Private Information and Design of Unemployment Insurance

Maksym Khomenko*

Job Market Paper

September, 2018

Abstract

Unemployment insurance (UI) programs around the world are predominantly government-provided with mandatory universal coverage. One explanation for the dominant adoption of mandatory UI is that private knowledge about unemployment risks might lead to a highly selected pool of insured individuals and generate large welfare losses. I use the institutional features of the Swedish UI system, which combines both voluntary and mandatory programs, to study the optimal design and regulation of UI. With detailed administrative data, I estimate a structural model of insurance choice that captures heterogeneity in preferences and quality of information about future unemployment risks. The model is used to study several alternative designs of the UI program. The results suggest that mandating UI would be a welfare-improving policy only if the government is willing to provide high subsidies. In contrast, an alternative two-part tariff contract results in 6.1% higher consumer surplus on average for all subsidy levels. Contracts with fixed length and enrollment timing dominate all other considered options and generate considerable consumer surplus gains from 83% to 106% on average depending on the contract duration.

*Department of Economics, University of Gothenburg, *maksym.khomenko@economics.gu.se.*
I thank Mikael Lindahl, Aico van Vuuren and Liran Einav for invaluable advice throughout the development of this project. This work has benefited particularly from the comments of Natalie Bachas, Timothy Bresnahan, Raj Chetty, Mark Duggan, Randi Hjalmarsson, Petra Persson, Emmanuel Saez, Måns Söderbom and seminar participants at the University of Gothenburg and Stanford University.

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] From a theoretical point of view, the presence of adverse selection might lead to welfare losses in such a system. However, moral hazard and heterogeneity of preferences might rationalize the adoption of voluntary UI. This ambiguity and the absence of conclusive empirical evidence raise the question regarding optimal regulations in UI. Therefore, this paper attempts to comprehensively study the consequences of mandates and alternative designs of UI programs.

The essence of adverse selection in the context of UI is that individuals tend to have private information about their overall unemployment risk types (e.g. working in a risky occupation, an industry or a firm), which might create an insurance pool of relatively high-risk individuals. This could 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).

On the other hand, the presence of moral hazard and heterogeneity of preferences for insurance may serve as a rationale for a voluntary system. Moral hazard in UI means that availability of insurance entails, for instance, a reduction in job search or on-the-job efforts, which raises the probability of unemployment. As a result, it might amplify the costs under a mandatory system and make such a policy suboptimal. Furthermore, in the case of preference heterogeneity, a mandate might impose the excess burden on low risk-aversion individuals who do not value insurance even in the presence of substantial risks. Therefore, a positive correlation between purchasing insurance and unemployment risks might not be sufficient to motivate the introduction of a mandate.

Adverse selection might potentially be intensified by unrestricted enrollment timing, which leads to selection not only on the overall risk but also based on the variation of risks over time.[2] The presence of time-selection was documented in, for example, dental (Cabral, 2016) and health insurance (Einav, Finkelstein, & Schrimpf, 2015) markets. In the context of UI, it means that individuals tend to buy insurance when they have higher unemployment risks but at the same time cost more to an insurer. Therefore, alternative contracts that restrict time-selection might be welfare-improving especially with a view to the inefficiencies arising from

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

[2] There is a membership eligibility condition that acts as a timing restriction but does not fully resolve the timing issue.

mandates. I study the effect of two contracts with such properties. First, I consider an "open enrollment" contract with fixed duration and timing of enrollment. Another alternative is an "entry costs" or "two-part tariff" contract, which in addition to monthly premiums charges entry fees upon the payment of first monthly premium (Cabral, 2016).[3,4]

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 regulations. An eligibility condition for the income-based coverage requires paying insurance premiums for at least twelve consecutive months. It enables studying time-selection and optimal contract design in addition to just a mandatory versus voluntary system trade-off.

This paper uses detailed individual-level administrative data, which allow observing exact dates of unemployment and insurance spells together with a variety of demographic and labor market characteristics for the period 1999 - 2014. 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 enrollment timing eligibility condition, I demonstrate that individuals are more likely to start unemployment spells with twelve months of UI enrollment. I show that this evidence is robust and persists for various subgroups.

The empirical strategy exploits detailed individual data, insurance price and generosity variation, and the observed time-selection patterns to estimate a dynamic insurance choice model. The purpose of estimating this structural model is to recover a distribution of risk preferences and individual information about future unemployment and 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 generosity of benefits over time primarily due to a UI reform in 2007. Another source of variation stems from cross-sectional differences in generosity of benefits as a result of benefits cap and differences in premiums across UI funds. The identification of individual information exploits patterns in timing of insurance purchase relative to timing of future unemployment and changes in unemployment risks. The results show a considerable variation in risk preferences and quality of information about future employment perspectives. I also estimate a choice inertia parameter that implies considerable choice persistence meaning that the insurance status in a previous period impacts future decisions.

The essence of insurance markets consists of an interplay of individual risk preferences, risks and private information about those risks. This complexity rationalizes the use of a model that

[3] Similarly, one can charge for the exit from an insurance pool.

[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 should presumably discourage exits to re-enter the insurance pool later in the presence of high unemployment risk.

enables combining those parts. Some of the existing works provide policy conclusions about UI based on a "reduced form" association between realized risks and insurance probabilities, using either observable characteristics or arguably exogenous institutional variations (Hendren, 2017; Landais, Nekoei, Nilsson, Seim, & Spinnewijn, 2017). In contrast, this paper attempts to augment the existing evidence with a comprehensive insurance choice model. This approach allows studying a broader spectrum of alternative regulations and welfare consequences at the expense of imposing a number of theory-based structural assumptions.

To evaluate welfare under current and alternative structures of UI, I use the model estimates to recover the UI demand functions and the distribution of willingness-to-pay (WTP) for insurance contracts under consideration. I also estimate a prediction model of unemployment risk to predict costs of covering individuals. The findings suggest that in the absence of a moral hazard response mandates would generate large welfare gains up to 115% in terms of consumer surplus compared to the current system if the government is willing to provide fairly large subsidies.[5] However, in the absence of large subsidies, the mandate is predicted to cause large welfare losses. The intuition is that a limited budget requires raising prices which might severely reduce consumer surplus since individuals who do not value insurance cannot unenroll.[6,7]

Therefore, in addition to mandates, I consider a number of alternative contract designs, which address the private information problem and allow for voluntary enrollment. I find that an alternative two-part tariff contract that charges extra fixed costs upon the first premium payment provides small consumer surplus gains of 6.1% on average compared to the status quo. Furthermore, large welfare benefits are predicted in the case of an open enrollment contract, which allows enrolling only in a specific period of time and provides fixed duration coverage of 18 and 24 months. The results suggest that the 24 months contract welfare-dominates all other considered alternatives. In addition, although dominated by the 24 months contract, the 18 months contract is welfare superior to other considered policies except for a mandatory system with very high subsidies. To put this into perspective, 18 and 24 months contracts yield 83% and 106% average consumer surplus gain, correspondingly. This considerable welfare improvement stems from the virtual removal of time selection and inertia. As a result, it leads to better choices that generate higher consumer surplus but at the same sufficiently restricts private information.

This paper contributes to a large literature on private information in social insurance and

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

[6]The welfare loss, in this case, does not result from low subsidies but from high prices that are required to sustain the system under these conditions.

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

insurance markets. Most attention regarding the importance of private information in designing social insurance systems has been dedicated to health insurance, annuity, and long-term care markets. In particular, a large literature documents the presence,[8] discusses sources,[9] and analyses consequences of asymmetric information,[10] as well as studies policies aimed at addressing inefficiencies in insurance markets.[11] The literature related to unemployment insurance has been primarily focused on the optimal UI theory[12] and on estimating labor supply responses to insurance benefits.[13] However, to the best of my knowledge, only a few empirical papers focus on the canonical private information problem in UI. Hendren (2017) shows that the absence of private UI markets is due to the excess mass of private information. In this paper, I do not focus on the existence of private information and the effect on the private markets but attempt to look at how contract designs can be used to address the problem and generate welfare gains.

Another paper studying the problem of private information in UI using the Swedish setup is Landais et al. (2017). The authors document that insured individuals on average have higher unemployment risk. It is argued that adverse selection must be an important component of the observed positive correlation between unemployment risks and insurance statuses. 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.[14] 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 policies that are difficult to study using the approach in Landais et al. (2017). The reason is that to interpret the effect of alternative insurance designs, one needs to take into account preferences, risks and risk perception. These parameters are difficult to recover without theoretical assumptions. Furthermore, the empirical approach in this paper allows for

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

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

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

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

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

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

[14] 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.

more comprehensive exploration of detailed data and rich variation not limited to price changes. Understanding complex individual behavior is required to analyze welfare and consider alternative policies. These fundamental differences in approaches account for discrepancies in some of the policy conclusions.

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. Although the focus is on an annuity market, the authors also use a comprehensive dynamic structural model of choice under uncertainty to recover policy-relevant dimensions of individual heterogeneity. A similar approach to estimating an insurance choice model is used by Cohen and Einav (2007).

Finally, the paper is related to a strand of the literature studying optimal design of insurance contracts.[15] Previous work emphasizes the importance of contract structure beyond pricing, which was a dominant focus of the literature. In particular, this paper contributes by adding a piece of evidence regarding the importance of a dynamic component of adverse selection. Similar time-selection evidence was documented in healthcare (Einav et al., 2015), dental care (Cabral, 2016) and annuity markets (Aron-Dine, Einav, Finkelstein, & Cullen, 2015; Einav et al., 2015; Einav, Finkelstein, & Schrimpf, 2017), which highlights the importance of this dimension of adverse selection for a wider range insurance markets. There are a number of papers that study the role of non-linear benefits schedules on the dynamics of unemployment. For instance, Kolsrud et al. (2018) study the role of a non-linear benefits schedule in Swedish UI but their work is more related to the literature on a labor supply response. Similarly, DellaVigna et al. (2017) analyze the role of a benefits structure in the presence of non-classical behavioral responses. Instead, I consider the role of non-linear time-based insurance eligibility and additional dimensions of private information that it creates.

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 empirical analysis. Section 4 describes a structural model and an estimation approach. Section 5 analyses welfare and alternative policies. Section 6 concludes.

[15] Azevedo and Gottlieb (2017) studies 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 mandatory and voluntary income-based programs. The basic compulsory insurance similarly to the mandatory systems grants a fixed daily amount of 320 SEK ($35) conditionally on meeting basic and work requirements.[16] In particular, 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, in addition to basic and working conditions, requires paying monthly fees to UI funds for at least 12 consecutive months.[17,18] Before 2007, fees for employed and unemployed individuals coincided. As a result of a labor market reform, which also altered the structure of UI, fees for employed individuals more than tripled on average. Figure 1 demonstrates average fees for employed and unemployed individuals over time. Generosity of benefits has also been varying as demonstrated in Figure 2

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.[19] Unemployment without a valid reason (voluntary or because of unacceptable behavior) results in an uncompensated period of up to 45 days.

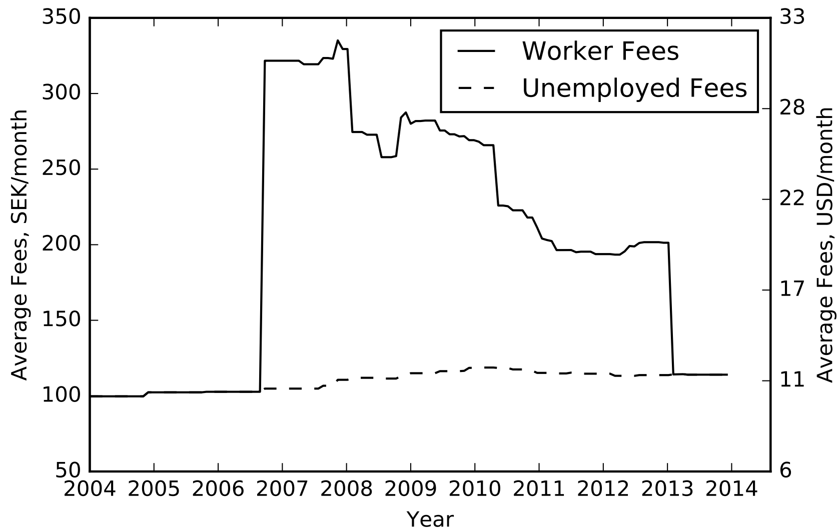
[16]The amount was raised to 365 SEK ($40) in September 2015. For more details regarding changes in 2015 see http://www.fackligtforsakringar.n.nu/a-kassan or http://www.regeringen.se/artiklar/2016/09/en-battre-arbetsloshetsforsakring/ .

[17]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.

[18]Enrollment requires working for 1 month.

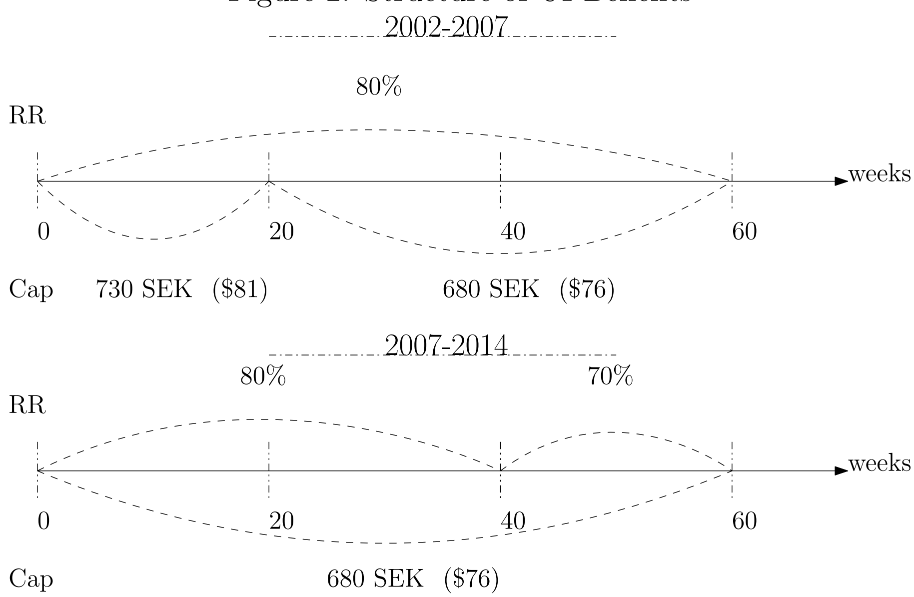
[19]If the accumulated unemployment duration exceeds 300 days, an individual is assigned to an intensified counseling program or can be granted with an extension of 300 days if the counseling is deemed to be unnecessary (but only once). This option disappeared after the reform in July 2007. For more 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 line represents a schedule of benefits for 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 80% replacement rate subject to the cap, which depended 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. At the same time, basic mandatory insurance benefits amount to 7040 SEK ($782) of monthly income. Average income in the sample used in the analysis, which I discuss in the next section, is approximately 28 000 ($3 111) SEK in 2008. It is almost 74% higher than the first cap and 87% higher than the second cap. A labor market reform introduced changes in both a replacement rate and 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.[20] The cap became constant for an entire 60 weeks period and amounted to 680 SEK ($76).[21]

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 membership records including insurance fund affiliations, which allows merging price data. The dataset contains 2 167 287 unique individuals[22] over the period 1999 - 2014. It is not representative to the population since it excludes individuals who have not claimed UI benefits.[23]

I match the IAF dataset to data from the Public Employment Service (PES), which provides information on all registered unemployment spells including dates and unemployment categories.[24] A rich set of annually observed individual characteristics comes from the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) including a

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

[21]Eligibility for income-based insurance is a prerequisite for even higher income compensation from a union without a cap. The analysis in this paper does not take it into account. Although the presence of additional fund-based insurance affects parameter estimates, it does not affect the comparative analysis of various UI designs.

[22]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.

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

[24]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. Different aggregation does not affect the estimates.

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.[25]

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 credibly 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 baseline sample used in the estimation differs from the full 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 introduces complications in the estimation since those individuals face an income stream that is a mix of wage and benefits. Therefore, to preserve a 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. The reason behind this sample restriction is that to construct monthly wages, I impute values for periods with income from social security using income from previous periods without social security payments. It allows accounting for the fact that wages are likely to be understated during benefits recipiency. It results in a baseline estimation sample that contains 865 363 individuals. Table 1 presents key descriptive statistics of the full sample and the selected baseline estimation sample in comparison with the full economically active population 16 - 64 years old.

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

Table 1: Descriptive Statistics and Unemployment Patterns

	Full Sample	Baseline Estimation Sample	Swedish population 16 - 64 years old
	(1)	(2)	(3)
Descriptive Statistics, 2008			
Income, *SEK/month*			
Mean	24 754	24 834	28 623
Median	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%
Number of Unemployment Months per Individual, 2002 - 2014			
Mean	12.72	12.68	8.69
P1	1	1	1
P10	3	3	2
P25	5	5	3
P50	9	9	5
P75	16	16	11
P90	27	27	19
P99	64	64	49
Always Employed	83.9%	83.8%	89.5%
N	2 167 287	865 363	7 811 784

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) presents a these details for 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 latter part describes distribution of a number of unemployment months that individuals accumulated during 2002 - 2014.

Table 1 shows that full and baseline samples are very similar in terms of observables. Slight differences are observed in a share of female, which is 51% in a baseline sample compared to 53% in the full sample. Also, a baseline 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 primary selection margin is the recipiency of UI benefits. As a result, individuals who are omitted from the full sample on average have higher income. This effect is mechanical since unemployed individuals should have less wage income. Apart from the differences in income, the selected sample contains slightly more individuals with higher education, which is also a mechanical effect since less relatively young individuals who are most likely have not finished higher education are included in the selected sample. Finally, a full sample is represented by a 4% smaller share of female individuals.

Although full and baseline samples are very similar in terms of unemployment risks, they, as expected differ considerably from a full population. Selected samples contain approximate 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 is shifted to the right for the selected samples.[26]

3 Descriptive Evidence

3.1 Sources of Private Information in UI

Unemployment insurance as any other selection market 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 the set-up of UI is that individuals tend to have more information about their overall unemployment risk. 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).[27] It would generate a qualitatively similar selection pattern but have different policy implications. For instance, a mandate can be a welfare-improving policy in the adverse selection case but it might also result in welfare losses. The reason is that the absence of choice imposes the excess burden on individuals who do not value insurance. In addition, the potential presence of moral hazard, which might generate similar positive correlation pattern, would result in reversed policy directions. Moral

[26] Although selection margin is benefits claiming, many individuals have not been unemployed even in insurance data samples. The reason is that individuals might have received compensations before 2002 or claimed other contributions not categorized as open unemployment.

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

hazard or ex-post selection is a behavioral response to being insured that increases unemployment probability. The intuition is that a lack of incentives due to lower financial stakes leads to less job-search or on-the-job efforts. Mandates might provoke such a response and raise costs of covering individuals, which outweighs risk-pooling benefits.

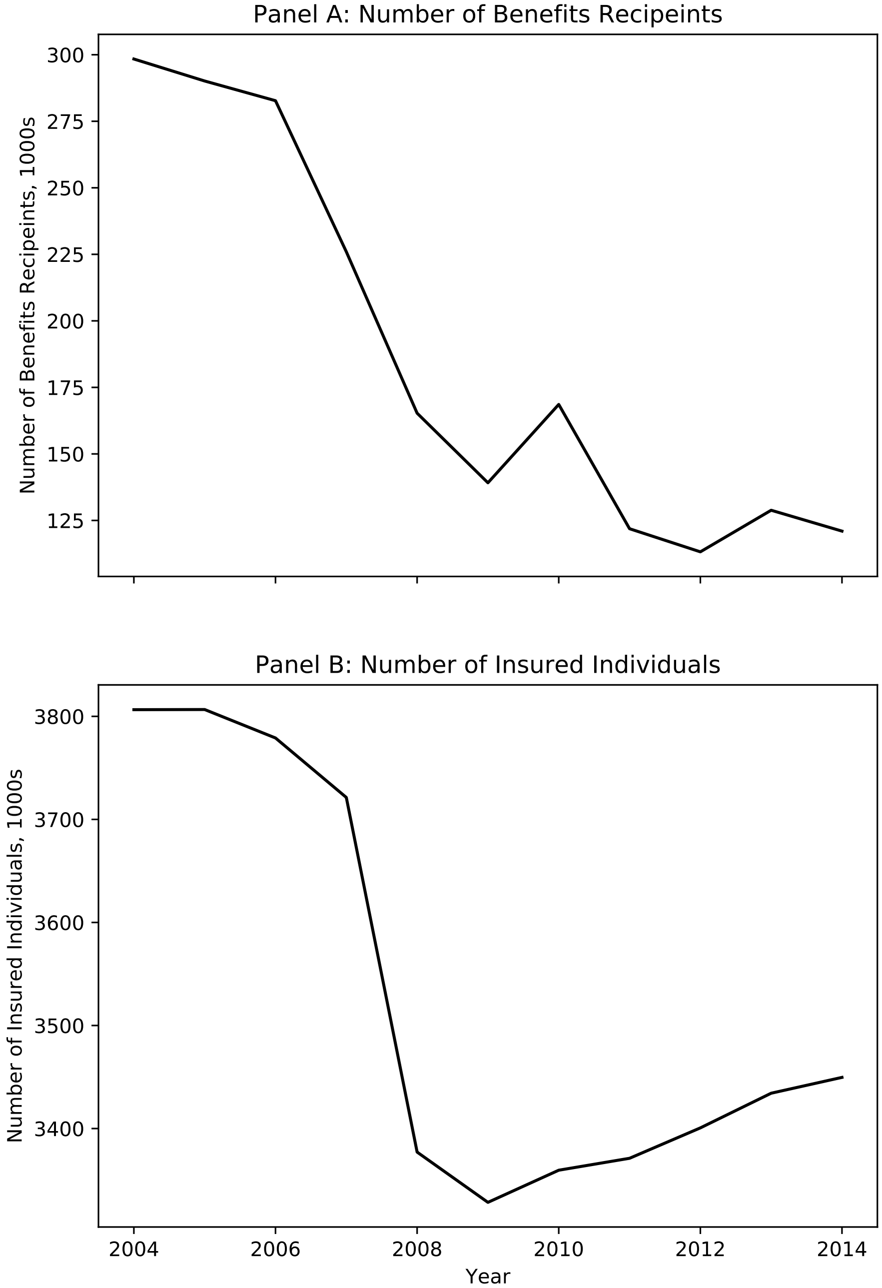
To sum up, there are many scenarios arising from the complexity of insurance decisions that fundamentally hinges on risk perceptions and preferences for risks exposure. This ambiguity might result in a need of the opposite policy measures while generating same "reduced form" patterns in the data. This section does not attempt to disentangle those forces since it might have a limited use for the welfare analysis. For a discussion and an attempt to disentangle 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 awareness of risks, which often requires a theoretical structure.

3.2 Behavioral Responses and Patterns in UI

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 a replacement rate as a result of a cap, which also differs depending on the duration of unemployment.

Another dimension of the variation is a result of a reform in 2007 which raised insurance premiums primarily for employed individuals and weakly reduced generosity of benefits. These changes created differences in incentive and caused behavioral responses illustrated in Figure 3.

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



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

The Figure shows that a reform had an effect on a number of aggregate indicators that 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).[28] However, this

[28]Note that a number of insured and a number of benefits recipients are not directly linked since one can receive

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. It implies that a dynamic nature of insurance eligibility introduces time-selection. The literature has documented these 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 likelihood of being insured. On the one hand, it can be argued that time-selection is a part of adverse selection and can be resolved by mandates. On the other hand, a potential presence of preference heterogeneity might justify milder regulations such as alternative contracts that restrict time-selection. Therefore, the model presented in the next section attempts to jointly recover distributions of risk preferences and information quality.

Table 2 describes the patterns of unemployment and insurance spells within the course of observed 192 months (1999 - 2014).

Table 2: Patterns of Employment and Insurance Spells, 1999 - 2014

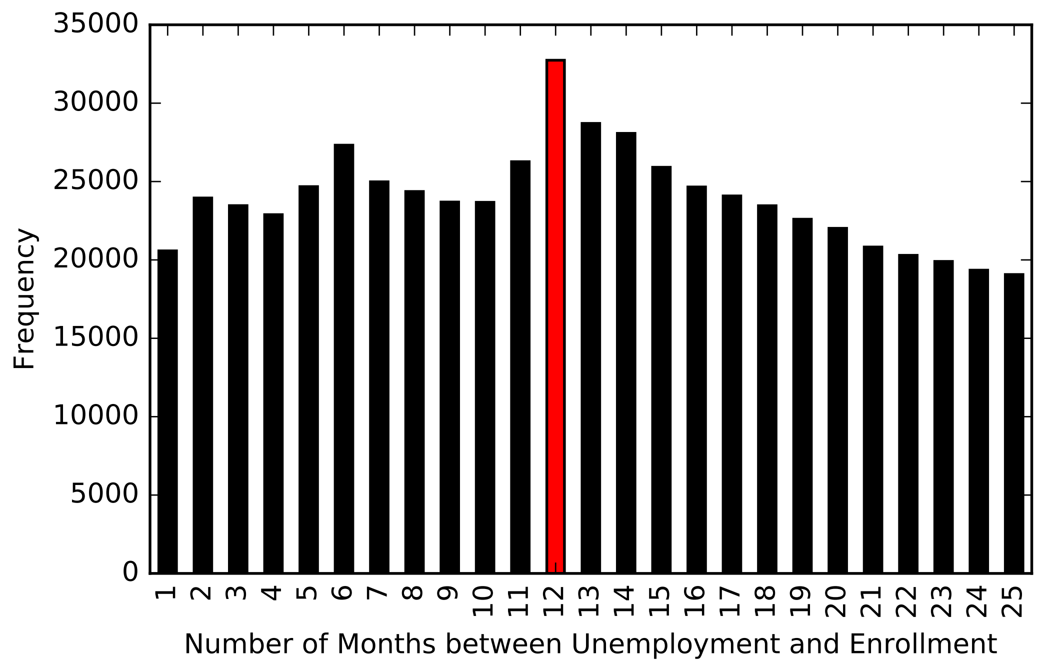
Statistics	Unemployment		Insurance	
	# of Periods	Duration of Periods	# of Periods	Duration of Periods
	(1)	(2)	(3)	(4)
Median	2	7	1	99
Min	0	1	1	1
Max	42	192	11	192

Notes: The Table describes basic patterns of unemployment and insurance spells in the data. Columns (1) and (2) correspond to unemployment. The former characterizes the distribution of a number of unemployment periods, whereas the latter describes a distribution of durations of unemployment spells. These durations might be truncated if started before 1999 or were over after 2014. A column (2) describes statistics conditionally on being unemployed, which excludes those who did not have any unemployment spells within the observed period. Columns (3) and (4) present the same evidence but for insurance periods.

basic insurance even without being a fund member.

The Table shows that many individuals tend to have only one insurance, which often covers the whole period under consideration. A maximum number of insurance sequences amount to eleven within the course of sixteen years. Conditionally on being unemployed (column 2), a median individual is unemployed for seven months. Most individuals have two unemployment spells within a period under consideration. It, firstly, suggests that individuals might display a considerable amount of inertia in fairly frequent monthly choices revealed by a dominance of long insurance spells. A potential competing force, which is exacerbated in the absence of the inertia would be the above-mentioned time-selection. The evidence of time-selection can be obtained from a distribution of a number of enrollment periods with which individuals start unemployment spells in the data displayed in Figure 4.

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

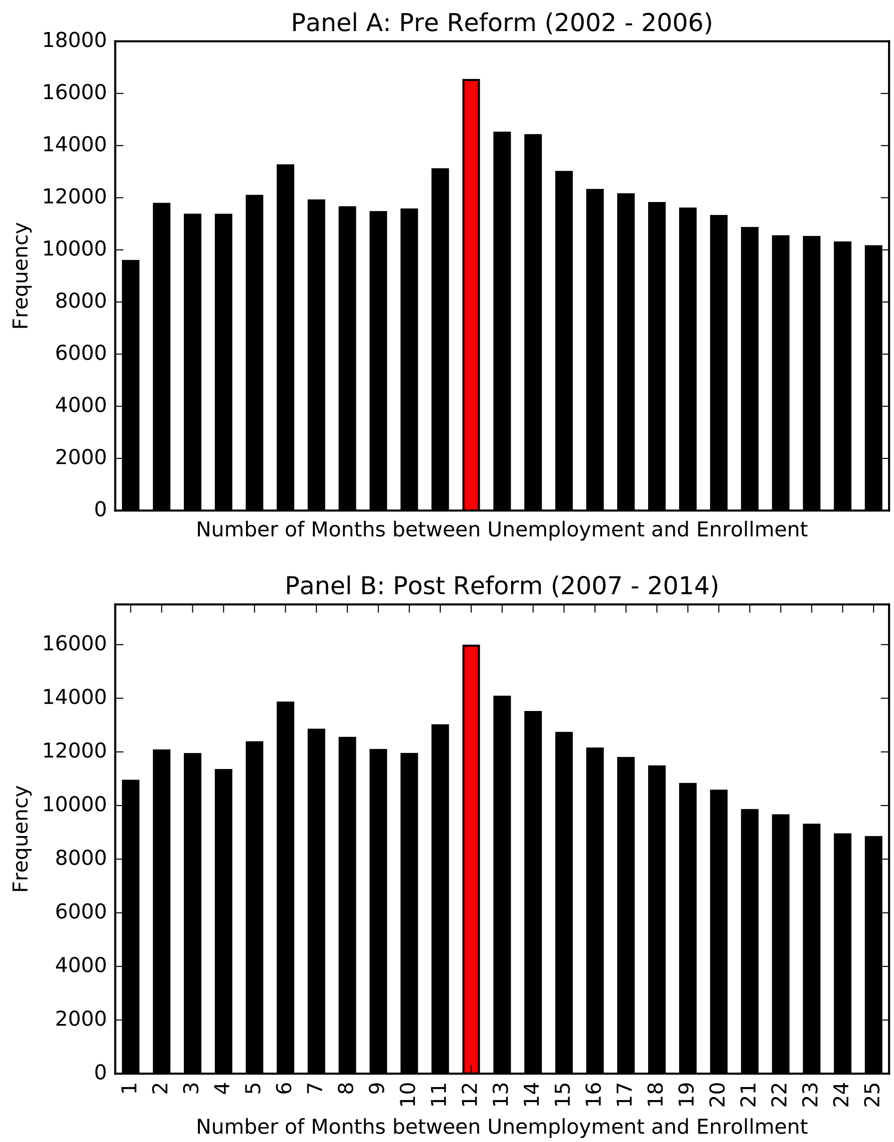


Notes: The Figure presents a discrete histogram of a distribution of a number of enrollment months before a commencement of unemployment spells. The Figure is truncated at one. The red bar denotes twelve consecutive months of enrollment required for eligibility. The histogram contains a clear spike exactly at the red bar implying that individuals are more likely to start an unemployment spell with exactly twelve months of enrollment to minimize the total premium amount. The Figure also shows a bunching after the eligibility threshold supporting the presence of the time-selection evidence.

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. The area to the right of the spike also contains a bunching with a missing mass to the left of the red bar. The region to the left of the red bar contains another smaller bunching area, which might be a result of layoff notices specified in employment contracts or differences in 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 structure about future employment outcomes. The additional evidence for various subgroups is presented in the Appendix B (Figures 13, 14 and 15) and shows identical patterns.

Figure 5 presents the same evidence separately before and after the reform in 2007, which weakly reduced the generosity of benefits and raised insurance premiums. As can be seen, the patterns are similar for both periods. The presence of considerable differences on those figures could alert about the time-selection effect accompanied with a moral hazard component. It would mean 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 of 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, some individuals who decide to quit a job accordingly to choosing 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. It is difficult to graphically see considerable differences in bunching patterns which can be explained by either the absence of this effect or a small scale of the reform, which did not induce such a behavioral response.

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 enrollment months before unemployment spells. It replicates the evidence in Figure 4 but separately for periods before (Panel A) and after the reform in 2007 (Panel B).

This section described main behavioral patterns observed in the data. Firstly, it shows that individuals seem to 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 or, in the unlikely case, the absence of risk variation over time. Finally, individuals display time-selection behavioral response. 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

This section describes a model of an individual decision to pay insurance premiums. The model has a purpose of recovering risk preferences, inertia, and quality of information about future employment outcomes, which determine insurance choices. The estimates are used to obtain individual willingness to pay for insurance, which together with estimated cost functions enables conducting a welfare analysis and counterfactual policy experiments. The structure of the model is motivated by a number of descriptive patterns. Firstly, individuals respond to changes in unemployment risks, benefits, and premiums. It naturally leads to viewing a decision-making process as a choice under uncertainty over employment outcomes. Another part of the model is motivated by the bunching evidence from the previous section. It suggests that individuals not only have more information about their average risks but also about the timing of unemployment and variation of risks over time. As a result, it allows minimizing insurance premiums by selecting the optimal timing of enrollment. The model remains agnostic about the sources of this information being it due to knowledge about the future, negotiations with an employer to delay layoff or any other alternative explanations. I assume that there are a number of discrete types that denote how many months before unemployment can be foreseen. It is motivated by temporary contracts with a fixed termination date, legally enforced layoff notice requirements and informal arrangements within a firm. Consequently, the probability of being a given type depends on a large set of relevant labor market characteristics. To put this into perspective, imagine an individual who always knows employment outcomes for four months in the future. She is aware that she will be employed with certainty for the next four months. The outcomes beyond this period are uncertain. If an individual instead faces unemployment in three months, it is only known that there is forthcoming unemployment in 3 months while having no information about how long it will take to find a job. Note that unemployment "in three months" will become unemployment "in two months" in the next insurance decision period. The model is dynamic, which is motivated by a forward-looking eligibility requirement. Bunching evidence presented before emphasizes the importance of dynamic incentives in this set-up.

Loosely speaking, to recover a parametrized distribution of heterogeneous risk preferences and discrete types, I mostly leverage variation in generosity of benefits and fees across individuals and over time, which allows identifying risk preference parameters. To identify a distribution of discrete types, the model exploits time-selection patterns in the data conditionally on labor market characteristics. The inertia parameters, which is an important component, is identified solely from a functional form assumption. Upon estimating the parameters, I conduct a number of counterfactual experiments to study the effect of mandating the system and the adoption of alternative contracts.

The remainder of this section formally describes the model and identification, discusses an estimation approach and presents the results.

4.1 Model

The model attempts to capture a decision-making process of an individual within the environment of voluntary UI in Sweden. I model an insurance choice as a forward-looking decision to maximize the expected utility of income. I refer to this action as "to be insured" from now on, which means paying insurance fees but not necessarily being eligible for benefits yet. All dynamics in the model is monthly. Each period an individual observes information about future employment, wages, insurance benefits and prices, and makes a decision to pay insurance fees to gain or keep the eligibility for income-based insurance in the future. The insurance decision has a dynamic effect through an eligibility condition that requires being insured for at least 12 consecutive months.

More formally, each individual i makes a decision each month t conditionally on observing the following information:

1. *A current number of accumulated enrollment periods* $\kappa_{it} \geq 0$.

2. *Known with certainty employment statuses for* $s > 0$ *periods or up to forthcoming unemployment.*[29] More formally, $s = \min\{s_e, s_u\}$ where s_e denotes a number of periods foreseen in the future defined by a type of a worker and s_u denotes a number of periods until next unemployment. I refer to the types while having in mind this information structure. The type is observed by an individual but unobserved to an econometrician. I assume that individuals can be one of twelve types $s \in \{1, ..., 12\}$ denoting how many periods in the future or before unemployment are observed. I limit the attention to twelve discrete types since it is required to capture behavior stemming from the enrollment eligibility requirement. At the same time, being a type $s > 12$ does not provide much additional information that allow benefiting more from a superior information.[30]

[29] I exclude $s = 0$ since it is unlikely that individuals do not know employment outcomes for the current month.

[30] There are government regulations that specify the layoff notice period depending on tenure. However, firms tend to include layoff conditions in the contract. Different regulations apply to temporary contract workers and probation periods. In the case of temporary workers, the layoff notice is a period to a termination date or layoff notice specified in the contract depending on which comes first. Probation period workers usually are regarded as permanent workers with two weeks of a layoff notice. In addition, individuals might have extra information

3. *Probability of unemployment outside of the "type information"*. An individual knows her true employment outcomes for next s periods and forms rational beliefs about a probability of being unemployed at a period after s within a planning horizon T, which is discussed later. I assume that there are two state-dependent probabilities of employment: p_t^0 - probability of employment at time t conditionally on being unemployed at $t-1$; and p_t^1 - probability of employment conditionally on being employed at time $t-1$, which creates an unemployment/employment persistence effect. It implies that individuals form expectations regarding n^{th} unknown period using Markov-type updating.

   Example 1. *Imagine an individual who is employed at period t with certainty. She believes that a probability of employment in the next period is p_{t+1}^0 if unemployed at period t and p_{t+1}^1 otherwise. Since an individual knows that she is employed at period t, a probability of being employed at time $t+1$ is p_{t+1}^1. Consequently, beliefs about the probability of being employed at $t+2$ are formed as $E[p_{t+2}] = p_{t+1}^1 p_{t+2}^1 + (1-p_{t+1}^1)p_{t+2}^0$. Similarly, the probability of being employed further at period $t+3$ is $E[p_{t+3}] = E[p_{t+2}]p_{t+3}^1 + (1-E[p_{t+2}])p_{t+3}^0$, where $E[p_{t+2}]$ is computed at the previous step. Therefore, an individual would roll such a probability chain for any unknown period in the future.*

   More generally, the expected probability of employment at period $t+n$ is constructed as follows:

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

4. *Expected wage, insurance fees and replacement rate for the entire planning horizon.* I assume that individuals have correct beliefs about those variable over the planning horizon $t = \{0, ..., T\}$. I discuss the choice of T later when I describe estimation and parametrization of the model.

An individual chooses whether to be insured or not $l_t \in \{0, 1\}$:

$$l_t = \begin{cases} 0, & \text{if uninsured} \\ 1, & \text{if insured} \end{cases}$$

The only state variable affected by a current insurance choice is a number of accumulated

beyond the scopes of contract conditions.

enrollment periods at $t+1$ (κ_{t+1}), which determines the eligibility status Λ_{t+1}:

$$\kappa_{t+1} = \begin{cases} \kappa_t + 1, & \text{if} \quad l_t = 1 \\ 0 \quad , & \text{if} \quad l_t = 0 \end{cases} \tag{2}$$

$$\Lambda_{t+1} = \begin{cases} 1, & \text{if} \quad \kappa_{t+1} >= 12 \\ 0, & \text{if} \quad \kappa_{t+1} < 12 \end{cases} \tag{3}$$

If an individual decides to continue or start paying insurance fees at time t, a number of accumulated enrollment periods increases by one in the next period and drops to or remains zero if fees are not paid. The eligibility comes after at least twelve periods are accumulated and lasts until a payment is missed.

A one-period payoff of an individual depends on insurance, employment and eligibility statuses:

$$\pi_{it} = \underbrace{(1-e_{it}) \cdot \left(\overbrace{(1-\Lambda_{it}) \cdot \underline{b}_{it}}^{\text{ineligible}} + \overbrace{(\Lambda_{it} \cdot \min\{b_{it} \cdot \bar{w}_{it}, B_{it}\}}^{\text{eligible}} \right)}_{\text{unemployed}} + \underbrace{e_{it} \cdot w_{it}}_{\text{employed}} - \underbrace{l_{it} \cdot \tau_{it}}_{\substack{\text{pay} \\ \text{premiums}}} \tag{4}$$

where e_{it} is employment status, which equals to one if an individual is employed; $\underline{b}_{it}$ - basic insurance amount that individuals would get if ineligible; b_{it} - wage replacement rate under voluntary insurance, B_{it} - voluntary insurance cap; $\bar{w}_{it}$ - mean income during twelve months preceding the current period based on which insurance value is constructed; w_{it} - actual wage received if employed.

The payoff of an insurance decision $l_{it} \in \{0,1\}$ is a sum of monthly incomes over the planning horizon T conditionally on optimally choosing a future insurance sequence:

$$\Pi_t = \overbrace{\pi_t(l_t)}^{current} + \overbrace{\sum_{n=t+1}^{T} \pi_n(l_n^*(l_t, \{e_k\}_{t=1}^T))}^{planned} \tag{5}$$

where $\pi_t(l_t)$ is a payoff from a current period when the actual decision about insurance is made; $\pi_n(l_n^*(l_t, \{e_k\}_{t=1}^T))$ is a planned payoff at some period n in the future conditionally on optimally choosing all l_n^* and conditionally on an expected employment sequence $\{e_k\}_{t=1}^T$.

There are three important clarifications about the model. Firstly, the intuition of equation (4) is that a decision to pay insurance fees has an effect on all the periods directly by affecting eligibility statuses and indirectly through forthcoming insurance decisions. Therefore, to decide whether to pay premiums now, an individual uses a backward induction procedure. It starts from a terminal period T and rolls back to select the optimal sequence of insurance decisions conditionally on a current insurance choice l_t. Secondly, the formulation of the equation (4) is loose for exposition purposes. I omit the notation that an optimal insurance sequence, a current decision, and payoffs are affected by time-varying state variables including a replacement rate, cap, fees and wages. Finally, $\{e_k\}_{t=1}^{T}$ denotes a particular sequence that an individual expects while planning. Recall that the information structure consists of a known number of employment outcomes for s periods in the future and uncertainty about the remainder of the planning horizon (from s to T). If an individual does not have any uncertainty, $\{e_k\}_{t=1}^{T}$ would be a unique known sequence. The lack of information introduces a multiplicity of potential sequences. As a result, each potential sequence implies a different optimal planning rule and payoff. The following example is intended to clarify the logic.

Example 2. *Imagine an individual who plans over T periods in the future to decide whether to pay an insurance premium now. The individual knows employment outcomes for all periods in the planning horizon except two final periods. For those periods, she forms beliefs that at the penultimate period she will be employed with probability 0.94 and with a probability 0.95 at the last period of the planning horizon. Note that this is a simplified example in which probabilities of unemployment are independent across periods for simplicity. The model instead, as discussed earlier, imposes more realistic Markov structure but this example suffices for the illustrative purposes. Table 3 summarizes the example.*

Table 3: An Example of Various Employment Sequences under Uncertainty

	Planning Horizon (T)													Sequence Probability
Probability	-	-	-	-	-	-	-	-	-	-	...	0.94	0.95	ξ_j
Sequence												0	0	$0.06 \cdot 0.05 = 0.003$
												0	1	$0.06 \cdot 0.95 = 0.057$
	1	1	1	1	1	1	1	1	1	1	...	1	0	$0.94 \cdot 0.05 = 0.047$
												1	1	$0.94 \cdot 0.95 = 0.893$

Notes: *The Table demonstrates an example of possible employment sequences, which creates outcomes uncertainty in the model. An individual knows true future outcomes except for the last two periods. It results in 4 possible combinations of different sequences that might be realized. Note that each of those sequences might generate a completely different optimal insurance sequence. Using the probabilities of unknown periods, it is possible to calculate the probabilities of each sequence in column* ξ_j.

Therefore, an individual has to solve 8 dynamic programming problems: twice for each of *those sequences to decide whether to pay fees at a current period t (for "insured" and "uninsured" choices at t) based on expected utility defined below.*

The example demonstrates that each individual i at time t solves dynamic programming problems for all possible sequences because of an uncertainty over employment statuses outside of a known interval defined by the individual type s. Note that a number of sequences is 2^{T-s}, which can become an extremely large number. I discuss how I address those complications in the next Appendix A. Note that each sequence apart from generating a different payoff also occurs with a different probability. The probability of a sequence j is more formally defined:

$$\xi_{jt} = \prod_{q=t+s+1}^{T} m_{qt} \tag{6}$$

where m_q is a probability of a specific outcome in the sequence at time q as demonstrated in the example.

An individual chooses to pay an insurance premium if her expected utility of paying now at time t is larger than expected utility of not paying defined as follows:

$$EV(l_t) = \sum_{j=1}^{2^{(T-s)}} \xi_{jt} \cdot u\left(\Pi_{jt}(l_t)\right) \tag{7}$$

where u - utility function; $\Pi_{jt}(l)$ - payoff of sequence j at a decision period t conditionally on choosing $l = \{0, 1\}$ defined in (5).

The individual i chooses to be insured at time t if $EV(l_{it} = 1) \geq EV(l_{it} = 0)$.

4.2 Parametrization and Estimation

I limit a planning problem to $T = 19$. A chosen T must be larger than 12 in order to capture time-selection behavior as a result of eligibility requirement. Since, it is required to solve a dynamic model many times for each individual, time, type, sequence and current insurance decision to compute payoffs of each action, it becomes computationally burdensome for a large T. I experiment with different values of T. Taking more periods into consideration after $T = 19$ does not considerably affect results but adds computational costs. In addition, a number of employment sequences grows exponentially with T. Therefore, I also restrict the sequences that have negligibly small probabilities. Appendix A discusses computational details.

I assume that individuals have CRRA utility over payoffs of sequences:

$$u(\Pi_j) = \frac{(\Pi_j)^{1-\rho}}{1-\rho} \tag{8}$$

Recall from the previous section that an individual chooses to buy insurance if $EV(l_{it} = 1) \geq EV(l_{it} = 0)$ or, alternatively, if $EV(l_{it} = 1) - EV(l_{it} = 0) \geq 0$. As noted in Apesteguia and Ballester (2018), such a utility difference has a unique value of a risk preference parameter ρ where $EV(l_{it} = 1) - EV(l_{it} = 0) = \Delta EV_{it} = 0$.[31] Denote a risk preference value ρ where $\Delta EV_{it}(\rho) = 0$ as λ. Any $\rho < \lambda$ would imply that individual should not buy insurance since she does not have sufficiently high risk preference. Similarly, if an individual has $\rho > \lambda$, she should buy insurance. It implies that if one could estimate both $EV(l_{it} = 1)$ and $EV(l_{it} = 0)$ for all individuals i, observed periods t and potential types s, it would be possible to numerically compute a value of risk preferences λ when (i, t, s) is indifferent between paying premiums or not. Note that the threshold would not only differ by individual and time but also by a potential type s, which is unknown but observed by an individual. The following example demonstrates the logic.

Example 3. *Assume for expositional purposes that an individual has the following true employment sequence in the future (always employed):*

$$e = \{1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1\}$$

In the absence of any information about the future, an individual would have the following beliefs about employment probabilities for each of those periods:

$$p = \{0.92 \quad 0.92 \quad 0.87 \quad 0.84 \quad 0.89 \quad 0.97 \quad 0.95 \quad 0.93 \quad 0.93 \quad 0.89 \quad 0.94 \quad 0.95 \quad 0.94 \quad 0.95\}$$

Depending on the individual type $s \in \{1, ..., 12\}$, there might be different threshold risk preference levels at which an individual is indifferent between buying insurance or not. Table 4 demonstrates such an example of a decision of an individual i at time t:

[31] Note that although ΔV_{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. I discuss this in more details in the identification section.

Table 4: Example of a Role of Types

(i,t)		Time										
		1	2	3	4	5	6	...	16	17	18	19	$\boldsymbol{\lambda_{its}}$
Types (s)	1	1	0.92	0.87	0.84	0.89	0.97	...	0.94	0.95	0.94	0.95	12.1
	2	1	1	0.87	0.84	0.89	0.97	...	0.94	0.95	0.94	0.95	8.7
	3	1	1	1	0.84	0.89	0.97	...	0.94	0.95	0.94	0.95	6.1
	4	1	1	1	1	0.89	0.97	...	0.94	0.95	0.94	0.95	4.2
	5	1	1	1	1	1	0.97	...	0.94	0.95	0.94	0.95	2.1
	6	1	1	1	1	1	0	...	0.94	0.95	0.94	0.95	1.2
	7	1	1	1	1	1	0	...	0.94	0.95	0.94	0.95	-0.1
	8	1	1	1	1	1	0	...	0.94	0.95	0.94	0.95	-1.2
	9	1	1	1	1	1	0	...	0.94	0.95	0.94	0.95	-3.4
	10	1	1	1	1	1	0	...	0.94	0.95	0.94	0.95	-5.2
	11	1	1	1	1	1	0	...	0	0.95	0.94	0.95	-6.7
	12	1	1	1	1	1	0	...	0	0	0.94	0.95	-12

Notes: The Table demonstrates an example of the role of types in insurance decisions. It shows that despite having the same beliefs about future employment probabilities, being a different type affects a minimum risk preference value required to rationalize being insured presented in the column λ_{its}.

An example shows various information structures that an individual might have depending on type $s \in \{1, ..., 12\}$. In addition, as the column λ_{its} suggests, even for an individual i and time t, there might be twelve unique different risk preference thresholds depending on a type.

As the example shows, individual conditions that are mapped to her decision to buy insurance can be summarized by a risk preference threshold, which formally is defined as follows:

$$\lambda_{its} = \rho_{its} : \Delta_{its}(\rho_{its}) = 0 \tag{9}$$

Note that everything in equations (9) is known or can be inferred from the data except λ_{its}. It implies that λ_{its} can be computed numerically by solving the model for each i,t,s repeatedly to find a value ρ_{its} that satisfies (9). The only object not observed from the data directly is beliefs about unemployment probabilities ($\{p_k\}_{k=t+s}^{t+T}$) outside of a type-specific known interval. To recover those values I assume that individuals have rational expectations about unemployment outcomes. Therefore, I recover beliefs about probabilities of employment using an equation:

$$Pr(Y_{it} = 0) = Logit(X, Y_{i,t-1})$$

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

Using recovered probabilities for each (i, t), I construct the probabilities ξ_j from (7) using equation (6). As a result, for each individual and period, I obtain twelve different threshold risk preference values where an individual is indifferent between buying insurance or not.

Note that the probability that an individual buys insurance is the probability that her risk preference value is at least as large as the estimated threshold. Although estimated risk preference thresholds do not provide a distribution of risk presences and types, which are fundamental model primitives required for further welfare analysis, it is a key step to recover them.

I assume that risk preferences are normally distributed in the population with a mean $\alpha X'$ and a standard deviation σ, where α is a vector of unknown parameters to be estimated and X is an array of individual characteristics to account for heterogeneity:

$$\rho_{it} \sim N(\alpha X', \sigma) \tag{10}$$

where X contains a constant, binned age, gender, family, higher education dummy, a dummy if there are children in a family, and binned income.

A probability that (i, t, s) with the risk preference threshold λ_{its} buys insurance is a probability that the actual risk preference value is larger than λ_{its}. Given a parametric distribution in (10), this probability can be expressed:

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

where $\Phi\left(\frac{\lambda_{its} - \alpha X'}{\sigma}\right)$ is a cumulative normal distribution denoting a probability that risk preferences are below the threshold λ_{its}.

I assume that a probability that (i, t) is type s has a multinomial logit form:

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

where β_s - vector of parameters affecting a type probability.

There are twelve vectors β_s corresponding to each type. I normalize the first vector by setting all elements to 0.3. Z_{it} is an array containing observables that are expected to affect type probabilities. 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 dummy variables such as industry, occupation type, education level, education specialization, etc. It results in a large set of variables, which makes estimation burdensome since there are eleven parameters in β (the first one is normalized) for each variable in Z. Furthermore, many of those variables in Z are highly correlated since, for instance, education and labor market affiliations are closely related to each other. Therefore, I cluster individuals based on these variables to reduce the dimensionality of variables in Z to five dummy variables denoted as cluster allocations. I also add binned age variables which together with a constant comprise a vector of eight parameters β for each type.

The probability that an individual i, t is insured is:

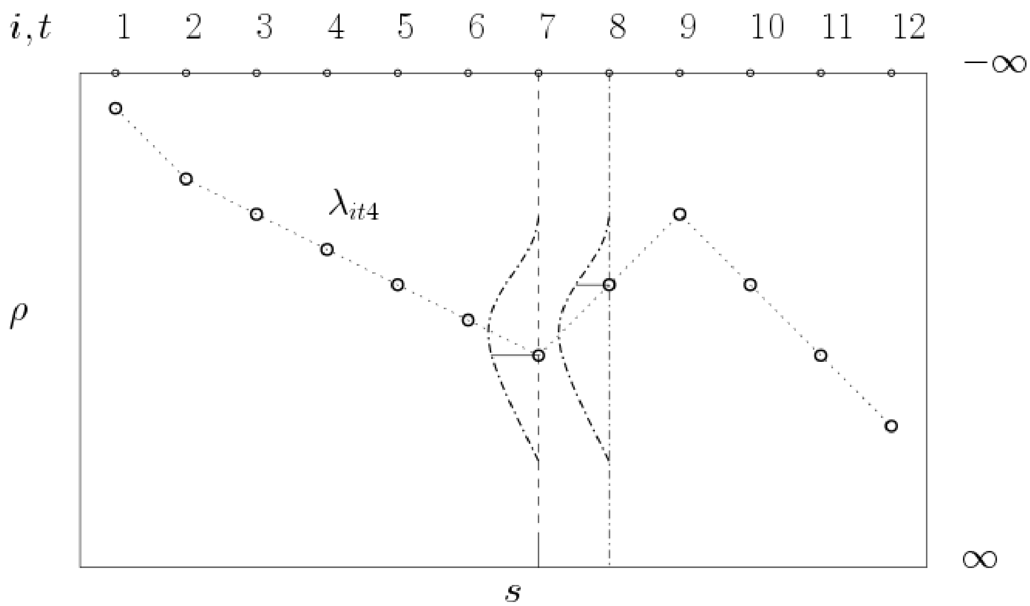
$$Pr_{it}(l = 1) = 1 - \sum_{s=1}^{d_{it}} \phi_{its} \Phi\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right) \tag{13}$$

It yields a likelihood function:

$$L = \prod_i \prod_t \left(\overbrace{1 - \sum_{s=1}^{12} \phi_{its} \Phi\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right)}^{\text{Insurance Probability}} \right)^{y_{it}} \left(\overbrace{\sum_{s=1}^{12} \phi_{its} \Phi\left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma}\right)}^{\text{1 - Insurance Probability}} \right)^{1-y_{it}} \tag{14}$$

To illustrate the logic behind the likelihood function and an estimation approach, consider Figure 6. The Figure illustrates that for a particular individual, time and type s, a probability of buying insurance is just the area to the bottom (towards infinity) of a density function, which illustrates the equation (11). The probability that an individual buys insurance is a weighted sum of those areas where type probabilities from (13) serve as weights.

Figure 6: Probability of Buying Insurance



Notes: The Figure graphically summarizes the estimation logic. A y-axis denotes a range of risks preference values of ρ from $-\infty$ to ∞. The x-axis denotes discrete types from one to twelve marked above the figure. Empty dots on the figure are λ_{its} points. The bell-shaped curves represent probability density functions of risk preference parameters (normal distribution). The area to the bottom of each perpendicular line from the curve to the dashed type-line denotes the probability that an individual of the corresponding type has risk preferences above the threshold denoted by the white-fitted point (cumulative distribution). The probability that (i,t) buys insurance is a weighted sum of each of those areas for each discrete type weighted by probabilities of types (ϕ).

The final piece of the model concerns inertia in insurance choices. Since an individual should make a decision fairly often, it is reasonable to expect considerable inertia in choices, which is supported by the evidence that many individuals are either never or always insured. A failure to take into account inertia might lead to misleading evaluation of willingness to pay. For instance, the model would suggest that an individual is always insured because of a particularly high willingness to pay, which in fact might be a consequence of inertia. Consequently, policy recommendations might wrongly suggest providing insurance at extra government costs. Despite an importance of accounting for inertia, it is challenging to identify from the data. The literature studying or attempting to take into account switching costs either look for exogenous variation where a group of individuals makes a decision without a default option (Handel, 2013) or impose additional parametric assumptions. Unfortunately, the institutional environment fails to provide compelling variation that would allow identifying inertia. Therefore, with a view to its importance, I opt for the second option. I augment a probability of insurance expression with an extra term $\Upsilon = \frac{1}{\gamma}(1 - l_{t-1}) + \gamma l_{t-1}$:

$$Pr_{it}(l = 1) = \left(\sum_{s=1}^{12} \phi_{its} \left(1 - \Phi \left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma} \right) \right) \right)^{\Upsilon} \tag{15}$$

where l_{t-1} - previous insurance status; γ - proportional inertia coefficient.

Such a specification allows for additional flexibility of the model. I do not allow heterogeneity in inertia. The intuition for such parametrization is that when insured individuals are more likely to keep being insured, Υ will be a large positive number, which moves probability $\Phi\left(\frac{\lambda_{its}-\alpha X'_{it}}{\sigma}\right)$ towards zero. In contrast, if previously uninsured individuals are more likely to keep being uninsured, Υ will be close to zero, which forces $\Phi\left(\frac{\lambda_{its}-\alpha X'_{it}}{\sigma}\right)$ to go to one and, thus the insurance probability to zero. I restrict Υ to be weakly larger than zero. It yields a modified likelihood function:

$$L = \prod_i \prod_t \left(\overbrace{1 - \left(\sum_{s=1}^{12} \phi_{its} \left(1 - \Phi \left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma} \right) \right) \right)^{\Upsilon}}^{\text{if insured}} \right)^{1-y_{it}} \cdot \left(\overbrace{\left(\sum_{s=1}^{12} \phi_{its} \left(1 - \Phi \left(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma} \right) \right) \right)^{\Upsilon}}^{\text{if uninsured}} \right)^{y_{it}} \tag{16}$$

To sum up, a modeling 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 needed 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).[32]

4.3 Empirical Identification

The model outlined in the section has two components. The first one is a computation of risk preference thresholds that denote the risk preference value at which an individual i, at time t and type s is indifferent between buying insurance or not. The second component is finding parameters of a parametric distribution of risk preferences, inertia, and a type distribution which

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

maximize the likelihood of observing actual insurance choices.

Therefore, identification of model parameters, firstly, requires uniqueness of risk preference thresholds for each i, t, s since it should be only one corresponding probability of being insured. The uniqueness results follow from Apesteguia and Ballester (2018). As shown by Apesteguia and Ballester (2018), for lotteries that are not first order stochastic dominance related, the utility difference $\Delta = EU_{insured}(\rho) - EU_{uninsured}(\rho)$ is increasing in some interval but since any CRRA utility function approaches to zero as $\rho \to \infty$, the difference also converges to zero. It implies that multiple risk preference parameters yield the same utility difference, which creates an identification problem. It also means that a probability of being insured does not monotonically rise with a degree of risk aversion because often used additive error terms start dominating the utility difference at large risk preference parameter values. Although the utility difference between two lotteries in CRRA function is non-monotonic in a risk preference parameter, which complicates the estimation of discrete choice models under uncertainty, the approach in this paper is immune to this issue. For the case of choices with a dominant option, there is no indifference point which implies that a dominant option should be chosen. For the most cases where being insured and uninsured might be preferred depending on the risk preference value, the indifference point is unique and computed numerically. Although the logic of estimation is in the spirit of a proposed solution in Apesteguia and Ballester (2018), it is more in line with the estimation approach of choice models under uncertainty in Cohen and Einav (2007) and Einav, Finkelstein, and Cullen (2010). The estimation approach only requires obtaining the points of indifference between being insured and uninsured uniqueness of which follows from Apesteguia and Ballester (2018).[33,34]

Since for each (i, t, s) there is a unique threshold value, I identify parameters (α, σ) of type-conditional insurance probabilities $P_{it}(l = 1|s) = 1 - \Phi(\frac{\lambda_{its} - \alpha X'_{it}}{\sigma})$ using over-time variation in insurance premiums and generosity of benefits, and cross-sectional variation due to non-linear benefits structure.

The identification of the parameters of a type distribution $\{\beta_s\}_{s=2}^{12}$ exploits patterns of a timing of insurance purchases relative to the timing of unemployment and changes in unemployment risks. The source of identification of type distribution parameters is similar to the time-selection evidence. Risk preference indifference points vary depending on how many periods are observed

[33]Apesteguia and Ballester (2018) does 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.

[34]Although the indifference point is theoretically unique, it is not computationally true 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 the Appendix A

in the future. Note that time-selection patterns are unrelated to a risk preference distribution since only superior information about the unemployment timing can rationalize observed time-selection patterns. More precisely, an individual (i, t) has risk preferences fixed across types but computed risk preference thresholds vary by type. Hence, it informs type distribution parameters conditionally on observed relevant labor market characteristics included in Z.

Finally, an inertia parameter is identified from a functional form assumption and should be viewed as a weakness of the model, which are the costs paid to capture choice persistence. Although this assumption has an important effect on the estimated parameters and derived objects such as willingness-to-pay, the policy conclusions are robust to a functional form of an inertia component or even to its inclusion in the model, which I discuss later in the paper.

4.4 Parameter Estimates and Model Fit

The model outlined in the previous section has 13 parameters of a risk preferences distribution, an inertia parameter, 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 A provides more details of estimation of parameters and standard errors. Table 5 presents risk preference and inertia parameters:

Table 5: Parameters of a Risk Preference Distribution and Inertia

	Coefficients	Std.Errors
α: Constant	-1.543	(0.105)
α: Age $(30; 40]$	4.263	(0.042)
α: Age $(40; 50]$	-3.919	(0.04)
α: Age > 50	2.307	(0.11)
α: Gender	1.08	(0.098)
α: Family	-1.628	(0.136)
α: Higher Education	-0.75	(0.105)
α: Has Children	1.389	(0.07)
α: Income $(25\%; 50\%]$	-5.169	(0.08)
α: Income $(50\%; 75\%]$	-2.456	(0.038)
α: Income $> 75\%$	-3.089	(0.2)
σ: Std. Deviation	39.9	(0.6)
Υ: Inertia	0.003	(< 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. 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.

The Table shows that higher income individuals tend to be less risk-averse while no clear monotonic pattern for age is observed. Being a female, unmarried, without higher education and having children is associated with higher risk aversion. It implies that those characteristics increase the probability of buying insurance since higher risk preferences are associated with a higher likelihood of buying insurance conditionally on unemployment risks and information structure. The model suggests a 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 and willingness-to-pay for the welfare analysis. Although there might be an interest to interpret parameters to shed light on risk preference heterogeneity, I view this discussion as having a limited use since those risk preferences are very specific to the context of UI and affected by inertia parameters that lack identifying variation in the data.

The model also shows an important role of inertia implied by the corresponding parameter that takes a value of 0.003. To put this into perspective, an individual who has a probability of buying insurance 0.8 in the absence of inertia has a probability 0.999 conditionally on being insured before and $4.97 \cdot 10^{-33}$ if uninsured before upon adjusting for inertia.

In addition to a risk preference distribution, the model generates 88 parameters of a type distribution. Table 8 with parameters and standard errors is included in the Appendix C. Table 6 summarizes the information from type parameters by presenting a resulted type probabilities in the estimation sample.

Table 6: Type Probabilities

Type	Probabilities	
	Mean	Std. Dev.
I	0.006	0.003
II	0.05	0.01
III	0.001	<0.001
IV	<0.001	<0.001
V	<0.001	<0.001
VI	<0.001	<0.001
VII	<0.001	<0.001
VIII	<0.001	<0.001
IX	0.006	0.003
X	0.003	0.002
XI	0.9	0.02
XII	0.027	0.016

Notes: The Table shows type probabilities, which are determined by the estimated type parameters. It suggests that most individuals have information about eleven periods in the future. Many people also have information about two and twelve periods. Remaining probabilities are considerably smaller.

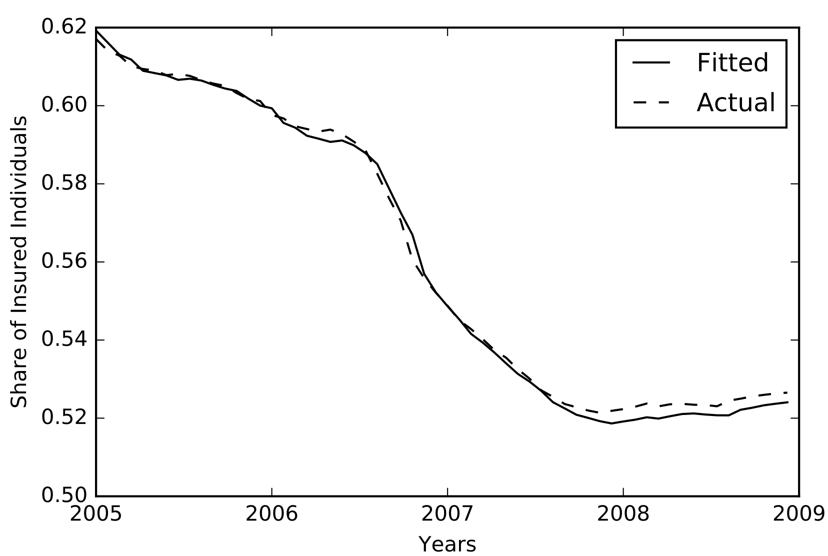
The type distribution estimates suggest that around 90% of individuals must know 11 periods in the futures based on their revealed behavior. Around 5% and 2.7% of individual-time choices are based on knowing two and twelve periods in the future while other information structures are considerably less common. An important clarification concerns the interpretation of these probabilities. On the one hand, they represent the formal labor market and legal requirements that restrict the timing of layoff. On the other hand, these probabilities are also affected by the presence of temporary contracts, probation periods, and, perhaps more importantly informal information sharing between firms and workers. Therefore, the recovered beliefs are a combination of both, legal requirements and informal information transmission.

Table 7: Model Fit - Share of Insured Individuals by Subgroups

	Shares of Insured Individuals	
	Actual	Predicted
Age ≤ 30	0.569	0.568
Age $(30; 40]$	0.561	0.561
Age $(40; 50]$	0.562	0.561
Age > 50	0.555	0.554
Gender	0.572	0.571
Family	0.562	0.561
Higher Education	0.558	0.557
Has Children	0.563	0.563
Income $\leq 25\%$	0.571	0.569
Income $(25\%; 50\%]$	0.563	0.563
Income $(50\%; 75\%]$	0.558	0.557
Income $> 75\%$	0.556	0.555

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

Figure 7: Model Fit - Demand



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

The Figure 7 and the Table 7 suggest that the model predicts insurance patterns that closely match actual evidence both over time and by subgroups. The quality of the in-sample fit, however, should be taken with a caution as a measure of the validity of the model since it is not surprising that such a rich model fits the data well.

5 Welfare

This section discusses how the estimates of the model are used to compare various regulations in UI. I focus on analyzing an effect of mandating the system and the introduction of alternative insurance contracts. 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 capture various contract design dimensions and target specific features of private information. Firstly, I consider a contract with fixed costs amounting to 6 times monthly fees to be paid when entering the insurance pool for the first month in a sequence. It should presumably 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 a specific month and has a prespecified duration. I look at 18 and 24 months contract durations. I do not consider a 12 months contract, for example, because model parameters suggest that some individuals might have information about up to 12 months in the future. As a result, this contract does not leave much uncertainty and should be avoided. An open enrollment contract is aimed at eliminating time-selection and reduce choice inertia. Since choice inertia should primarily come from a high frequency of monthly choices, open enrollment contracts, in contrast to a current system and an entry costs contract, should be free from inertia. Welfare analysis is based on the pooled sample of individuals over years 2005-2008 (48 months).[35]

[35] I drop 2009 since to consider 18 and 24 months contract I need to have a number of periods covering the contract length. The price for those contracts is a monthly price paid each month of the duration of the contract to make it comparable to the current system. Therefore, an open enrollment 24 months contract contents two choices that individuals make during the period under consideration, whereas 18 months contract includes three choices where the outcomes under the third choice are truncated at first 12 months.

5.1 Measuring Welfare

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

To understand the effect on consumers, it is required to recover individual willingness to pay for a particular insurance contract, which, in other words, is the maximum price that she is willing to pay. Consumer surplus (CS) can then be measured as a difference between WTP and paid price. In a typical consumer behavior model a WTP measure also directly translates to the demand function since individuals should buy a product only when the willingness to pay is above an offered price. This is not the case in this setup since the presence of inertia implies that individuals might not behave fully optimally. It means that even if WTP is lower than a price, individuals might still keep being insured, or similarly, keep being uninsured when WTP is larger than a price. Therefore, in addition to WTP, one needs to recover a demand function under various contract structures to characterize a take-up response to various policy changes.

Analysis of the effect of government budget requires two main components: demand function 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 can be obtained, I formally define how I construct the required 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 was observed for "worker" premiums and most of the time individuals are employed, the latter 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 fix the actual price for unemployed. Note also that all components required for welfare analysis are contract/regulation-specific and should be separately obtained for each considered policy k.

Since a key sufficient statistics in the model is a threshold risk preference parameter 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 potential 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 6. Using the same procedure as before, I compute an array of risk preference thresholds $\lambda_{itk}(g)$.

Given obtained thresholds and risk preference parameters, I calculate a share of insured individuals under some policy k and price g as follows:

$$\zeta_k(g) = \sum_i \sum_t \zeta_{itk}(g) = \sum_i \sum_t \left(1 - \Phi\left(\frac{\lambda_{itk}(g) - \alpha X'_{it}}{\sigma}\right)\right)^{\Upsilon} \tag{17}$$

The expression follows directly from an insurance probability formula defined in the estimation section and sums probabilities up over individuals and periods.

The last element required to evaluate consumer welfare is consumer surplus, which is a difference between willingness to pay and a price when an individual decides to buy insurance. I assume that consumer surplus is zero for uninsured except for the case of a mandatory system when it might become negative since everyone is forced to buy insurance.[36] As discussed above, take-up probabilities are not directly linked to WTP because of inertia, which implies that, for example, an individual might have high WTP but low probability of buying insurance.

To calculate an expected WTP for each (i, t), I use the following approach. A threshold recovery procedure allows obtaining maximum risk preference values at which insurance would be bought under each policy k and price g for each individual and time period (λ). Since the threshold 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:[37]

$$E[WTP_{itk}] = \int_{\hat{\lambda}} \hat{g}_{itk}(\hat{\lambda}) dF(\hat{\lambda}; \alpha X'_{it}, \sigma) \tag{18}$$

where $F(\hat{\lambda}; \alpha X'_{it}, \sigma)$ 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

[36] Conceptually, consumer surplus might also become negative in the case of a voluntary system due to inertia. It is, however, theoretically unclear if those negative values should be truncated at zero since the surplus loss stems not from the government actions, which is a focus of the paper, but from optimization errors. In the analysis here I do not take into account consumer welfare loss as a result of inertia because of this conceptual ambiguity and, since the identification of inertia is problematic in this case, which raises concerns regarding the use of recovered inertial costs.

[37] I use 100 quadratures to obtain the integral numerically. Instead of integrating from $-\infty$ to ∞, for each case I find the risk preferences that correspond to 0.1% an 99.9% percentiles. Then I construct equally spaces bins and integrate within this interval with 100 quadratures after reweighing bin probabilities to ensure that they sum up to 1. Since computational procedure allows obtaining $\lambda_{itk}(g)$ on a grid of values g because the function cannot be derived analytically, I use linear interpolation to fill the values between grid points in the integration.

average of WTPs resulted from all potential risk preference values weighted by a probability of having each of those values. Consumer surplus is defined as:

$$CS_k(g) = \sum_i \sum_t t_k \left(\zeta_{itk}(g) \cdot (E[WTP_{itk}] - g)\right) \tag{19}$$

where $t_k(\bullet)$ is a contract-specific function that truncates the difference to zero in the case of a voluntary system and does not affect the expression in the case of a mandatory system.

To analyze the consequences for a government budget, I need take-up probabilities obtained above and expected costs of insuring each i, t. To recover expected costs under voluntary and basic insurance denoted as H_{itk}^{vol} and H_{itk}^{out}, I use 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. Predicted probabilities of employment together with information about benefits structure allows constructing expected costs of covering each individual under both basic and supplementary insurance systems. More precisely, the expected costs of insuring an individual (i, t) are unemployment probability times a benefits level. The benefit levels depend on income, replacement rate, and a cap, or fixed benefits in the case of the basic insurance. I also adjust the amount for benefits taxation.[38] I take into account basic income expenditures since a counterfactual mandatory system would remove current basic insurance and replace a voluntary system with one mandatory system. Therefore, to adequately compare costs, it is required to take into account the costs of providing basic insurance.

A total cost function can be obtained:[39]

$$TC_k(g) = \sum_i \sum_t \left[\zeta_{itk}(g)(H_{itk}^{vol} - g) + (1 - \zeta_{itk}(g))H_{itk}^{out}\right] \tag{20}$$

The welfare analysis requires choosing a welfare criterion. One of the approaches for insurance markets is described in Einav, Finkelstein, and Cullen (2010). It aims at finding a point where a demand function (willingness to pay) intersects a marginal cost function. In other words, the optimal price would be at the point where it is not optimal to reduce price any further since an additional individual does not value insurance more than she costs to an insurer. Welfare costs

[38] In the welfare comparison later in this section, I scale the costs down by 30% to account for the fact that insurance benefits are taxable. It is difficult to exactly determine the actual tax that would be paid on those monthly benefits since individual income taxation in Sweden has a non-linear structure with 18 000 SEK annual exemption and around 32% tax rate above the exemption rate up to 468 700. Another two brackets add 20% and 5% to marginal tax value. As a result, determining the exact tax on benefits is complicated. Therefore, I use 30%. Note that the choice of this constant does not have any effect on the policy analysis since it just rescales axes but leaves curves unchanged relative to each other in Figures 11 and 12.

[39] Note that in the case of mandates the costs of basic coverage will be zero since everyone is automatically enrolled in the voluntary system.

can then be measured with using optimal and actual prices. This approach, however, implies a number of complications for my case. Firstly, the specific structure of the model allows obtaining only expected willingness to pay, which is not directly related to the demand. It stems from the uncertainty over actual risk preferences and the presence of inertia, which leads to the fact that some individuals might be insured while having a low value of insurance. Such a welfare criterion would suggest excluding those individuals from the insurance pool since they cost more to the government in comparison with their willingness to pay. As a result, in the extreme case this criterion would suggest raising the price up to the point where as insurance pool is empty.

Instead, I opt for another approach. I evaluate the welfare by comparing systems under various government expenditure levels in terms of the total consumer welfare. Recall that the model allows obtaining total consumer surplus $CS_k(g)$ and total government costs $TC_k(g)$ under a system k and price g defined in (19) and (20), correspondingly. It implies that those functions can be combined into the correspondence:

$$CS_k(g) \widehat{=} TC_k(g) \tag{21}$$

The equation (21) is a correspondence since it is not guaranteed that each price gives a unique pair of total costs and consumer surplus.[40] 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. 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 high 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 the 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 $\forall \varepsilon_k(\chi)$ 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 of budget costs for the government even within the same system. At the same time, it also takes

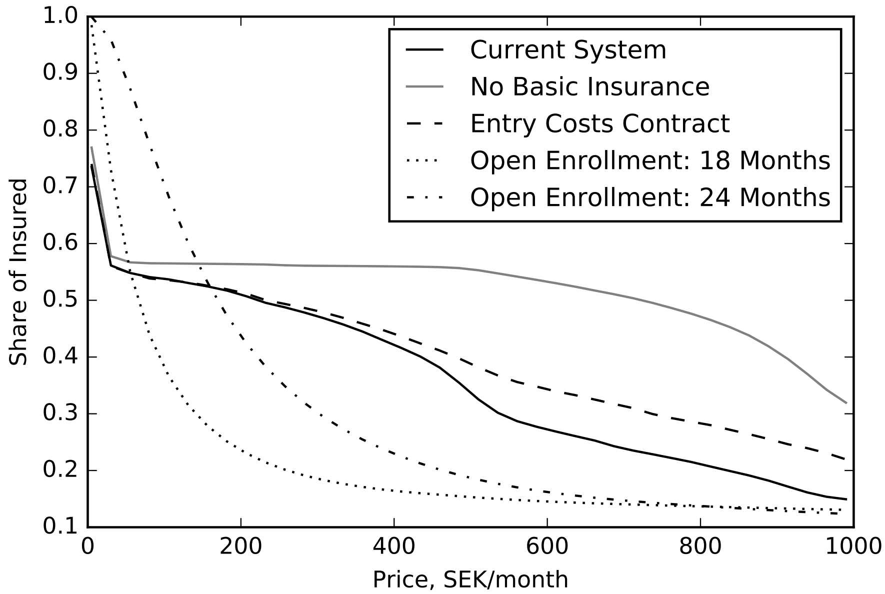
[40] The reason is that a change in prices affects both probabilities of insurance, which also translates into changes in risk composition among insured individuals, and government revenues through the sum of collected premiums.

into account the fact that some budget costs are unattainable for some systems. It implies that systems can be directly compared only when they both can be used to reach some budget balance. It is especially important when analyzing mandates since these policies should theoretically be able to support a wider range of χ because of restrictions on individual responses.[41] Finally, this framework takes into account a complication arising from inertia by considering a total consumer surplus instead of minimum individuals willingness to pay among insured.

5.2 The Welfare Consequences of Alternative Designs

As discussed in the previous section, potential changes in the structure of the contract and prices have a number of channels through which welfare is affected. Firstly, individuals react to those changes by enrolling or leaving an insurance pool. Figure 8 demonstrates a counterfactual demand under various considered policies.

Figure 8: Counterfactual Policies Demand



Notes: The Figure demonstrates the demand function of a current system, a system without basic insurance, which would correspond to a mandatory insurance case, a system with an entry costs contract and open enrollment contracts with 18 and 24 months durations.

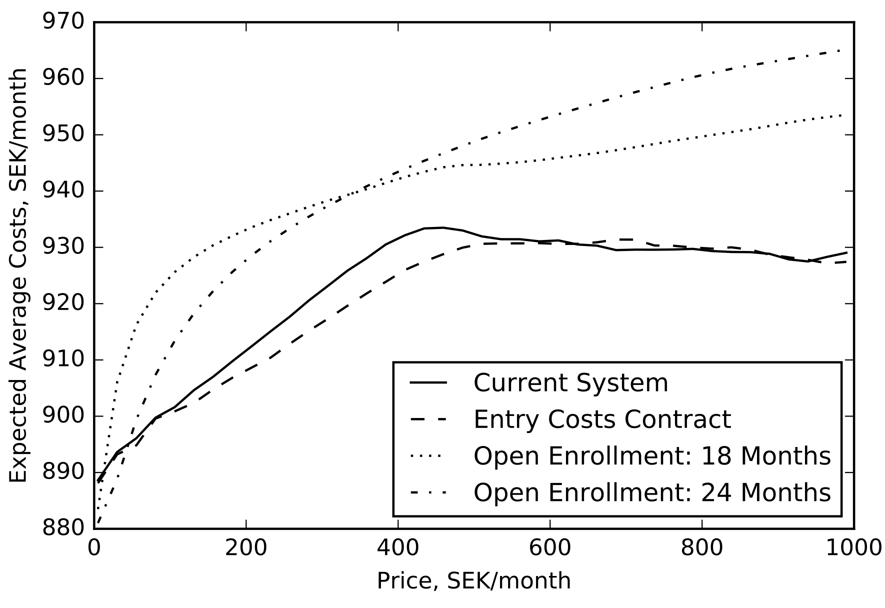
Figure 8 suggests that, firstly, a presence of basic insurance plays an important role in individuals' decisions to buy insurance and a current system would attract more individual in the absence of this outside option. This observation comes from the comparison of the gray

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

curve for the "No Basic Insurance" case and a black solid line for the "Current System" case. Secondly, an entry costs contract does not differ much for low and high prices because entry costs are proportional to monthly fees and they do not matter for small prices. For high prices when a current contract take-up falls to about 10% coverage, entry costs stop playing a role since individuals who buy insurance unambiguously benefit from coverage. Those individuals are, for example, currently unemployed and losing insurance results in unambiguous monetary losses. Finally, demand functions for open enrollment contracts, in contrast to other curves, are smooth and very steep primarily because of the absence of inertia. In addition, a 24 months contract demand function is less steep since it involves more uncertainty for individuals. In other words, an 18 months contract would allow them to buy new insurance sooner, thus it is less risky to be uninsured now.

The second policy-relevant dimension is budget costs. Figure 9 plots average cost functions obtained from the expression (20) but disregarding the costs of providing the mandatory insurance and taxes on the benefits for expositional purposes:

Figure 9: Average Cost Functions



Notes: The Figure demonstrates average costs of insuring individuals under a voluntary system disregarding the presence of the costs of basic insurance and benefits taxation. The curve is obtained by dividing each value in a cost function by an expected number of insured individuals.

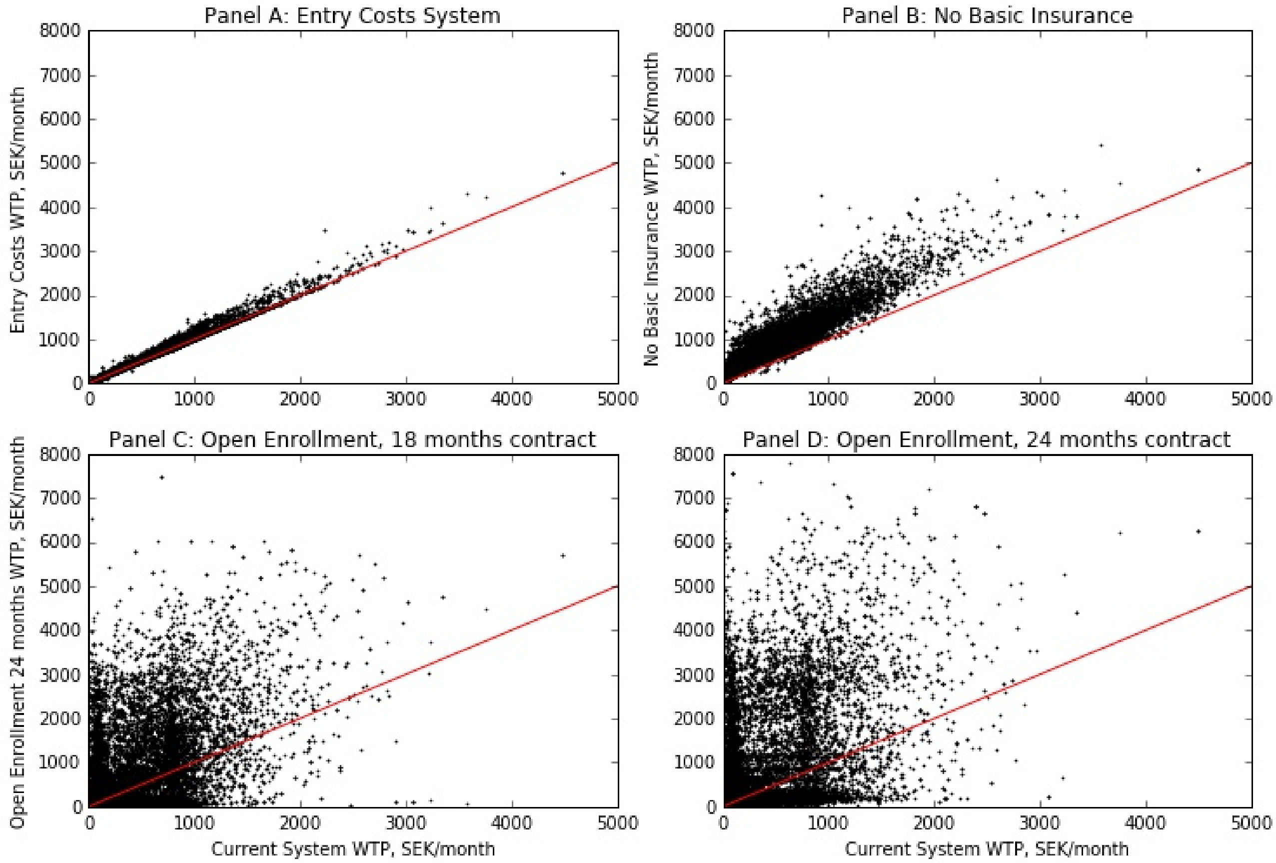
Presented cost functions show mixed evidence regarding selection. Current and entry costs contracts are associated with clear and fairly steep positive slope up to a price of around 400

SEK/month.[42] After this price, the average cost curves become fairly flat and even with a slightly negative slope. This region after 400 SEK/month corresponds to a part of the demand functions fall more sharply for current and fixed costs contracts, and almost reach the bottom for open enrollment contracts. Slight reversal of slopes for current and fixed contract curves, however, is attributed to the composition of expected costs and an interplay with the enrollment responses. Average costs for open enrollment contracts display a steady increasing trend that becomes less steep around 400 SEK/month for an 18 months contract and around 600 SEK/month for 24 months contract. In both cases, those relatively flat regions are associated with flat regions of the corresponding demand curves where only slightly more than 10% of individuals remain insured.

Recall that presented demand functions not only take into account willingness to pay but also inertia, which jointly affect the resulted insurance status. However, for calculating consumer surplus, only willingness to pay is relevant. Figure 10 demonstrates a joint distribution of expected WPTs of alternative systems in comparison with a current system. The Figure summarized the attractiveness of alternative contracts in comparison with a current system. The red line allows comparing distributions of WTP. Points that lie above the red line mean that a given individual value an alternative contract more than a current system since she is willing to pay more. Panel A demonstrates that current and entry costs contracts are very similar, which is in line with the evidence from demand functions. However, almost all points lie slightly above a red 45^0 line, which means that entry costs contract is valued more. The reason is that a contract imposes costs on future re-enrollment thus individuals have more value to pay premiums now to avoid paying entry costs later. As expected, individuals would value the current system much more in the absence of a basic insurance, which is displayed on Panel B. Panel C and D show considerable differences compared to a current system. Such differences are primarily explained by the fact that open enrollment contracts change the timing of decisions which imply that choices are made under different information structures. Overall, individuals value such open enrollment contracts more primarily because they remove waiting periods and thus provide eligibility from the beginning of a contract.

[42] A positive slope corresponds to a negative slope if plotted in terms of a number of insured as, for instance in Einav, Finkelstein, and Cullen (2010)

Figure 10: Comparison of WTP under various 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 an individual. If a given point lies above the red line, the corresponding alternative contract is valued more by this individual.

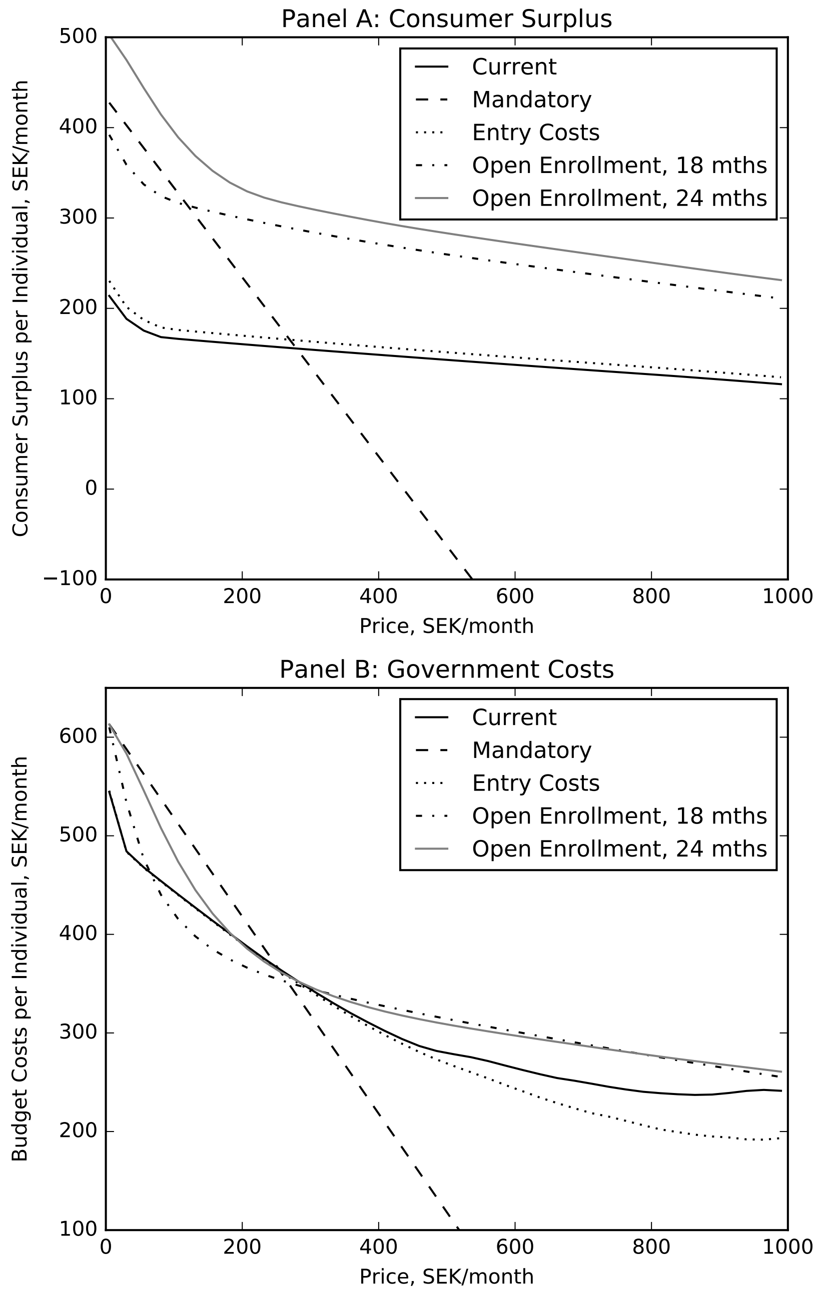
Recovered willingnesses to pay, demand and cost functions are the elements required to compare various contract structures. To do that, I plot the consumer surpluses resulted from a set of considered prices against budget costs that would be needed under these prices. For system or a contract structure to dominate another one under some budget cost level, it must be that it always yields higher consumer surplus.

I firstly plot consumer surplus and resulted budget costs presented in Figure 11 Panels A and B, correspondingly. Panel A shows that consumer surplus is mostly affected by a price raise at low price levels since an increase in prices have a small effect when a small share of people remains insured. Under a mandatory system, any price increase has a constant effect on consumer surplus since individuals are not allowed to unenroll. A similar pattern is presented on Panel B. Mandatory system is capable of considerably reducing budget costs buy increasing price since individuals are locked in and cannot unenroll under no moral hazard assumption. All alternative voluntary system contracts would allow reaching a cost level up to 200 SEK/month per individuals (entry costs contract) within a considered price range. There are two reasons why it is more difficult to control costs under a system with consumer choice. Firstly, the government still needs to finance basic insurance under which premiums are not paid. Secondly, since every price increase results in a lower number of individuals who pay benefits, spending stop falling at some price level and even start growing. This is the example a correspondence between budget costs and consumer surplus levels. As Panel B shows, budget costs curve for a current system has several points where different prices yield the same budget costs levels. Since consumer surplus is monotonically increasing on a Panel A, it implies that there would be several consumer surplus levels corresponded to the same budget level. Total costs are lower for current and entry costs contracts compared to the open enrollment contracts due to inertia.

To conclude which contract structure welfare dominates other competing designs at some government costs, one should compare the resulted consumer surpluses at various budget costs. It shows which system generates higher benefits to individuals while requiring the same government expenditures. 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 might have different support for various systems in terms of costs.

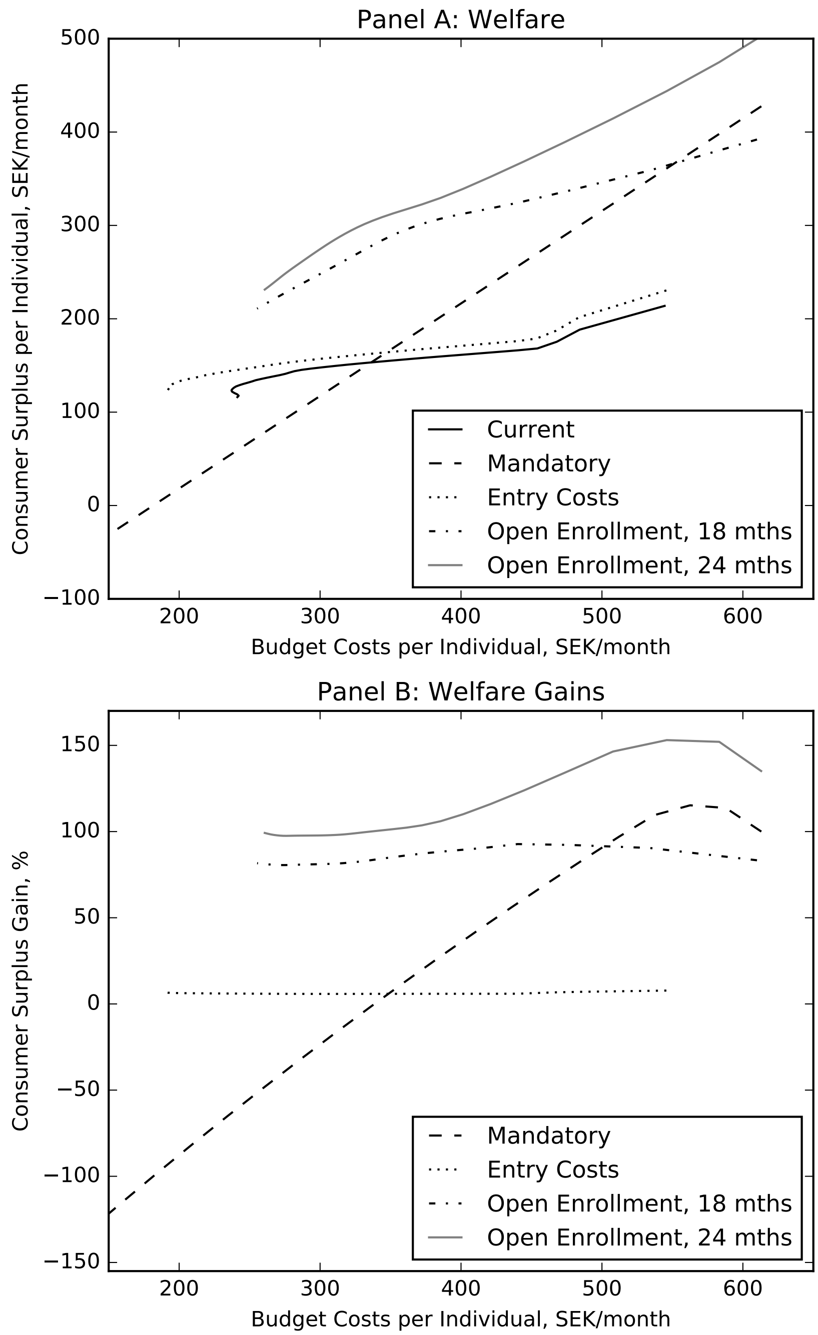
Figure 12 summarizes welfare analysis. Panel A demonstrates relationships between government costs and generated consumer surplus under considered policies. This evidence follows from Figure 11. In other words, Panel A represents the y-axis of Panel A plotted against the y-axis of Panel B from Figure 11. The main point of the Figure is to show which policies result in higher consumer surplus while requiring the same subsidy levels. This approach allows being agnostic about optimal pricing and just compare various system based on the required budget costs. 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. Before interpreting the policy insights of Figure 12, it is important to stress that all but mandatory system curves cover only a part of the x-axis range. It stems from the fact that only limited cost levels can be reached without mandates as described in Figure 11. In addition, since these are the correspondences, multiple consumer surplus levels might correspond to the same budget expenditures. 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 to be the case only in very rare cases for some budget levels of a current contract around 250 SEK/month on the x-axis. Furthermore, it does not affect any part of the analysis since even in the case of multiplicity, the gatherings of points always lie above or below other curves.

Figure 11: Effect of Premiums on Consumer Surplus and Budget Costs



Notes: Panel A plots a monthly price against consumer surplus. I divide total costs by a number of "active" individual-months observations for expositional purposes instead of presenting the sum of the expected costs over this period. Panel B presents relationship between premiums and resulted budget costs.

Figure 12: 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 as percentage welfare gains compared to a current system at the corresponding budget level. The interpretation of the plots is that the system dominates another one under some government cost level if it lies above another one 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 yields small welfare improvement. Although a mandate can generate sizable welfare gains for high UI spendings, it might be very detrimental for scarce budget cases. The results suggest that an open enrollment contract would be the best option for nearly all levels of expenditures. To put these findings into perspective, consider Panel B, which replicates the same evidence as on Panel A but in terms of consumer surplus gains in comparison with a current system. It suggests that within the considered range of the government costs, entry contract results in roughly 6.14% higher consumer surplus, on average along the line. The intuition is that a lower price is required for the entry costs contract to achieve each budget balance or/and more people remain in the insurance pool, which increases overall consumer surplus. Both current and entry costs contracts dominate the mandatory system if the government aims at reducing the costs below approximately 350 SEK per individual-month (but cannot reduce lower than 200 SEK per individual-month within the range of considered prices). At the same time, if the government is willing to increase insurance spendings, a mandatory system would welfare dominate those contracts. To be more precise, mandatory contract on average results in a large consumer surplus loss of 168%. It is primarily driven by the presence of large losses under high prices that dominate in magnitude gains under low prices. Gains from a mandate reach 115%, whereas losses pick at 572% under the lowest consider budget cost level. The reason is that a mandatory system does not allow individuals to sort themselves based on the insurance needs, thus price changes have a large effect on consumer surplus. If the government is willing to contribute more and a fairly low price suffices, benefits of risk pooling generate higher consumer welfare. In contrast, a smaller budget requires raising prices further, which under reduces welfare for all consumers. Under the voluntary system, individuals without a need for insurance have an option to unenroll, which mitigates this surplus loss. It is important to note that a potential moral hazard effect is not taken into account here. I view this considered price range being fairly modest and should not cause a large moral hazard response.

Finally, the results suggest that open enrollment contracts would virtually welfare dominate all other options. The average gains amount to 83% for 18 months and 106% for 24 months contracts compared to the current structure, correspondingly. An 18 months contract is less efficient than a mandatory system under high subsidies (budget costs more than 500 SEK/month per individual) but 24 months contact always welfare dominates mandatory system within a budget range under consideration. There are two features of open enrollment contracts that make them an attractive option from the welfare point of view. First, this design is free from inertia and, as a result, individuals follow their preferences more closely while choosing insurance. Secondly, although there is a danger of exacerbating adverse selection as a result of an inertia reduction, a particular structure of contracts provides sufficient selection restrictions.

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 those 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 actual risk preference parameters must be larger in the missing population, those individuals have less information about their employment perspectives (types) or even stronger inertia. To examine the importance of sample selection for the welfare analysis, I pretend that all the missing individuals are always insured or always uninsured, which does not change welfare conclusions qualitatively. I also use obtained model parameters to simulated insurance decisions for a missing sample. The welfare analysis remains robust to these changes. This discussion does not take into account the fact that a missing population might have different preferences and information. However, at least within the scopes of estimated parameters the welfare conclusions are robust.

In the model, I assume that a planning horizon T is limited to 19 periods. Experimenting with different options around the chosen value does not affect results considerably. I also use various specifications for the estimation of beliefs, which are fed into the decision model. Variables entering the prediction model also have a minor effect on the estimates since estimates of type distribution absorb those changes.

The most concerning assumption is a functional form of inertia because there is no variation in the data that separately identified this parameter. There is a multitude of alternative functional forms that could be used to incorporate switching costs. I choose a given option because of its restrictive nature to limit the impact that it has on other identified parameters. Changes in this assumption do have an impact on estimates and willingness to pay. However, since the main purpose of the analysis is to provide a comparison between current and alternative regulations in UI, this assumption does not qualitatively change conclusions. To be more precise, excluding inertia, allowing for heterogeneity in inertia, introducing separate inertia coefficients for previously insured and uninsured and assuming that only previously insured individuals are affected does not change any of the conclusions presented in this section. Therefore, although this functional form assumption is crucial for quantitative results, it does not affect the policy suggestions based on the empirical analysis in this paper.

Finally, a bigger picture concern is the validity of such a neoclassical-type model that to a large extent disregards more sophisticated behavioral mechanisms. However, data shows that individuals react to incentives as expected (e.g. higher prices, less generous insurance, and lower risks reduce the demand for insurance). All the 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 might include contract design regulations, which 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, such an 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 potential policies and considering contract design as an alternative to widely-discussed pricing regulations and mandates. Moreover, the results suggest that those 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 without taking into account the potential moral hazard response to mandates, it might 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 opened 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.

Cohen, A., & Einav, L. (2007). Estimating risk preferences from deductible choice. *American economic review*, *97*(3), 745–788.

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 Estimation Details

The estimation procedure in this paper consists of two steps: computation of risk preference thresholds and estimation of parameters. I firstly compute risk preference thresholds where individuals are indifferent between buying insurance or not. To do that I need to solve a dynamic programming problem for each individual i, time t, type s, each potential employment sequence and each duration ($l \in \{0, 1\}$). A major complication arises with 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 a duration of a planning horizon to 19 periods. It does not fully resolve the issue with a number of sequences 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 proceed as follows. 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 99%. It allows keeping computation feasible. I use the bisection method to compute thresholds where the expected utility of buying insurance equals to the expected utility of being uninsured. Although the bisection method is slower than, for example, Brent method, it is safer for this type of non-monotonic problems. It requires imposing bounds, which I set to very large and very low-risk preference values. This also allows solving the issue that although the utility difference has the unique value of risk preferences, where it equals zero, it might become zero because of numerical constraints. It comes from the fact that this difference has a limit of zero as a risk aversion coefficient goes to infinity.

For some cases, the CRRA utility function yields non-finite values. For those cases, I create a grid of large and small values. In this case of non-finite values, next value in the grid is tried until the value that yields a finite value is found. In the case when both acceptable bounds yield negative or positive utility differences, I code a difference as never-insured or always-insured, correspondingly. Those values then coded with the largest or the smallest bound, correspondingly.

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 servers with the data. I pre-compile all computationally intensive parts of the code using a just-in-time compiler, which provides a significant speed-up. I use 50 cores in the estimation. As a result, computing thresholds for the 25% 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 the model at each call of an optimization algorithm.

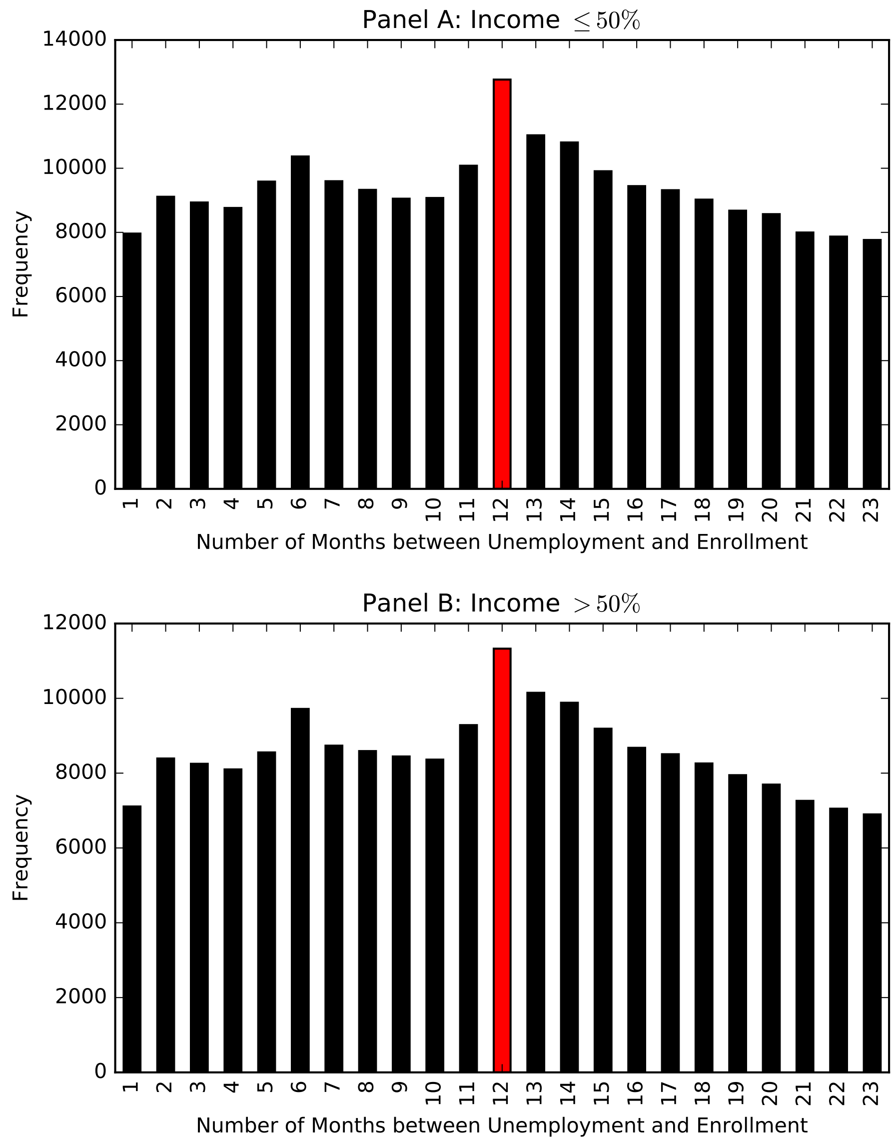
The second stage concerns computing parameters based on the estimated thresholds using a maximum likelihood procedure described in the main text. I use a L-BFGS-B algorithm with bounds on parameters and a user-defined analytical gradient function. I do not use asymptotic standard errors since some of the type parameters are supposed to be as large or as small as possible. As a result, a parameters search algorithm climbs up along the likelihood curve until the derivative becomes close to zero where it stops. The issue is that the second derivative of the likelihood function does not exist in this case since the function is just a zero-slope line with respect to this parameter around the optimum. Therefore, an attempt to obtain a numerical or analytical Hessian matrix fails since it contains non-finite values and thus is not invertible, which is required for asymptotic standard errors.[43]

Therefore, I opt for the bootstrap. I use 100 draws with replacement and estimate the model in parallel on 20 cores. I use this obtained distribution of parameter estimates to calculate standard errors presented in this main text. I use only 100 draws since it takes around 8 hours for an optimizer to converge.

[43] Note that to get close to zero probability in a logit CDF function it is sufficient for βZ to be not very far below zero since the value enters the function with the exponent.

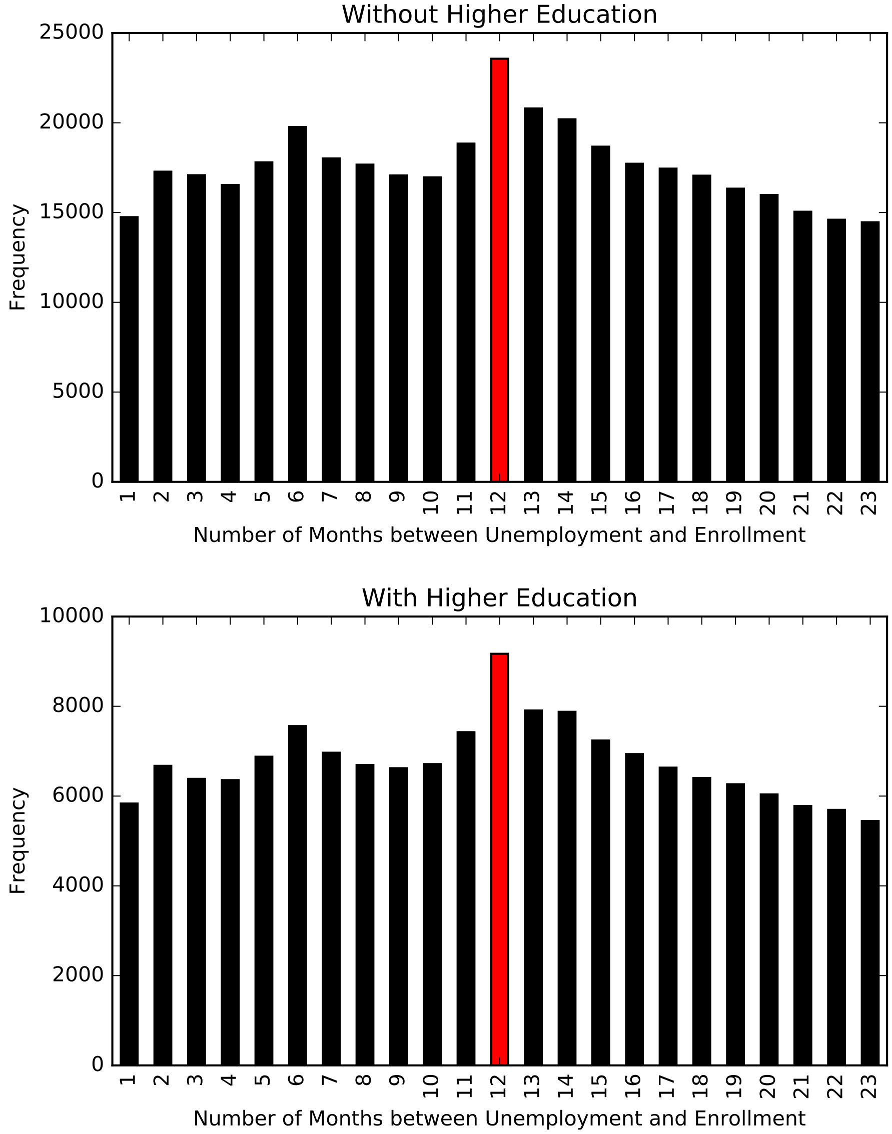
B Supplementary Figures

Figure 13: Bunching Around the Eligibility Requirement By Income



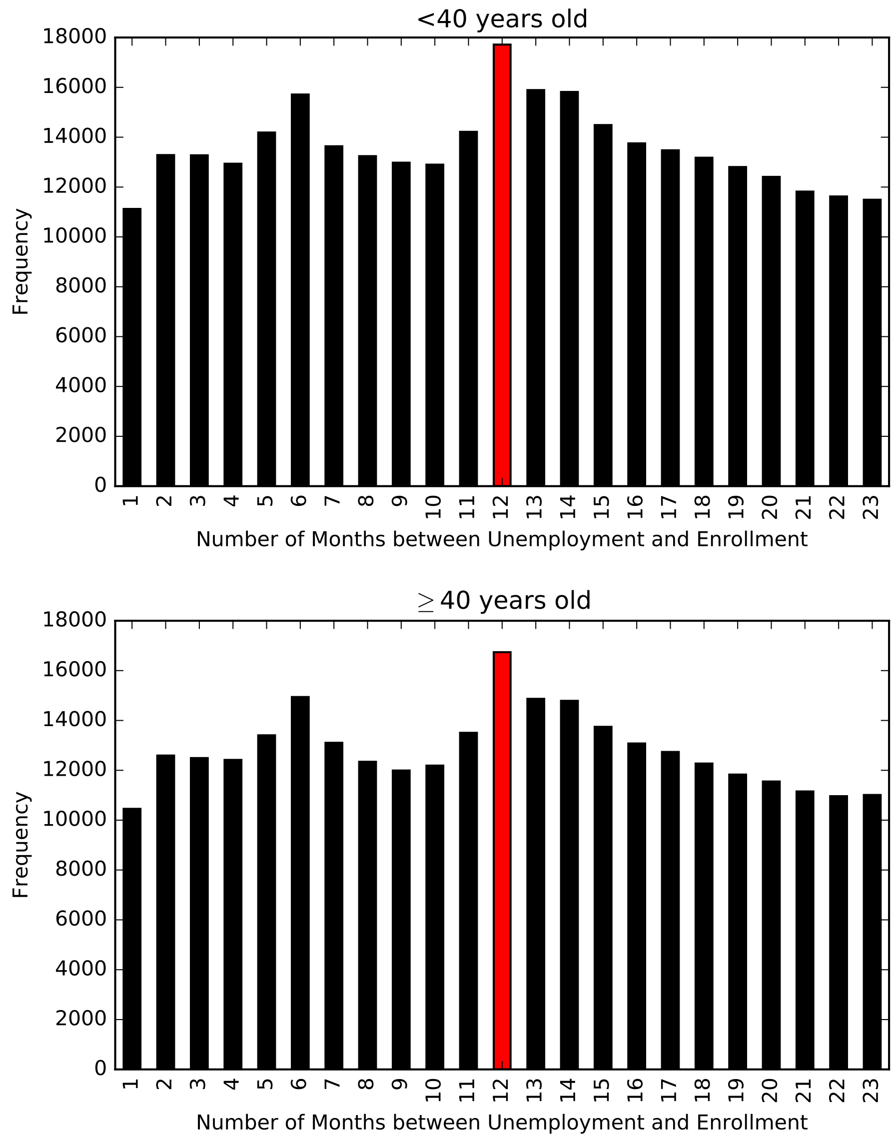
Notes: The Figure presents a discrete histogram of a distribution of a number of enrollment months before the start of unemployment spells similarly to the evidence in the main text but separately for individuals with below the median income (Panel A) and above the median income (Panel B).

Figure 14: 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 15: 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.

C Supplementary Tables

Table 8: Types Parameters

	Types											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Constant	0.3	-1.543	-5.168	0.305	0.143	-0.195	-0.212	-0.26	-0.373	-0.503	-0.264	-0.101
	(—)	(0.341)	(0.012)	(0.016)	(0.015)	(0.026)	(0.048)	(0.015)	(0.121)	(0.073)	(0.128)	(0.293)
Cluster 1	0.3	4.263	-2.545	1.098	-0.203	-0.348	-0.41	-0.541	-0.391	-0.551	0.208	-0.033
	(—)	(0.095)	(0.009)	(0.015)	(0.006)	(0.007)	(0.02)	(0.013)	(0.034)	(0.038)	(0.132)	(0.111)
Cluster 2	0.3	-3.919	-3.098	0.286	-0.427	-0.077	-0.152	-0.397	-0.261	-0.311	0.317	-0.038
	(—)	(0.066)	(0.015)	(0.026)	(0.018)	(0.007)	(0.024)	(0.014)	(0.101)	(0.059)	(0.132)	(0.053)
Cluster 3	0.3	2.307	39.891	1.04	-0.276	-0.187	-0.41	-0.331	-0.446	-0.504	0.007	-0.118
	(—)	(0.22)	(0.008)	(0.009)	(0.015)	(0.022)	(0.033)	(0.013)	(0.122)	(0.118)	(0.208)	(0.073)
Cluster 4	0.3	1.08	1.377	-1.247	-1.452	-1.532	-1.804	-1.466	-1.649	-0.084	-0.263	4.355
	(—)	(0.175)	(0.02)	(0.014)	(0.007)	(0.006)	(0.009)	(0.014)	(0.065)	(0.035)	(0.081)	(0.034)
Age (30; 40]	0.3	-1.627	0.843	-0.448	-0.197	-0.31	-0.374	-0.259	-0.088	0.154	-0.32	0.933
	(—)	(0.151)	(0.026)	(0.018)	(0.02)	(0.015)	(0.04)	(0.017)	(0.193)	(0.154)	(0.206)	(0.096)
Age (40; 50]	0.3	-0.75	-0.039	-0.303	-0.242	-0.427	-0.421	-0.133	-0.53	0.395	0.138	0.187
	(—)	(0.052)	(0.011)	(0.014)	(0.007)	(0.011)	(0.012)	(0.005)	(0.072)	(0.066)	(0.09)	(0.036)
Age > 50	0.3	1.389	0.982	-0.365	-0.264	-0.387	-0.489	-0.685	-0.515	-0.042	< 0.001	0.875
	(—)	(0.062)	(0.008)	(0.008)	(0.006)	(0.004)	(0.014)	(0.011)	(0.104)	(0.049)	(0.095)	(0.097)