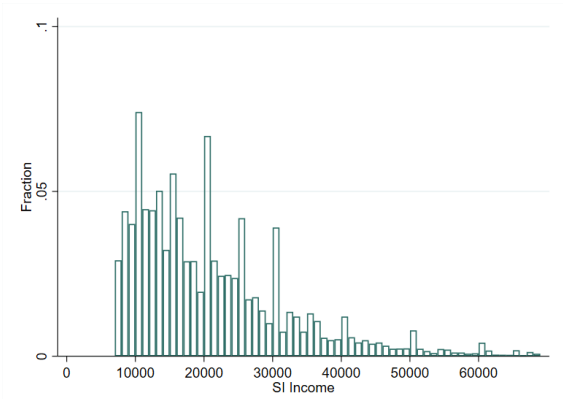
Notes: These figures show the percent of individuals that switched ownership status from below 30% to above 30%, or and from above 50% to below 50%. In both cases, only approximately 2.5% of business owners changed their ownership status enough to move from treatment to control group or vice versa.

Figure A.2: SI Incomes

(a) SI Income: Y vs T owners



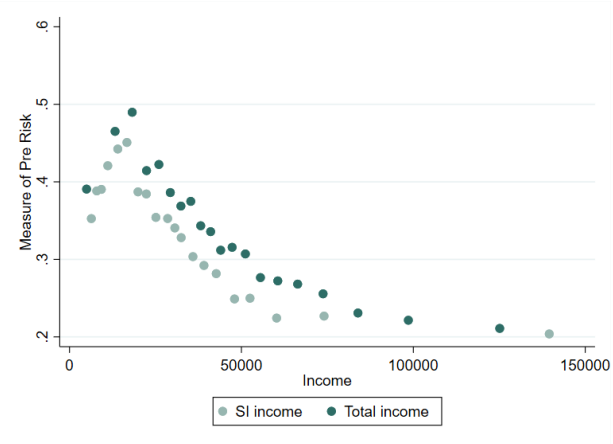
(b) Distribution of SI Income of Y owners



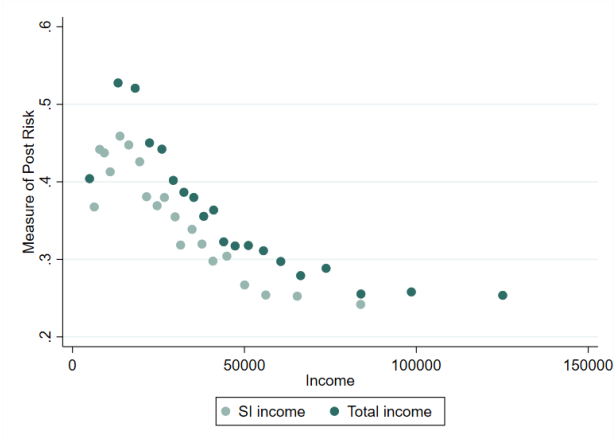
Notes: Figure (a) shows social insurance income with respect to labour income for the always restricted and always unrestricted groups. Each point contains 2.5% of observations with respect to their group. Figure (b) shows the distribution of SI income in 2010 for the always unrestricted individuals.

Figure A.3: Pre-Reform and Post-Reform Claims by Income

(a) Total Benefits Pre-Reform



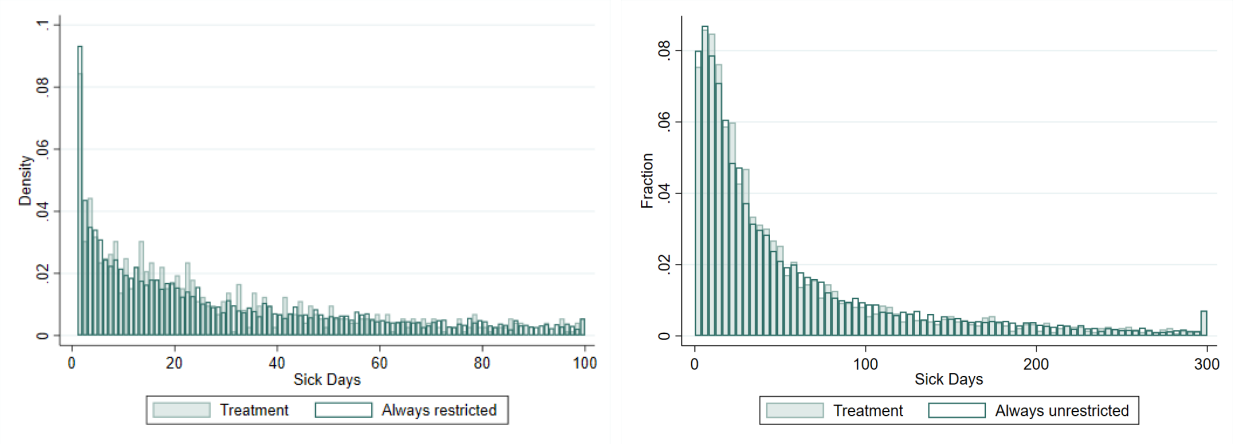
(b) Total Benefits Post-Reform



Notes: These figures show the relationship between pre- and post-reform benefit claims and individuals' total income.

Figure A.4: Sick Pay Claims Distributions

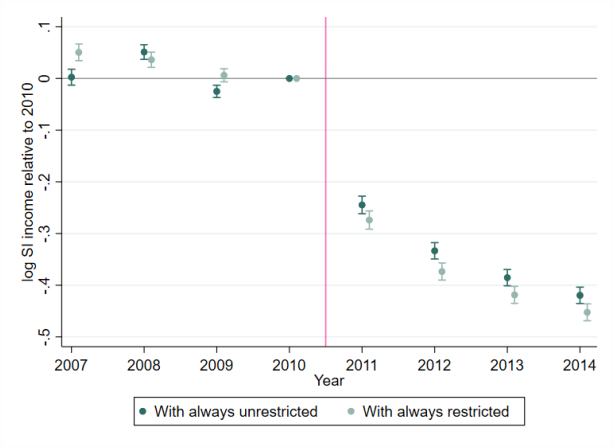
(a) Pre-reform: Treated vs Always Restricted (b) Post-reform: Treated vs Always Unrestricted



Notes: These figures show the distribution of sick day claims (a) before the reform (in 2007-2010) by the treated and always restricted individuals, and (b) after the reform (in 2011-2014) by the treated and always unrestricted individuals.

Figure A.5: Changes to Reported SI Income

(a) SI Income Change: Event Study Estimates



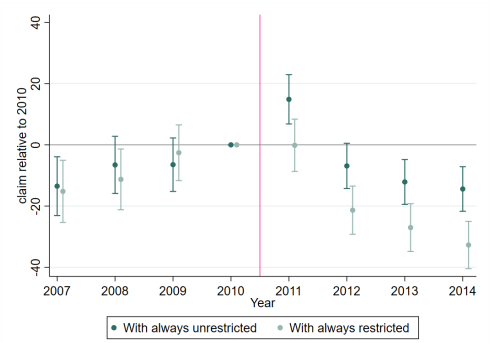
Notes: Figure shows coefficient estimates from an event study specification, which compares social insurance income of treated individuals to that of the always restricted and always unrestricted business owners. Standard errors that are clustered at the individual level, 95% confidence intervals shown.

B Additional Empirical Evidence

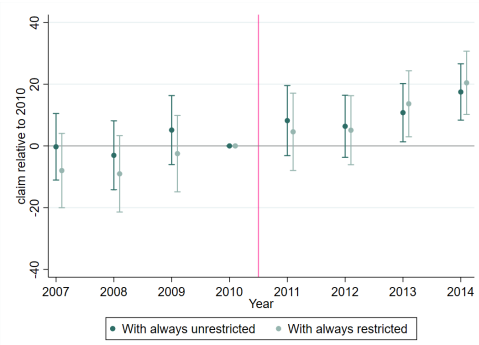
B.1 Moral Hazard – Intensive and Extensive Margins

Figure B.6: Moral Hazard Responses – Number of Days Claimed, Conditional on Claiming (Intensive Margin)

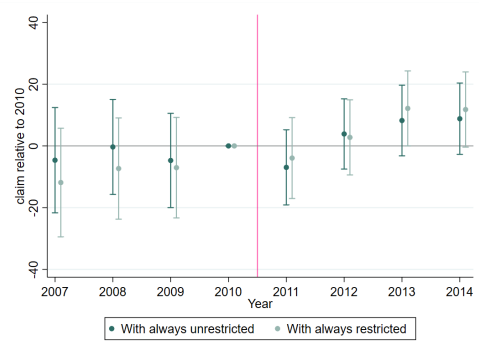
(a) Total Absent Days



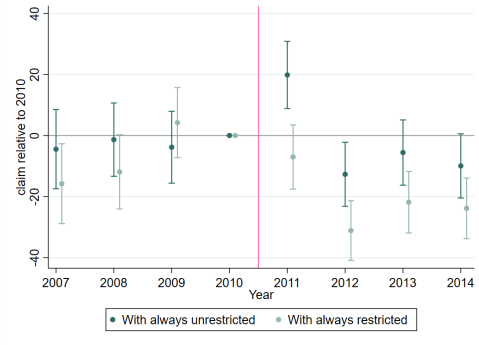
(b) Sick Pay Days



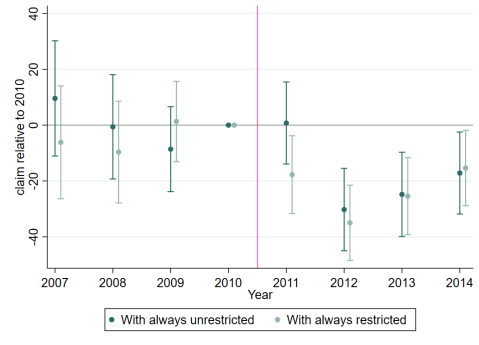
(c) Parental Pay Days



(d) Basic UI Pay Days



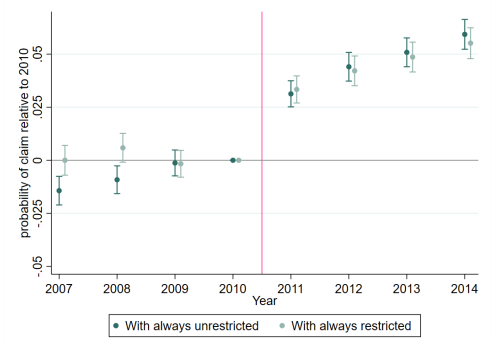
(e) Earnings-Based UI Pay Days



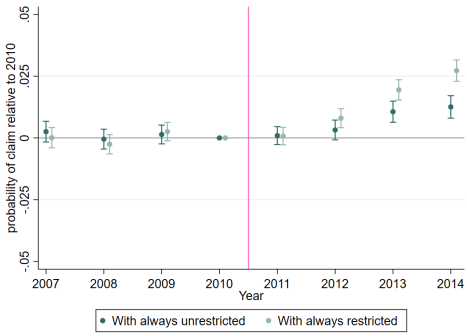
Notes: These figures show the results of estimating (4), where the outcome variable measures whether the individual, and the control group consists of all always restricted individuals (those with ownership share between 10% and 30%) or the all always unrestricted individuals (those with ownership share between 50% and 70%). In Figure (a), the outcome is the total number of absent days, including absences due to sickness, parental leave, and unemployment. In Figures (b), (c), (d) and (e), the outcome is the number of absent days due to sickness, parental leave, basic unemployment, and earnings-based unemployment, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown.

Figure B.7: Moral Hazard Responses – Probability of Claiming (Extensive Margin)

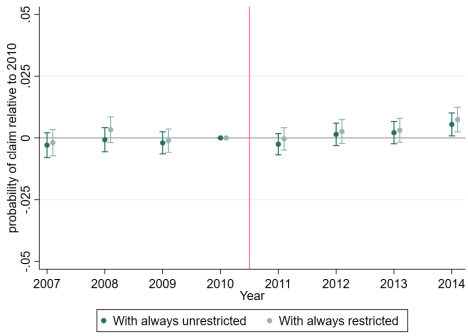
(a) Total Absent Days



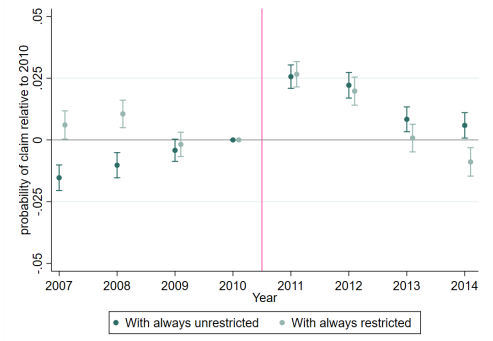
(b) Sick Pay Days



(c) Parental Pay Days



(d) UI Pay Days



Notes:These figures show the results of estimating (4), where the outcome variable measures the number of days per year an individual claims absence, conditional on claiming, and the control group consists of all always restricted individuals (those with ownership share between 10% and 30%) or the always unrestricted individuals (those with ownership share between 50% and 70%). In Figure (a), the outcome is a dummy on absence, including absences due to sickness, parental leave, and unemployment. In Figures (b), (c), (d) and (e), the outcome takes value 1 if an individual is absent due to sickness, parental leave, basic unemployment, and earnings-based unemployment and 0 otherwise, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown.

B.2 Adverse Selection – Intensive and Extensive Margin

Figure B.8: Adverse Selection – Number of Days Claimed, Conditional on Claiming

Panel A: Pre-reform Claims **Panel B: Post-reform Claims**

(a) All claims (b) All claims

(c) Sick Pay (d) Sick Pay

(e) Parental Pay (f) Parental Pay

Figure B.8: Adverse Selection – Number of Days Claimed, Conditional on Claiming

Panel A: Pre-reform Claims **Panel B: Post-reform Claims**

(g) Unemployment Pay (Basic) (h) Unemployment Pay

(i) Unemployment Pay (Union) (j) Unemployment Pay

Notes: These figures show the average number of benefit days claimed by the treated individuals in 2006-2010 (Panel A) and in 2011-2014 (Panel B), conditional on claiming, by bins of SI/LI in 2011-2014. Each SI/LI bin contains 5% of 2011-2014 person-year observations. The y-axis shows the average number of days claimed, conditional on claiming: in Figures (a) and (b) the outcome is the total number of absent days, in (c) and (d) – the number of sick pay days, in (e) and (f) – the number of parental pay days, in (g) and (h) – the number of basic unemployment pay days, and in (i) and (j) – the number of earnings-based unemployment days. The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

B.3 Learning Over Time

Figure B.9: Adverse Selection – Number of Days Claimed by the Always Unrestricted

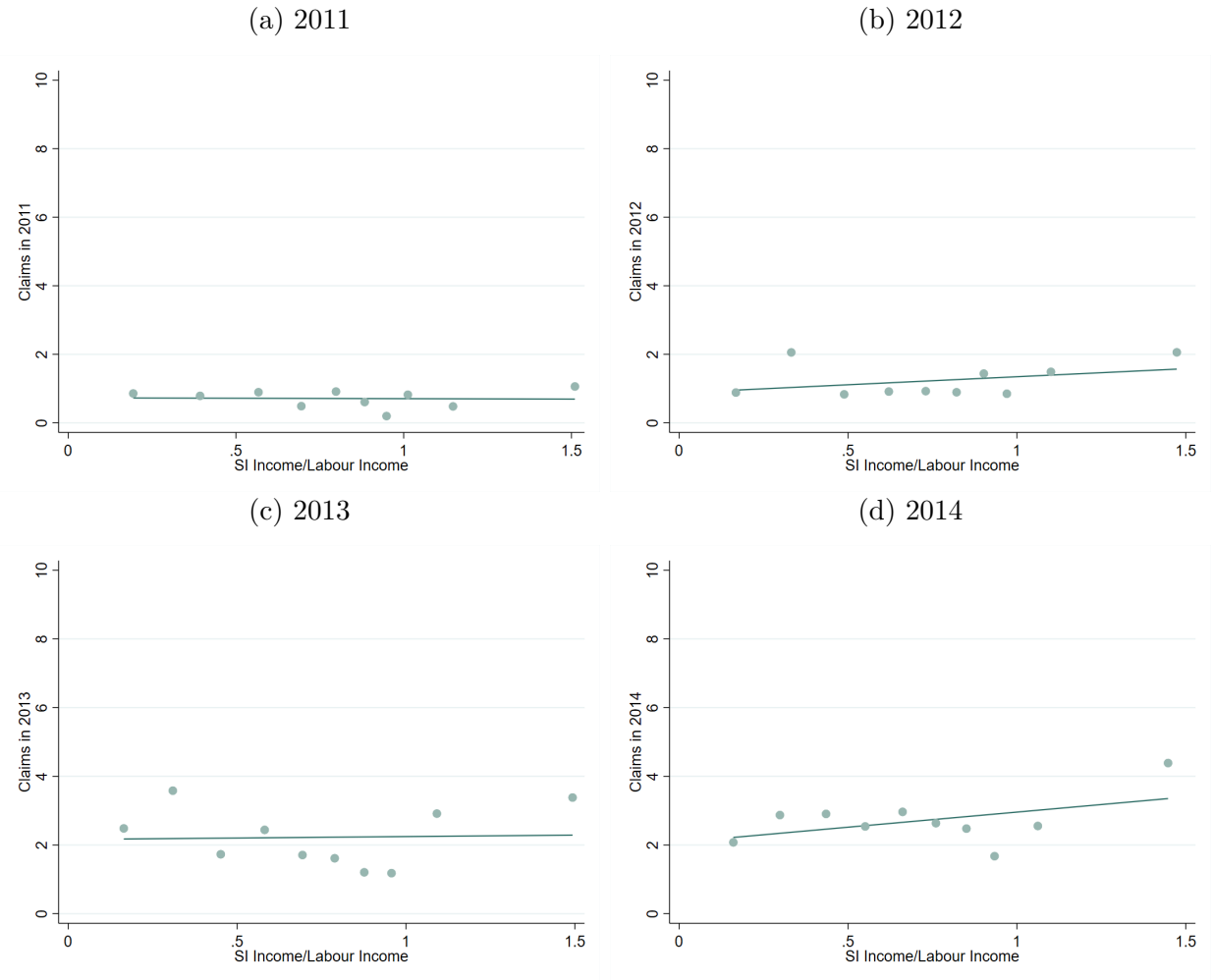
(a) All claims

(b) Sick Pay (c) Parental Pay

(d) Unemployment Pay (e) Unemployment Pay

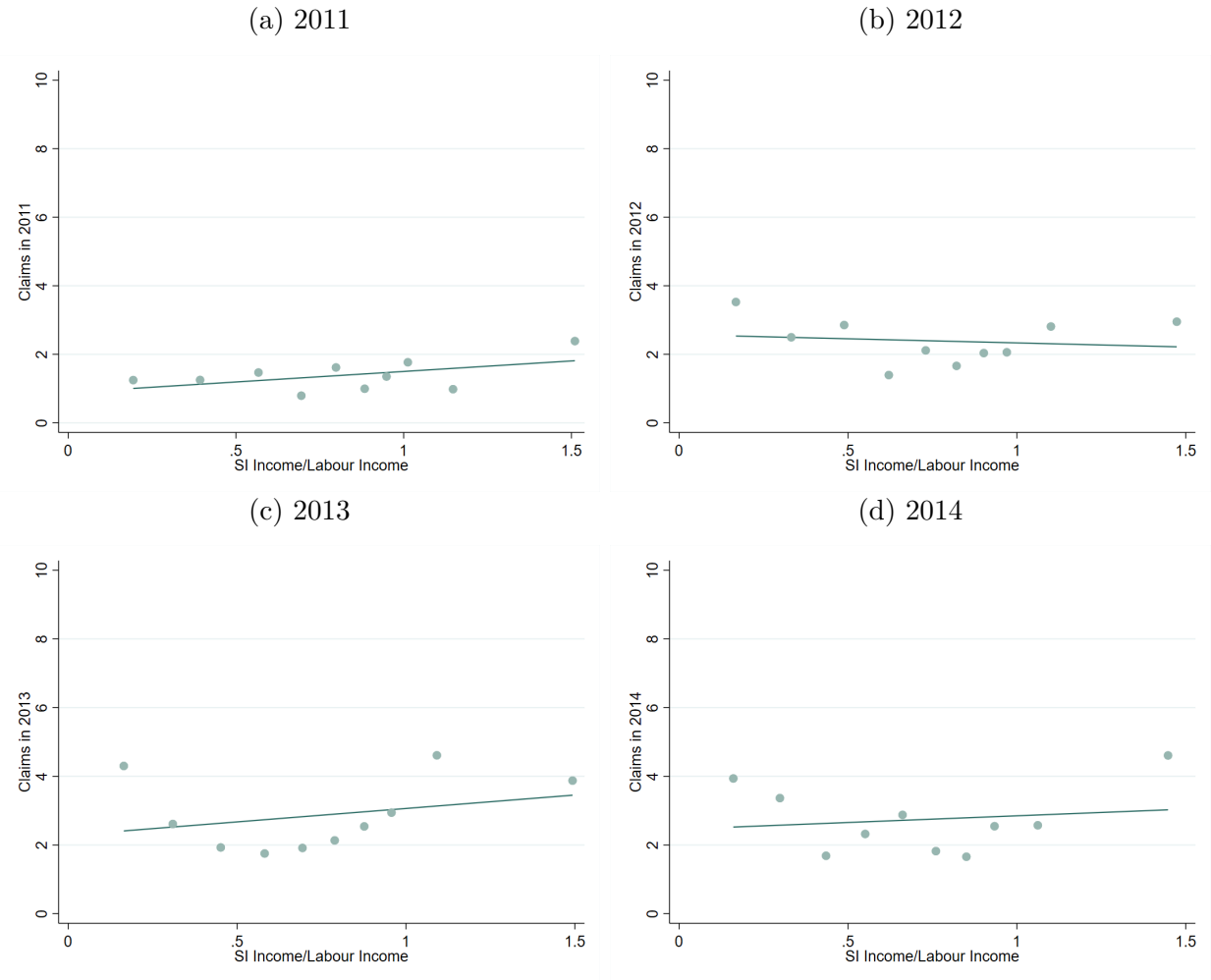
Notes: These figures show the average number of benefit days claimed by the always unrestricted individuals in 2011-2014 (Panel B) by bins of SI/LI in 2011-2014. Each SI/LI bin contains 5% of 2011-2014 person-year observations. The y-axis shows the average number of days claimed (including zeros): in Figures (a) and (b) the outcome is the total number of absent days, in (c) and (d) – the number of sick pay days, in (e) and (f) – the number of parental pay days, in (g) and (h) – the number of basic unemployment pay days, and in (i) and (j) – the number of earnings-based unemployment days. The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

Figure B.10: Learning Over Time – Sick Pay Claims



Notes: These figures show the average number of benefit days claimed by the treated individuals in 2011, 2012, 2013, and 2014, by bins of SI/LI in 2011-2014. Each SI/LI bin contains 10% of 2011-2014 person-year observations. The y-axis shows the average total number of days claimed (including zeros). The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

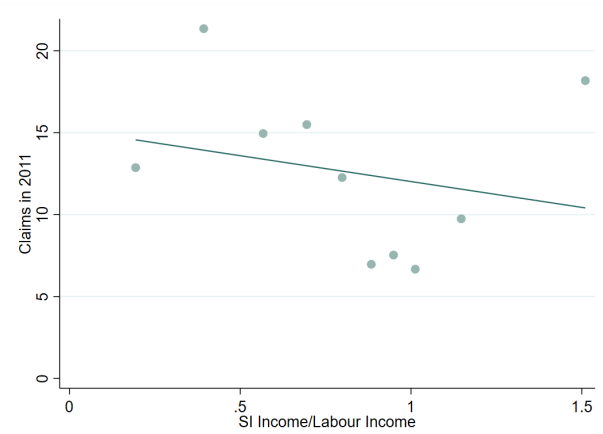
Figure B.11: Learning Over Time – Parental Leave Claims



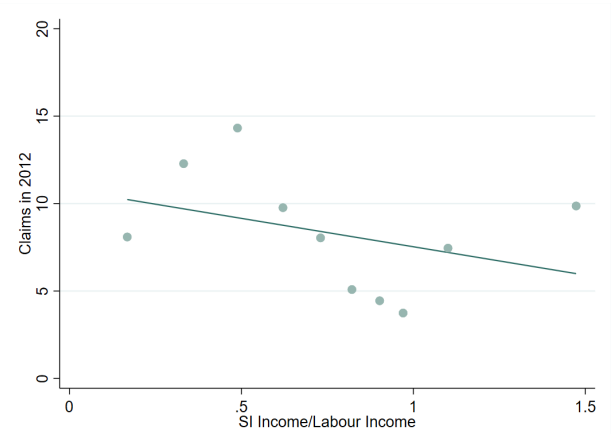
Notes: These figures show the average number of benefit days claimed by the treated individuals in 2011, 2012, 2013, and 2014, by bins of SI/LI in 2011-2014. Each SI/LI bin contains 10% of 2011-2014 person-year observations. The y-axis shows the average total number of days claimed (including zeros). The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

Figure B.12: Learning Over Time – Unemployment Claims

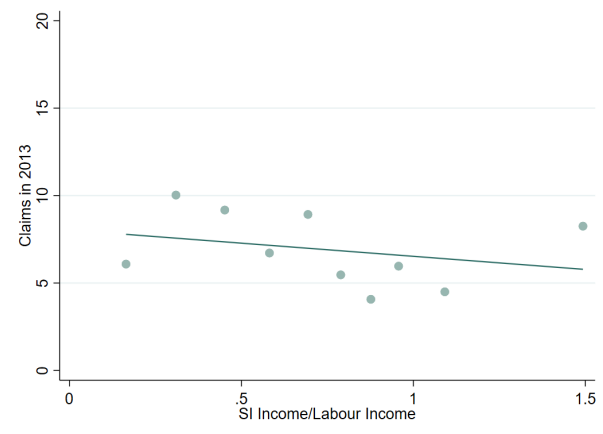
(a) 2011, Basic



(b) 2012, Basic



(c) 2013, Basic



(d) 2014, Basic

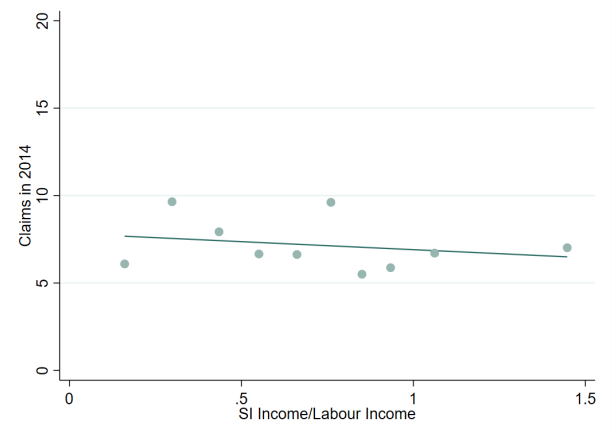
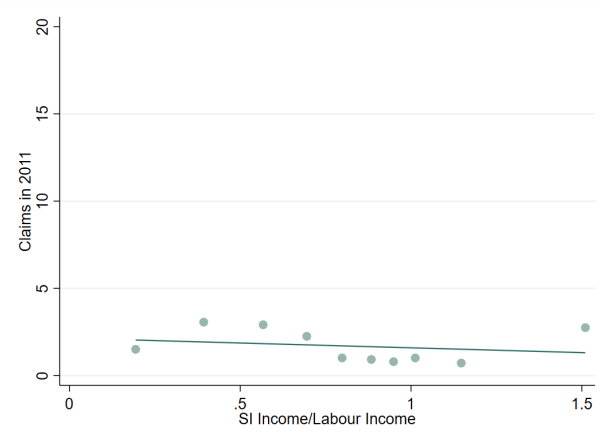
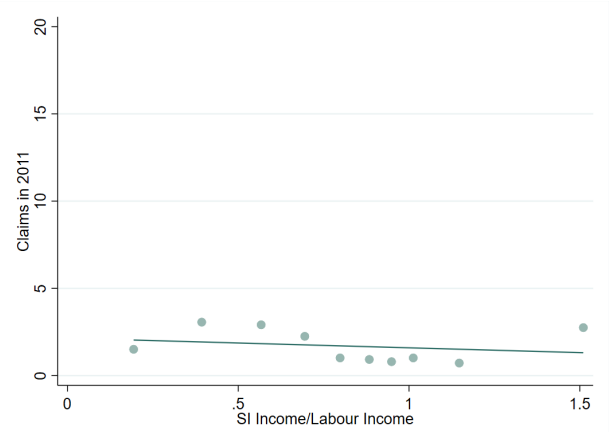


Figure B.12: Learning Over Time – Unemployment Claims

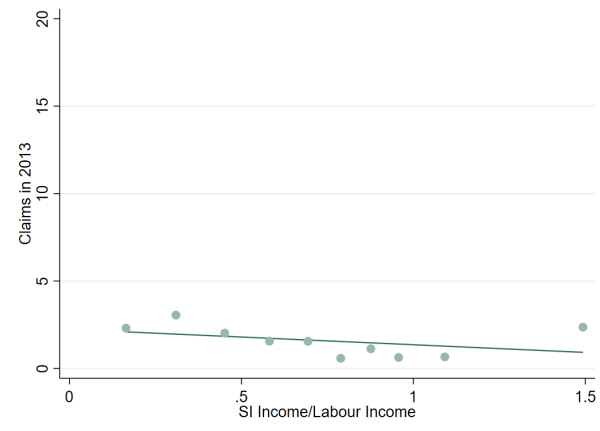
(e) 2011, Union



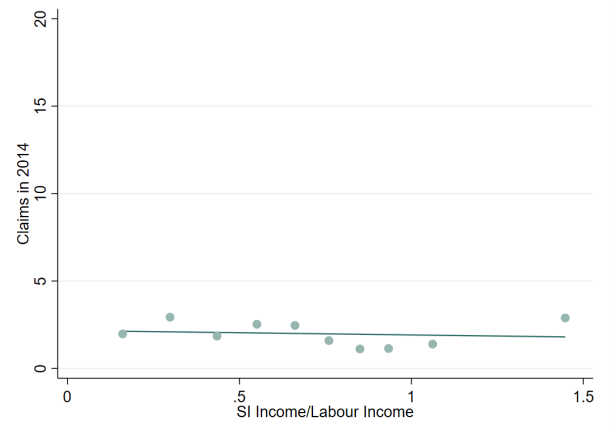
(f) 2012, Union



(g) 2013, Union



(h) 2014, Union



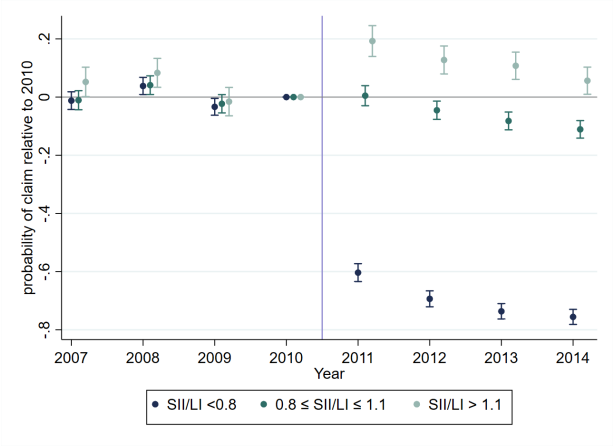
Notes: These figures show the average number of benefit days claimed by the treated individuals in 2011, 2012, 2013, and 2014, by bins of SI/LI in 2011-2014. Each SI/LI bin contains 10% of 2011-2014 person-year observations. The y-axis shows the average total number of days claimed (including zeros). The fitted line shown is estimated on the raw data (i.e. person-year observations), and includes the following controls: age, gender, municipality and industry.

C Robustness Checks

C.1 Using Always Unrestricted As a Control Group

Figure C.13: Heterogeneity for Change of SI Income

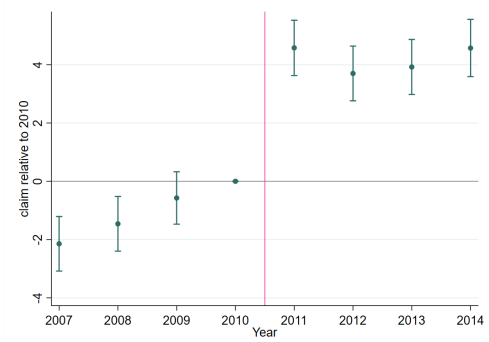
(a) SI Income Change: Event Study Estimates



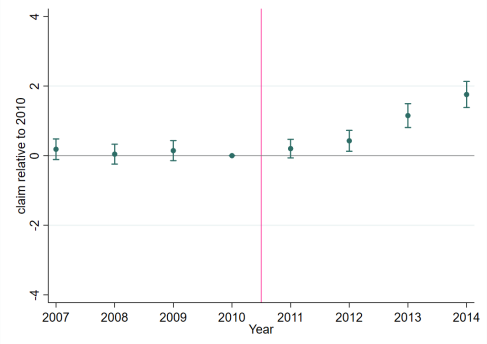
Notes: This figure shows coefficient estimates from an event study specification (4), which compares the ratio of SI income to labour income of treated individuals to that of the always unrestricted business owners (those with ownership share between 50% and 70%), relative to 2010. Treated and control individuals are separated into three groups based on the average SI/LI in post-reform years (2011-2014): those with SI/LI below 0.8, between 0.8 and 1.1, and above 1.1. Equivalent results but using the always restricted individuals as a control group are shown in Figure 2(d).

Figure C.14: Moral Hazard Responses – Number of Days Claimed

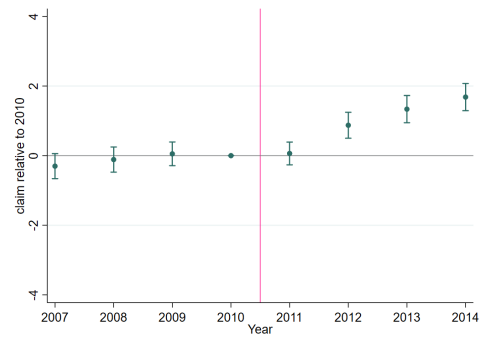
(a) Total Absent Days



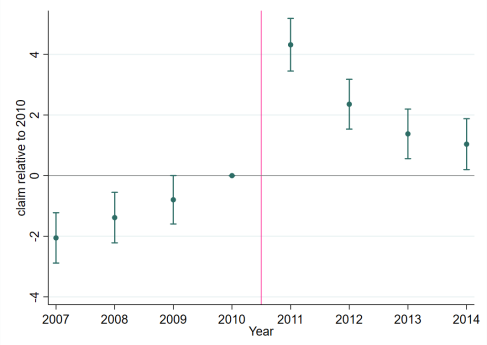
(b) Sick Pay Days



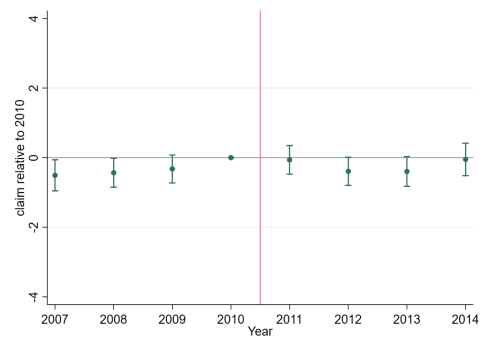
(c) Parental Pay Days



(d) Basic UI Pay Days



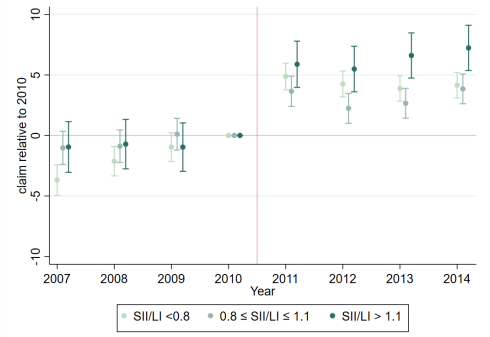
(e) Earnings-Based UI Pay Days



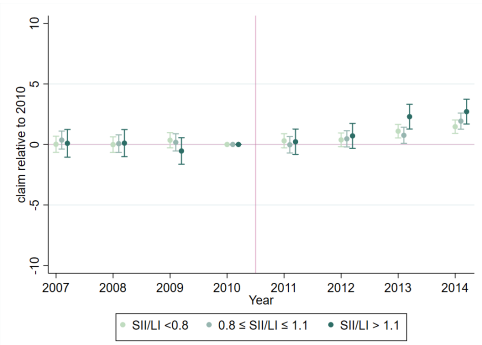
Notes: These figures show the results of estimating (4), where the outcome variable measures the number of days per year an individual claims absence (including zeros), and the control group consists of all always unrestricted individuals (i.e. those with ownership share between 50% and 70%). In Figure (a), the outcome is the total number of absent days, including absences due to sickness, parental leave, and unemployment. In Figures (b), (c), (d) and (e), the outcome is the number of absent days due to sickness, parental leave, basic unemployment, and earnings-based unemployment, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown. Equivalent results but using the always restricted individuals as a control group are shown in Figure 3.

Figure C.15: Moral Hazard Response Heterogeneity – Number of Days Claimed

(a) Total Absent Days

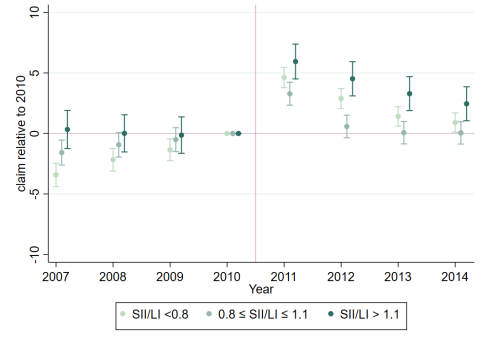


(b) Sick Pay Days



(c) Parental Pay Days

(d) Basic UI Pay Days



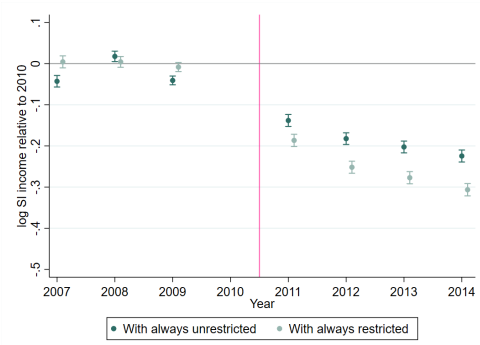
(e) Earnings-Based UI Pay Days

Notes: These figures show the results of estimating (4), where the outcome variable measures the number of days per year an individual claims absence. In each figure, the treated individuals are separated into three groups based on the average SI/LI in post-reform years (2011-2014): those with SI/LI below 0.8, between 0.8 and 1.1, and above 1.1. The control group consists of always unrestricted individuals (i.e. whose ownership share is between 50% and 70%) with corresponding levels of average SI/LI in post-reform years (2011-2014), i.e. below 0.8, between 0.8 and 1.1, or above 1.1. In Figure (a), the outcome is the total number of absent days, including absences due to sickness, parental leave, and unemployment. In Figures (b), (c), (d) and (e), the outcome is the number of absent days due to sickness, parental leave, basic unemployment, and earnings-based unemployment, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown. Equivalent results but using the always restricted individuals as a control group are shown in Figure 4.

C.2 Including Individual Fixed Effects

Figure C.16: Moral Hazard Responses – Number of Days Claimed, Including Individual Fixed Effects

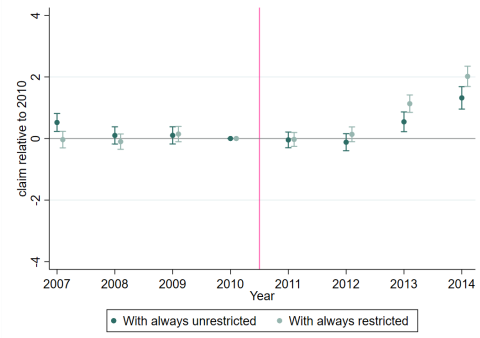
(a) Income reported

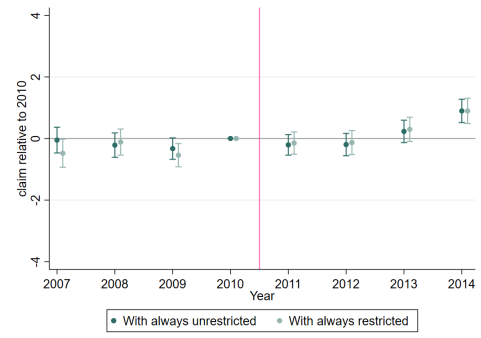


(b) Total Absent Days

(c) Sick Pay Days

(d) Parental Pay Days





(e) UI Pay Days

(f) Earnings-Based UI Pay Days

Notes: These figures show the results of estimating (4) with individual fixed effects, where the outcome variable measures the logarithm of SI income reported or the number of days per year an individual claims absence, and the control group consists of all always restricted individuals (i.e. those with ownership share between 10% and 30%), or all always unrestricted individuals (i.e. those with ownership share between 50% and 70%). In Figure (a), the outcome is the logarithm of SI income. In Figure (b) the outcome is total number of absent days, including absences due to sickness, parental leave, and unemployment. In Figures (c), (d), (e) and (f), the outcome is the number of absent days due to sickness, parental leave, basic unemployment, and earnings-based unemployment, respectively. Standard errors are clustered at the individual level, 95% confidence intervals shown.