Private Information and Design of Unemployment Insurance

Maksym Khomenko*

Abstract

Unemployment insurance (UI) programs around the world are predominantly government-provided with universal coverage. One explanation for the dominant adoption of mandatory UI is that private knowledge about unemployment risks might lead to a selected pool of insured individuals and generate welfare losses. At the same time, mandates might have a detrimental effect on welfare because of fully restricted individual choices. This ambiguity motivates a need to consider alternative designs of UI that allow for the individual choice but restrict selection into insurance based on risks. I use institutional features of the Swedish voluntary UI system and detailed administrative data to study the optimal design of UI. To evaluate welfare under various alternative regulations, I estimate a structural model of insurance choice that captures heterogeneity in preferences and private information about future unemployment risks. The results suggest that mandating UI would unambiguously reduce welfare by on average 49% in terms of consumer surplus compared to a current system. In contrast, appropriate designs with voluntary enrollment generate large welfare gains. In particular, contracts with fixed enrollment timing and predetermined duration improve welfare by 58% - 95% in terms of the consumer surplus. A "two-part tariff" contract that fails to sufficiently restrict risk-based selection results in average consumer surplus loss of 3%.

Keywords: unemployment insurance, private information, contract design, mandate

JEL classification: J65, D82, D81, G22, H55

*Department of Economics, University of Gothenburg, *maksym.khomenko@economics.gu.se.*
This work has benefited from the comments of Natalie Bachas, Timothy Bresnahan, Raj Chetty, Mark Duggan, Andreas Dzemski, Liran Einav, Randi Hjalmarsson, Mikael Lindahl, Arash Nekoei, Petra Persson, Emmanuel Saez, Aico van Vuuren and seminar participants at the University of Gothenburg, Stanford University and Winter Meeting of Econometric Society.

1 Introduction

Unemployment insurance (UI) is a part of a broader spectrum of social insurance programs in many countries. A typical UI program is state-provided and tax-financed with compulsory enrollment. At the same time, a few developed countries including Sweden have introduced a voluntary UI system.[1] On the one hand, the presence of adverse selection might lead to welfare losses in such a system. On the other hand, moral hazard and heterogeneity of preferences might rationalize the adoption of voluntary UI. This ambiguity and the absence of conclusive empirical evidence motivate a need to consider alternative regulations which preserve an individual choice but restrict selection into insurance based on risks. Therefore, this paper attempts to comprehensively study the optimal design of UI.

The essence of adverse selection in the context of UI is that individuals tend to have private information about their unemployment risks (e.g. working in a risky occupation, an industry or a firm). Consequently, it might lead to an insurance pool of relatively high-risk individuals and even result in a classic example of the "market for lemons" unraveling (Akerlof, 1978). Alternatively, above-optimal prices might generate welfare losses and require large subsidies to sustain a program (Einav, Finkelstein, & Cullen, 2010).

At the same time, the presence of heterogeneity of preferences for insurance may serve as a rationale for a voluntary system. In this case, a mandate might impose the excess burden on low risk-aversion individuals who do not value insurance even in the presence of substantial risks. It also implies that a positive correlation between the likelihood of purchasing insurance and unemployment risks might not be sufficient to motivate the introduction of a mandate since it might be driven by a correlation between risks and risk preferences.[2]

Given these concerns regarding both voluntary and mandatory systems, it might be worth considering designs of UI contracts that address selection and at the same time allow for voluntary enrollment. For example, when adverse selection is primarily driven by unrestricted enrollment timing, alternative contracts that restrict time-selection might be welfare-improving.[3] In the context of UI, it means that individuals tend to buy insurance when they have higher unemployment risks, which vary over time. The presence of such selection was documented in, for example, dental (Cabral, 2016) and health insurance markets (Aron-Dine, Einav, Finkelstein, & Cullen,

[1]Similar voluntary UI systems exist in Finland, Norway, and Iceland.

[2]Moral hazard in UI means that the availability of insurance entails, for instance, a reduction in job search or on-the-job efforts, which raises probabilities or durations of unemployment. As a result, it might amplify the costs under a mandatory system and make such a policy suboptimal. However, moral hazard is not a focus of this paper but its implications are discussed in robustness selection.

[3]There is a membership eligibility condition that acts as a timing restriction but does not completely remove the possibility of time-selection.

2015; Einav, Finkelstein, & Schrimpf, 2015). Therefore, I study potential consequences of two contracts that restrict the selection of enrollment timing. First, I consider an "open enrollment" contract with fixed enrollment timing and predetermined duration. Another alternative is an "entry costs" or "two-part tariff" contract which in addition to monthly premiums charges entry fees upon enrollment of previously uninsured (Cabral, 2016). In contrast to the open enrollment contract, this design affects time-selection by discouraging unenrollment when unemployment risks are low to enroll later when risks are high.[4]

The context of Swedish voluntary unemployment insurance provides an appropriate set-up to understand the interaction between risks, private information, and individual preferences that should guide the choice of policy measures. This paper uses detailed individual-level administrative data, which allow observing dates of unemployment and insurance spells together with a variety of demographic and labor market characteristics. I start by augmenting the existing evidence of a positive correlation between insurance and unemployment probabilities by showing the presence of time-selection patterns. Using the eligibility condition for the income-based coverage that requires paying insurance premiums for at least twelve consecutive months, I demonstrate that individuals are more likely to start unemployment spells with exactly twelve months of UI enrollment. This evidence is robust and shows the presence of private information about unemployment timing.

To study welfare consequences of designs of UI, I estimate a dynamic insurance choice model that exploits the variation in insurance premiums and benefits generosity as well as time-selection patterns. It enables recovering distributions of risk preferences and private information about future unemployment risks, which jointly determine insurance decisions. To identify risk preferences, I leverage two sources of variation. The first one is a result of differences in premiums and the generosity of benefits over time primarily due to a UI reform in 2007. Another source of variation stems from cross-sectional differences in premiums across industry-specific UI funds and replacement rates due to a benefits cap. The identification of private information types exploits patterns of timing of insurance purchase relative to the timing of future unemployment or changes in unemployment risks. To separately identify risk preferences and information about unemployment, I assume that changes in the attractiveness of UI do not affect the structure of private information about unemployment conditionally on the observed determinants of this information. The assumption is in line with the evidence from the data.[5] The results show

[4]In other words, if an individual interrupts the sequence by leaving the insurance pool even for one month, new entry requires paying entry fees again. As a result, this design discourages exits to re-enter the insurance pool later when needed.

[5]For example, I assume that although the UI reform in 2007 changed the generosity of benefits and premiums, it did not affect the labor market itself such that individuals did not become more or less informed about their

considerable variation in risk preferences and quality of information about future employment perspectives. I also estimate inertia parameters that suggest considerable choice persistence. It means that the insurance status in a previous period impacts future decisions. To identify the inertia parameters I assume that individuals who are aware of the forthcoming unemployment make inertia-free decisions.[6]

The efficiency of insurance programs is determined by an interplay between individual risk preferences, risks and private information about those risks. This complexity rationalizes a use of such a model that combines those parts to provide policy recommendations. Some of the existing works provide policy conclusions about UI based on a "reduced form" association between realized risks and insurance probabilities using observable characteristics, survey responses or arguably exogenous institutional variation (e.g. Hendren, 2017; Landais, Nekoei, Nilsson, Seim, & Spinnewijn, 2017). Instead, the approach in this paper allows not only studying a broader spectrum of alternative regulations but also exploring richer variation and behavioral patterns to understand the consequences of various policies at the expense of imposing a number of theory-based assumptions.

To evaluate welfare under current and alternative structures of UI, I use the model estimates to recover UI demand functions and distributions of willingness-to-pay (WTP) for corresponding insurance contracts. The findings suggest that mandates would generate considerable welfare losses amounting to 243 SEK/month ($27 or 49%) per individual compared to the current system.[7] The intuition is that a mandate restricts selection not only on risks but also on preferences, which generates a consumer surplus loss.[8]

In contrast, appropriate contract design regulations are predicted to generate large welfare gains. I find that an alternative two-part tariff contract that charges extra fixed costs upon the payment of the first premium would perform slightly worse than the status quo. The reason is that it does not sufficiently restrict selection on risks but imposes additional fixed costs burden on individuals. However, an open enrollment contract with 18 months duration is predicted to generate welfare improvement of 545 SEK/month per individual ($61 or 95%) on average. In comparison with the entry costs design, it virtually removes time-selection without imposing large additional costs on consumers. In contrast to mandates, it restricts undesirable selection

future employment perspectives. I show that time-selection patterns did not change as a result of the reform in 2007.

[6]I investigate the sensitivity of welfare analysis to this assumption. I find that welfare conclusions are robust to various formulations of inertia.

[7]This number applies to the range of subsidy levels considered in the welfare analysis.

[8]However, as I discuss in the section dedicated to the welfare analysis, a mandatory system in the absence of a moral hazard response allows achieving any reasonable budget balance. In contrast, the voluntary system is very limited in terms of which subsidy levels are feasible because of behavioral responses to price changes.

without severe choice restrictions. A similar design of the open-enrollment contract but with 24 months duration leads to smaller welfare gains of 337 SEK/month ($36 or 58%) per individual on average. Smaller welfare gains stem from higher risk-exposure due to a longer contract duration.

This paper contributes to a large literature on private information in insurance programs and markets. Most attention to the importance of private information in insurance has been dedicated in health insurance, annuity, and long-term care markets. In particular, a large literature documents the presence,[9] discusses sources[10], analyses consequences of asymmetric information[11] as well as studies policies aimed at addressing inefficiencies in insurance markets.[12] The literature related to unemployment insurance has been primarily focused on the optimal UI theory[13] and on estimating labor supply responses to insurance benefits.[14] However, to the best of my knowledge, only a few empirical papers focus on the canonical private information problem in UI such as Hendren (2017), who shows that the absence of private UI markets is a result of the excess mass of private information. In this paper, I do not focus on the existence of private information and an effect on private markets but primarily attempt to look at how contract design can be used to address the problem.

Another paper studying private information in UI using the Swedish setup is Landais et al. (2017). The authors document that insured individuals on average have higher unemployment risks. It is argued that adverse selection must be an important component of the observed positive correlation between unemployment risks and insurance take-up. The paper concludes that mandating the system would not be an optimal policy because individuals who are not covered under the current system value insurance less than expected costs of covering them.[15] Instead, the combination of subsidies and a minimum basic insurance mandate is suggested to be a welfare-improving policy. In this paper, I attempt to look deeper into insurance decision-making by imposing a structure of the model. It allows examining a broader set of counterfactual

[9]See e.g. Chiappori and Salanie (2000); Finkelstein and Poterba (2004).

[10]See e.g. Barsky, Juster, Kimball, and Shapiro (1997); Abbring, Chiappori, and Pinquet (2003); Abbring, Heckman, Chiappori, and Pinquet (2003); Finkelstein and McGarry (2006); Cutler, Finkelstein, and McGarry (2008); Fang, Keane, and Silverman (2008).

[11]See e.g. Spence (1978); Einav, Finkelstein, and Cullen (2010); Hendren (2013).

[12]See e.g. Einav, Finkelstein, and Schrimpf (2010); Handel, Hendel, and Whinston (2015); Handel, Kolstad, and Spinnewijn (2015).

[13]See e.g. Baily (1978); Hopenhayn and Nicolini (1997); Holmlund (1998); Card and Levine (2000); Fredriksson and Holmlund (2001); Autor and Duggan (2003); Chetty (2006, 2008); Kroft (2008); Shimer and Werning (2008); Spinnewijn (2015); Landais, Michaillat, and Saez (2018b, 2018a); Kolsrud, Landais, Nilsson, and Spinnewijn (2018).

[14]See e.g. Moffitt (1985); Meyer (1990); Lalive, Van Ours, and Zweimüller (2006); Schmieder, Von Wachter, and Bender (2012); Card, Johnston, Leung, Mas, and Pei (2015); Landais (2015); DellaVigna, Lindner, Reizer, and Schmieder (2017).

[15]The findings are based on the estimates of WTP and expected costs from extrapolation of points observed before and after a reform in 2007, which changed insurance premiums and generosity of benefits.

policies that are difficult to study using the approach in Landais et al. (2017). The reason is that to analyze alternative insurance designs, one needs to take into account preferences, risks and private information about these risk. However, these parameters are difficult to recover without theoretical assumptions. Furthermore, such a structural model is necessary to study policies that have not been observed in this context before. Finally, the empirical approach in this paper allows for more comprehensive exploration of detailed data and rich variation not limited to price changes to understand complex insurance choices.

The model used in the empirical analysis is in the spirit of Einav, Finkelstein, and Schrimpf (2010) who evaluate the costs associated with private information and corresponding gains of mandates in an annuity market. The authors also use a comprehensive dynamic structural model of choice under uncertainty to recover policy-relevant dimensions of individual heterogeneity.

Finally, the paper is related to a strand of the literature studying the optimal design of insurance contracts.[16] Previous works emphasize the importance of a contract structure beyond pricing, which was a dominant focus of the literature. This paper contributes by adding a piece of evidence of the importance of a dynamic component of adverse selection. Similar time-selection evidence was documented in healthcare (Aron-Dine et al., 2015; Einav et al., 2015; Einav, Finkelstein, & Schrimpf, 2017) and dental care markets (Cabral, 2016). There are a number of papers that study the role of a non-linear benefits schedule on the dynamics of unemployment. For instance, Kolsrud et al. (2018) study the role of duration-dependent UI benefits but this work is more related to the literature on labor supply responses. Similarly, DellaVigna et al. (2017) analyze the role of a benefits structure in the presence of non-classical behavioral responses. Instead, I consider non-linear time-based insurance eligibility and additional dimensions of adverse selection that it creates instead of looking at how UI benefits affect the duration of unemployment.

The paper is organized as follows. Section 2 introduces institutional details of UI in Sweden and describes the data. Section 3 presents descriptive evidence that motivates the empirical analysis and modeling choices. Section 4 describes a structural model and an estimation approach. Section 5 analyzes welfare under current and counterfactual policies. Section 6 concludes.

[16] Azevedo and Gottlieb (2017) study perfect competition in selection markets with the endogenous contract formation. They show that mandates may cause distortions associated with lower prices for low-coverage policies, which results in adverse selection on the intensive margin.

2 Institutional Setting and Data

2.1 UI in Sweden

A vast majority of developed countries have adopted centrally provided and mandatory unemployment insurance systems. Such systems are typically funded through taxes and cover all eligible individuals. In contrast, unemployment insurance in Sweden is divided into basic and voluntary income-based programs. The basic compulsory insurance similarly to mandatory systems grants a fixed daily amount of 320 SEK ($35) conditionally on meeting basic and work requirements.[17] Individuals are required to be registered at the Public Enrollment Service (PES), carry out a job-seeking plan and work at least 80 hours per month over six uninterrupted months during the preceding year.

Eligibility for voluntary income-based insurance also requires paying monthly fees to UI funds for at least 12 consecutive months.[18,19] Before 2007, fees for employed and unemployed individuals coincided. As a result of a labor market reform, fees for employed individuals more than tripled on average. Figure 1 demonstrates average fees for employed and unemployed individuals over time.

Benefits recipiency is limited to the period of 300 days (60 weeks or 14 months) of interrupted or uninterrupted unemployment after which eligibility requires fulfilling the working conditions from the beginning.[20] Unemployment without a valid reason results in an uncompensated period of up to 45 days. The reform in 2007 also reduced the generosity of benefits displayed in Figure 2.

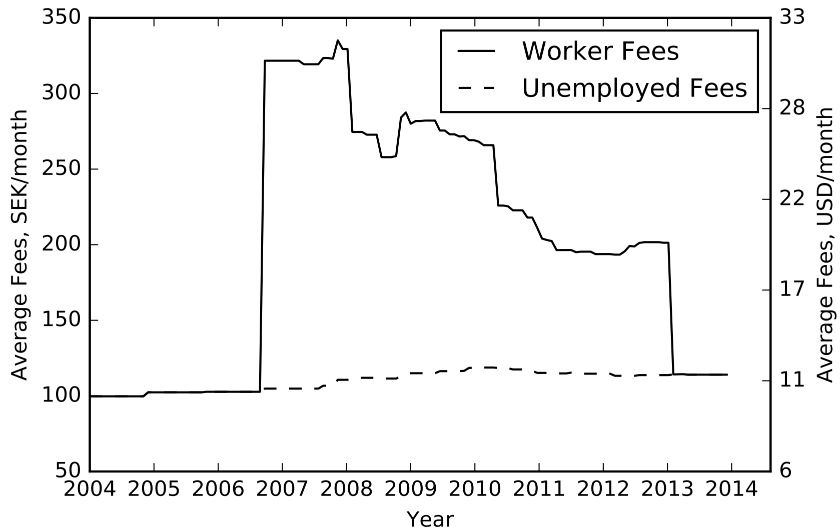
[17] The amount was raised to 365 SEK ($40) in September 2015. For more details regarding changes in 2015 see http://www.fackligtforsakringar.n.nu/a-kassan or http://www.regeringen.se/artiklar/2016/09/en-battre-arbetsloshetsforsakring/ .

[18] There are 29 UI funds that were active during the period under consideration. Individuals are often enrolled in a UI fund based on an industry or a type of employment since funds are linked to labor unions. Therefore, there is virtually no competition among funds.

[19] Enrollment requires working for 1 month.

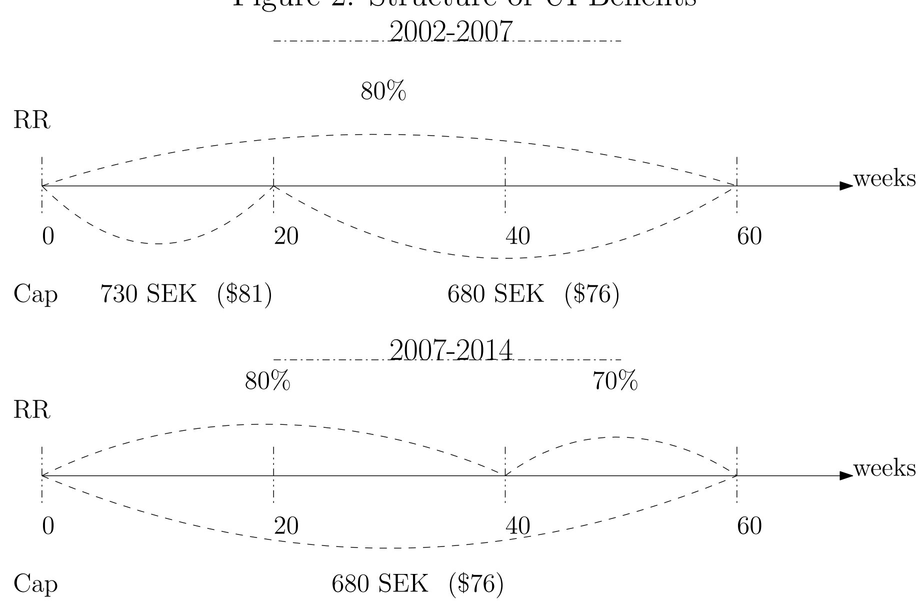
[20] If the accumulated unemployment duration exceeds 300 days, an individual is assigned to an intensified counseling program or can be granted with an extension of 300 days if the counseling is deemed to be unnecessary (but only once). This option disappeared after the reform in July 2007. For more information see https://handels.se/akassan/arbetslos1/regler1/forandringar-i-a-kassan-sedan-2007/ .

Figure 1: Voluntary Insurance Fees, SEK/month



Notes: The Figure demonstrates changes in monthly insurance fees during the period 2004 - 2014. The lines represent average over insurance funds premiums, which vary slightly. Two lines correspond to fees paid by employed and unemployed individuals, correspondingly. Those lines coincide during 2004 - 2007 and after 2013. Fees for employed individuals were considerably higher during 2007 - 2014.

Figure 2: Structure of UI Benefits



Notes: The Figure presents the structure of UI benefits before and after the reform in 2007. The lines with arrow represent schedules of benefits for a maximum of 60 weeks of accumulated unemployment covered by UI. Replacement rate (RR) is presented above the corresponding line. The cap is displayed below the corresponding line.

Before the reform in 2007, voluntary UI provided an 80% replacement rate subject to a cap, which depends on a number of accumulated unemployment weeks. For individuals who accumulated less than 20 weeks of unemployment, the cap was 730 SEK ($81) and 680 SEK ($76) for those with more accumulated weeks. To put this into perspective, the insurance caps correspond to approximately 16 060 SEK ($1 784) and 14 960 SEK ($1 662) of monthly income, respectively. Basic mandatory insurance benefits amount to 7 040 SEK ($782) of monthly income. Average income in the sample used in the analysis, which I discuss in the next section, is approximately 24 834 ($2 759) SEK in 2008. It is almost 54% higher than the first cap and 66% higher than the second cap. A labor market reform introduced changes in both a replacement rate and the cap structure in January 2007. The replacement rate for the first 40 weeks remained 80% and was reduced to 70% for the following 20 weeks.[21] The cap became constant for an entire 60 weeks period and amounted to 680 SEK ($76).[22]

2.2 Data

The empirical analysis in this paper is based on Swedish administrative data from a number of sources. A core dataset comes from a public authority that administers unemployment insurance funds (Inspektionen för arbetslöshetsförsäkringen - IAF). It contains monthly membership records including insurance fund affiliations and premiums. The dataset contains 2 167 287 unique individuals[23] over the period 1999 - 2014. It is not representative of the population since it does not contain individuals who have not claimed UI benefits.[24]

I match the IAF dataset to the data from the Public Employment Service (PES), which provides information on all registered unemployment spells including dates and unemployment categories.[25] A rich set of annually observed individual characteristics comes from the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) including a

[21] Parents with children, younger than 18 are eligible for additional 150 days of 70% replacement rate benefits. Those who are not eligible for additional benefits and continue under the job and activity guarantee program have 65% replacement rate.

[22] Eligibility for income-based insurance is a prerequisite for even higher income compensation from a union that removes the cap. The analysis in this paper does not take it into account. Although the presence of additional fund-based insurance affects parameter estimates, it should not affect the comparative analysis of various UI designs.

[23] In fact, the dataset contains 2 199 941 unique individuals but 32 654 individuals were missing in the longitudinal dataset, which provides individual labor market characteristics. Therefore, those individuals, which are a negligible share of the dataset, are excluded.

[24] Legal restrictions do not allow disclosing membership information about individuals who have not claimed unemployment benefits.

[25] The structural model presented later in this paper has monthly dynamics. I aggregate daily employment and insurance data to monthly. For the cases when, for instance, unemployment duration covers only a part of a month, I code this month as unemployment. Another option would be to round months off.

wide range of demographic characteristics, education, income from various sources (e.g. wage, profit, capital income, social security payment), unemployment, social insurance participation and many others.[26]

Although the data span a period 1999 - 2014, I limit attention to 2002 - 2014 to present the evidence in the next section while using the data for 1999 - 2001 to construct state variables that affect eligibility (e.g. previous enrollment, basic insurance eligibility, a number of accumulated unemployment weeks). The descriptive evidence in the next section is based on this sample to which I refer as "full sample".

A sample used in the estimation differs from the full initial sample due to a number of restrictions that primarily exclude individuals who might not make active unemployment insurance decisions. For computational reasons, I restrict the data used in the estimation to 2005 - 2009 to capture a period containing the reform at the beginning of 2007, which provides important identifying variations for model parameters. I exclude individuals who at least once during 2005 - 2009 were registered at PES with categories that are unrelated to unemployment and usually not administered by the UI authority (e.g. training and educational programs, programs for people with disabilities). It reduces the sample by 672 890 individuals. I also exclude part-time unemployed since they have different budget sets not captured within the scopes of the empirical model. Accounting for part-time unemployment would introduce complications in estimation since those individuals face an income stream, which is a mix of wage and benefits. Therefore, to preserve model tractability, I omit those individuals. It reduces the sample further by 185 321 individuals. I exclude individuals who were constantly either older than 64 or younger than 24 years old during the estimation period 2005 - 2009. A final restriction affects individuals who were always receiving social insurance benefits (e.g. disability, unemployment, sickness) during 2005-2009. It results in a baseline estimation sample that contains 865 363 individuals.[27] Table 1 presents key descriptive statistics of the full sample and the selected baseline estimation sample in comparison with the economically active population of 16 - 64 years old.

[26]Wage data comes from annual records. I divide yearly wage by a number of employment months in a given year to calculate monthly wages.

[27]I randomly split estimation sample into two equally sized samples. I use a 5% random sample of the first sample in estimation and welfare analysis for computational reasons. I use second sample to investigate the quality of the model fit.

Table 1: Descriptive Statistics, 2008

	Full Sample	Estimation Sample	Swedish population 16 - 64 years old
Employment Income			
Mean, *SEK/month*	24 754	24 834	28 623
Median, *SEK/month*	23 233	23 308	25 317
Married	87%	87%	88%
With Children	54%	54%	54%
Nr. of Children, *median*	1	1	1
Age, *median*	40	40	40
Female	53%	51%	49%
With Higher Education	28%	27%	25%
N	2 167 287	865 363	-

Notes: Column (1) shows descriptive statistics and unemployment patterns for the full sample. Column (2) represents the sample used in the empirical analysis. Column (3) describes the full Swedish population for the comparison purposes. The upper part of the Table shows descriptive statistics for 2008, which is one of the years used in estimation. The lower part describes a distribution of a number of unemployment months that individuals accumulated during 2002 - 2014.

Table 1 shows that full and estimation samples are very similar in terms of observables. Slight differences are observed in a share of female, which is 51% in the estimation sample compared to 53% in a full sample. Also, an estimation sample contains 27% of individuals with higher education, whereas 28% of individuals in the full sample have higher education. Both of these samples differ slightly from a full population. The main selection margin is the recipiency of UI benefits. Consequently, individuals who are omitted from the full sample on average have higher employment income not adjusted for work intensity. This difference is mechanical since unemployed individuals should have less wage income. The selected sample contains slightly more individuals with higher education, which is also mechanical since it contains less relatively young individuals who are most likely have not finished higher education. Finally, a full sample is represented by a 4% lower share of female individuals.

Although full and estimation samples are very similar in terms of unemployment patterns, they, as expected, differ from a full population. Selected samples contain a 6% larger share of those who at least once during 2002 - 2014 were unemployed. Similarly, conditionally on being unemployed at least once, a distribution of a number of accumulated unemployment months is shifted to the right for the selected samples.

3 Descriptive Evidence

Unemployment insurance is at risk of private information problem, which might have non-negligible welfare costs. The term private information typically includes adverse selection and moral hazard. The essence of adverse selection in UI is that individuals tend to have more information about their overall unemployment risks. It usually leads to a positive correlation between insurance probabilities and unemployment risks. However, such a positive correlation might not only be driven by adverse selection.

Another alternative theoretical explanation, which is unrelated to private information, is a correlation between risk-preferences and risks (e.g. more risk-averse individuals have higher risks).[28] It would generate a qualitatively similar selection pattern but have different policy implications. The reason is that the absence of a choice imposes the excess burden on individuals who do not value insurance. In addition, the presence of moral hazard might generate a similar positive correlation pattern but require different policy measures. Moral hazard or ex-post selection is a behavioral response to being insured that increases unemployment probabilities. The intuition is that a lack of incentives due to lower financial stakes leads to less job-search or on-the-job efforts.

It implies that there are many scenarios arising from the complexity of insurance decisions that fundamentally hinges on risk perceptions and preferences for risks exposure. This ambiguity might result in a need of the opposite policy measures while generating the same "reduced form" patterns in the data. This section does not attempt to disentangle those forces since it might have limited use for the welfare analysis. For a discussion and an attempt to separate those scenarios using institutional variation, one should consult Landais et al. (2017). The main point of this discussion is that policy conclusions aimed at maximizing welfare rely on being able to disentangle risk preferences and information about risks, which often requires a theoretical structure. More importantly, in order to study alternative contract design regulations, it is required to identify the sources of selection to be targeted by the contract features.

In this section, I present a number of descriptives patterns in the data that motivate modeling choices in the next section. There are several sources of variation that play a key role in the empirical analysis. Firstly, I leverage cross-sectional variation in incentives to be insured. This variation stems from differences in insurance premiums across occupation-specific UI funds and in

[28] De Meza and Webb (2001) show that multiple levels of heterogeneity might also result in advantageous selection.

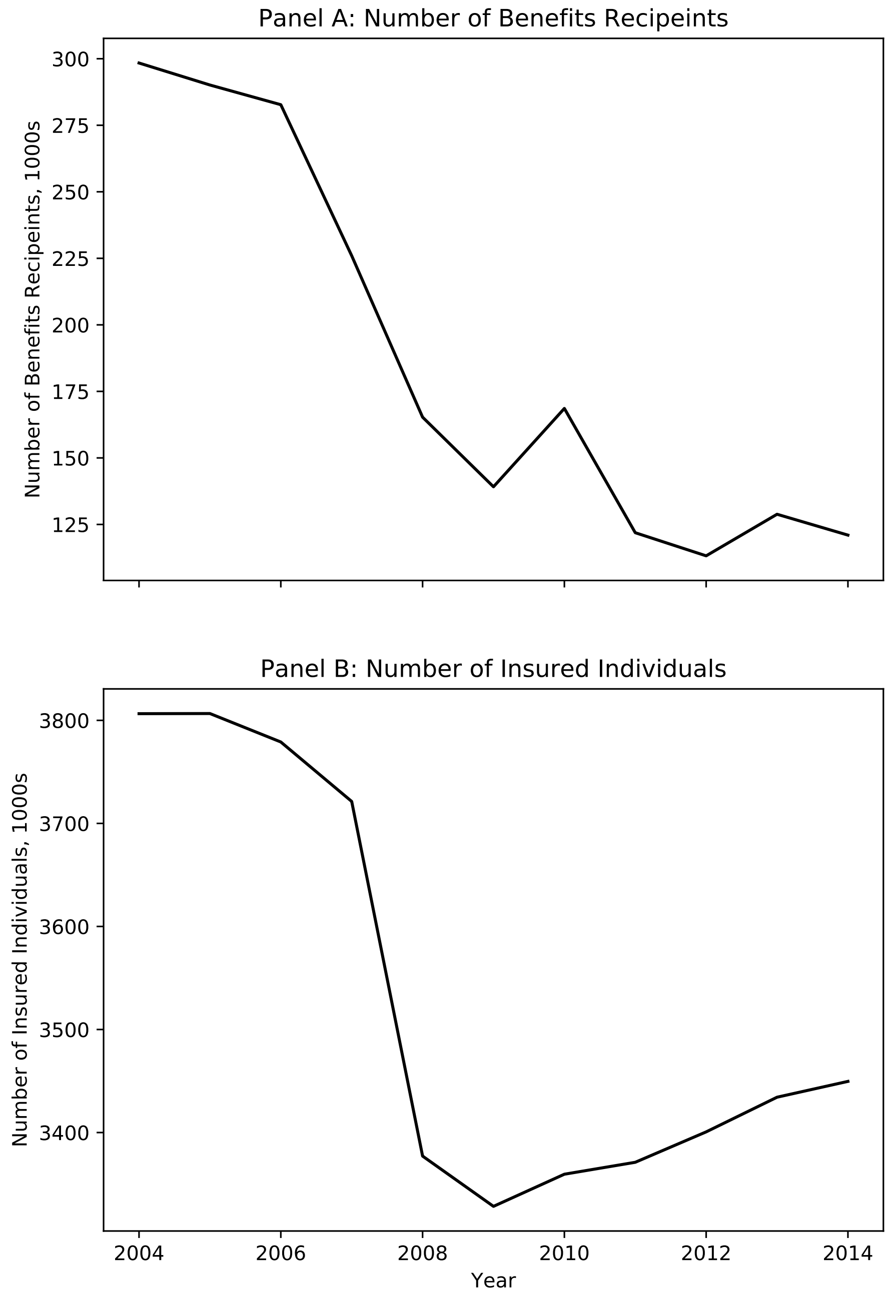
a replacement rate due to a cap, which varies with unemployment duration. Another dimension of the variation is a result of a reform in 2007, which raised insurance premiums primarily for employed individuals and weakly reduced the generosity of benefits. These changes caused behavioral responses illustrated in Figure 3.

The Figure shows that reform is associated with changes in a number of aggregate indicators, which might be driven by individual responses to the reform. More precisely, a number of benefits recipients and insured dropped in 2007 (Panels A and B, correspondingly).[29] However, this aggregate evidence cannot be solely attributed to changes in the structure of UI. The reason is that insurance decisions and aggregate outcomes are jointly determined by individual preferences, insurance structure, and labor market conditions.

Apart from an important role of adverse selection and moral hazard discussed in Landais et al. (2017), another dimension of private information might stem from the specific structure of insurance contracts. One of the eligibility conditions for voluntary UI requires being insured for at least twelve consecutive months. In this case, individuals with superior information about employment outcomes should start paying insurance fees exactly twelve months before the unemployment date, which would lead to time-selection. The literature has documented similar behavioral patterns in, for example, health insurance (Aron-Dine et al., 2015; Einav et al., 2015, 2017) and dental markets (Cabral, 2016). The presence of this phenomenon also contributes to a positive correlation between unemployment risks and the likelihood of being insured. Although, it can be argued that time-selection is a part of adverse selection and can be resolved by mandates, alternative contracts that specifically restrict time-selection might be welfare-improving. The presence of time-selection can be shown with a distribution of a number of enrollment periods with which individuals start unemployment spells in the data displayed in Figure 4.

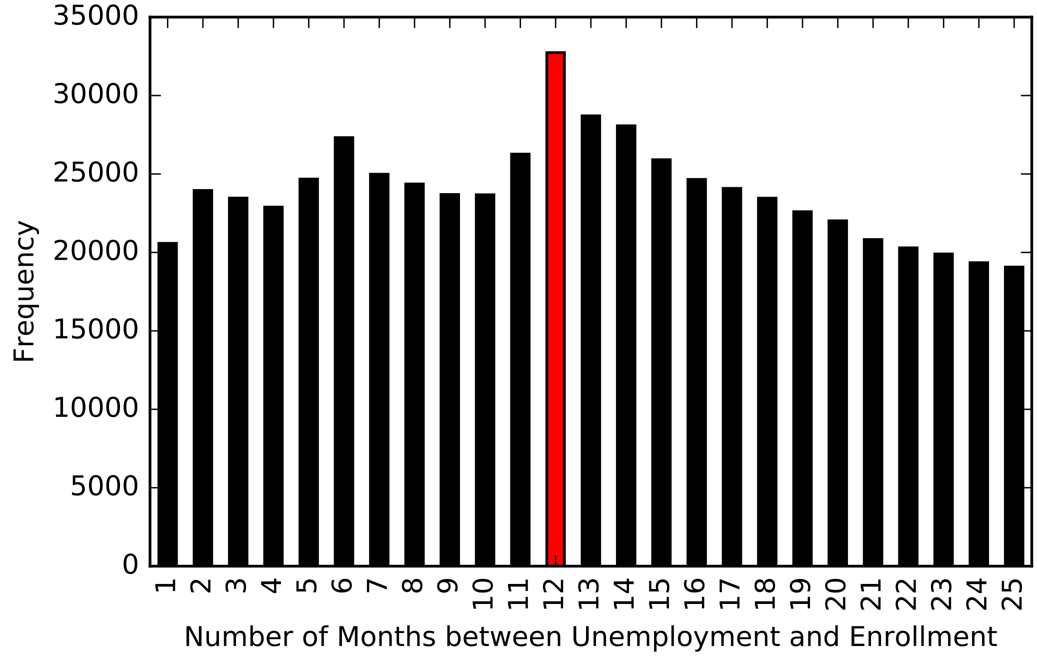
[29]Note that a number of insured and a number of benefits recipients are not directly linked since one can receive basic insurance even without being a fund member.

Figure 3: Unemployment Insurance and Benefits Recipiency, 2004 - 2014



Notes: The Figure presents aggregate indicators over time. The source is *Inspektionen för arbetslöshetsförsäkringen.*

Figure 4: Distribution of Accumulated Enrollment Months at the Beginning of Unemployment



Notes: The Figure presents a distribution of a number of accumulated enrollment months before the commencements of unemployment spells. The red bar denotes twelve consecutive months of enrollment required for eligibility. The histogram contains a spike exactly at the red bar, which implies that individuals are more likely to start unemployment spells with twelve months of enrollment.

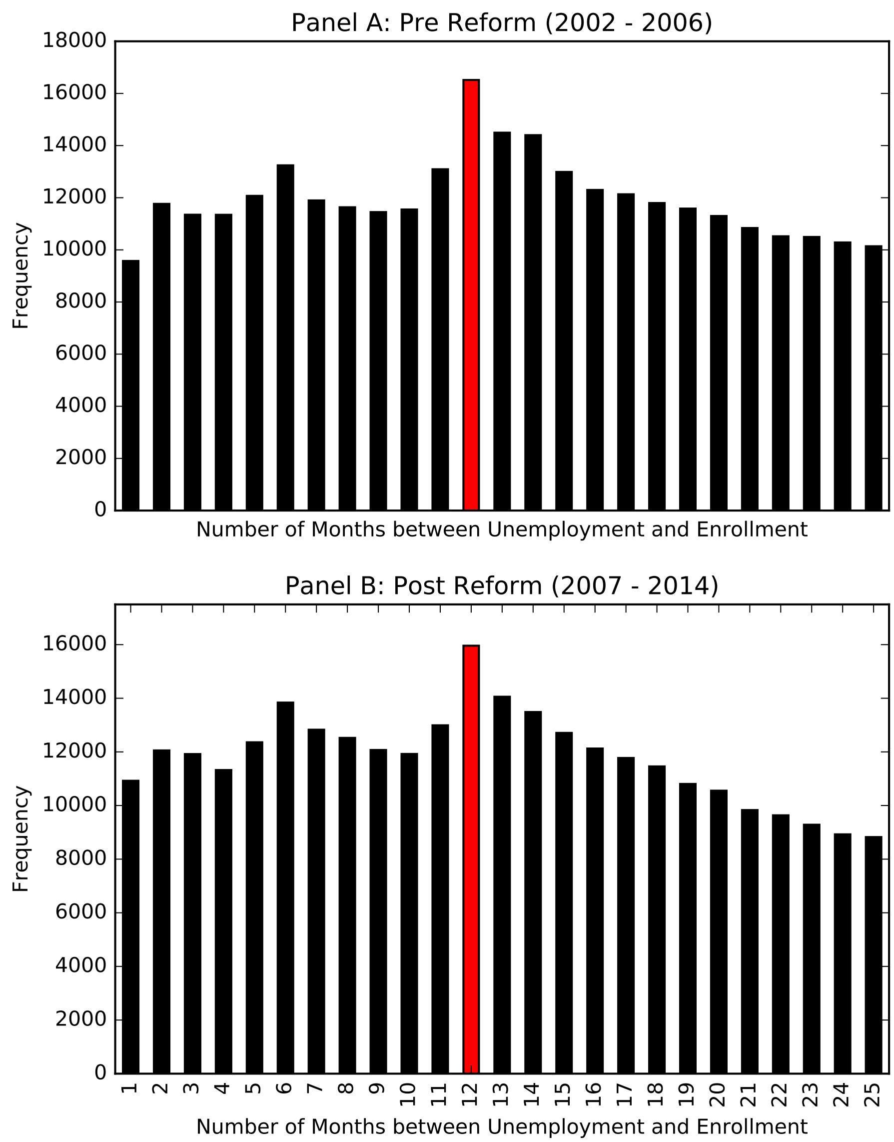
The distribution in the Figure has a spike (red) at exactly twelve months of enrollment, which suggests that individuals are more likely to start paying insurance premiums twelve months before unemployment. It allows being eligible for benefits exactly at the commencement of an unemployment spell, which minimizes the total amount of premiums to get eligibility. An area of the distribution to the left of the red spike is non-uniform and non-monotonic, due to differences in private information about future employment outcomes. These differences are a result of various layoff notification specified in employment contracts, individuals informal knowledge about unemployment or the presence of probation contracts that often last for 6 months. The model in the next section systematically exploits those patterns and attributes them to the differences in the information about future employment outcomes. It is important to note that the model is agnostic about the source of private information since only its existence is welfare-relevant. Time-selection evidence for various subgroups is presented in the Appendix C (Figures 17, 18 and 19) and shows identical patterns.

The key identification assumption that will allow using changes in the generosity of benefits and premiums to separately identify distributions of risk preferences and private information is that changes in insurance conditions do not affect private information about unemployment. The example of the violation of this assumption would be, for example, if reform in 2007 not only changed the attractiveness of insurance but also information about future unemployment. It would imply that changes in insurance decisions are not only driven by changes in attractiveness of insurance but also by changes in private information structure. I investigate a potential violation of the identification assumption in the identification section. In this section, I present the time-selection evidence but separately for the periods before and after the reform in 2007 in Figure 5.

As can be seen, the patterns are similar for both periods. However, this evidence should be viewed as neither necessary nor sufficient to ensure the validity of the assumption. The presence of considerable differences on those figures could alert about both changes in information and time-selection accompanied by a moral hazard response. The latter means that individuals not only select the timing of insurance but also choose if and when to become unemployed. The intuition is that the reform in 2007 weakly reduced the generosity of benefits and raised premiums, which implies that it costs more to qualify for less generous benefits. In the absence of the changes in information about future unemployment, the reform did not change bunching incentives for individuals who just knew about forthcoming unemployment. Those individuals should still prefer being covered even for one month compared to not paying fees and being ineligible. However, individuals who decide to facilitate a layoff and choose enrollment timing are affected since insurance becomes less generous. It might encourage them to keep being employed or switch a job without relying on benefits. Those individuals would exclude themselves from the bunching area and reduce the spike. The fact that it is difficult to graphically see considerable differences in bunching patterns can be also explained by a relatively small scale of the reform, which did not induce such institutional changes and behavioral responses.

Another important pattern of insurance decisions is that many individuals tend to have only one insurance spell, which often covers the entire observed period. A maximum number of insurance sequences in the course of observed period 1999 - 2014 amount to eleven. The median duration of an insurance sequence is 99 months. It might suggest that individuals display a considerable amount of inertia in fairly frequent monthly choices.

Figure 5: Distribution of Accumulated Enrollment Months: Before and After the Reform



Notes: The Figure presents a discrete histogram of a distribution of a number of accumulated enrollment months before the commencement of unemployment spells. It replicates the evidence in Figure 4 but separately before (Panel A) and after the reform in 2007 (Panel B).

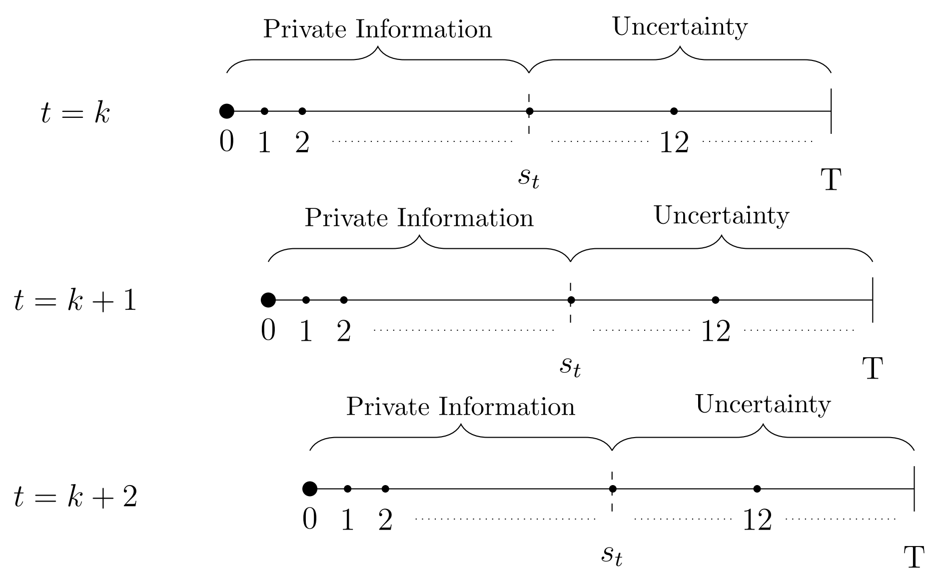
This section described the main descriptive patterns observed in the data. Firstly, it has been shown that individuals react to changes in premiums and benefits generosity. Secondly, the fact that many individuals have long insurance sequences might suggest a presence of choice inertia. Finally, the data display the signs of time selection. The model presented in the next section attempts to incorporate those elements in a framework that enables addressing the question of optimal regulations in UI.

4 Empirical Model

4.1 Model

I model a forward-looking decision of an individual who faces the risk of unemployment and maximizes the expected utility of income. The insurance decisions are monthly, which corresponds to the timing of premium payments. The model resembles an overlapping individual structure depicted in Figure 6.

Figure 6: Structure and Timing of Insurance Decisions

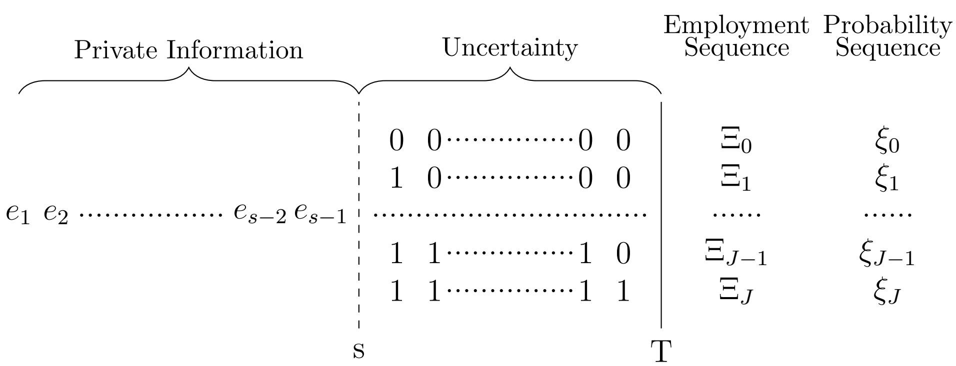


Notes: The Figure illustrates the overlapping-individual structure of the dynamic decision in the model. It shows that each period t an individual solves a new dynamic optimization problem of length T to decide whether to pay monthly insurance premiums at t.

The Figure suggests that an individual solves a new optimization dynamic optimization problem each period t to decide whether to pay insurance premiums $l_t \in \{0, 1\}$. The information structure at the time of each decision consists of two parts. The first one denoted "Private Information" means that an individual can perfectly foresee employment outcomes in the next s periods. This knowledge might come from multiples sources, e.g. lay-off notifications, informal information sharing with an employer. I refer to the length of a perfect foresight period s

as *private information type*.[30] According to the institutional details, lay-off notifications are restricted to maximum of 12 months. Therefore, I assume that individuals can be any of the types from 1 to 12.[31] The model presented in this paper is agnostic about the sources of this private information since only its existence is important for the welfare analysis presented later in the paper. Another part of individual's information is denoted by "Uncertainty" and imply that after the window of perfect foresight, a worker is uncertain about employment outcomes for the remaining part of the planning horizon from s to T. In the model, this uncertainty is treated as a collection of all potential employment sequences that might happen from s to T illustrated in Figure 7.

Figure 7: Structure and Timing of Insurance Decisions



Notes: Figure illustrates the essence of uncertainty and private information in the model. Private information is modeled as a perfect foresight for s periods in the future meaning that an individual can perfectly observe whether she is employed or unemployed in each of these periods. $e \in \{0, 1\}$ denotes the realization of employment/unemployment outcomes. Since an individual does not know if she is employed in periods from s to T, there might be multiple potential employment sequences $j \in J$ spanning all possible combinations of zeros and ones depicted in the Figure. Each sequence is denoted by Ξ_j and can occur with a probability ξ_j. The number of sequences $J = 2^{T-s}$.

The Figure shows that individual's information consists of a perfect foresight about the employment outcomes from time of a decision to s months in the future $e = \{e_k\}_{k=1}^{s}$, and

[30]More formally, $\hat{s} = \min\{s, s_u\}$ where $\hat{s}$ denotes the number of periods that can actually be foreseen in the future, s_u is a number of periods until next unemployment and s is a number of periods that can be observed in the future in the absence of earlier unemployment, to which I refer as private information type. This formulation means that an individual can perfectly know future employment outcomes for s periods unless there is forthcoming unemployment. It reflects the fact that individuals cannot observe the end of the unemployment spell. In this case, the information is limited to only one period ahead in the unemployment spell.

[31]Temporary worker can foresee the layoff even further ahead. I discuss the implication of this the robustness section of the paper.

uncertainty that lead to $J = 2^{T-s}$ possible $\{0, 1\}$ sequences of length $T - s$. Each sequence is denoted Ξ_j and can occur with probability ξ_j. A current period is indexed by 0 and is already observed.

This overlapping structure is adopted because of the features of UI in Sweden. Each individual decision has a critical impact only on the next 12 months since it determines whether or not she will be eligible for UI benefits during this period. It stems from the discontinuity in membership eligibility condition that requires paying monthly insurance premiums for 12 uninterrupted months. A current decision will still have an impact on future outcomes after 12 months but only through on decisions in the future. In other words, even if an individual decides not to pay an insurance premium, she still can become eligible for benefits in any period after 12 months if she acts correspondingly in the future. In addition, such a representation of a decision-making process has a considerable computational advantage discussed later in this paper. This structure, however, is not a typical dynamic agent model. In particular, this overlapping structure does not force individuals to be committed to the optimal choices computed before and, instead solves for an optimal choice each new period upon arrival of new information. This structure results in more realistic assumptions about state variables in the model, which I discuss in the remainder of this section.

I assume that an individual i decides to pay insurance premiums at time t if it maximizes the expected utility of a sum of incomes over the next T periods.[32]

$$l_t^* = \arg\max_{l_t \in \{0,1\}} U(l_t; \rho_t, s_t) = \arg\max_{l_t \in \{0,1\}} \sum_{j=1}^{J=2^{T-s_t}} \xi_{tj} \cdot \frac{\left(\overbrace{\Pi_j\left(l_t; \{\hat{l}\}_{k=t+1}^{t+T}, \Xi_{tj}\right)}^{\text{Payoff}}\right)^{1-\rho_t}}{1-\rho_t} \tag{1}$$

where $\hat{l}$ is an insurance decision that an individual plans to make in the future.

The formulation in (1) means that an individual faces the uncertainty of employment outcomes over the next T periods. Although she can perfectly foresee the outcomes of next s periods because of private information, she is uncertain about the outcomes after periods s and up to T. It leads to $J = 2^{T-s}$ potential employment sequences in the future. Each of these sequences leads to different payoff, which is a sum of income streams over the planning horizon of T months conditionally on optimally planning future insurance decisions:

[32] I drop index i from the equation for convenience.

$$\Pi_j\left(l_t;\{\hat{l}\}_{k=t+1}^{t+T},\Xi_j\right)=\sum_{n=t}^{T}\pi(l_n,\kappa_n(l_{n-1}),e_{nj},\Gamma_n) \tag{2}$$

where e_{nj} is an employment status from one of employment sequences Ξ_j; κ_n - a number of enrollment periods that has been accumulated; Γ_n - a collection of state variables in each period n not affected by individual decisions. It includes wage w_n, replacement rate b_n, cap B_n, insurance premiums τ_n and basic insurance amount in the case of not being eligible $\underline{b}_{it}$.

The number of enrollment periods (κ) is the only state variable affected by an individual choice and evolves as follows:

$$\kappa_{t+1}=\begin{cases}\kappa_t+1, & \text{if} \quad l_t=1\\ 0 \quad , & \text{if} \quad l_t=0\end{cases}$$

In turn, κ_n determines whether an individual is eligible for income based insurance:

$$\Lambda_{t+1}=\begin{cases}1, & \text{if} \quad \kappa_{t+1}>=12\\ 0, & \text{if} \quad \kappa_{t+1}<12\end{cases}$$

One-period payoff of an individual can be expressed:

$$\pi_t=\underbrace{(1-e_t)\cdot\left(\overbrace{(1-\Lambda_t)\cdot\underline{b}_t}^{\text{ineligible}}+\overbrace{(\Lambda_t\cdot\min\{b_t\cdot\bar{w}_t,B_t\}}^{\text{eligible}}\right)}_{\text{unemployed}}+\underbrace{e_t\cdot w_t}_{\text{employed}}-\underbrace{l_t\cdot\tau_t}_{\substack{\text{pay}\\ \text{premiums}}}$$

An individual chooses to pay premiums if it maximizes the expected utility in (1). The model described so far represents an individual decision to pay insurance each period as a sequence of static choices under uncertainty coming from the absence of perfect information about employment outcomes between s and T.

Although the decision is static, it nests a sequence of dynamic decisions captured in Equation (2). In other words, to determine the optimal insurance path under sequence Ξ_j, she has to solve the dynamic programming exercise. It means that each individual at each period has to solve J number of dynamic programming exercises to decide whether to pay insurance premiums according to (1).

Assumption 4.1. *Each decision period t individuals have perfect foresight about state variables in $\{\Gamma_n\}_{n=t}^{t+T}$.*

Assumption 4.1 implies that individuals can perfectly foresee all state variables that are not affected by an insurance decision. The sequential structure of the problem makes the assumption being not far from reality since it only has to hold for T periods in the future. Therefore, the assumption seems reasonable for not very large T. I postpone the discussion of the choice of the planning horizon T for later. In contrast, more standard models in which individuals would commit to an optimal strategy and plat until some common terminal period (e.g retirement), this assumption would be questionable since it has to hold for the entire planning horizon.[33]

Assumption 4.2. *Individuals have rational state-dependent beliefs about probabilities of employment in the periods of uncertainty conditionally on individual and labor market conditions.*

Assumption 4.2 implies two restrictions on indivuduals' beliefs about unemployment probabilities outside of private information. Firstly, the beliefs should be rational in the sense that if identical individual in terms of relevant individuals and labor market characteristics face a probability p of unemployment in some period t, she should believe that with the probability p she also will be unemployed in a period t. Secondly, the assumption allows these beliefs to be state-dependent. It means that unemployment probability p_t is conditional on the employment status in the preceding period e_{t-1}. Given that an individual has conditional beliefs about each of the periods $p_t(e_{t-1}, X_t)$, where X_t is set of relevant individual characteristics, individuals form unconditional employment beliefs for any period in the future after s as a Markov sequence:

$$E\left[p_{t+n}\right] = E\left[p_{t+n-1}\right] \cdot p^1_{t+n} + \left(1 - E\left[p_{t+n-1}\right]\right) \cdot p^0_{t+n} \qquad \forall n > t+s+1 \tag{3}$$

where $E\left[p_{t+s+1}\right]$ is $p_t(e_{t-1}, X_t)$ since the outcome of the period $t+s$ to be observed because of the private information.

Individual's beliefs about a probability of each sequence j can be expressed:

$$\xi_{jt} = \prod_{q=t+s+1}^{T} m_{qt} \tag{4}$$

where m_{qt} is a probability that the outcome e_{jq} from sequence Ξ_j is true. If $e_{jq} = 1$, $m_q = E\left[p_q\right]$ and $m_q = 1 - E\left[p_q\right]$ otherwise.[34]

The model presented in this section contains two objects that are observed to individuals but not to us and hence have to be recovered using the model and data. These parameters are the

[33] Another modeling alternative would be to assume rational expectations about these state variables, in which case multidimensional integration is required. It complicates the process of solving the model and, again, raises concerns regarding the assumption is realistic.

[34] Note that for the periods from the private information region $q \leq t+s$, $m_q = 1$ if e_q is true and $m_q = 0$ otherwise.

distribution of individual risk preferences ρ_{it} and the distribution of individual types s_{it}. The next section discusses the identification of these parameters.

4.2 Identification

Identification of the empirical model outlined in this section concerns separately recovering distributions of risk preferences and types denoted as $F(\rho)$ and $\Phi(s)$. Note that an individual buys insurance ($l_{it} = 1$) if $U(l_{it} = 1; \rho_{it}, s_{it}) > U(l_{it} = 0; \rho_{it}, s_{it})$ from (1). As shown by Apesteguia and Ballester (2018), for a class of utility functions that include CRRA, there is a unique risk preference parameter denoted by λ_{its} where $U(l_{it} = 1; \rho_{it}|s_{it}) = U(l_{it} = 0; \rho_{it}|s_{it})$ conditionally on type s.[35,36] If an individual risk preference value is above λ_{its}, she should pay premiums conditionally on being a type s. Conditional probability of paying insurance premium is:

$$Pr\left(l_{it} = 1|s_{it}\right) = Pr\left(\rho_{it} > \lambda_{its}\right) = \int_{\lambda_{its}}^{\infty} dF(\rho) \tag{5}$$

Unconditional probability of paying insurance premiums can be written:

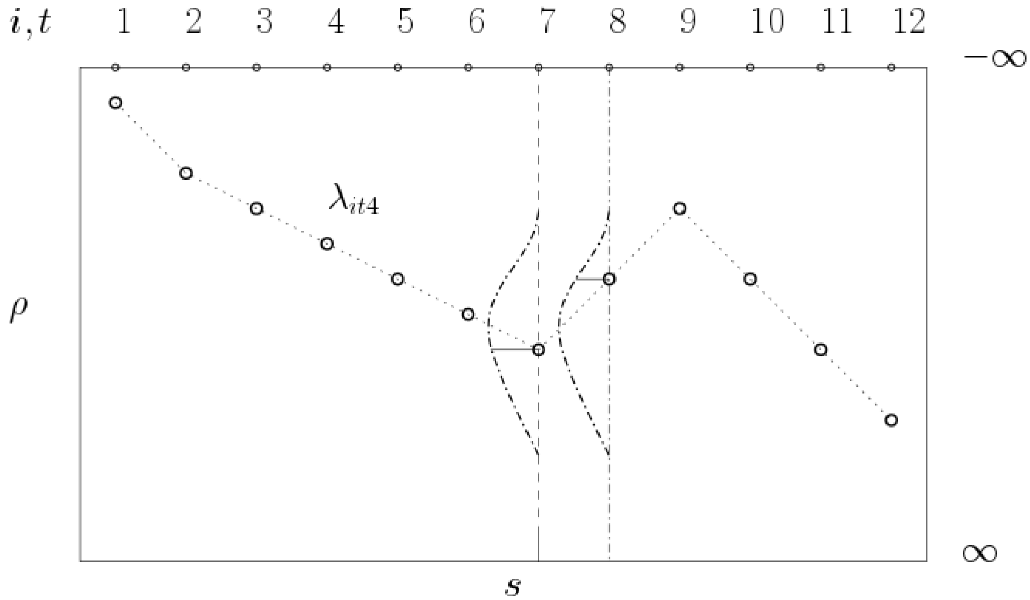
$$Pr\left(l_{it} = 1\right) = \int_{s} Pr\left(\rho_{it} > \lambda_{its}|s\right) d\Phi(s) = \int_{s}\int_{\lambda_{its}}^{\infty} dF(\rho) d\Phi(s) = \sum_{s=1}^{12} \phi_{its} \int_{\lambda_{its}}^{\infty} dF(\rho) \tag{6}$$

Note that λ_{its} is obtained from the data by solving the model without any information about the distributions of risk preferences or information types. Although it is irrelevant for the identification discussion, it might be useful to mention that the model does not have a closed form solution but λ_{its} can be obtained numerically by solving (1) for each i, t and $s \in \{1, ..., 12\}$. Therefore, λ_{its} can be viewed as sufficient statistics from the outlined model that fully describes its outcomes. It is useful to view the identification problem as illustrated in Figure 8.

[35] Apesteguia and Ballester (2018) do not prove the uniqueness of an indifference point directly but they prove that the upper bound of an interval, where the difference is monotonic, converges to this unique indifference point as $t \to \infty$ where $t > 0$ multiplies the outcomes of the lottery.

[36] Although the indifference point is theoretically unique, it is not true numerically since because of computer precision constraints, a limit of a utility difference that approaches zero actually becomes zero at some point. I discuss how I deal with the computation of thresholds in Appendix B.

Figure 8: Graphical Interpretation of the Identification Problem



Notes: The Figure illustrates the graphical interpretation of the identification of model parameters. The vertical axis denotes support of risk preference parameters of CRRA utility function. The horizontal axis denotes discrete support of types that can take any value between 1 and 12 assumed in the model. The empty dots denote λ_{its} and depict risk preference thresholds where an individual is indifferent between being insured and uninsured. The bell-shaped curves denote a distribution of risk preferences $F(\rho)$. Conditionally on type s_{it}, a probability to buy insurance is a probability that ρ_{it} is higher than λ_{its}, which corresponds to the area above the solid line that connects a bell-shaped curve and a dashed-dotted "type-line". Each of these areas is a graphical representation of Equation (5). The unconditional probability from (6) is a sum of these individual type-conditional probabilities (5) weighted by type probabilities ϕ_{its}.

To summarize, I observe a distribution of insurance outcomes and a distribution of state variables that affect the insurance decisions. Using a structure of the model, it can be summarized in the distribution of risk preference threshold λ. Both distributions of unknown parameters cannot be identified non-parametrically. Below I make an argument for the case in which the model could be identified non-parametrically. Then gradually proceed to our case and discuss which additional assumptions are necessary for identification.

The challenge to identify the model non-parametrically stems from the absence of continuous contract choice since individuals face only a binary contract offer, and fixed duration of the pre-enrollment period (12 months of membership). With continuous contract choice that maps generosity of benefits (g) and duration of pre-eligibility period (d) to $\tau(g, d)$, observing individual choices would allow pinpointing each individual risk preference value and duration of lay-off notification under some regularity conditions for the function $\tau(g, d)$.[37]

Our case differs in two ways. Firstly, contracts do not vary by in pre-enrollment period duration. Therefore, it is required to impose a parametric assumption on the distribution of individual types $\Phi(s)$. Institutional details of Swedish UI described in Appendix A suggest that displacement rules often vary based on a number of observed characteristics such as tenure in the firm and age. It can also vary by firm or industry. Since the types in the model not only represent formal lay-off notification requirements but also information sharing, it is reasonable to expect that type distribution but also conditional on other relevant individual characteristics such as education. Therefore, in the absence of the variation in pre-enrollment conditions, one has to impose a deterministic functional form of a distribution of types conditionally on relevant labor market characteristics, which is believed to be a reasonable assumption based on the institutional background.

Although the model lacks the variation in pre-enrollment conditions, there is variation in the generosity of UI and insurance premiums. The variation mainly results from the reform in 2007, which reduced the generosity of UI benefits for some individuals and raised insurance premiums for all.[38] One additional assumption is needed at this point.

Assumption 4.3. *Changes in UI generosity and premiums as a result of UI reform in 2007 do not affect type distribution* $\Phi(s)$.

Assumption 4.3 requires that a parametric distribution of types $\Phi(s)$ is fixed and do not change together with UI conditions. Otherwise, one cannot disentangle the response to changes in UI conditions from the response to changes in private information. Figure 8 suggests that given that a distribution of types $\Phi(s)$ is fixed, one can use responses to changes in UI conditions to identify risk preferences. More precisely, these differences in the labor market and insurance conditions are translated to differences in risk preference thresholds λ_{its}. Since the individual type is now fixed, I return to the case in Equation (5) or an individual point in Figure 8 disregarding the possibility of being one of 12 types. As a result of the variation due to the reform in 2007, the same individual faces two different insurance choices, which should lead to different λ_{its} conditionally on fixed s that can be obtained from the model. The individual responses to these changes allow recovering bounds of individual risk preferences. For example, let a risk preference

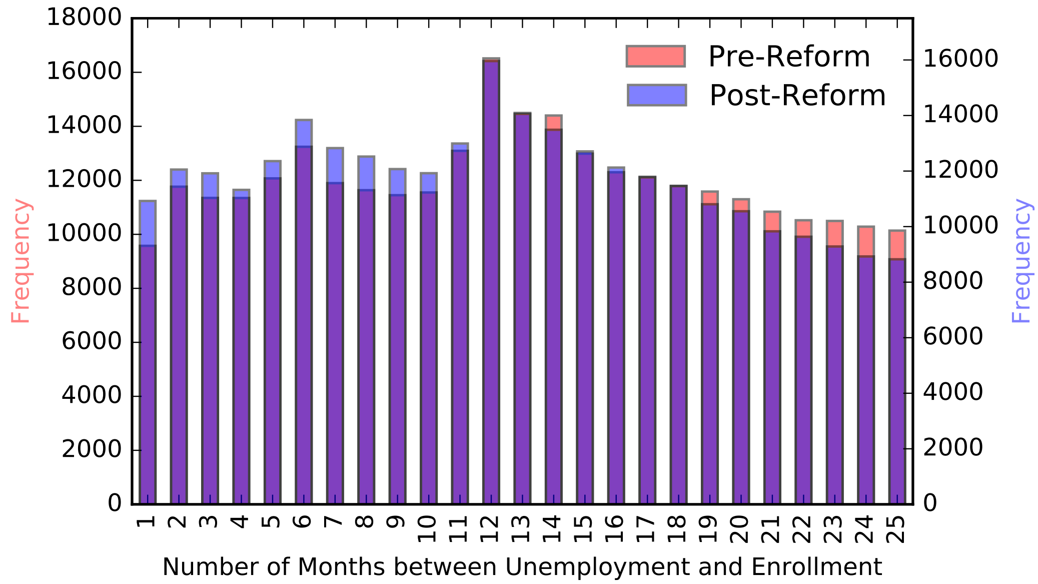
[37] I do not show formally these conditions since this scenario does not correspond to the environment of UI in Sweden and this example serves as an exposition of the identification logic and a need for additional assumptions. However, one can show that identification would require that $\tau(g, d)$ is continuous and strictly increasing in both arguments. It ensures that based on the individual risk preference value and the private information type, there exists a unique contract that is preferred.

[38] Another less salient variation is cross-section and stem from variation in premiums over insurance funds and generosity of benefits due to differences in duration of unemployment and employment risks.

threshold for buying insurance is 10 before the reform and I observe her being insured. After changes in the insurance structure, she should have risk preferences at least as large as 20 to buy insurance and I observe that she stops paying insurance premiums. It allows concluding that her risk preference value should be between 10 and 20.

To investigate how reasonable Assumption 4.3 is, I plot distributions of time selection before and after the reform in Figure 4.2.

Figure 9: Time-Selection Before and After the Reform in 2007



Notes: Figure plots time-selection patterns similar to the evidence presented earlier in the paper separately for the period before and after the reform in 2007.

The Figure illustrates time-selection patterns before and after the reform in 2007. Assumption 4.3 requires that the reform does not have an effect on the type distribution. The evidence that would falsify the assumption is differences in time-selection patterns before and after the reform. However, Figure 4.3 suggests that time selection patterns are overall very similar. Although it does not allow concluding that the assumption should hold, at least descriptive patterns in the data do not convincingly reject it.

Note that the model and assumptions I made only allow pinpointing bounds of risk preference parameters. It is a result of a binary choice that individuals face. Therefore, an additional parametric assumption on the distribution of risk preferences is required for point identification.

4.3 Parametrization and Estimation

Before I discuss parametrization of the model, I make one additional assumption regarding the duration of the planning horizon T. I limit the length of a planning problem to $T = 18$. A chosen T must be larger than 12 in order to capture time-selection behavior as a result of the eligibility requirement. Since, it is required to solve a dynamic model many times for each individual, time, type and sequence to compute payoffs of each action, it becomes computationally burdensome for a large T. In addition, Assumptions 4.1 and 4.2 start being more questionable as T grows. I discuss the importance of a choice of T in the robustness section at the end of the paper.

Assumption 4.4. *Parametric assumptions:*

1. *Individual risk preferences are normally distributed with mean being a function of individuals characteristics $\mu_{it} = \alpha X'_{it}$ and a common standard deviation σ*

$$\rho_{it} \sim N(\alpha X'_{it}, \sigma) \tag{7}$$

2. *Individuals can be one of 12 types $s \in \{1, ..., 12\}$ with probability ϕ_{its}. Types are drawn from multinational logit discrete distribution:*

$$\phi_{its} = \frac{\exp(\beta_s Z'_{it})}{\sum_{k=1}^{12} \exp(\beta_k Z'_{it})} \tag{8}$$

Assumption 4.4 contains key parametric restrictions of the model that stem from the identification discussion. A set of model parameters contains a vector of parameters α, a parameter σ and vectors of type distribution parameters $\{\beta_k\}_{k=2}^{12}$.[39]

I estimate the parameters of the model in three steps. Firstly, one of the model assumptions states that individuals have rational state-dependent beliefs about employment probabilities outside of private information. Using the data on employment outcomes and a large set of demographic and labor market characteristics, I estimate the following model.

$$Pr(e_{it} = 1 | e_{i,t-1}) = Logit(Q_{it} | e_{i,t-1}) \tag{9}$$

where Q_{it} includes observed labor market and individual characteristics and year fixed effects; $e_{i,t-1}$ - previous employment status.

[39] One vector of parameters in multinational logit distribution has to be normalized. I set elements of a vector of type 1 to 0.3.

Equation (9) suggests that it is estimated separately for those who have been employed in the previous period and for those who were unemployed. Upon obtaining parameters of the model, one can predict probabilities of employment for each individual and period in the data conditional on previous employment status. To construct probabilities of sequences in J for each individual and period, I use Equation (4).

On the second stage, I compute risk preference thresholds λ_{its} for each individual i, time t and type s from Equation (1). Recall that an individual chooses to buy insurance if $U(l_{it} = 1; \rho_{it}, s_{it}) \geq U(l_{it} = 0; \rho_{it}, s_{it})$ or, alternatively, if $U(l_{it} = 1; \rho_{it}, s_{it}) - U(l_{it} = 0; \rho_{it}, s_{it}) \geq 0$. As noted by Apesteguia and Ballester (2018), such a utility difference has a unique value of a risk preference parameter ρ_{it} where $U(l_{it} = 1; \rho_{it}, s_{it}) - U(l_{it} = 0; \rho_{it}, s_{it}) = \Delta_{it} = 0$.[40] It means that an individual with a risk preference value that yields $\Delta_{it} = 0$ is indifferent between buying insurance or not. I denote this risk preference value where $\Delta_{it} = 0$ as λ. Any $\rho < \lambda$ would imply that an individual should not buy insurance since she is "sufficiently risk-loving". Similarly, if an individual has $\rho > \lambda$, she should buy insurance. Since the data provide all information required to estimate both $U(l_{it} = 1; \rho_{it}, s_{it})$ and $U(l_{it} = 0; \rho_{it}, s_{it})$, it is possible to numerically compute a value of risk preferences λ at which $\{i, t, s\}$ is indifferent between paying premiums or not. This cutoff not only differs across individuals and time but also by a type s, which is unknown but observed by an individual. As discussed in the identification section, an insurance decision can be summarized by a risk preference threshold, which formally is defined as follows:

$$\lambda_{its} = \rho_{its} : \Delta_{its}(\rho_{its}) = 0 \tag{10}$$

On the third stage, I estimate parameters of the model $\Omega = \{\alpha, \sigma, \{\beta_k\}_{k=2}^{12}\}$. Note that the probability that an individual pays premiums is the probability that her risk preference value is at least as large as the estimated threshold.[41] Given a parametric distribution in (7), this probability can be expressed:

$$Pr(l_{its} = 1) = Pr(\rho_{it} \geq \lambda_{its}) = 1 - F\left(\frac{\lambda_{its} - \alpha X'}{\sigma}\right) \tag{11}$$

[40] Note that although Δ_{it} has a unique intersection with a zero line for a finite value of ρ, the function is not monotonic in ρ, which creates complications in the estimation of discrete choice models under uncertainty. The approach used in this paper does not suffer from this issue.

[41] The normality assumption on risk preference distribution is common in the insurance markets literature (e.g. Einav, Finkelstein, & Schrimpf, 2010; Handel, 2013). Cabral (2016) uses log-normal distribution of risk preference, which rules out the possibility of negative risk preference values. Computed distribution of thresholds presented in Appendix C suggests that the model should allow for risk loving individuals at the expense of imposing the symmetry. I do not allow σ to vary to restrict the model, which already has many parameters. However, heterogeneity in σ might allow restricting a number of predicted risk-loving individuals but is uncommon in the literature.

where $F\left(\frac{\lambda_{its}-\alpha X'}{\sigma}\right)$ is a cumulative normal distribution denoting a probability that an actual risk preference value is below λ_{its}.

As discussed earlier, the institutional details suggest that the probability depends on the labor market affiliations and demographic variables such as age. Therefore, I include a large set of labor market variables (e.g industry, occupation type, education level, education specialization). It generates a large set of parameters, which makes estimation burdensome since there are eleven vectors in β (the first one is normalized) for each matrix of characteristics in Z. However, many variables in Z are highly correlated since, for instance, education and labor market affiliations are closely related to each other. Therefore, I use Principal Component Analysis to reduce the dimensionality of variables in Z to five dummy variables denoted as cluster allocations.[42] I also add binned age variables which together with a constant comprise a vector of eight parameters in β for each type.

The probability that an individual pays premiums is:

$$Pr_{it}(l=1) = 1 - \sum_{s=1}^{12} \phi_{its} \Phi\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right) \tag{12}$$

Since individuals face complicated and frequent choices. I expect an important role of inertia, which has to be included in the model. The identification of inertia parameters requires an inertia-free group of individuals (Handel, 2013). To account for inertia, I make the following assumption.

Assumption 4.5. *Individuals, who because of private information observe forthcoming unemployment or were unemployed in the period before, are not affected by inertia ($\eta = 0$) and face choice affected by inertia otherwise ($\eta = 1$).*

I augment the choice probability equation with the inertia component:

$$Pr_{it}(l=0|s) = F\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right)^{\Upsilon_{it}}$$

$$\Upsilon_{it} = \underbrace{\eta_{it} \cdot \left[\overbrace{(1 - l_{i,t-1}) \cdot \gamma_0}^{\substack{\text{previously}\\\text{uninsured}}} + \overbrace{l_{i,t-1} \cdot \gamma_1}^{\substack{\text{previously}\\\text{insured}}}\right]}_{\substack{\text{with}\\\text{inertia}}} + \underbrace{(1 - \eta_{it}) \cdot 1}_{\substack{\text{without}\\\text{inertia}}}$$

where l_{t-1} - previous insurance status; $\{\gamma_0, \gamma_1\}$ - inertia parameters.

[42]These five components explain approximately 60% of variation.

The intuition for such parametrization is that when insured, individuals are more likely to keep being insured. Hence, Υ will be a large positive number, which moves probability $F\left(\frac{\lambda_{its}-\alpha X'_{it}}{\sigma}\right)$ towards zero. Similarly, if previously uninsured individuals are more likely to keep being uninsured, Υ will be close to zero, which forces $F\left(\frac{\lambda_{its}-\alpha X'_{it}}{\sigma}\right)$ to go to one and, thus the insurance probability to zero. When an individual is affected by inertia, Υ is one, which leaves the insurance probability unchanged.

It yields a likelihood function:

$$L = \prod_i \prod_t \left(\overbrace{1 - \sum_{s=1}^{12} \phi_{its} F\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right)^{\Upsilon_{its}}}^{\text{if insured}} \right)^{y_{it}} \cdot \left(\overbrace{\sum_{s=1}^{12} \phi_{its} F\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right)^{\Upsilon_{its}}}^{\text{if uninsured}} \right)^{1-y_{it}} \tag{13}$$

A modeling and estimation approach described in this section has a number of advantages. Firstly, it is computationally attractive since to search for parameters which maximize the likelihood function, it is not required to recompute the model with a computationally intensive dynamic programming. Instead, pre-estimated thresholds λ_{its} are sufficient to estimate parameters of a likelihood function and allow for rich model heterogeneity. Secondly, the likelihood function is smooth and has an analytical gradient, which makes it computationally attractive to optimize using fast gradient-based non-linear optimizers. Furthermore, it does not require simulation methods, which are prone to the simulations bias (Train, 2009).[43]

4.4 Parameter Estimates and Model Fit

The model outlined in the previous section has 13 parameters of a risk preferences distribution, two inertia parameters, and 88 type distribution parameters. I estimate a model using maximum likelihood. I obtain standard errors of the parameters using bootstrap with 100 draws with replacement. Appendix B provides more details of the estimation of parameters and standard errors.

[43]Note that although the likelihood function treats the insurance decisions i, t as independent, the interdependence is introduced indirectly through the estimation of thresholds.

Table 2: Parameters of a Risk Preference Distribution and Inertia

		Coefficients	Std. Errors
α:	Constant	50.535	(0.11)
α:	Age $(30; 40]$	-2.581	(0.033)
α:	Age $(40; 50]$	-2.661	(0.346)
α:	Age > 50	-1.5	(0.555)
α:	Gender	0.234	(0.337)
α:	Family	0.642	(0.529)
α:	Higher Education	3.99	(0.831)
α:	Has Children	1.308	(0.218)
α:	Income $(25\%; 50\%]$	-56.148	(0.241)
α:	Income $(50\%; 75\%]$	-68.218	(0.372)
α:	Income $> 75\%$	-35.370	(0.652)
σ:	Std. Deviation	113.757	(0.169)
γ_1	Inertia	178.251	(0.2)
γ_0		0.006	(< 0.001)

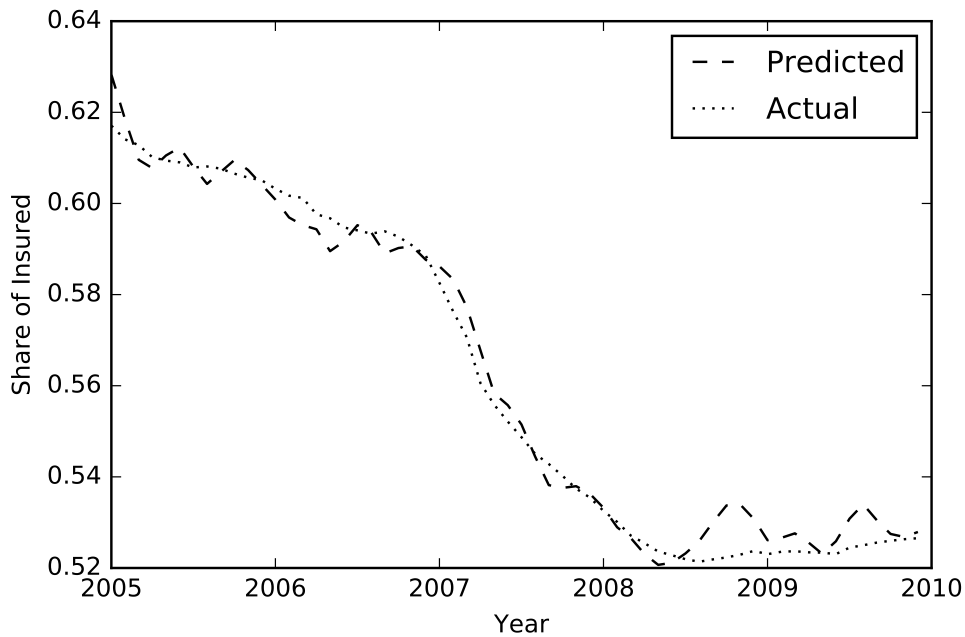
Notes: The Table presents parameter estimates of a risk preference distribution and inertia together with bootstrapped standard errors in the brackets in the corresponding column. An income variable is binned into groups according to the percentiles of the distribution. For example, a variable Income $(50\%; 75\%]$ denotes if an individual has an income within 50% - 75% percentiles of a distribution.

Table 2 presents risk preference and inertia parameters. The Table shows that older and higher income individuals tend to be less risk-averse. Being a female, married, with higher education and having children is associated with higher risk aversion. It implies that those characteristics increase the probability of buying insurance conditionally on unemployment risks and private information. The model displays considerable unobserved risk preference heterogeneity implied by a fairly large standard deviation in a risk preference distribution. I do not provide an extensive discussion of the model parameters since their main use is to recover demand, willingness-to-pay and cost functions for the welfare analysis.

The model also shows an important role of inertia implied by the corresponding parameters that take a value of 178.251 for previously insured and 0.006 for previously uninsured. To put this into perspective, an individual who has a probability of buying insurance 0.9 in the absence of inertia has a probability 0.999 conditionally on being insured before and 0.001 if uninsured before upon adjusting for inertia.

In addition to a risk preference distribution, the model generates 88 parameters of a type distribution. Table 7 with parameters and standard errors is included in Appendix D. Table 8 summarizes the information from type parameters and present type probabilities. The model results suggest that 73% individuals have information only about one period ahead while 7% can perfectly foresee employment outcomes for two periods in the future. In line with the time-selection evidence, a considerable share of individuals (15%) have private information about twelve periods in the future, which allows them to perfectly time enrollment. The remaining types are uncommon (around 5% in total). Such a sharp model prediction of a type distribution is to a large extent driven by a limited number of variables included in vector Z for computational reasons. It implies that allowing for richer heterogeneity by including more relevant characteristics will most likely produce higher probabilities for uncommon types. However, the distribution of probabilities is still in line with the priors based on the anecdotal evidence.

Figure 10: Model Fit - Demand



Notes: The Figure demonstrates actual (dotted) and predicted demand (dashed) demand functions for 2005 - 2009 on the external sample. The y-axis represents a share of insured individuals. External sample means that the estimation sample was initially split into two equally sized samples. The estimation is conducted on the first sample and the fit of the model is illustrated on the second to investigate the performance of the model on the part of the sample not used in the estimation.

Finally, Figure 10 shows the model fit on the external sample which was not used in estimation.[44] The model predicts insurance patterns that closely match actual evidence. In-sample fit using the estimation sample is presented in Appendix D.

5 Welfare

This section describes how the estimates of the model are used to compare various regulations in UI. Although a mandate is one of the most widely discussed regulations in insurance markets and can be viewed as a policy that eliminates adverse selection, it also imposes a burden on those who prefer being uninsured. Therefore, alternative contracts, which also restrict the scopes of private information but impose milder choice restrictions, might be preferred to traditional pricing mechanisms and mandates. While there are many potential counterfactual contracts, I focus on two alternatives that target specific features of private information. Firstly, I consider a contract with fixed costs of six times monthly premiums to be paid when entering the insurance pool.[45] It should discourage time-selection by creating a value of long-term fund enrollment.

Secondly, I consider an often called "open enrollment period" contract that allows entering a fund only at the specific month and has the prespecified duration. I look at 18 and 24 months contracts. I do not consider a 12 months contract, for example, because estimates suggest that some individuals might have private information up to 12 months in the future. As a result, this contract does not leave enough uncertainty and should be avoided. An open enrollment contract is aimed at directly eliminating time-selection. Welfare analysis is based on the pooled sample of individuals over the years 2005-2009 (60 months).

5.1 Measuring Welfare

Welfare analysis requires obtaining a number of components using estimated parameters to construct relevant welfare-metrics. There are two dimensions in which various regulations have an impact: consumer welfare and government budget costs.

To understand the effect on consumers, one needs to recover willingness to pay for a particular insurance contract, which is the maximum price that would be paid. Consumer surplus (CS) can then be measured as a difference between WTP and actual price. It determines demand for insurance since it should be purchased only if WTP is larger than price. Such a relationship might not be true in the presence of inertia. In this case, insurance premiums might still be paid even

[44]Before estimation, I split the sample randomly into two equal parts. The estimation procedure is conducted using the first half of the sample. I simulate the outcomes for the second sample, which are presented in the Figure. Since the whole sample is sufficiently large, there is no need to perform cross-validation.

[45]I study similar 3 and 9 months contracts that produce similar results.

if insurance is valued less than it costs because of choice persistence.[46] Although I attempt to recover inertia in the estimation, I do not take it into account in the welfare analysis. The reason is that the focus of the paper is on how contract design can generate welfare gains by restricting risk-based selection and preserving insurance choices. In addition, it is theoretically unclear how welfare analysis should be conducted when comparing contracts that clearly prone to inertia (current and entry costs contracts) with alternatives that presumably should not exhibit such properties because of considerably less frequent choices (open enrollment contracts). Finally, to which extent the government should internalize welfare costs of sub-optimal choices is a controversial question.

The effect on a government budget comes from two main components: demand and total cost functions. In summary, the essence of the welfare analysis in this set-up involves understanding how various changes affect insurance take-up, consumer surplus, and government costs. Before defining how exactly welfare conclusions are obtained, I formally define how I construct these components using the model and parameter estimates.

Recall that the voluntary part of UI in Sweden has two different prices: for employed and unemployed individuals. Since most price variation is observed for premiums for employed, the former price is a more important strategic variable, to which I refer as g. Therefore, I choose it to be varied in the counterfactual analysis and keep the price for unemployed being actual price. Note that the components necessary for welfare analysis are contract/regulation-specific and should be separately obtained for each considered policy k. Also, to even up the comparison of voluntary contracts and mandates, I consider voluntary contracts in the absence of basic insurance since it would be unavailable under the mandatory system.[47] It implies that all the computed objects required for welfare analysis correspond to the systems with no basic insurance.

Since a key sufficient statistics in the model is risk preference thresholds described in the previous section, all counterfactual price or policy changes require reestimating those thresholds, which is the most computationally intensive part of the model. To be more precise, for each counterfactual policy I solve the model to obtain an array of thresholds for each individual i at each time t and policy k on a grid of prices $g \in [\underline{g}; \bar{g}]$. The computational procedure described in the previous section does this also for each unknown type $s \in \{1, ..., 12\}$. It means that the only object obtained from model parameters needed for recovering counterfactual thresholds are types. To overcome a need to carry out this exercise twelve times for each type, I take a random draw of types using probabilities recovered from the model and summarized in Table 4.

[46] Similarly, insurance might not be bought even if it is valued more than it costs.

[47] The basic insurance system is a mandatory system and the introduction of an alternative universal mandate will automatically remove this basic coverage.

Using the same procedure as before, I compute an array of risk preference thresholds $\lambda_{itk}(g)$ for $g \in [\underline{g}; \bar{g}]$.

To calculate an expected WTP for each $\{i, t\}$, I use the following approach. A threshold recovery procedure allows obtaining maximum risk preference values (λ) at which insurance would be bought under each policy k and price g for each individual and time period. Since the indifference level function $\lambda_{itk}(g)$ must be smooth and monotonically increasing in price g, it can be inverted to obtain $\hat{g}_{itk}(\hat{\lambda})$, which would represent a maximum price that an individual with risk preferences $\hat{\lambda}$ would be willing to pay. Therefore, I can calculate expected WTP by integrating over risk preferences:[48]

$$E[WTP_{itk}] = \int_{\hat{\lambda}} \hat{g}_{itk}(\rho) dF(\rho; \alpha X'_{it}, \sigma) \tag{14}$$

where $F(\rho; \alpha X'_{it}, \sigma)$ is an individual-specific risk preference normal CDF that depends on recovered parameters α and σ, and individuals-specific vector of characteristics X_{it}.

The intuition of this formula is that an expected individual willingness to pay is a weighted average of WTPs resulted from all potential risk preference values weighted by probabilities of having each of those values. Using identical logic, one could obtain consumer surplus for each $\{i, t\}$ as follows:

$$CS_{itk}(g) = \begin{cases} \int_\rho \left[\overbrace{(\hat{g}_{itk}(\rho) - g) \cdot \mathbb{1}[\hat{g}_{itk}(\rho) - g > 0]}^{\text{if buys insurance}} \right] dF(\rho; \alpha \cdot X'_{it}, \sigma), & \text{if voluntary system} \\ \int_\rho \left[\overbrace{(\hat{g}_{itk}(\rho) - g)}^{\text{always buys insurance}} \right] dF(\rho; \alpha \cdot X'_{it}, \sigma), & \text{if mandatory system} \end{cases} \tag{15}$$

To recover expected costs, I start by using detailed unemployment data to predict probabilities of being unemployed for all individuals i at all periods t in the sample as a function of labor market characteristics denoted as ζ_{it}. The costs of covering (H_{it}) in the case of unemployment are determined by observed income, cap and a replacement rate. Expected costs of covering individual $\{i, t\}$ are:

[48] I use 100 knots to obtain the integral numerically. Instead of integrating from $-\infty$ to ∞, for each case I find risk preferences that correspond to 0.1% an 99.9% percentiles. Then I construct equally spaces bins and integrate within this interval with 100 knots after reweighing bin probabilities to ensure that they sum up to 1. Since a computational procedure allows obtaining $\lambda_{itk}(g)$ on a grid of values g, I use linear interpolation to fill the values between grid points in the integration.

$$TC_{itk}(g) = \begin{cases} \int_\rho \left[\overbrace{\left(\zeta_{it} \cdot \left(H_{it} - \underline{g}\right) - (1-\zeta_{it}) \cdot g\right) \cdot \mathbb{1}\left[\hat{g}_{itk}(\rho) - g > 0\right]}^{\text{if buys insurance}} \right] dF(\rho; \alpha \cdot X'_{it}, \sigma), & \\ & \text{if voluntary system} \\ \overbrace{\left(\zeta_{it} \cdot \left(H_{it} - \underline{g}\right) - (1-\zeta_{it}) \cdot g\right)}^{\text{always buys insurance}} & \text{if mandatory system} \end{cases} \tag{16}$$

I evaluate the welfare by comparing systems are contracts under various government expenditure levels in terms of total generated consumer surplus. Recall that the model allows obtaining total consumer surplus $CS_k(g)$ and total government costs $TC_k(g)$ under all systems k and price g defined in (15) and (16), correspondingly. It implies that those functions can be combined into the correspondence:

$$CS_k(g) \widehat{=} TC_k(g) \tag{17}$$

Equation (17) is a correspondence since it is not guaranteed that each price gives a unique pair of total costs and consumer surplus.[49] As a result, it is possible that there is a set of prices that yield the same value of budget costs χ and consumer surplus levels. At the same time, it is possible that there are no prices that allow sustaining a given budget level χ. For example, the government might not be able to achieve profit from a voluntary system if it requires a considerable rise in prices since it would force all individuals out of the insurance pool. It would imply that for this budget balance χ the set of prices is empty.

I define a set of prices that yields total costs χ under system k as $\varepsilon_k(\chi)$. The system k is said to be welfare-dominant with respect to a system m under a budget balance χ if under all prices $g \in \varepsilon_k(\chi)$ and $q \in \varepsilon_m(\chi)$ a system k always leads to higher consumer surplus than under m. More formally:

Definition 1. *A system k welfare-dominates a system m under a budget balance χ if* $\forall g \in \varepsilon_k(\chi)$ *and* $\forall q \in \varepsilon_m(\chi)$*:*

$$CS_k(g) > CS_m(q)$$

This definition embraces a number of desired properties of a welfare criterion for this case. Firstly, it takes into account that there might be a number of prices that require the same level

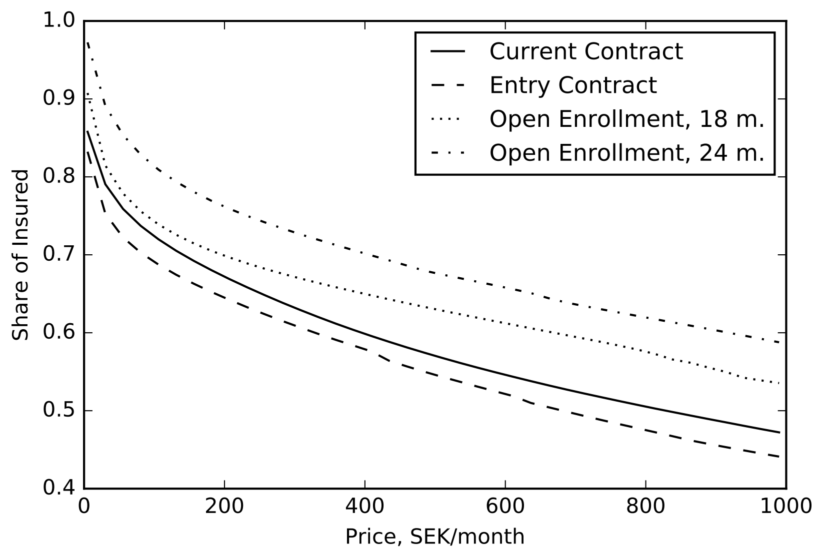
[49]The reason is that a change in prices affects both probabilities of insurance, which also translates into changes in a risk composition among insured individuals, and government revenues through the sum of collected premiums.

of budget costs for the government even within the same system. At the same time, it also takes into account that some subsidy levels are unattainable for some systems. It implies that systems can be directly compared only under reachable budget balances. It is especially important when analyzing mandates since these policies should theoretically be able to support a wider range of costs because of restrictions on individual responses.[50]

5.2 The Welfare Consequences of Alternative UI Designs

As discussed in the previous section, changes in the structure of the contract and prices affect welfare through a number of channels. Firstly, individuals react to those changes by enrolling or leaving an insurance pool. Figure 11 demonstrates counterfactual demand functions under various considered policies.

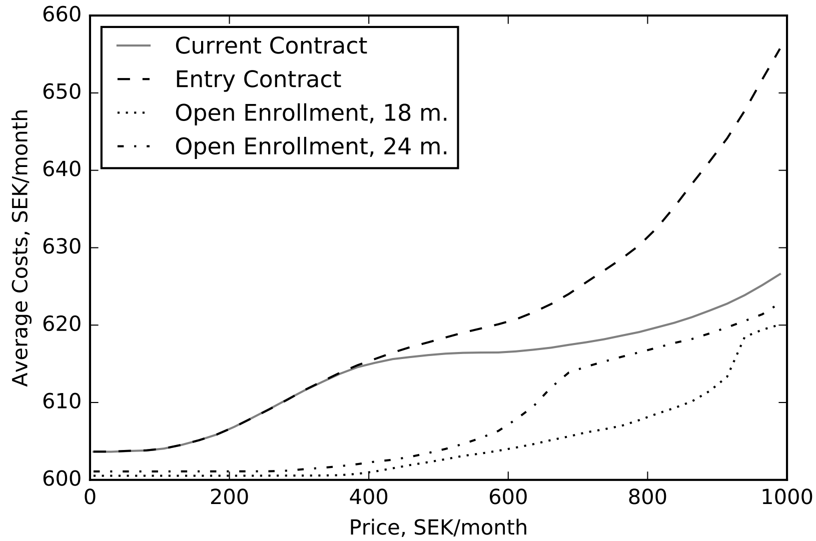
Figure 11: Counterfactual Policies Demand



Notes: The Figure demonstrates the demand function of a current system, a system with an entry costs contract and open enrollment contracts with 18 and 24 months durations.

[50]This statement might not be true if there is a large moral hazard response to mandates.

Figure 12: Average Cost Functions



Notes: The Figure demonstrates average costs of insuring individuals under voluntary systems. The curves are obtained by dividing each value in a cost function by an expected number of insured individuals.

Figure 8 suggests that an entry costs demand function is downward-shifted in comparison to the current contract. Demand functions for open enrollment contracts are less steep on average and are shifted upwards compared to other designs. The demand for the 24 months contract is slightly upward-shifted compared to the 18 months contract since it involves more uncertainty and hence is less attractive.

The second policy-relevant dimension is budget costs. Figure 12 plots average cost functions. Presented cost functions show upward slopes in prices. It corresponds to downward-sloping cost curves in a number of insured individuals, which signals the presence of adverse selection (Einav, Finkelstein, & Cullen, 2010). The average cost curves for open enrollment contracts are less steep and shifted down compared to other curves. It signals that these contracts allow restricting selection compared to a current system or an entry cost contract. An interesting feature of the entry cost contract is that it actually results in more selection. The intuition is that entry fees keep high-risk individuals who expect to benefits from insurance, whereas it does not provide benefits from holding low-risk individuals in the pool and discourages new enrollments.

Before looking at more formal welfare analysis, the evidence presented in Figures 11 and 12 suggests a number of important insights regarding the welfare consequences of the contracts under consideration. Open enrollment contracts attract more individuals and, at the same time, cost less per individual. In contrast, entry costs contracts attract fewer individuals but cost weakly more per individual. It suggests potentially large welfare gains of open enrollment contracts and welfare losses associated with entry costs contracts in comparison to a current system.

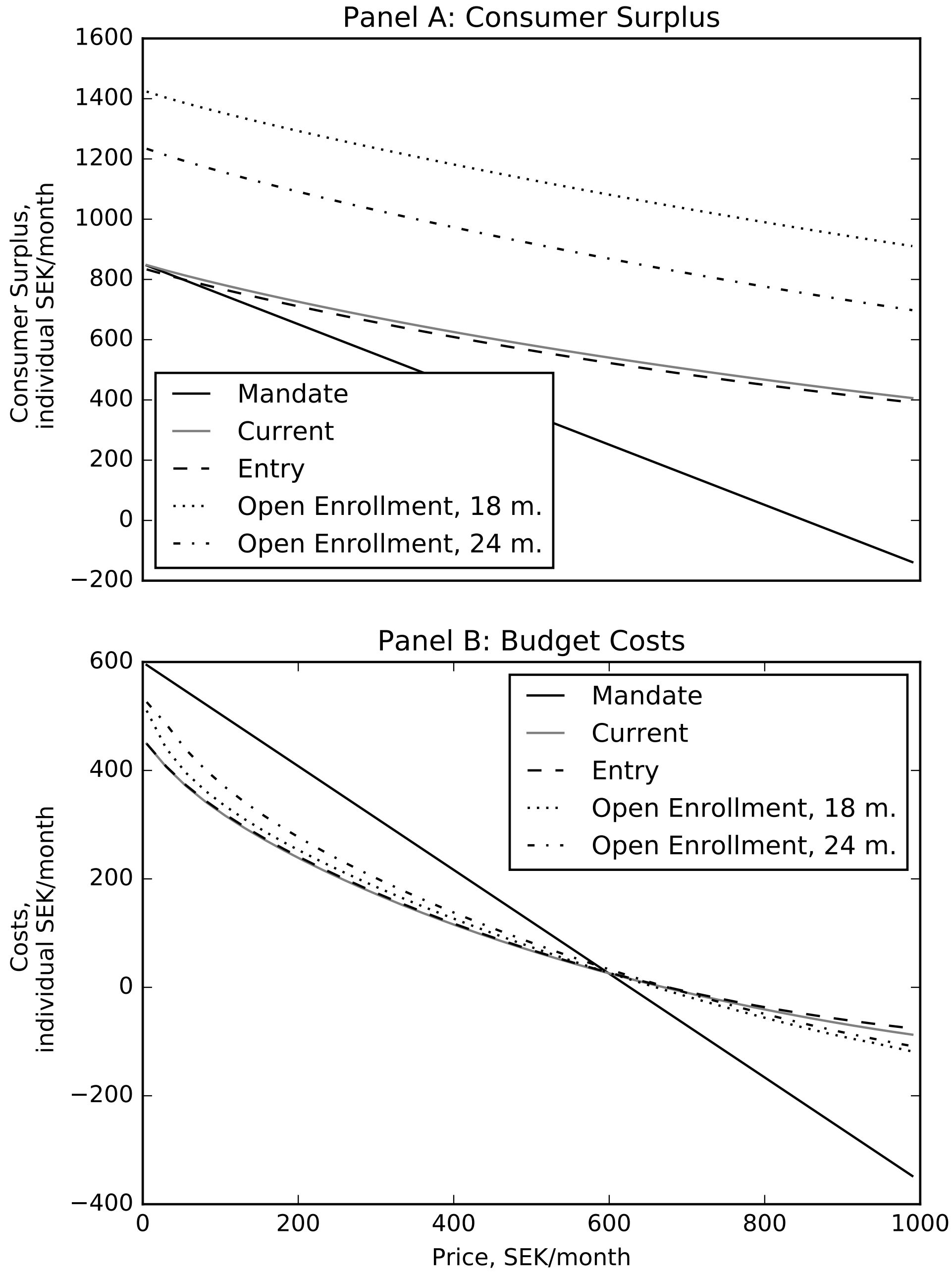
Figure 13 presents consumer surplus and budget costs under various prices. Panel A shows that the open enrollment contract with 18 months duration generates higher consumer surplus under all considered price levels due to its features described above. Other voluntary contracts are similar in terms of consumer surplus. Under a mandatory system, price increases have more a pronounced negative impact on consumer surplus since individuals are not allowed to respond to a price increase by leaving the insurance pool. Therefore, a mandatory system is most detrimental for consumer surplus.

However, a mandatory system is capable of considerably reducing budget costs since individuals are locked in and cannot unenroll as demonstrated in Panel B. All voluntary contracts have similar performance in terms of cost reduction. For high prices, current and entry costs systems allow reaching lower expenditure levels compared to the open enrollment contracts.

To conclude whether a contract structure welfare dominates a competing design at some government costs, one should compare the resulted consumer surpluses at various budget costs from Figure 13. It also takes into account the fact that some systems might not allow sustaining some levels of government expenditures at least within a considered interval of prices. It implies that the correspondences from (17) might have different support for various systems in terms of costs.

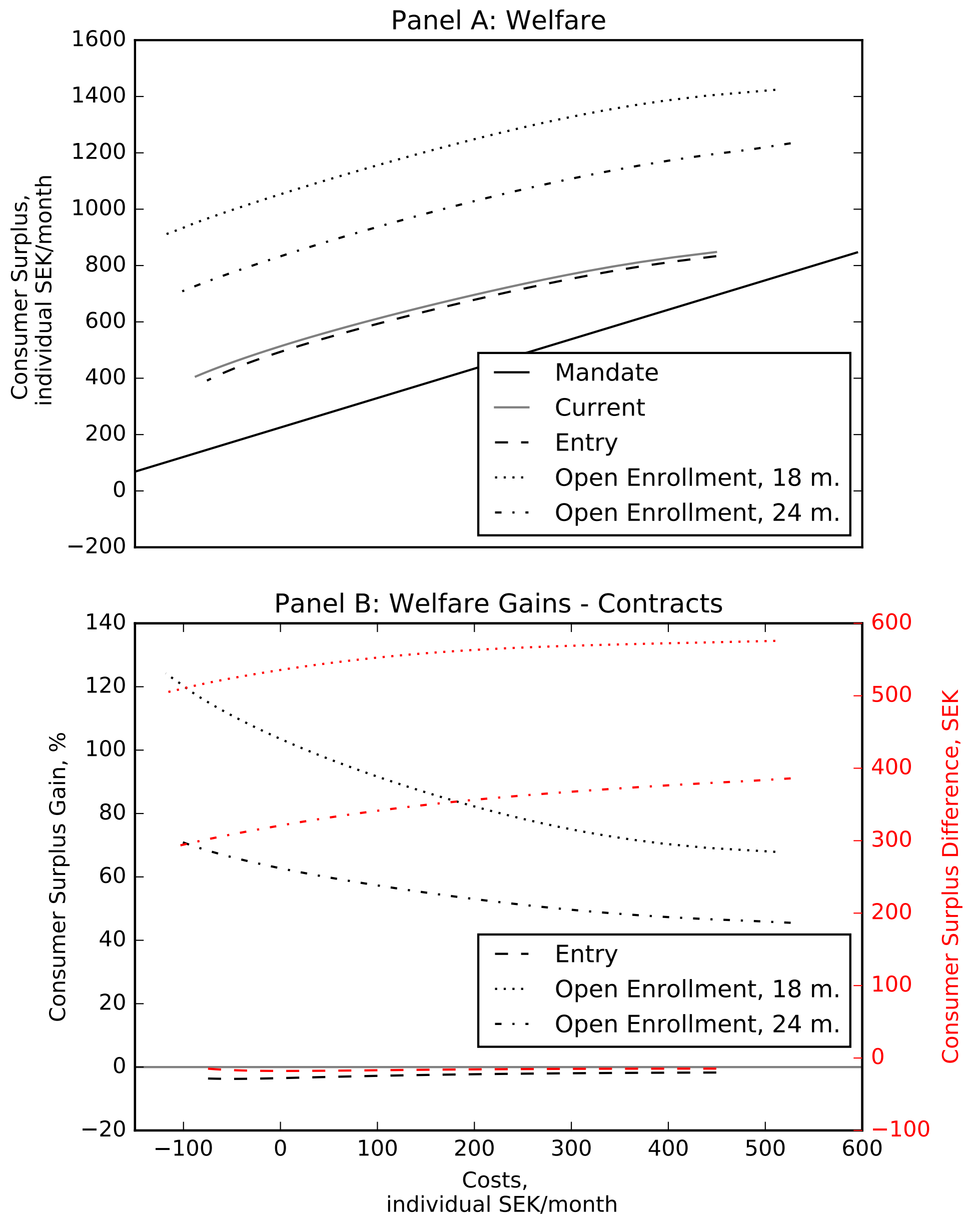
Figure 14 summarizes the welfare analysis. Panel A demonstrates relationships between government costs and generated consumer surplus under considered policies. In other words, Panel A represents the y-axis of Panel A plotted against the y-axis of Panel B from Figure 13. The main point of the Figure is to illustrate which policies lead to higher consumer surplus while requiring the same subsidy levels. This approach allows being agnostic about optimal pricing. If the system is located above on the y-axis, it should be preferred since it yields higher consumer surplus at the same cost level. Although a part of the previous section was devoted to emphasizing and clarifying the fact that it is theoretically possible to have multiple consumer surpluses, it appears not to be the case in practice.

Figure 13: Effect of Premiums on Consumer Surplus and Budget Costs



Notes: Panel A plots a monthly price against consumer surplus. Panel B presents the relationship between premiums and resulted budget costs. I divide total consumer surplus and total costs by a number of "active" individual-months observations for expositional purposes instead of presenting the sums over individuals and observed months. Note that in contrast to average cost curves, this normalization is constant and does not vary with a number of insured individuals for all prices.

Figure 14: Welfare



Notes: The Figure demonstrates the main results of welfare analysis. Panel A plots government costs per individual-month for 2005-2008 against the resulted consumer surplus. I divide total costs and consumer surplus by a number of "active" individual-months observations for expositional purposes. Panel B presents the same evidence as on Panel A but in terms of percentage welfare gains compared to a current system at the corresponding budget level and in terms of a difference on the right y-axis colored in red. The interpretation is that a system dominates another one under some government cost level if it lies above on both Panels. It implies that it results in higher consumer surplus at the same cost level.

Panel A suggests that an entry costs contract is very close to a current system but cause small welfare losses. The Figure indicates that mandates generate sizable welfare losses. Finally, the results suggest that open enrollment contracts would be the best option for nearly all achievable levels of expenditures.

To put it into perspective, a mandate would lead to 48.8% or 243 SEK/month per individual consumer surplus loss compared to a current system on average over the considered price levels. The reason is that, as demonstrated in Panel B of Figure 13, mandates are effective to reduce costs only at high prices. At lower prices, all voluntary systems are less expensive. At the same time, a mandate is the worst system in terms of the effect on the consumer surplus. Therefore, the results suggest that it is the least favorable design of UI among considered options.

Panel B similarly compares various voluntary designs. It suggests that within the considered range of the government costs, entry contract results in 2.9% lower consumer surplus on average along the line. The intuition is that entry costs contract is worse for consumers since it is more expensive at the same premium levels. At the same time, average cost curves show that it leads to even more selection, especially at high prices.

Finally, the results suggest that open enrollment contracts would welfare dominate all other options. The average gains amount to 95% (545 SEK) for 18 months and 58% (338 SEK) for 24 months contracts compared to the current system, correspondingly. There are two features of open enrollment contracts that make them an attractive option from the welfare point of view. First, this contract structure virtually removes a time-selection part of risk-based selection. Secondly, the estimates of WTP shows that individuals often value this contract more than a current one primarily because of the absence of the 12 months pre-eligibility period.

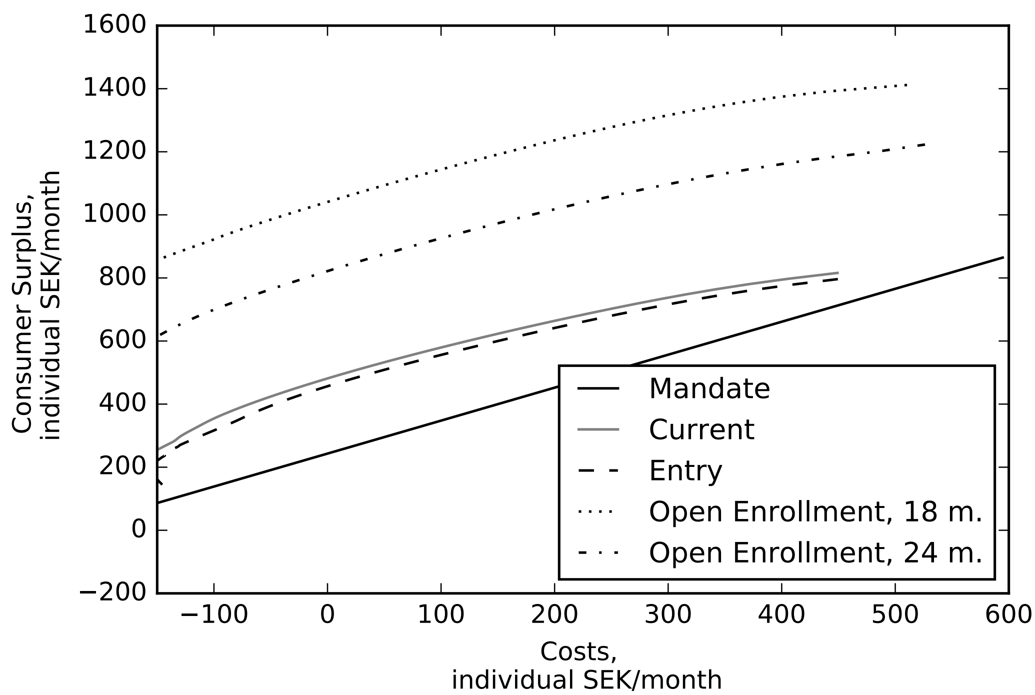
To sum up, the results of this section show that in line with the concerns regarding the effect of mandates, it is predicted to be the least desirable policies among the considered options. Instead, appropriately chosen alternative contract designs tailored to remove harmful selection without considerable distortion to individual choices are predicted to generate sizable welfare gains.

5.3 Robustness and Discussion

The model presented in this paper requires many assumptions that might raise concerns regarding the validity and sensitivity of the welfare analysis. Therefore, it is important to discuss the role of these assumptions.

The first point, which is, however, unrelated to the model and analysis directly, is a sample selection. The insurance data lack information for those individuals who have not received insurance benefits. It is not a random sample despite similarities with a general population in terms of observables. Most likely, a sample contains a relatively risky part of the population. At the same time, the share of insured individuals is smaller in the sample compared to a full population by roughly 10%. It implies that a missing population is risk-averse, has less information about employment perspectives (types) or displays more inertia. To examine the importance of the sample selection for the welfare analysis, I use the model parameters to simulate the choices of individuals whose actual choices are not observed. Figure 15 replicates the results of the welfare analysis from Panel A in Figure 14. In contrast to the results in Figure 14, Figure 15 pretends that individual preferences and type parameters are not affected by sample selection and shows how welfare conclusions after including missing individuals in the sample.

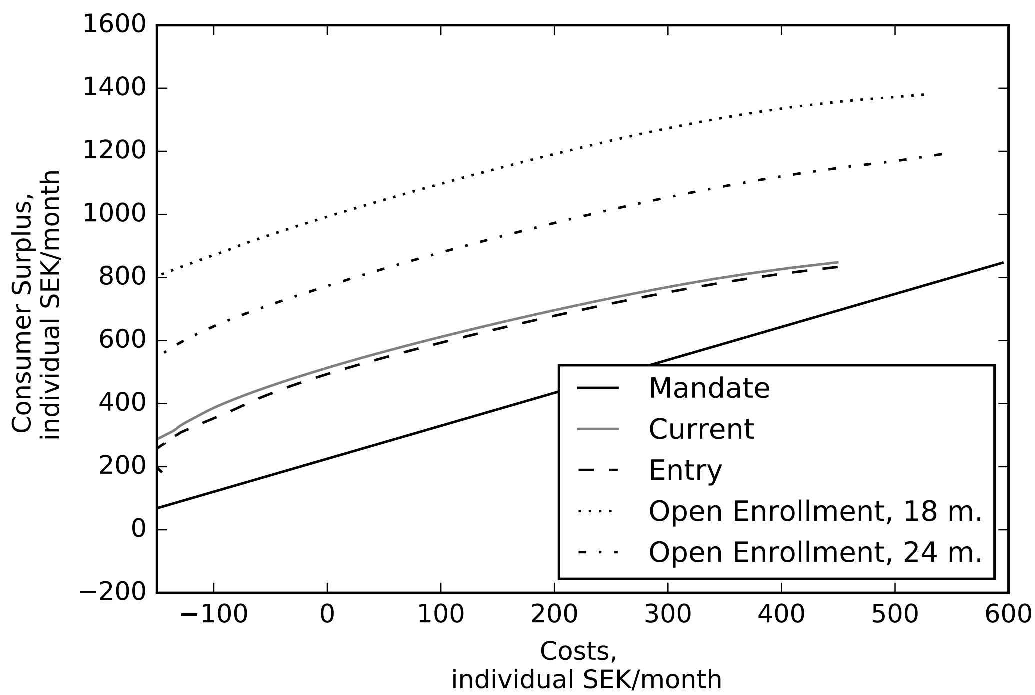
Figure 15: Robustness of Welfare Analysis - Sample Selection



The Figure shows that although the levels of curves on the y-axis change, they are located relative to each other similarly to the main results of the welfare analysis. It implies that if estimated model parameters are not considerably affected by the sample selection, the welfare conclusions are robust to omitting a part of the sample. This approach does not take into account the fact that a missing population might have different preferences and information structure. However, at least within the scopes of estimated parameters, the welfare conclusions are robust.

Another important feature of the model is the way temporary contracts are treated. A type distribution assumed in this paper can be viewed as being truncated at 12 meaning that individuals should not have information about more than 12 months ahead. Although institutional details suggest that it is most likely must be true, there are concerns associated with the presence of temporary contracts in which case individuals might have more than 12 months knowledge. The data contains the variable that denotes the category of unemployment including if an individual is currently on the limited-term employment. However, it is difficult to determine the duration of the contract from this information. Therefore, to assess a potential effect on the welfare conclusions, I exclude all individuals who report being on the temporary contract. Figure 16 illustrates the effect of excluding these individuals on welfare results.

Figure 16: Robustness of Welfare Analysis - Temporary Contracts



The Figure shows similar results to Figure 14. Although curves shift, their ordinal positions with respect to each other remain the same.

In the model, I assume that a planning horizon T is limited to 18 periods. Experimenting with different options around the chosen value does not affect results considerably. A number of employment sequences grows exponentially with T. To make the computation feasible and not to solve the dynamic programming for each sequence, I limit the attention only to those sequences, which have non-zero probabilities.[51] Therefore, a choice of T remains crucial for computational costs. Appendix B discusses these computational details. To illustrate the effect of this choice on the results, I present a distributions of risk preference indifference points in Figures 22 and 23

[51]Theoretically, all sequences have non-zero probability but practically they do because of the computer precision limit.

in the Appendix C using $T = 19$ and $T = 21$ in comparison to $T = 18$, which is used to obtain main results in this paper. The results suggest that a choice of T proportionally affect how spread the distribution of thresholds is and hence should not be critical for the welfare analysis.

The counterfactual analysis does not take moral hazard into account. The main concern associated with that would be that counterfactual policies not only change insurance decisions but also risks. To minimize concerns associated with this model abstraction, I consider modest price changes that should not create large labor market responses.

Finally, a bigger picture concern is the validity of such a neoclassical-type model that to a large extent disregards more sophisticated behavioral mechanisms such as the role of family in income insurance or borrowing. The data show that individuals react to incentives as expected (e.g. higher prices, less generous insurance, and lower risks reduce the demand for insurance). All other potential behavioral components are falling under the risk preferences and an inertia parameter. An implicit assumption in the dynamic model is the absence of a discount factor since it is not identified. The assumption does not seem to be extreme since I model monthly dynamics in which case future-discounting should not play an important role. It also should not have any effect on the observed bunching patterns since even sizable variation in time preferences will not affect the bunching incentives in the presence of information about the future.

6 Conclusions

This paper attempts to provide one of the first comprehensive analyses of the optimal regulations in unemployment insurance. Existing literature documents a positive correlation between insurance and unemployment risks often attributed to risk-based selection. I augment this evidence by showing the importance of understanding an interplay among risks, private information structure and preferences to analyze the effect of alternative counterfactual policies. I conclude that potential regulations are not limited to mandates and pricing policies but also should include contract design regulations. These regulations either encourage long-term enrollment or mechanically restrict time-based selection.

One of the key messages of this paper is a difficulty to provide welfare suggestions using just correlation evidence that often arise from multiple dimensions of individual heterogeneity (Finkelstein & McGarry, 2006; Einav, Finkelstein, & Ryan, 2013). This paper develops a model and a computationally attractive estimation approach that attempts to recover some of those dimensions of heterogeneity. Even taking all the model and parametric assumptions with a grain of salt, this approach allows more comprehensive exploration of the interplay among various forces affecting individual decisions. As a result, it enables recovering welfare-relevant indicators to illustrate the outcomes of alternative policies. Furthermore, it allows widening the spectrum of available policies and considering the contract design as an alternative to widely-discussed pricing regulations and mandates. Moreover, the results suggest that appropriate contract designs would provide relatively large welfare gains.

The results of this paper should not be directly extrapolated outside of the context because of a sample selection and considerable differences among labor markets in Sweden and other countries. However, the analysis provides a number of insights applicable to a broader audience. Firstly, despite a considerable heterogeneity in estimated willingness to pay, individuals do value insurance. It might suggest that individuals in countries with weaker social security and less stable labor markets have even more need for unemployment insurance. At the same time, private markets are unlikely to play this role due to a considerable amount of private information. Therefore, apparently, UI will remain a part of government policies. Secondly, the results imply at the very least an ambiguous impact of mandates that are widely adopted around the world. Even in the absence of a moral hazard response, it is predicted to be an undesirable policy because of the burden imposed on individuals who have low insurance value. Instead, alternative contracts such as restricted enrollment timing seem to provide considerable gains by reducing private information without imposing excess costs on individuals. It raises concerns regarding a nearly universal adoption of mandatory UI, which suggests that the optimal regulation in UI is an open policy-relevant issue for future research.

References

Abbring, J. H., Chiappori, P.-A., & Pinquet, J. (2003). Moral Hazard and Dynamic Insurance Data. *Journal of the European Economic Association*, *1*(4), 767–820.

Abbring, J. H., Heckman, J. J., Chiappori, P.-A., & Pinquet, J. (2003). Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish? *Journal of the European Economic Association*, *1*(2-3), 512–521.

Akerlof, G. (1978). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. In *Uncertainty in Economics* (pp. 235–251).

Apesteguia, J., & Ballester, M. (2018). Monotone Stochastic Choice Models: The Case of Risk and Time Preferences. *Journal of Political Economy*, *126*(1), 74–106.

Aron-Dine, A., Einav, L., Finkelstein, A., & Cullen, M. (2015). Moral Hazard in Health Insurance: Do Dynamic Incentives Matter? *Review of Economics and Statistics*, *97*(4), 725–741.

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

Azevedo, E., & Gottlieb, D. (2017). Perfect Competition in Markets with Adverse Selection. *Econometrica*, *85*(1), 67–105.

Baily, M. N. (1978). Some Aspects of Optimal Unemployment Insurance. *Journal of Public Economics*, *10*(3), 379–402.

Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, *112*(2), 537–579.

Cabral, M. (2016). Claim Timing and Ex-Post Adverse Selection. *Review of Economic Studies*, *84*(1), 1–44.

Card, D., Johnston, A., Leung, P., Mas, A., & Pei, Z. (2015). The Effect of Unemployment Benefits on the Duration of Unemployment Insurance Receipt: New Evidence from a Regression Kink Design in Missouri, 2003–2013. *American Economic Review*, *105*(5), 126–130.

Card, D., & Levine, P. (2000). Extended Benefits and the Duration of UI Spells: Evidence from the New Jersey Extended Benefit Program. *Journal of Public Economics*, *78*, 107–138.

Chetty, R. (2006). A General Formula for the Optimal Level of Social Insurance. *Journal of Public Economics*, *90*(10), 1879–1901.

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

Chiappori, P.-A., & Salanie, B. (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy*, *108*(1), 56–78.

Cutler, D., Finkelstein, A., & McGarry, K. (2008). Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance. *American Economic Review*, *98*(2), 157–162.

DellaVigna, S., Lindner, A., Reizer, B., & Schmieder, J. (2017). Reference-Dependent Job Search: Evidence from Hungary. *Quarterly Journal of Economics*, *132*(4), 1969–2018.

De Meza, D., & Webb, D. (2001). Advantageous Selection in Insurance Markets. *RAND Journal of Economics*, 249–262.

Einav, L., Finkelstein, A., & Cullen, M. (2010). Estimating Welfare in Insurance Markets Using Variation in Prices. *Quarterly Journal of Economics*, *125*(3), 877–921.

Einav, L., Finkelstein, A., & Ryan, S. (2013). Selection on Moral Hazard in Health Insurance. *American Economic Review*, *103*(1), 178–219.

Einav, L., Finkelstein, A., & Schrimpf, P. (2010). Optimal Mandates and the Welfare Costs of Asymmetric Information: Evidence from the UK Annuity Market. *Econometrica*, *78*(3), 1031–1092.

Einav, L., Finkelstein, A., & Schrimpf, P. (2015). The Response of Drug Expenditure to Non-Linear Contract Design: Evidence from Medicare Part D. *Quarterly Journal of Economics*, qjv005.

Einav, L., Finkelstein, A., & Schrimpf, P. (2017). Bunching at the Kink: Implications for Spending Responses to Health Insurance Contracts. *Journal of Public Economics*, *146*, 27–40.

Fang, H., Keane, M. P., & Silverman, D. (2008). Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. *Journal of Political Economy*, *116*(2), 303–350.

Finkelstein, A., & McGarry, K. (2006). Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. *American Economic Review*, 938–958.

Finkelstein, A., & Poterba, J. (2004). Adverse Selection in Insurance Markets: Policyholder Evidence from the UK Annuity Market. *Journal of Political Economy*, *112*(1), 183–208.

Fredriksson, P., & Holmlund, B. (2001). Optimal Unemployment Insurance in Search Equilibrium. *Journal of Labor Economics*, *19*(2), 370–399.

Handel, B. (2013). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review*, *103*(7), 2643–2682.

Handel, B., Hendel, I., & Whinston, M. (2015). Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk. *Econometrica*, *83*(4), 1261–1313.

Handel, B., Kolstad, J. T., & Spinnewijn, J. (2015). Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets.

Hendren, N. (2013). Private Information and Insurance Rejections. *Econometrica*, *81*(5), 1713–1762.

Hendren, N. (2017). Knowledge of Future Job Loss and Implications for Unemployment Insurance. *American Economic Review*, *107*(7), 1778–1823.

Holmlund, B. (1998). Unemployment Insurance in Theory and Practice. *Scandinavian Journal of Economics*, *100*(1), 113–141.

Hopenhayn, H., & Nicolini, J. P. (1997). Optimal Unemployment Insurance. *Journal of Political Economy*, *105*(2), 412–438.

Kolsrud, J., Landais, C., Nilsson, P., & Spinnewijn, J. (2018). The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden. *American Economic Review*, *108*(4-5), 985–1033.

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

Lalive, R., Van Ours, J., & Zweimüller, J. (2006). How Changes in Financial Incentives Affect the Duration of Unemployment. *Review of Economic Studies*, *73*(4), 1009–1038.

Landais, C. (2015). Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design. *American Economic Journal: Economic Policy*, *7*(4), 243–278.

Landais, C., Michaillat, P., & Saez, E. (2018a). A Macroeconomic Approach to Optimal Unemployment Insurance: Applications. *American Economic Journal: Economic Policy*, *10*(2), 182–216.

Landais, C., Michaillat, P., & Saez, E. (2018b). A Macroeconomic Approach to Optimal Unemployment Insurance: Theory. *American Economic Journal: Economic Policy*, *10*(2), 152–81.

Landais, C., Nekoei, A., Nilsson, P., Seim, D., & Spinnewijn, J. (2017). Risk-Based Selection in Unemployment Insurance: Evidence and Implications.

Meyer, B. (1990). Unemployment Insurance and Unemployment Spells. *Econometrica*, *58*(4), 757–782.

Moffitt, R. (1985). Unemployment Insurance and the Distribution of Unemployment Spells. *Journal of Econometrics*, *28*(1), 85–101.

Schmieder, J., Von Wachter, T., & Bender, S. (2012). The Effects of Extended Unemployment Insurance over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years. *Quarterly Journal of Economics*, *127*(2), 701–752.

Shimer, R., & Werning, I. (2008). Liquidity and Insurance for the Unemployed. *American Economic Review*, *98*(5), 1922–42.

Spence, M. (1978). Product Differentiation and Performance in Insurance Markets. *Journal of Public Economics*, *10*(3), 427–447.

Spinnewijn, J. (2015). Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs. *Journal of the European Economic Association*, *13*(1), 130–167.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Appendix

A Displacement Rules

Time-selection is studied in this paper as an important dimension of private information. Therefore, this section discusses the main rules regarding worker displacement initiated by an employer. I do not discuss job separation initiated by an individual since voluntary employment is not covered within UI at least for the first 45 days.

If a firm decides to displace an employee, it has to provide layoff justification. The absence of work is the most common reason for worker displacement. The employer has to prove the lack of work. If more than five but less than 25 workers are to be displaced based on the absence of work, the firm should inform Employment Agency at least 2 months in advance. If 25 - 100 or more than 100 workers are to be displaced, the agency should be informed 4 or 6 months in advance, correspondingly. In this case, the order of displacement is determined by the tenure in the firm.[52] For more details see Landais et al. (2017). Another reason for worker displacement is unsuitability of a worker for the occupied position. The examples are unsatisfactory performance, threats of violence, theft, refusal to work, unlawful absence, etc. When such personal reasons are a ground of displacement, the displacement procedure is conducted according to the conditions and layoff notification rules specified in the law or in the contract. In the case of particularly serious violation of rules, the individual can be displaced immediately despite the layoff conditions.

In the absence of collective agreement, tenure in the firm determines the layoff notification duration:[53]

- *less than 2 years*: 1 month
- *2 - 4 years*: 2 month
- *4 - 6 years*: 3 month
- *6 - 8 years*: 4 month
- *8 - 10 years*: 5 month
- *more than 10 years*: 6 month

Often collective agreements or employment contracts overrule these regulations. Therefore, the information on tenure does not allow to determine actual time-related private information.[54] Special rules apply to individuals who have reached the age of 55 or older and have more than 10 years of continuous employment in the firm.

Another special employment form, which is fairly common is a trial employment contract. Such employment implies that before being granted a permanent contract, an individual has several months of employment with a particularly short layoff notice. The duration of trial

[52]Source: https://www.unionen.se/rad-och-stod/varsel-om-uppsagningar

[53]Source: https://www.unionen.se/rad-och-stod/uppsagningstider-om-din-arbetsgivare-sager-upp-din-anstallning

[54]Collective agreements that specify layoff notification agreements can be found https://www.unionen.se/rad-och-stod/om-kollektivavtal/sok-kollektivavtal

employment periods varies but cannot be more than 6 months. Most often used notification period is one month.[55]

Temporary contracts with predetermined layoff date are also widely used. Although the exact contract termination date is specified in the employment agreement, the contract can be terminated earlier but a specified layoff notice requirement applies. The maximum length of temporary employment shall not exceed 24 months in last five years.[56]

B Estimation Details

The estimation procedure in this paper consists of two steps: computation of risk preference indifference points and estimation of parameters. I firstly compute risk preference thresholds where individuals are indifferent between buying insurance or not. To do that, I solve a dynamic programming problem for each individual i, time t, type s and each potential employment sequence j. A major complication arises from a large number of employment sequences since it amounts to 2^{T-s}, where T is a length of an optimization horizon and s is a number of periods observed in the future. As can be seen, a number of sequences grows exponentially. Therefore, I make two restrictions to keep the estimation feasible.

Firstly, I limit the duration of a planning horizon to 18 periods. It does not fully resolve the issue but linearly reduces computational time and still dramatically decreases the number of sequences. Although the number is still extremely large, a vast majority of sequences have a probability close to zero. Therefore, I calculate probabilities for all potential sequences, which would be impossible without the restrictions on T. I rank the sequences in the descending order of likelihood. Then I select top 750 sequences or up to a point when sequence probabilities sum up to 0.99.

I use the bisection method to compute thresholds where the expected utility of buying insurance equals to expected utility of being uninsured. Although the bisection method is slower than, for example, the Brent method, it is safer for this type of non-monotonic problems. It requires imposing bounds, which I set to very high and very low-risk preference values. This also allows solving the issue with the zero limit of utility differences. More precisely, although the utility difference has the unique value of risk preferences where it equals zero, it might become actual zero at the limit as $\rho \to \infty$ because of numerical constraints.

The part that computes thresholds is written in Python due to requirements of Statistics Sweden, which does not allow using ahead-in-time compiled languages (e.g. C/C++) on their

[55]Source: https://www.unionen.se/rad-och-stod/provanstallning

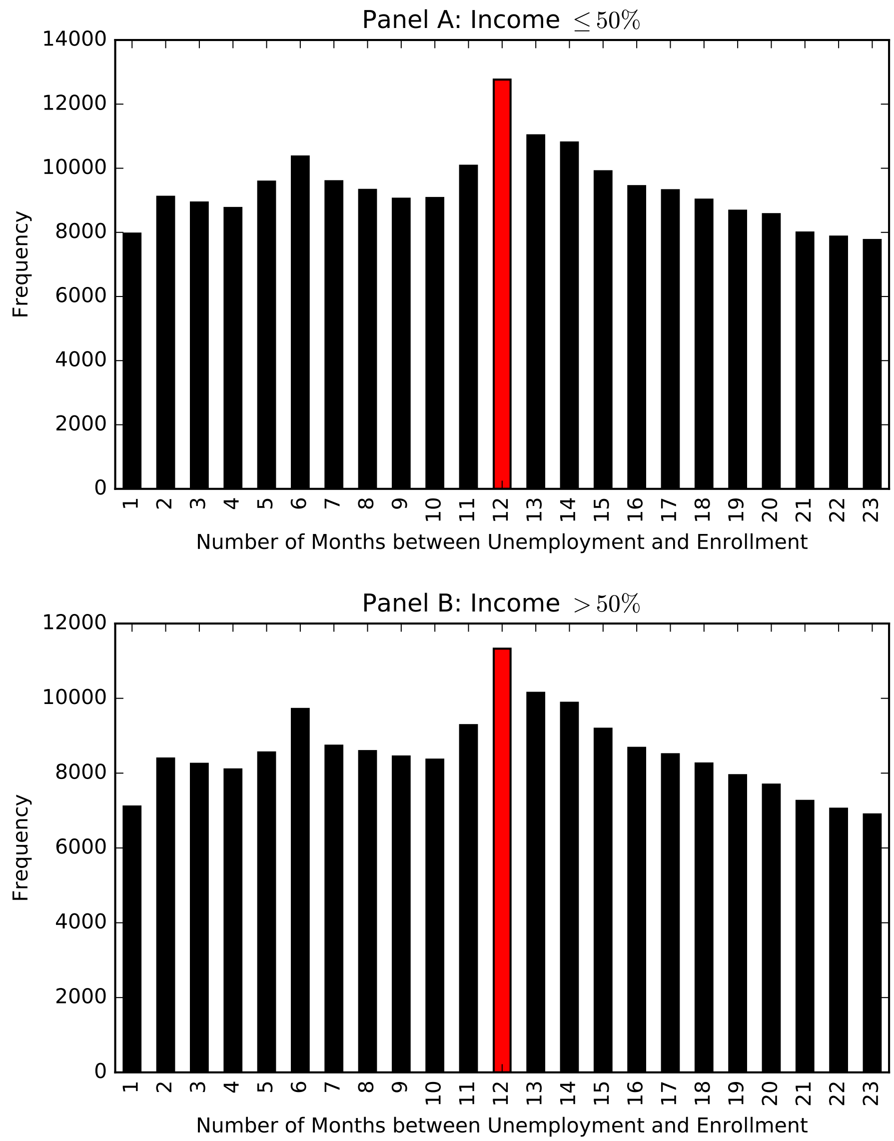
[56]Source: Paragraph 5 of the Law of Employment Protection.

servers with the data. I pre-compile all computationally intensive parts of the code using a just-in-time compiler, which provides significant speed-up. I use 50 cores in the estimation. As a result, computing thresholds for a 5% random sample takes approximately 5 hours. It is, however, much more computationally efficient compared to the estimation of parameters jointly with solving the model, which would require reestimating it at each optimization iteration. The second stage is parameter estimation based on the computed thresholds using a maximum likelihood procedure described in the main text. I use the L-BFGS-B algorithm with bounds on parameters and a user-defined analytical gradient function.

Bootstrap is used to calculate standard errors. I use 100 draws with replacement and estimate the model in parallel on 20 cores. Such a fairly low number of draws is chosen for computational feasibility reasons since it requires around 8 hours for the optimizer to converge.

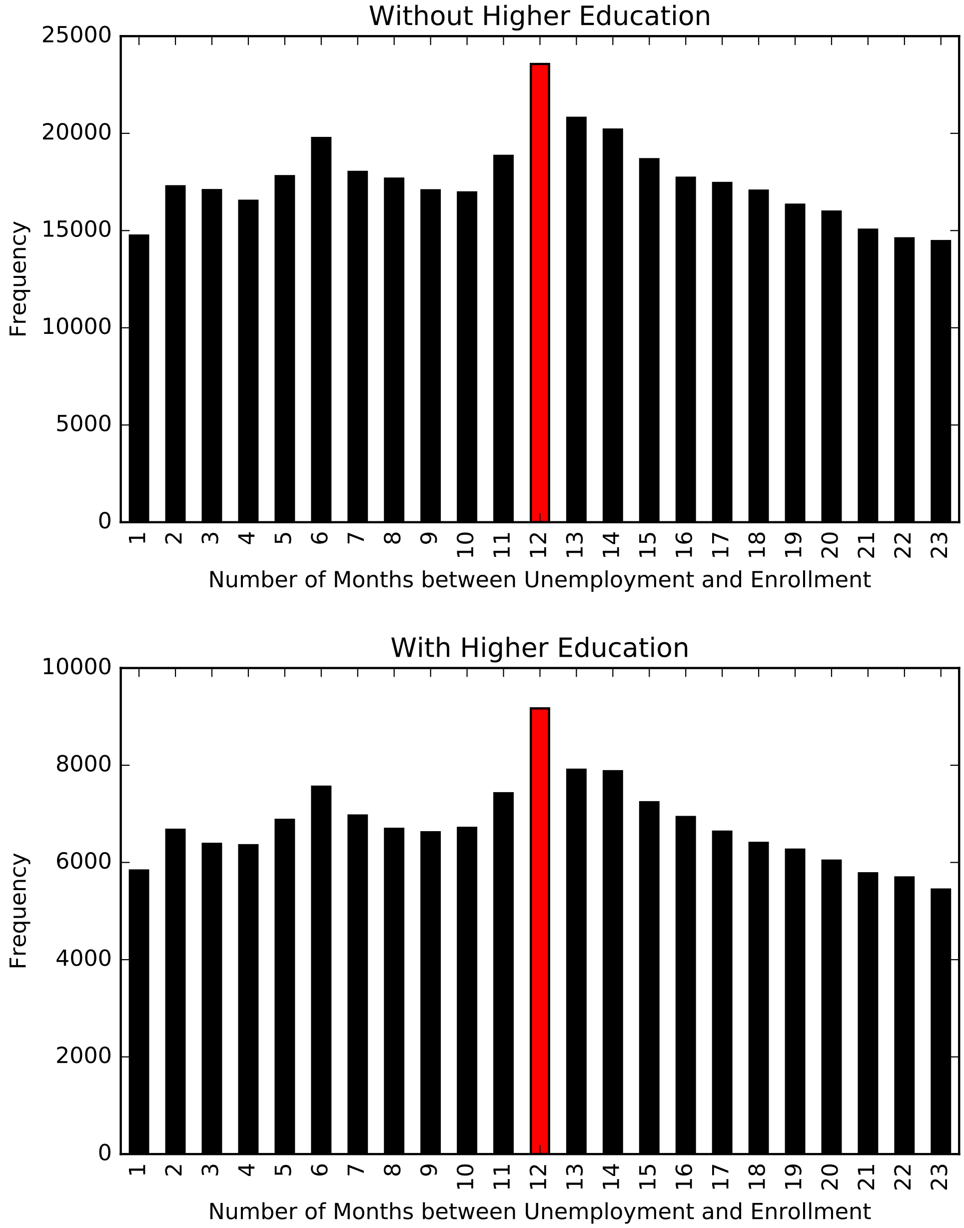
C Supplementary Figures

Figure 17: Bunching Around the Eligibility Requirement By Income



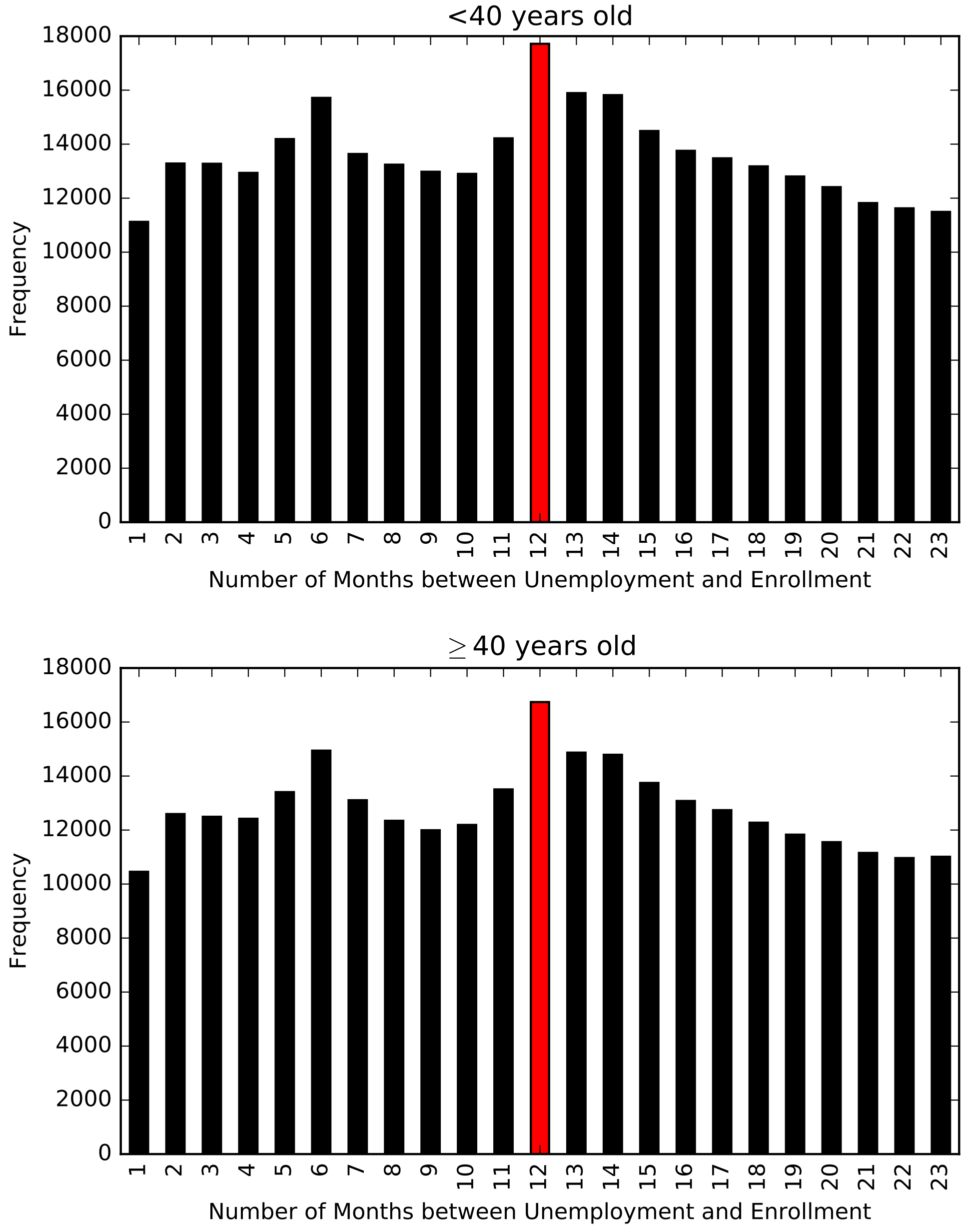
Notes: The Figure presents a discrete histogram of a distribution of a number of enrollment months before the start of unemployment spells similarly to the evidence in the main text but separately for individuals with below the median income (Panel A) and above the median income (Panel B).

Figure 18: Bunching Around the Eligibility Requirement by Education



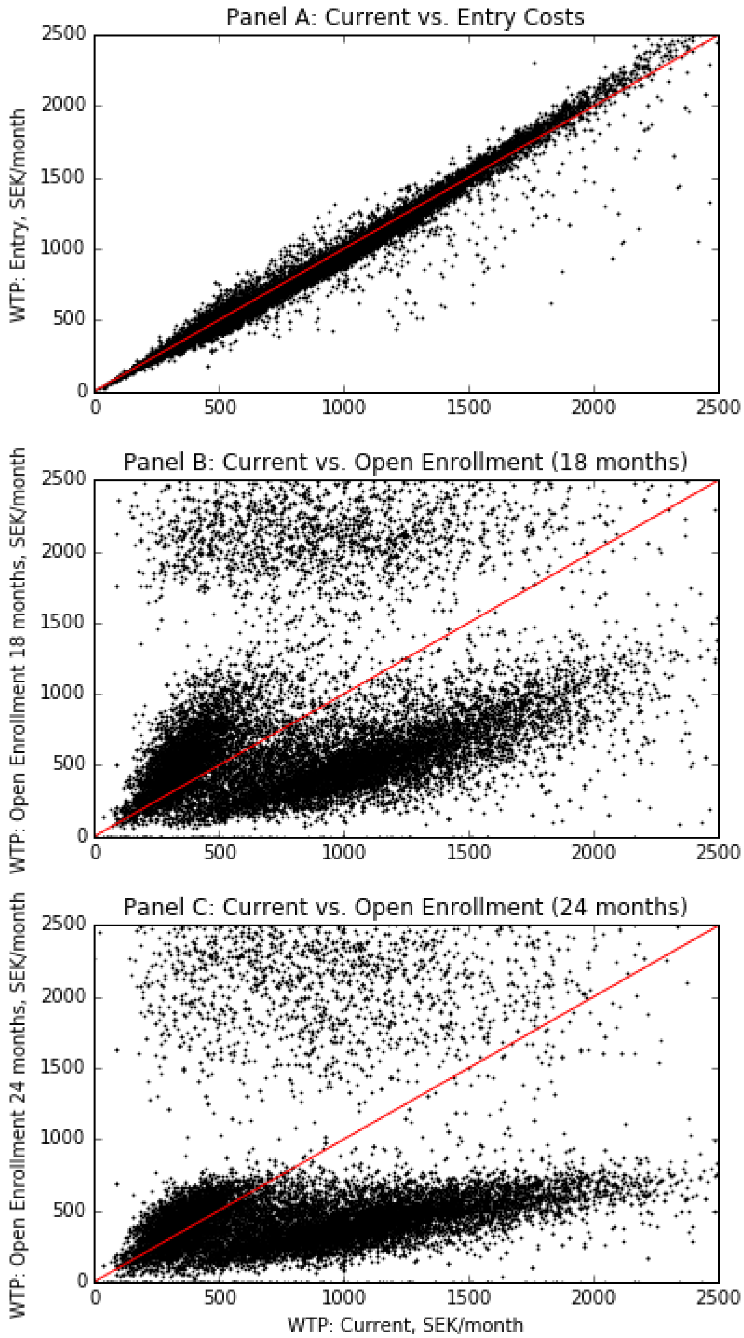
Notes: The Figure presents a discrete histogram of a distribution of a number of enrollment months before the start of unemployment spells similarly to the evidence in the main text but separately for individuals without higher education (Panel A) and with higher education (Panel B).

Figure 19: Bunching Around the Eligibility Requirement by Age



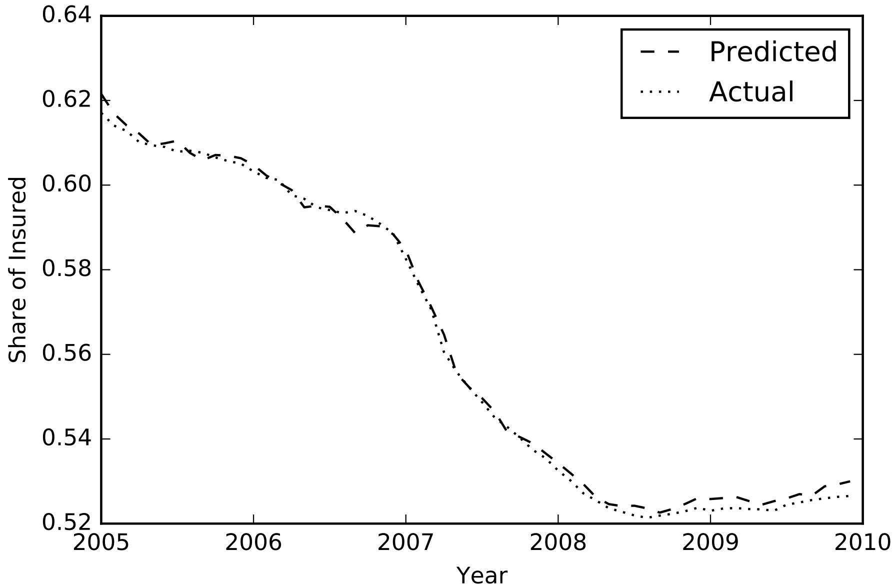
Notes: The Figure presents a discrete histogram of a distribution of a number of enrollment months before the start of unemployment spells similarly to the evidence in the main text but separately for individuals younger (Panel A) and older (Panel B) than 40 years old.

Figure 20: Comparison of WTP under Alternative Systems



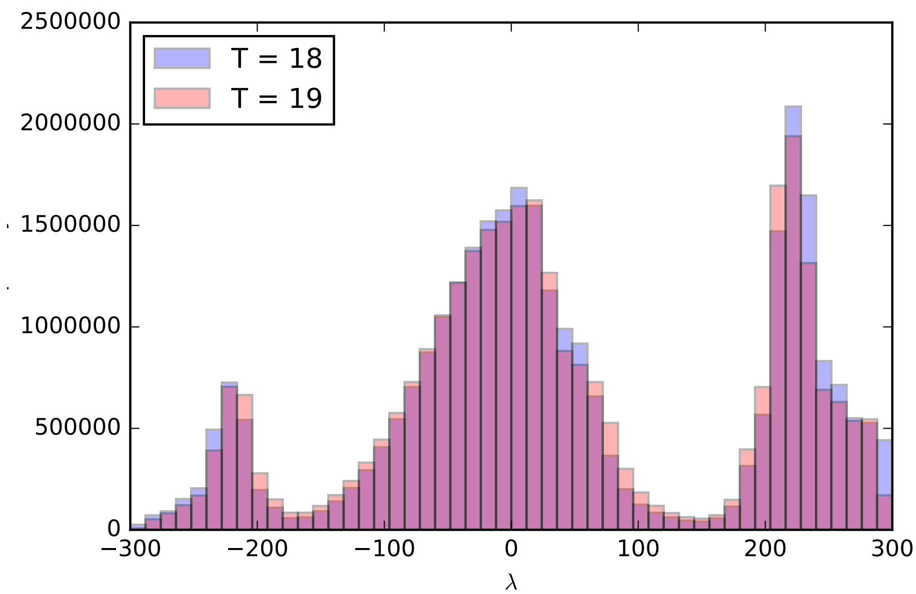
Notes: The Figure demonstrates WTP for counterfactual insurance systems (y-axis) against WTP for a current insurance system (x-axis). Red lines have 45^0 angle and allow seeing whether the corresponding system is more valued by individuals. Each point represents average willingness to pay for each individual within the considered time periods. If a given point lies above the red line, the corresponding alternative contract is on average valued more by this individual.

Figure 21: Model Fit - Demand



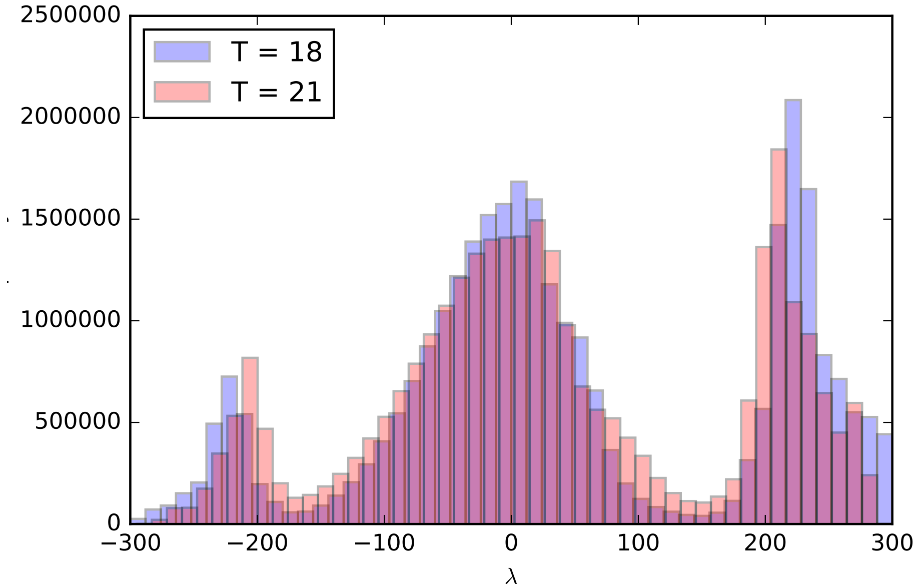
Notes: The Figure demonstrates actual (dashed) and predicted demand (solid) demand functions for 2005 - 2009. The y-axis represents a share of insured individuals.

Figure 22: Distribution of Risk Preference Indifference Points - 18 months vs. 19 months



Notes: The Figure demonstrates a distribution of thresholds under $T = 18$ and $T = 19$. Y-axis denotes a frequency of the distribution.

Figure 23: Distribution of Risk Preference Indifference Points - 18 months vs. 21 months



Notes: The Figure demonstrates a distribution of thresholds under $T = 18$ and $T = 21$. Y-axis denotes a frequency of the distribution.

D Supplementary Tables

Table 3: Types Parameters

	Types											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Constant	0.3	-16.150	-17.053	-8.729	-10.437	-4.759	-18.600	-9.208	-10.003	-10.486	-18.143	-1.253
	(—)	(1.137)	(0.693)	(1.016)	(0.036)	(1.950)	(0.470)	(0.103)	(0.154)	(1.070)	(0.305)	(0.062)
Cluster 1	0.3	0.233	6.090	4.909	0.632	-7.971	-3.830	-2.831	-0.863	-6.258	16.239	0.053
	(—)	(1.336)	(0.783)	(1.039)	(0.026)	(0.095)	(0.168)	(0.002)	(0.031)	(0.629)	(0.305)	(0.124)
Cluster 2	0.3	0.028	5.479	-11.025	-8.332	-3.430	9.123	-3.607	-7.397	-3.573	-9.267	0.175
	(—)	(1.443)	(0.726)	(0.010)	(0.000)	(0.627)	(0.471)	(0.003)	(0.000)	(0.628)	(0.309)	(0.059)
Cluster 3	0.3	9.292	-11.637	-0.086	-0.161	1.474	-2.122	-0.552	0.027	2.144	-1.170	0.206
	(—)	(2.527)	(0.004)	(1.075)	(0.015)	(1.236)	(0.168)	(0.025)	(0.040)	(1.026)	(0.309)	(0.088)
Cluster 4	0.3	-25.240	-4.194	-0.958	-0.978	-3.249	-11.460	-0.704	0.121	7.120	-10.282	-0.185
	(—)	(0.057)	(0.003)	(0.625)	(0.016)	(0.607)	(0.168)	(0.022)	(0.082)	(0.915)	(0.309)	(0.090)
Age (30; 40]	0.3	15.572	-9.704	-0.294	0.762	-9.110	4.487	-1.846	-0.753	-19.911	-2.773	0.421
	(—)	(1.280)	(0.004)	(1.837)	(0.028)	(0.095)	(0.445)	(0.011)	(0.019)	(0.629)	(1.414)	(0.094)
Age (40; 50]	0.3	4.797	-2.338	0.895	-3.468	-3.215	-0.226	-1.061	0.568	0.924	-2.707	0.153
	(—)	(1.506)	(0.008)	(1.761)	(0.001)	(0.340)	(0.249)	(0.034)	(0.106)	(1.364)	(1.354)	(0.071)
Age > 50	0.3	-1.690	5.842	-3.831	-5.903	2.103	-6.244	-8.028	-1.111	0.495	-2.719	0.278
	(—)	(0.080)	(0.670)	(1.190)	(0.000)	(1.300)	(0.168)	(0.000)	(0.036)	(1.400)	(1.264)	(0.056)

Table 4: Type Probabilities

Type	Predicted Share
I	73%
II	7%
III	
IV	
V	
VI	
VII	5%
VIII	
IX	
X	
XI	
XII	15%

Notes: The Table shows mean predicted type probabilities in the estimation sample determined by estimated type parameters.

Table 5: Model Fit - Share of Insured Individuals by Subgroups

	Shares of Insured Individuals	
	Actual	Predicted
Age ≤ 30	0.569	0.569
Age $(30; 40]$	0.561	0.563
Age $(40; 50]$	0.562	0.563
Age > 50	0.555	0.556
Gender	0.572	0.571
Family	0.562	0.563
Higher Education	0.558	0.558
Has Children	0.563	0.565
Income $\leq 25\%$	0.571	0.574
Income $(25\%; 50\%]$	0.563	0.563
Income $(50\%; 75\%]$	0.558	0.556
Income $> 75\%$	0.556	0.558

Notes: The Table demonstrates the actual and predicted shares of insured individuals by subgroups of individuals based on income, family, gender and education characteristics.