# ONLINE APPENDIX FOR:

Risk-based Selection in Unemployment Insurance: Evidence and Implications

by Landais, Nekoei, Nilsson, Seim & Spinnewijn

# Appendix A Observable Risk & Predicted Risk Model: Further Material

In this appendix:

(i) we present the relationship in the Swedish labor market, between a series of observable risk shifters, that are credibly exogenous to individuals' own actions, and individuals' unemployment risk.

(ii) we then present a model of predicted unemployment risk that combines all risk shifters together.

## A.1 Observable Risk Shifters in the Swedish Labor Market

The institutional context and the richness of the Swedish registry data allows us to observe determinants of unemployment risks that are arguably beyond the control of individuals. We present here three such observable sources of risk variation and show how they correlate with individuals' realized unemployment risk.

**Average Firm Layoff Rates** The first observable source of variation in unemployment risk stems from firm level risk. Firm level risk can vary cross-sectionally, due to permanent differences in turnover across firms, or over time, due to firms experiencing temporary shocks. We focus, to start with, on the permanent component of firm level risk, and explore how this permanent component correlates with an individual's displacement probability.

For each individual $i$ working in firm $j$, we define average firm displacement risk $\bar{\pi}_{-i,j}$ as the average probability of displacement of all other workers within the firm excluding individual $i$ over all years where the firm is observed active between 1990 and 2015.[48] We then plot, in Figure A.1 panel A, our measure of average firm risk $\bar{\pi}_{-i,j}$ in 20 bins of equal population size, against $\pi_{i,j}$, the individual probability of displacement in $t+1$, for all individuals ever employed during the period 2002-2007. The figure shows first that there is significant heterogeneity in firms' average separation rates. Second, the figure provides clear evidence that individuals' unemployment risk is very strongly correlated with average firm level risk.

**Layoff Notifications** The second observable source of variation in unemployment risk stems from variation in firm level risk over time. We leverage the fact that under Sweden's employment-

[48] For this purpose we match the employer-employee registry (RAMS) from 1985 to 2015 with the Public Employment Service (PES) registry for all years 1990-2015.

protection law, firms subject to a shock and intending to displace 5 or more workers simultaneously must notify the Public Employment Service in advance. Once a notification is emitted, employers need to come up with the list and dates for the intended layoffs. These layoffs may happen up to 2 years after the original notification has been sent. The layoff notifications provide a source of observable variation in firm displacement risk. In Figure A.2 we report the evolution of the displacement probability of workers around the first layoff notification event in the history of the firm. We define event year $n$ as the year to/since the firm emits its first layoff notification and follow an event study approach around that event. Our sample is the panel of workers who are employed in the firm at the date this layoff notification is emitted to the PES. The graph shows that a layoff notification is indeed associated with a sudden and large increase in the displacement probability of workers. Immediately following the layoff notification, the displacement rate of workers jumps by 6 percentage points compared to pre-notification levels, and remains high for about two years, before decreasing and converging back to pre-notification levels.

Because the panel of workers is selected based on being employed in the firm in year $n = 0$, one may worry that this surge in displacement rates is mechanical, as displacement can only increase after year 0 conditional on all workers being employed in year 0. To mitigate this concern, we follow a matching strategy and create a control panel of workers selected along the same procedure as the original panel. We use nearest-neighbor matching to select a set of firms that are similar, along a set of observable characteristics, to the firms emitting a layoff notification, but never emit a layoff notification.[49] We allocate to the matched firm in the control group a placebo event date equal to the layoff notification date of her nearest-neighbor in the treated group of firms. We then select workers that are in the control firm at the time of the placebo event date to create our matched control panel. Results in Figure A.2 show that, pre-event, the displacement risk is very similar in the control and treated groups, and that it evolves smoothly in the control firms around the event.

This evidence suggests that layoff notifications are a significant shifter of individuals' unemployment risk, as they immediately double the baseline displacement probability of workers.

**Relative Tenure** The third source of observable (and credibly exogenous) risk variation is at the individual level and stems from the strict enforcement of the Last-In-First-Out (LIFO) principle. When a firm wants to downsize within an establishment, the legal system prescribes that displacement occurs by descending order of tenure within the establishment. In practice, workers are divided into groups, defined by collective bargaining agreements in the establishment, and then a tenure ranking within each group is constructed. The tenure ranking of an individual within her establishment and collective bargaining agreement (CBA) group directly determines her probability to be separated. A limitation of our data is that workers' collective bargaining agreements are not directly observed. Instead, we use detailed occupation codes as proxies for the CBAs, which is a good approximation as most CBAs are done at the occupation level.

Figure A.1 panel B plots the probability of being displaced in $t + 1$ among individuals working

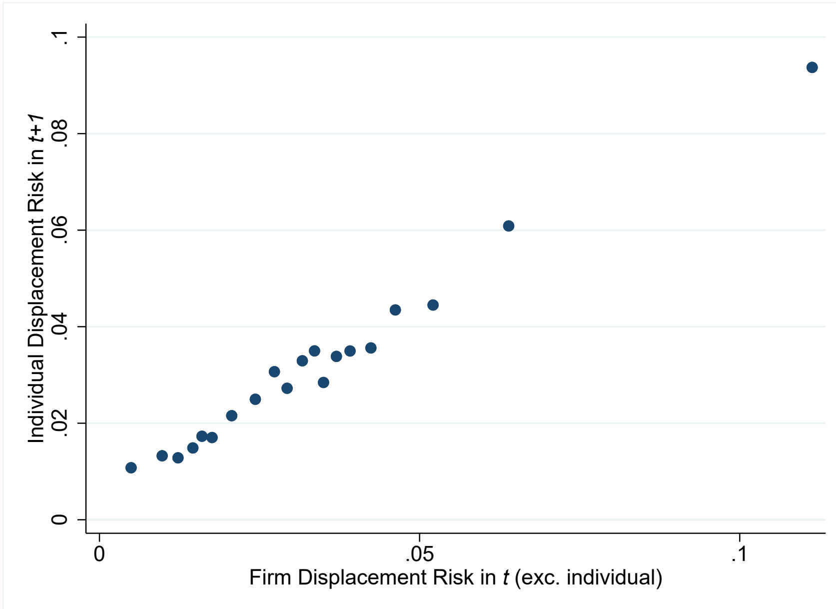
[49] The covariates used for matching are the number of employees, the 4 digit sector codes of the firm, the average earnings and average years of education of workers in the firm.

in firms that emit a layoff notification in $t + 1$, as a function of relative tenure ranking within establishment and occupation. The Figure provides clear evidence of a strong negative correlation between relative tenure ranking and individuals' displacement probability. Individuals within the lowest 10 percent of tenure rankings have a probability of being displaced in $t+1$ larger than .1; this probability declines steadily as tenure ranking increases, and then stays below .02 for individuals in the highest 50 percent tenure rankings.

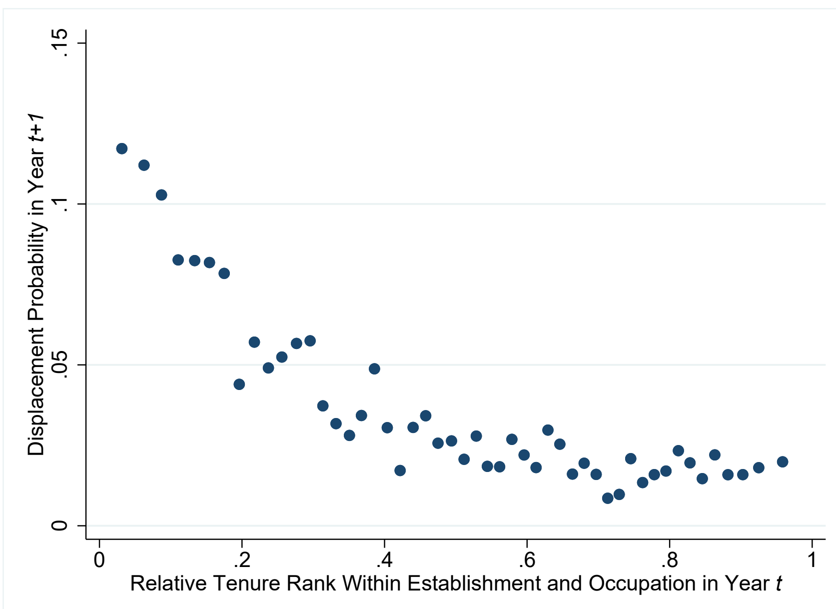
For firms with multiple establishments, one layoff notification needs to be sent for each establishment intending to layoff workers and the LIFO principle applies at the level of the establishment. While the institutionalization is specific to Sweden, the LIFO principle is used for determining redundancy in many countries (e.g., Netherlands, Poland, UK, etc).

Figure A.1: Risk Shifters: Firm Displacement Risk & Last-In-First-Out Principle

A. Firm Displacement Risk vs Individual Displacement Probability in $t+1$

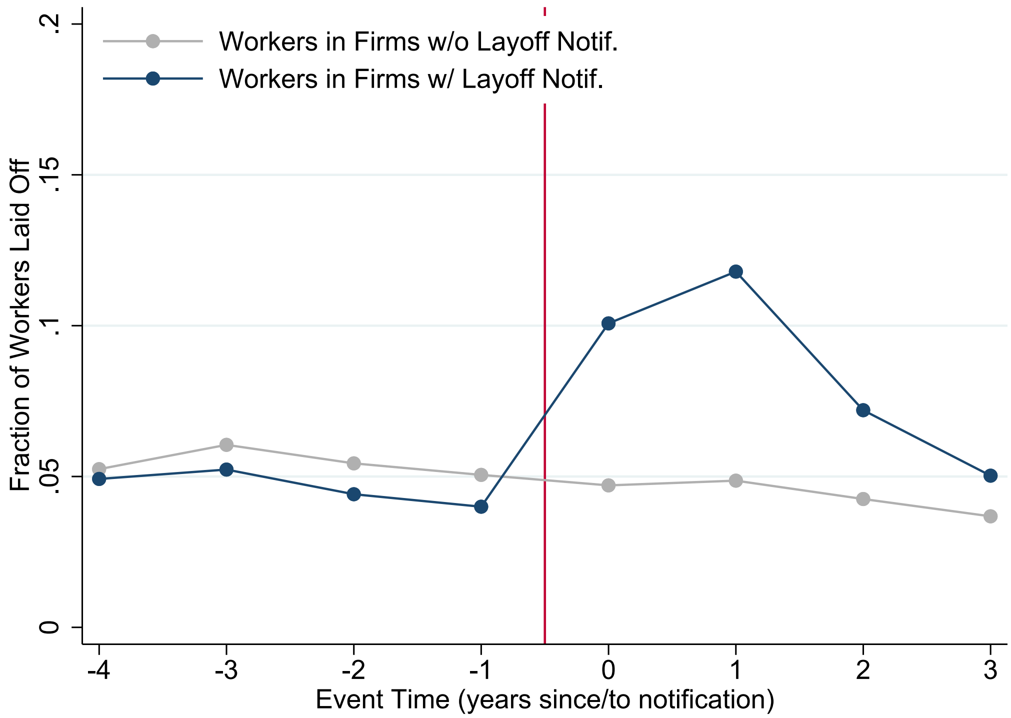



B. Relative Tenure Ranking in Year $t$ vs Displacement Probability in Year $t+1$




**Notes:** The Figure provides evidence of the role of firm level risk and of the Last-In-First-Out (LIFO) principle in creating variation in individuals' unemployment risks. In panel A, we provide evidence of the role of firm layoff risk as a shifter of individuals' own displacement probability. For each individual $i$ working in firm $j$, we define average firm displacement risk as the average probability of displacement of all other workers within the firm excluding individual $i$, $\bar{\pi}_{-i,j}$ over all years where the firm is observed active in our sample years. We then plot the average firm displacement risk in 20 bins of equal population size, against the individual probability of displacement in $t+1$. The Figure shows that there is significant heterogeneity in firms' separation rates, and that individuals' unemployment risk is very strongly correlated with firm level risk. Panel B plots the probability of being displaced in $t+1$ among individuals working in firms that emit a layoff notification in $t+1$, as a function of relative tenure ranking within establishment and occupation in year $t$. See Section 2.1 for institutional details. The Figure provides clear evidence of a strong negative correlation between relative tenure ranking and individuals' displacement probability.

Figure A.2: Layoff Notification and Displacement Risk




**Notes:** The Figure reports estimates of the evolution of the displacement probability of workers around the first layoff notification event in the history of the establishment. We define event year $n = 0$ as the year in which an establishment emits its first layoff notification, and focus on the panel of workers who are employed in the establishment at the date this layoff notification is emitted to the PES. The graph shows that a layoff notification is indeed associated with a sudden and large increase in the displacement risk of workers. Because the panel of workers is selected based on being employed in the firm in year $n = 0$, one may worry that this surge in displacement rates is mechanical, as displacement can only increase after year 0 conditional on all workers being employed in year 0. To mitigate this concern, we follow a matching strategy and create a control panel of workers selected along the same procedure as the original panel. We use nearest-neighbor matching to select a set of firms that are similar, along a set of observable characteristics, to the firms emitting a layoff notification, but never emit a layoff notification. We allocate to the matched firm in the control group a placebo event date equal to the layoff notification date of her nearest-neighbor in the treated group of firms. We then select workers that are in the control firm at the time of the placebo event date to create our matched control panel.

## A.2 Predicting Risk Using Observable Risk Shifters: a Zero-Inflated-Poisson Model of Unemployment Risk

We now present the model we use to compute the best predictor of future unemployment risk given all currently observed individual characteristics. To do so, we leverage the rich set of observables available in the Swedish registry data, and the various institutional features of the Swedish labor market.

**Setup** The measure of unemployment risk $\pi$ that we model is the number of days an individual is expected of spending unemployed in $t+1$. This is the relevant measure of risk to the UI system given insurance choices made in year $t$.[50]

The distribution of days spent unemployed is defined only over non-negative integers, and exhibits a significant mass at zero. To account for these facts, we model $\pi$ using a zero-inflated Poisson model. The expected number of days unemployed conditional on a vector of characteristics $X$ therefore takes the following form:

$$E(\pi|X) = (1 - f(0|X^I)) \exp(X^{C'} \beta^C)$$

For the zero-inflated part of the process, we parametrize the probability $f(0)$ using a logit: $f(0|X^I) = \exp(X^{I'}\beta^I)/(1+\exp(X^{I'}\beta^I))$. We will allow for the set of risk predictors $X^I$ and $X^C$ entering respectively the inflated part and the count part, to differ.

To account for moral hazard, we allow the risk of individuals with similar characteristics $X$ to differ if they are observed under the basic coverage or under the comprehensive coverage. To this purpose, we estimate separately two models of predicted risk. The first model is the predicted risk given $X$ under the basic coverage $\hat{\pi}_0 = E(\pi_0|X)$. This model is estimated on individuals who are observed under the basic coverage in $t$. The second model is the predicted risk given $X$ under the comprehensive coverage $\hat{\pi}_1 = E(\pi_1|X)$, which we estimate on individuals who are observed under the comprehensive coverage in $t$.

**Lasso Penalization** In terms of model selection, we discipline the choice of the many potential regressors by using the adaptive Lasso procedure for a zero-inflated Poisson (ZIP) model proposed by Banerjee et al. [2018]. The ZIP Log Likelihood function with LASSO penalty works in the following way.

Let $X^C = \{x_1^C, ..., x_K^C\}$ be the set of K regressors associated with predicting the number of days unemployed, conditional on some unemployment, according to a Poisson distribution. The corresponding coefficients are: $\{\beta_1^C, ..., \beta_K^C\}$. Let $X^I = \{x_1^I, ..., x_J^I\}$ be the set of J regressors associated with predicting some unemployment, according to a logistic distribution. The corresponding

[50]If an individual has bought the comprehensive coverage throughout year $t$, then the days she spends unemployed in year $t+1$ will be covered by the comprehensive benefits. In that sense, the relevant risk to determine the cost of providing the comprehensive coverage to an individual buying that coverage in year $t$ is the expected number of days she will spend unemployed in year $t+1$.

coefficients are: $\{\beta_1^I, ..., \beta_J^I\}$. The number of days (integers) spent unemployed by individual $i$ is $\pi_i$. Then we can write down the ZIP Log Likelihood function with LASSO penalty as follows:

$$L = L_1 + L_2 - L_3 - P_C - P_I$$

Where each of the $L_1, L_2, L_3$ components are defined as follows:

$$L_1 = \sum_{i:\pi_i=0}^{n} \log[\exp(X_i'^I \beta^I) + \exp(-\exp(X_i'^C \beta^C))]$$

$$L_2 = \sum_{i:\pi_i>0}^{n} \{\pi_i X_i'^C \beta^C + \exp(-X_i'^C \beta^C) - \ln(\pi_i!)\}$$

$$L_3 = \sum_{i=1}^{n} \log[1 + \exp(X_i'^I \beta^I)]$$

And the $P_C, P_I$ components are defined as follows,

$$P_C = \lambda_C \sum_{k=1}^{K} |\beta_k^C|$$

$$P_I = \lambda_I \sum_{j=1}^{J} |\beta_j^I|$$

**Estimation** We can then estimate the model for various levels of penalization for the count part $\lambda_C$ and the inflated part of the model $\lambda_I$. In practice, we draw 50 pairs of $\lambda_C$ and $\lambda_I$. The largest pair is chosen so that all variables except the constant are set to zero. This corresponds to the largest penalization.

We then randomly select a subset of observations from our sample to obtain a training sample, the rest of the observation is considered our test sample. We estimate the model on the training sample for all 50 pairs of $\lambda_C$ and $\lambda_I$. We then compute the MSE on the test sample for all 50 models, and select the lambda pair associated with the smallest MSE on the test sample.

**Predictors** The regressors we allow to initially enter the model are individual log earnings, family type, nine age bins, gender, twelve dummies for education level, year fixed effects, region fixed effects, industry fixed effects, dummies for the past layoff history of the individual, dummies for the layoff notification history of the firm, the leave-out mean of firm layoff risk, union membership, tenure rank, interactions between tenure ranking and firm layoff risk and interactions between tenure ranking and layoff notification history of the firm. We allow all these predictors to enter in both the count and inflated part.

When varying the level of penalization in the model, starting from the highest penalization, we can see what variables are the strongest predictors of the inflate and count part of unemployment risk. For the inflate component, the first variables to become significant are firm layoff risk, layoff notification dummies, and relative rank. This confirms the important role played by the institutional features of the Swedish labor market in determining unemployment risk. For the count component, the first variables to become significant (by order) are those associated with age, gender, education level, income, regions and years

The results show that the optimal penalization factors $\lambda$ associated with the count component are smaller while those associated with the zero component are higher, thus penalizing the inclusion of variables in the latter more. As a result, in our preferred model of predicted risk, in the zero component, a large share of variables have a coefficient set to zero.

**Model fit** As explained in section 2.3, we first assess the quality of the model fit in Figure 1 by plotting bin scatters of the relationship between predicted risk under basic (resp. comprehensive) coverage and actual realized risk for individuals under basic (resp. comprehensive) coverage. In both panels, the relationship is close to the 45 degree line indicating that the model does a good job at predicting the average realization of unemployment risk. In Table A.1 below, we provide further summary statistics on the distribution of predicted risk according to our model. In Panel A, we focus on individuals observed under the basic coverage, and compare the distribution of their realized risk $\pi_0$ to the distribution of their predicted risk under basic coverage $\hat{\pi}_0$. The average risk predicted by the model (2.95) is very close to the average realized risk (2.83). In Panel B, we do a similar exercise, focusing on individuals observed under the comprehensive coverage. We compare the distribution of their realized risk $\pi_1$ to the distribution of their predicted risk under basic coverage $\hat{\pi}_1$. We find again that the average risk predicted by the model (5.90) is very close to the average realized risk (5.65). In both panels, we find that there is much less dispersion in predicted risk than in realized risk. The standard deviation of predicted risk is roughly six times smaller than that of realized risk. This confirms that there still remains a significant dimension of idiosyncratic unemployment risk beyond what can be predicted by even our very rich set of observables.

Table A.1: Distribution of Realized and Predicted Risk for Individuals under Basic and under Comprehensive Coverage

Panel A. Predictable Risk Under Basic

| | $\pi_0$ | $\hat{\pi}_0$ |
|---|---|---|
| P25 | 0 | 2 |
| P50 | 0 | 3 |
| P75 | 0 | 3 |
| P90 | 0 | 4 |
| P99 | 107 | 13 |
| P99.9 | 346 | 32 |
| Mean | 2.84 | 2.95 |
| s.d. | 22.54 | 2.47 |
| N | 2,2296,727 | 2,232,136 |

Panel B. Predictable Risk Under Comprehensive

| | $\pi_1$ | $\hat{\pi}_1$ |
|---|---|---|
| P25 | 0 | 4 |
| P50 | 0 | 5 |
| P75 | 0 | 6 |
| P90 | 0 | 7 |
| P99 | 184 | 30 |
| P99.9 | 365 | 80 |
| Mean | 5.65 | 5.91 |
| s.d. | 33.17 | 5.86 |
| N | 15,003,779 | 14,879,543 |

**Notes:** The table reports moments of the distribution of predicted risk and realized risk from our sample of workers for years 2002 to 2006. Panel A focuses on individuals who are observed under the basic coverage in $t$. The first column reports moments of the distribution of their realized risk $\pi_0$ while the second column reports moments of the distribution of our measure of predicted risk under basic coverage $\hat{\pi}_0$. Panel B focuses on individuals who are observed under the comprehensive coverage in $t$. The first column reports moments of the distribution of their realized risk $\pi_1$ while the second column reports moments of the distribution of our measure of predicted risk under comprehensive coverage $\hat{\pi}_1$.

# Appendix B Positive Correlation Tests: Further Results

In this appendix we present further evidence regarding the positive correlation between unemployment risk and UI choices:

(i) we present positive correlation tests using alternative risk outcomes.

(ii) we present robustness analysis for the PCT using alternative specifications and non-parametric approaches

Table B.3 provides the summary statistics for our main sample broken down by UI coverage.

## B.1 Positive Correlation Tests: Alternative Risk Outcomes

We start by showing that the strong correlation between realized unemployment risk and UI choices documented in Figure 2 extends to using alternative measures of realized unemployment risk.

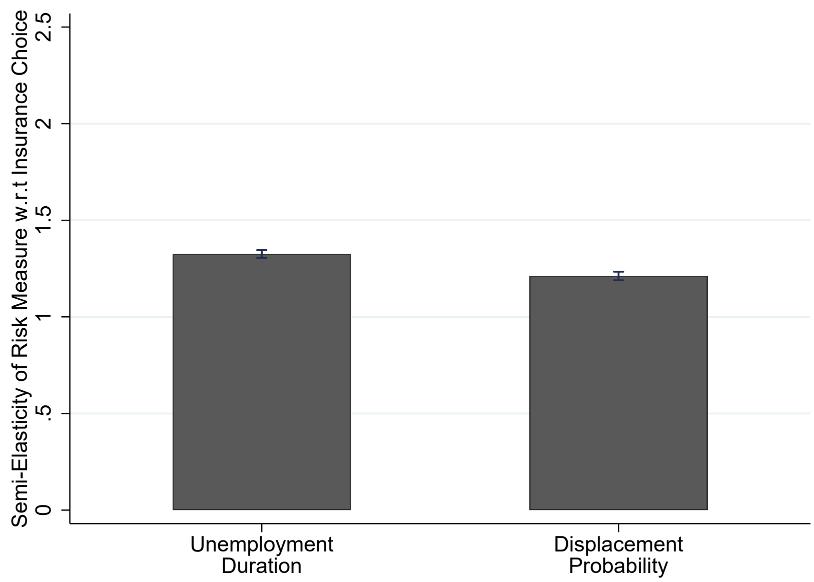
**Displacement Probability** We start by investigating the robustness of our PCT results to using the probability of displacement in $t+1$ as our measure of risk $\pi$. To control for observables $Z$, we model the probability of displacement as a probit:

$$E(\pi|Z) = \Phi(Z'\beta + \alpha \cdot \mathbb{1}[D=1]) \tag{20}$$

where $\Phi(.)$ is the standard normal c.d.f. The second bar of Figure B.1 reports the semi elasticity defined in (13) estimated from this model. The first bar of the figure reports our estimate of the PCT for our baseline measure of risk, that is the number of days spent unemployed in $t+1$. The graph confirms the presence of a strong positive correlation between UI choices and unemployment risk: Individuals who buy the comprehensive coverage in $t$ are 125% more likely to be displaced in $t+1$ than individuals who do not buy the comprehensive coverage.

We note that different measures of unemployment risks are subject to different types of moral hazard. Comparing the magnitude of the correlations across the different realized risk outcomes already sheds light on some margins of moral hazard. A large body of literature has for example documented that higher unemployment benefits increase the duration of unemployment spells conditional on becoming unemployed (see Schmieder and von Wachter [2016] for a recent review). Such moral hazard conditional on displacement will increase the correlation between unemployment duration in $t+1$ and insurance coverage in $t$ (first bar in Figure B.1). The probability of displacement, while immune to moral hazard once displaced, is potentially affected by moral hazard "on the job" (second bar in Figure B.1). An example of this would be collusion between employers and employees to qualify actual voluntary quits as "quits following a valid reason", which are eligible for unemployment benefits.

Figure B.1: Positive Correlation Tests - Displacement Risk vs Total Unemployment Risk




**Notes:** The first bar of the figure reports our estimate of the PCT for our baseline measure of risk, that is the number of days spent unemployed in $t+1$. The second bar reports the PCT for the displacement risk in $t+1$. This estimate is the semi elasticity defined in (13) estimated from probit specification (31).

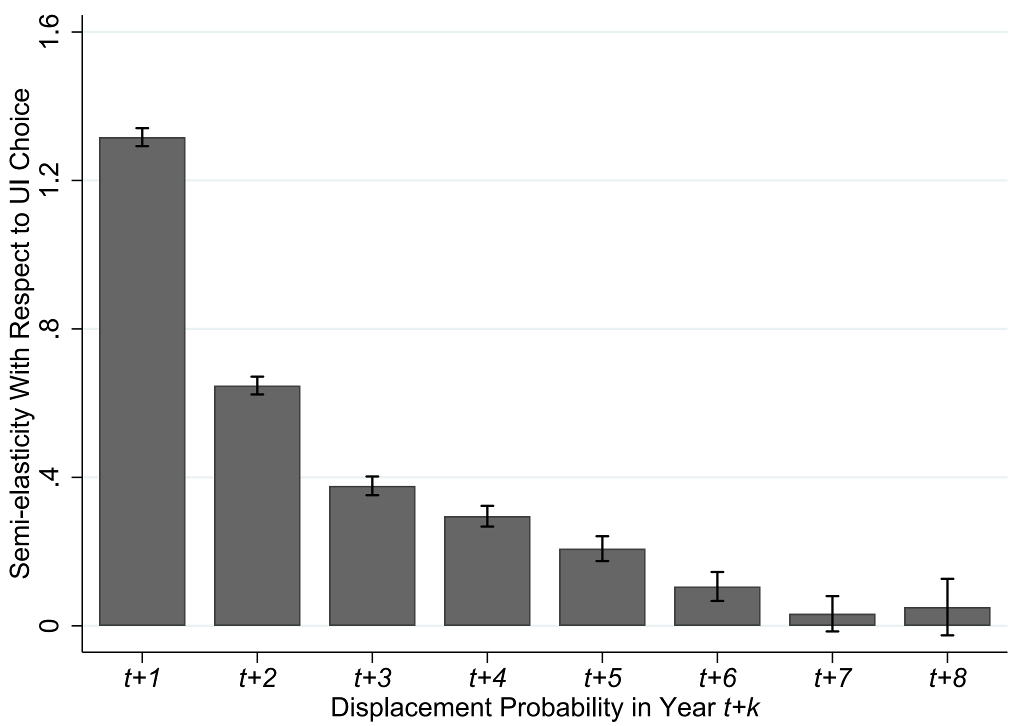
**Risk Dynamics** Our correlation tests use the risk outcomes in $t+1$, reflecting the idea that workers need to contribute for a year to be able to get the comprehensive coverage. However, the risk realization in $t+1$ may fail to fully capture the unemployment risk faced by an individual as she is making her coverage choice at time $t$, which justifies using risk realizations further into the future. In Figure B.2 we report the correlation of the insurance choice in $t$ with displacement outcomes in $t+1$, $t+2$,... up to $t+8$. For each displacement outcome, the chart displays $\hat{\alpha_k}/\bar{\pi}$, that is the semi-elasticity of the realized risk outcomes in $t+k$ with respect to insurance choices in $t$, from a simple linear specification where we also control for all displacement outcomes in previous years ($t+k-1$, $t+k-2$, etc.):

$$\pi_{i,t+k} = \alpha_k D_i + Z_i'\beta + \sum_{l=0}^{k} \pi_{i,t+k-l} + \epsilon_i, \tag{21}$$

The first thing to note is that the estimated PCT for displacement risk in year $t+1$, using the linear specification (21) is equivalent to the PCT of Figure B.1 above, estimated from the non-linear specification (20). This is indicative that the PCT results are robust to functional form specifications.

The Figure also reveals an interesting dynamic pattern. The positive correlation between insurance choice and risk decreases rapidly as we consider displacement risk further in the future, but remains statistically significant up to six years. This pattern could indicate that workers' insurance choices incorporate private information about unemployment risk further into the future (albeit to a decreasing extent), but it may also be affected by moral hazard responses.

Figure B.2: Positive Correlation Tests - Dynamics




**Notes:** Risk realization in $t+1$ may fail to fully capture the unemployment risk faced by an individual as she is making her coverage choice at time $t$, which justifies using risk realizations for that individual further into the future. This Figure reports the correlation of insurance choice in $t$ with displacement outcomes in $t+1$, $t+2$,... up to $t+8$. The Figure displays estimates of positive correlation tests following specification (12) estimated over the period 2002-2006. For each outcome, the chart displays $\hat{\alpha_k}/\bar{\pi}$, that is the semi-elasticity of the realized displacement rate in $t+k$ with respect to insurance choices in $t$. For each displacement outcome in year $t+k$, we control for displacement outcomes in previous years ($t+k-1$, $t+k-2$, etc.), for year fixed effects and for the limited set of characteristics $Z$ that affect the unemployment insurance coverage available to individuals. See text for details.

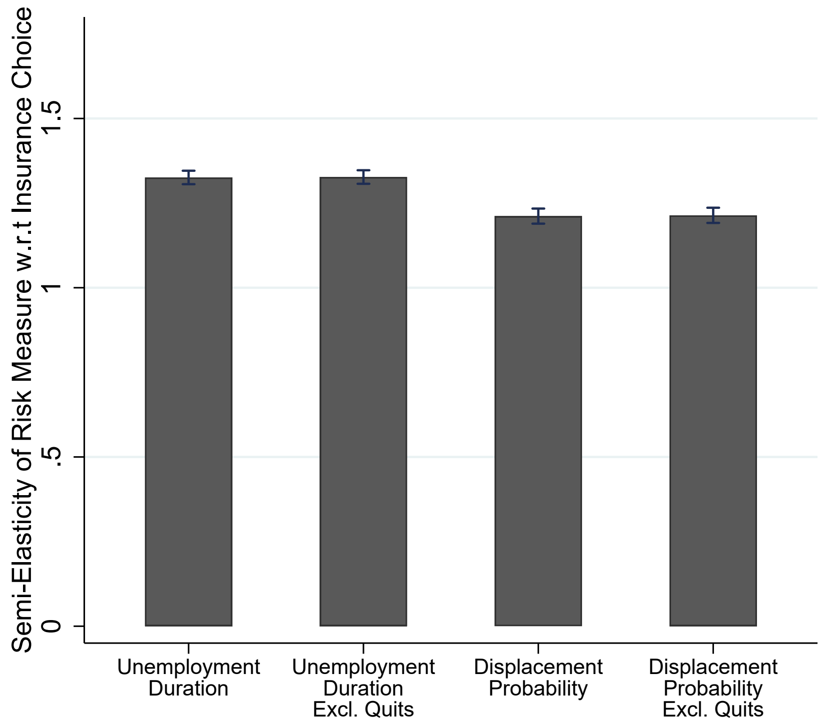
**Unemployment Risk Excluding Involuntary Quits** In the Swedish UI system, "quits following a valid reason" are eligible for unemployment benefits. They are therefore included in our measure of unemployment risk. The fact that involuntary quits are eligible to UI may raise the possibility of collusion between employers and employees to qualify actual voluntary quits as "quits following a valid reason". To understand to what extent this type of moral hazard drives the positive correlation between UI choices and realized unemployment risk, we exclude quits from the definition of unemployment. To do this, we use the fact that in the IAF data, a variable indicates whether an unemployment spell starts following a "quit for a valid reason".

We use again a simple linear specification:

$$\pi_i = \alpha D_i + Z_i' \beta + \epsilon_i, \tag{22}$$

In Figure B.3, we report $\hat{\alpha}/\bar{\pi}$, that is the semi-elasticity of the realized risk outcome in $t+1$ with respect to the insurance choice in $t$ from this specification. We first use as an outcome $\pi$ the total number of days spent unemployed in $t+1$, when including quits (first bar). Then, in the second bar, we report results where we use as an outcome $\pi$ the total duration spent unemployed in $t+1$ when excluding involuntary quits from the definition of unemployment risk. We then replicate this exercize using as an outcome the probability of displacement in $t+1$ when including quits (third bar) and when excluding involuntary quits (fourth bar) from the definition of displacement risk. The figure shows that the positive correlation between unemployment risk and UI choices is almost unaffected by the inclusion or exclusion of involuntary quits.

Figure B.3: Positive Correlation Tests: Risk Outcomes Including and Excluding Involuntary Quits




**Notes:** This Figure reports the correlation of insurance choice in $t$ with risk outcomes in $t+1$. The Figure displays estimates of positive correlation tests following specification (22) estimated over the period 2002-2006. For each outcome, the chart displays $\hat{\alpha}/\bar{\pi}$, that is the semi-elasticity of the realized risk with respect to insurance choices in $t$. For each outcome, we control for year fixed effects and for the limited set of characteristics $Z$ that affect the unemployment insurance coverage available to individuals. See text for details.

## B.2 Bivariate Probit & Non-parametric Tests

We now further investigate functional form restrictions and provide correct inference for the correlation tests.

First, we provide results of bivariate probit tests, popularized by Chiappori and Salanié [2000]. We specify both the choice of insurance coverage and the realization of our binary measure of unemployment risk (i.e., the probability of displacement) as probit models:

$$\begin{aligned} \mathbf{u}_i &= \mathbb{1}[Z'\alpha_1 + \epsilon > 0] \\ \pi_i &= \mathbb{1}[Z'\alpha_2 + \eta > 0] \end{aligned} \tag{23}$$

where $\mathbf{u}_i = u_{i,1} - u_{i,0}$ is the short-hand notation for the difference in indirect expected utility for individual $i$ between being in the comprehensive plan and being in the basic plan. We allow for correlation $\rho$ between the two error terms $\epsilon$ and $\eta$. The vector of controls $Z$ contains the same variables as in specification (12). We provide in Table B.1 estimates of $\rho$ and formal tests of the null that $\rho = 0$. Results confirm the presence of a strong and significant correlation between insurance choices and realized unemployment risk.

The functional forms involved in the bivariate probit tests are still restricted to the latent models being linear and the errors normal, excluding cross-effects or more complicated non-linear functions of the variables in $Z$. We therefore also produce results from non-parametric tests as in Chiappori and Salanié [2000]. The procedure of the test consists in partitioning the data into cells where all observations in a given cell have the same value for the variables in $Z$. The procedure then computes within each cell a Pearson's $\chi^2$ test statistic for independence between $\mathbf{u}$ and $\pi$. This test statistic is asymptotically distributed as a $\chi^2(1)$ under the null hypothesis that $\mathbf{u}$ and $\pi$ are statistically independent (within the cell). We report in the first column of Table B.2 results from this non-parametric procedure when cells are defined using the same controls $Z$ as in specification (12) and where our risk measure $\pi$ is the probability of displacement. Results again strongly confirm the presence of a positive correlation between insurance choices and unemployment probability. In Figure B.4 panel 1, we display the empirical distribution of the Pearson's $\chi^2$ test statistics computed from all the cells to allow for comparison with a theoretical $\chi^2(1)$ distribution. Taking the largest absolute difference between the theoretical and the empirical distribution gives the Kolmogorov-Smirnov test statistic reported in Table B.2.

In columns (2) to (4) of Table B.2, we explore the robustness of the positive correlation test to adding more observable characteristics in the vector $Z$. In other words, we want to explore how much positive correlation would remain if the UI policy was allowed to differentiate coverage or prices along obvious observable dimensions that do not currently enter the UI policy schedule (such as age, gender, etc.). To this effect, we reproduce the non-parametric Kolmogorov-Smirnov test adding sequentially more observable characteristics to the vector $Z$ when partitioning the data into cells. We start in column (2) of Table B.2 by adding demographic controls : age, then gender, and marital status. The Kolmogorov-Smirnov test statistic increases sharply, indicating that demographics may offer advantageous selection. Yet, we can still strongly reject the null

Table B.1: Positive Correlation Tests: Bivariate Probits

| | | | Test $\rho = 0$ | |
|---|---|---|---|---|
| | $\rho$ | s.d. | $\chi^2$ | P-Value |
| Proba. of displacement | .3047 | .0030 | 8842.4 | 0.00 |
| Proba. of displacement excl. quits | .3056 | .0031 | 8493.9 | 0.00 |

**Notes:** The Table reports positive correlation estimates between insurance and risk using bivariate probit models. We specify both the choice of insurance coverage and the probability of displacement as probit models allowing for correlation $\rho$ between the two error terms $\epsilon$ and $\eta$. The Table reports estimates of $\rho$ and its standard error. We also report results of formal tests of the null that $\rho = 0$. In the first row, we consider the probability of displacement. In the second row we consider the probability of displacement excluding quits, as some quits may be eligible for UI after a waiting period. See text for details.

of no positive correlation between risk and insurance choice. In column (3), we add controls for education (four categories), and industry (1-digit code). The Kolmogorov-Smirnov test statistic does not seem to be affected much by the inclusion of these controls for skills and other labor market characteristics. In column (4), we finally add controls for past unemployment history (dummies for having been unemployed in $t-1$, $t-2$ and up to $t-8$). The Kolmogorov-Smirnov test statistic decreases as a result, suggesting that past unemployment history creates significant adverse selection.

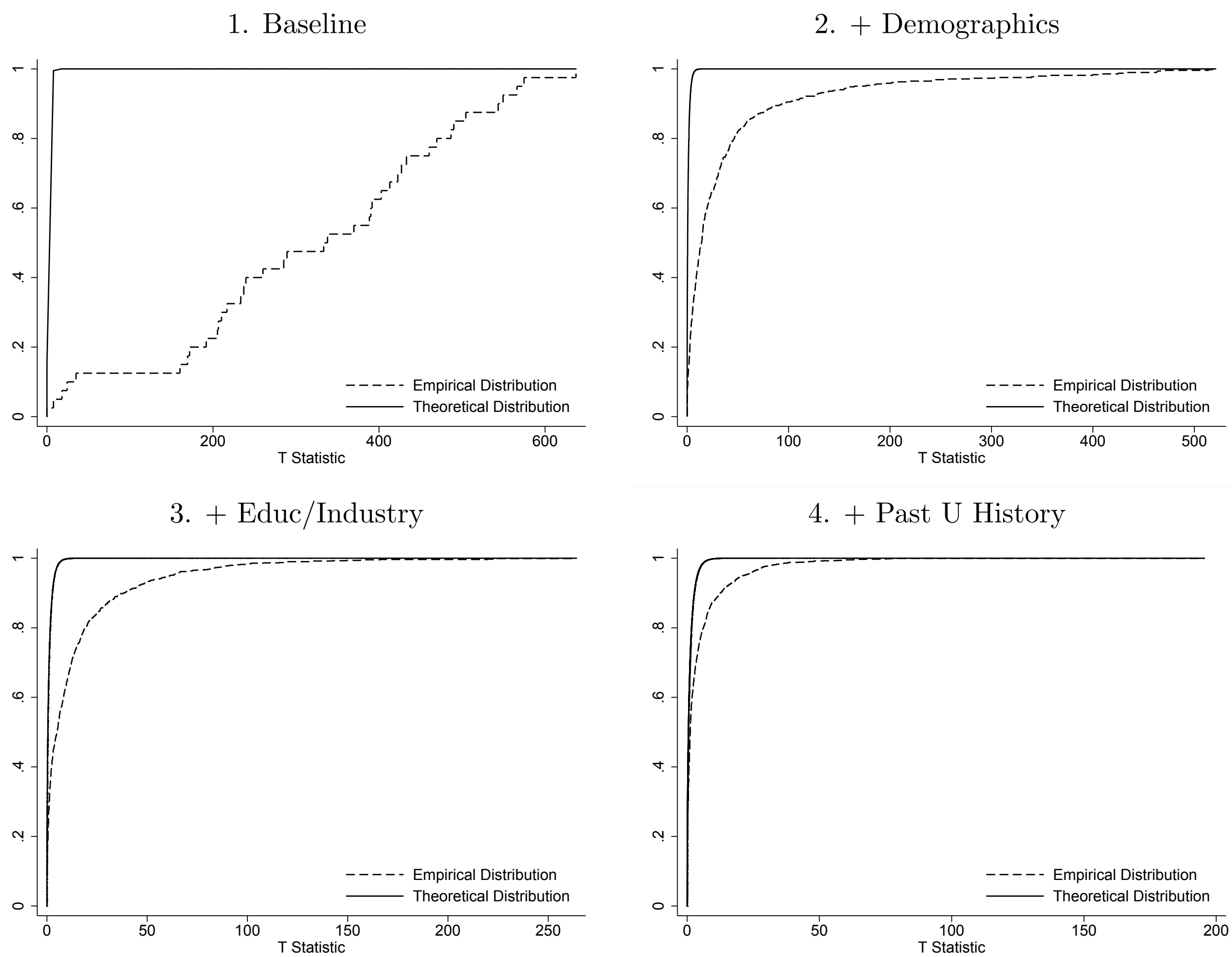
For all specifications of columns (1) to (4) of Table B.2, the corresponding panels 1 to 4 of Figure B.4 display the empirical distribution of the Pearson's $\chi^2$ test statistics computed from all the cells to allow for comparison with a theoretical $\chi^2(1)$ distribution.

Table B.2: Positive Correlation Tests: Non-Parametric Tests

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | **Variables included in partitioning the data in cells** | | | |
| | Baseline | + Demographics | + Educ & Industry | + Past U History |
| # of cells | 40 | 484 | 1,124 | 1,923 |
| Average cell size | 50,903 | 3,181 | 958 | 415 |
| Median cell size | 35,275 | 1,270 | 346 | 141 |
| Minimum cell size | 14,202 | 88 | 6 | 5 |
| Fraction of cells too granular | 0% | 24% | 65% | 80% |
| Fraction of rejected cells | 98% | 74% | 53% | 28% |
| Kolmogorov-Smirnov stat. | 5.98 | 15.37 | 16.20 | 10.47 |
| Binomial p-value | 0% | 0% | 0% | 0% |

**Notes:** The Table reports results from non-parametric tests of correlation between insurance choices in $t$ and probability of displacement in $t+1$. The procedure of the test consists in partitioning the data into cells where all observations in a given cell have the same value for the variables in $Z$. Columns (1) to (4) differ in the control variables included in $Z$ and used to partition the data. The procedure then computes within each cell a Pearson's $\chi^2$ test statistic for independence between $\mathbf{u}$ and $\pi$. This test statistic is asymptotically distributed as a $\chi^2(1)$ under the null hypothesis that $\mathbf{u}$ and $\pi$ are statistically independent (within the cell). The critical values of this statistic for 95% and 99% confidence are 1.36 and 1.63 respectively. The reported Kolmogorov-Smirnov test statistic is scaled by $\sqrt{n}$ where $n$ is the number of cells. When adding a lot of controls to the vector $Z$, some cells can become too granular to compute the test statistic (division by zero). We therefore also report in the Table the number of cells that are too granular.

Figure B.4: Positive Correlation Tests - Distribution of $\chi^2$ test statistics from all cells vs Theoretical $\chi^2(1)$ distribution - Additional Controls




**Notes:** The Figure displays the empirical distribution of the Pearson's $\chi^2$ test statistics for independence between $\mathbf{u}$ (UI choices) and $\pi$, the probability of layoff in $t+1$, computed from all the cells where we split individuals in cells corresponding to various observable characteristics. In panel 1, we only use priced characteristics (baseline controls of the positive correlation tests), corresponding to the test implemented in column (1) of Table B.2. In panel 2, we add controls for demographics (cf. column (2) of Table B.2). Panel 3 and 4 add education, industry and past unemployment history controls (cf. column (3) and (4) of Table B.2). The $\chi^2$ test statistic is asymptotically distributed as a $\chi^2(1)$ under the null hypothesis that $\mathbf{u}$ and $\pi$ are statistically independent (within the cell). We therefore compare this distribution with a theoretical $\chi^2(1)$ distribution. Taking the largest absolute difference between the theoretical and the empirical distribution gives the Kolmogorov-Smirnov test statistic reported in Table B.2.

Table B.3: Summary Statistics - By Coverage

| | A. Under Basic | | | | B. Under Comprehensive | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | P10 | P50 | P90 | Mean | P10 | P50 | P90 |
| | **I. Unemployment** | | | | | | | |
| Displacement probability | 1.96% | - | - | - | 3.21% | - | - | - |
| Displacement probability (exc. quits) | 1.81% | - | - | - | 3% | - | - | - |
| Unemployment probability | 2.29% | - | - | - | 3.85% | - | - | - |
| Days unemployed | 2.84 | 0 | 0 | 0 | 5.65 | 0 | 0 | 0 |
| Predicted days unemployed under Basic | 2.96 | 1.89 | 2.58 | 3.68 | 3.66 | 2.18 | 2.74 | 4.09 |
| Predicted days unemployed under Comprehensive | 5.34 | 3.45 | 4.67 | 6.82 | 5.91 | 3.66 | 4.78 | 7.24 |
| Unemployment spell duration (days) | 137.57 | 26 | 90 | 283 | 148.26 | 22 | 91 | 307 |
| Fraction receiving layoff notification | .04 | - | - | - | .06 | - | - | - |
| Fraction switching firms | .1 | - | - | - | .09 | - | - | - |
| | **II. Union and UI Fund Membership** | | | | | | | |
| Union membership | .13 | - | - | - | .84 | - | - | - |
| Switch from coverage 0 to 1 | - | - | - | - | .02 | - | - | - |
| Switch from coverage 1 to 0 | .01 | - | - | - | - | - | - | - |
| | **III. Demographics** | | | | | | | |
| Age | 35.52 | 25 | 33 | 55 | 41.7 | 27 | 42 | 55 |
| Years of education | 12.97 | 11 | 12 | 16 | 12.84 | 11 | 12 | 16 |
| Fraction men | .63 | - | - | - | .51 | - | - | - |
| Fraction married | .32 | - | - | - | .46 | - | - | - |
| | **IV. Income and Wealth, SEK 2003(K)** | | | | | | | |
| Gross earnings | 233.8 | 65.3 | 186.7 | 416 | 251.5 | 115 | 234.9 | 385.4 |
| Net wealth | 649.6 | -195.1 | 25.7 | 1521.7 | 343.5 | -155.5 | 102.4 | 1083.5 |
| Bank holdings | 73.9 | 0 | 0 | 135.8 | 45.3 | 0 | 0 | 120.4 |
| N | 2,296,727 | | | | 15,003,779 | | | |

**Notes:** The Table breaks down the summary statistics by UI coverage for our main sample of interest over the period 2002 to 2006. See Table 1.

# Appendix C Impact of Predictable Risk on UI Choice: Further Evidence

In this appendix we present further evidence regarding the relationship between predictable risk and UI choices:

(i) we present non-parametric evidence of the relationship linking UI choices to both $\hat{\pi}_0$ and $\hat{\pi}_1$

(ii) we present quasi-experimental evidence showing how the various institutional risk shifters detailed in Appendix A, which enter our predicted risk model, separately affect selection into coverage.

## C.1 Non-Parametric Evidence on the Relationship between Predicted Risk and Insurance Choice

The positive correlation tests between predicted risk and UI choice from section 4.2 shows clearly that individuals buying the supplemental coverage have a higher predictable risk on average in both the basic coverage ($E_1[\hat{\pi}_0] > E_0[\hat{\pi}_0]$) and the comprehensive coverage ($E_1[\hat{\pi}_1] > E_0[\hat{\pi}_1]$). We provide here more detailed non-parametric evidence on the relationship between insurance choice and predicted risk in both coverages.

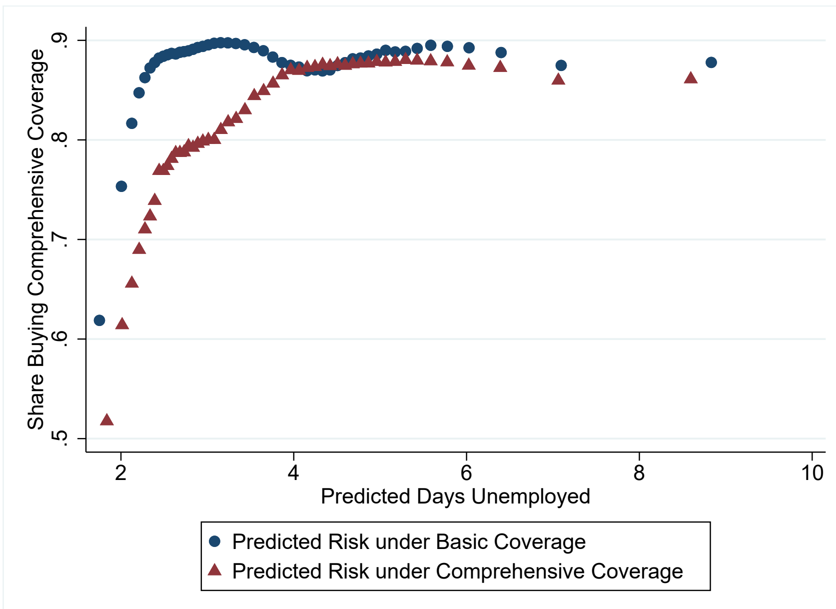
For this purpose, Figure C.1 offers a bin scatter correlating the probability to buy the comprehensive UI coverage in year $t$ with the predicted number of days unemployed of individual $i$, respectively under the basic coverage $\hat{\pi}_0$ and under the comprehensive coverage $\hat{\pi}_1$, based on her observable characteristics year $t$. The graph confirms evidence from the positive correlation tests of a strong positive correlation between individuals' predictable risk and their probability to buy the comprehensive UI coverage. Interestingly, the graph also suggests that the strong positive correlation between risk and insurance coverage is mostly driven by what happens at the bottom of the predicted risk distribution. Only about a half of individuals at the bottom of the predicted risk distribution ($\hat{\pi}_1 < 2$ days) buy the comprehensive coverage. But this fraction quickly rises as the predicted risk increases. It is then very stable, at around 85 to 90% for individuals with predicted risk $\hat{\pi}_1 < 5$ days. Note finally that conditional on the fraction buying the comprehensive coverage, the difference between predicted risk under basic and under comprehensive coverage captures the presence of moral hazard.

## C.2 Risk Shifter & UI Choices I: Average Firm Layoff Risk

The previous evidence focuses on risk measures from our predicted risk model. This model folds all sources of variations of observable risk together into a unique measure of predictable risk. We now also shed light on how the various institutional risk shifters that enter the predicted risk model individually affect selection into coverage.

The first source of risk variation is average firm level risk. We define again the average firm displacement risk $\bar{\pi}_{-i,j}$ of worker $i$ working in firm $j$ as the average probability of displacement

Figure C.1: Predicted Risk and UI Coverage Choice




**Notes:** The Figure displays a bin scatter correlating the probability to buy the comprehensive UI coverage in year $t$ with the predicted number of days unemployed of individual $i$, respectively under the basic coverage $\hat{\pi}_0$ and under the comprehensive coverage $\hat{\pi}_1$, based on her observable characteristics year $t$. The measures of predictable unemployment risk under basic and comprehensive coverage are from the model presented in Section 2.3. The model combines flexibly all observable sources of risk, including institutional shifters of risk such as the full history of the firm layoff notifications, and the relative tenure ranking of the individual. Model selection is based on the Lasso approach for zero-inflated poisson suggested by Banerjee et al. [2018]. To allow for moral hazard, we estimate a model of risk for individuals under the basic coverage, and a separate model of risk for individuals under the comprehensive coverage. The model predicts the number of days spent unemployed in year $t+1$ based on observable characteristics in year $t$.

of all other workers within firm $j$ excluding individual $i$ over all years where the firm is observed active in our sample years.

In section Appendix A we showed that there is significant heterogeneity in these average firms' separation rates, and that individuals' unemployment risk is very strongly correlated with this average firm level risk (panel A of Figure A.1).

We now investigate how average firm level risk correlates with unemployment insurance choices.

**Cross-Sectional Evidence** The first strategy consists in simply using the cross-sectional variation in displacement risk across firms. In Figure C.2 panel A, we group individuals in 50 equal size bins of firm layoff risk, and plot their average firm layoff risk against their average probability of buying supplemental coverage, residualized on the same vector $Z$ of baseline controls affecting UI contracts used in the positive correlation test of Section 4.1.

The graph displays a strong positive correlation between firm layoff risk and individuals' probability to buy the comprehensive UI coverage

We then estimate the correlation between average firm level risk $\bar{\pi}_{-i,j}$ and willingness-to-pay by running the following two-stage least square specification:

$$\begin{aligned} D_i &= \beta_{2SLS} \cdot \pi_i + Z_i' \alpha_1 + \epsilon \\ \pi_i &= \zeta \cdot \bar{\pi}_{-i,j} + Z_i' \alpha_2 + \eta \end{aligned} \tag{24}$$

where $D_i$ is our indicator variable for buying the supplemental coverage. This specification instruments individual realized risk by the average firm layoff risk and therefore exploits only variation in predictable risk coming from average firm layoff risk. For useful comparison, we also report the coefficient estimate $\beta_{OLS}$ of the following OLS specification correlating $D$ with individual risk:

$$D_i = \beta_{OLS} \cdot \pi_i + Z_i' \alpha + \nu \tag{25}$$

We estimate both models on our baseline sample of workers pooling all observations for years 2002 to 2006. We use as a measure of realized risk $\pi_i$ the realized displacement risk excluding quits in year $t+1$. We find a positive and strongly significant coefficient $\beta_{2SLS} = .50$ (.01) indicating that workers who work in firms that exhibit higher turnover rates are significantly more likely to buy the comprehensive coverage.

We also find that $\beta_{2SLS}$ is much larger than $\beta_{OLS}$, which is also informative. Clearly, the two-stage least square procedure removes potential attenuation bias from measurement error in $\beta_{OLS}$. But the two-stage least square, by projecting choices only on the average firm layoff dimension of displacement risk introduces some potential selection, if $\text{Cov}(\bar{\pi}_{-i,j}, \epsilon) \neq 0$. In other words, if workers who self-select into riskier firms are different along observed or unobserved characteristics correlated with willingness-to-pay for insurance, $\beta_{2SLS}$ will capture this additional selection effect.

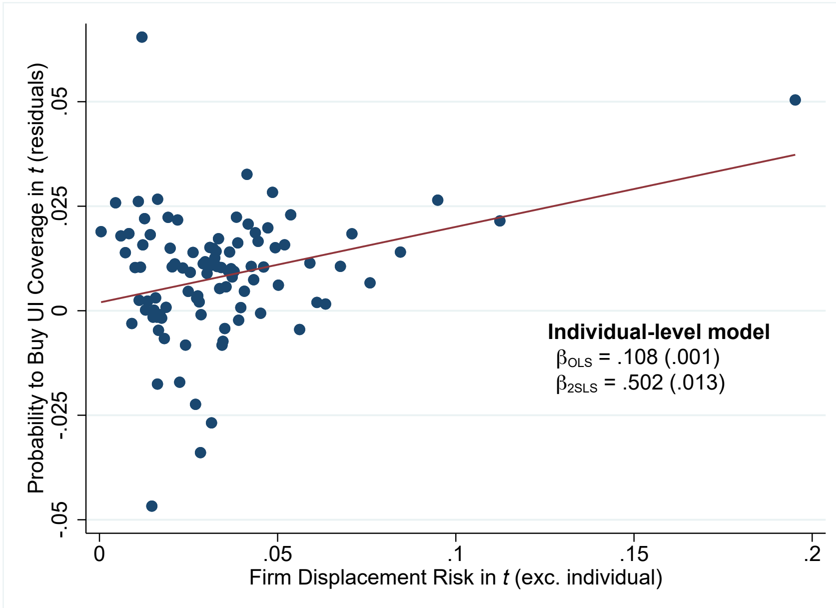
In panel B of Figure C.2, we explore the importance of such selection along observable characteristics in explaining the magnitude of $\beta_{2SLS}$. We introduce in the vector $Z$ of specifications (24) and (25) a rich set of additional controls: age, gender, marital status, education (four categories), industry (1-digit code), occupation (1-digit code), wealth level (quartiles) and past unemployment history (dummies for having been unemployed in $t-1$, $t-2$ and up to $t-8$). We still find a strong positive correlation between insurance choices and firm layoff risk ($\beta_{2SLS} = .245$). But adding these controls decreases the magnitude of the correlation between risk and UI choices significantly.

Even with this rich set of controls, $\beta_{2SLS}$ might still be picking some correlation between average layoff risk and *unobserved* characteristics affecting UI choices. This will be the case if workers who select to work in riskier firms have different preferences for insurance and/or if the there is an unobserved effect of riskier firm environments on insurance choices: firms with high turnover may have different prevalence of collective bargaining, different firm cultures that can affect individuals' UI choices for instance.

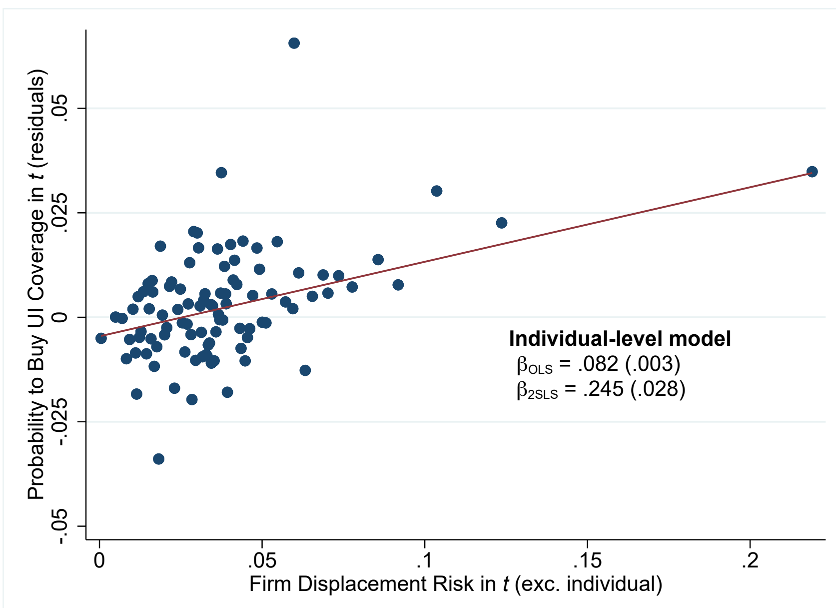
Decomposing the error term in specification (24) $\epsilon = \kappa_i + \rho_j$ into an individual specific component $\kappa_i$ and a firm specific component $\rho_j$, we can think of the the selection introduced by average layoff risk as the combination of individual fixed effects and firm fixed effects. We first move to a firm switcher design that allow us to control more directly for the unobserved individual specific component $\kappa_i$. In subsection C.3 we then show how to deal with both the individual specific ($\kappa_i$) and firm specific ($\rho_j$) sources of potential selection.

Figure C.2: Firm Level Risk and UI Coverage Choice

A. Baseline Controls for Contract Space




B. With Additional Demographic Controls




**Notes:** The Figure uses cross-sectional variation in displacement risk across firms as a risk shifter to estimate how UI coverage choices react to variation in risk that is not driven by individual moral hazard. Panel A groups individuals in 50 equal size bins of firm layoff risk, and plot their average firm layoff risk against their average probability of buying supplemental coverage, residualized on the same vector $X$ of baseline controls affecting UI contracts used in the positive correlation test of Section 4.1. We report on the graph the coefficient $\beta_{OLS}$ from an OLS regression of specification (25) and then the estimated coefficient $\beta_{2SLS}$ from our two-stage least square model (24) where we use $Z = \pi_{-i,j}$ as a risk shifter. In panel B, we replicate the same procedure, but now add to the regression the same rich set of demographic controls used in Figure 6, and find a similar strong positive correlation between insurance choices and firm layoff risk.

**Firm Switcher Design** In this strategy, we use the panel dimension of the data to control for the selection introduced by individual specific heterogeneity $\kappa_i$.

To this end, we focus on individuals who change firms ("switchers"). The employer-employee matched data (*RAMS*) registers all existing labor contracts on a monthly basis. We define a switch as moving from having a labor contract with firm $j$ (the origin firm) to having a contract with firm $k$ (the destination firm), without any recorded non-employment spell between these two contracts. We focus on individuals with more than 1 year of tenure in the origin firm. Switchers experience a change in their layoff risk coming from underlying variation in two risk shifters: their tenure ranking changes, and so does their underlying firm layoff risk.

First, switchers experience a reduction in their relative tenure ranking, as they become the "last-in" when they move to the destination firm. To document the magnitude of the variation in relative tenure ranking and corresponding layoff risk, following a firm switch, we define year $n = 0$ as the year of a firm switch, and run, on the sample of firm switchers, event studies of the form:

$$T_{i,n} = \sum_k \delta_k \cdot \mathbb{1}[n = k] + \mathbf{Z_i'}\alpha + \epsilon_{i,n} \tag{26}$$

where $T_n$ denotes the tenure ranking of individual $i$ in event year $n$, $\mathbb{1}[n = k]$ are a set of event time dummies, and $Z$ is the vector of baseline controls affecting UI contracts defined in section 4.1. Figure C.3 displays the evolution of relative tenure ranking of switchers as a function of event time by plotting the coefficients $\delta_k$, taking event time $n = -1$ as the omitted category. The graph confirms that relative tenure ranking decreases sharply at the moment of the firm switch.

Figure C.4 panel A explores how this variation in relative tenure ranking affects the probability of displacement over event time $n$. To this effect, we estimate a similar event study specification as in (26) where we use the probability of displacement $\pi_i$ in year $t + 1$ as an outcome. The graph shows that the displacement rate one year ahead increases sharply and significantly at the time of the firm switch.

In Figure C.5 panel A, we run a similar event study specification with $D_i$, a dummy for buying the comprehensive UI coverage as an outcome. The figure shows that the probability of buying the comprehensive coverage increases sharply by about 2.2 percentage points at the time of the firm switch. On the graph, we also display the coefficient from the following two-stage least square specification:

$$\begin{aligned} D_{i,n} &= \kappa_i + \beta_{2SLS} \cdot \pi_{i,n} + \mathbf{Z'_{i,n}}\alpha_\mathbf{1} + \epsilon_{i,n} \\ \pi_{i,n} &= \nu_i + \zeta \cdot \mathbb{1}[n \geq 0] + \mathbf{Z'_{i,n}}\alpha_\mathbf{2} + \eta_{i,n} \end{aligned} \tag{27}$$

where we use a dummy for having switched firm ($\mathbb{1}[n \geq 0]$) as risk shifter for individual displacement probability $\pi_{i,n}$ and control for individual fixed-effects. This specification is estimated on the sample of all workers who ever experience a firm switch between 2002 and 2006 and who have more than 1 year of tenure in the origin firm. $\beta_{2SLS}$ is positive and strongly significant, which again indicates that the positive correlation tests are not simply picking up moral hazard responses to insurance coverage.

While these event study specifications control for fixed underlying heterogeneity across individuals that may affect their UI choices ($\kappa_i$), one concern with this original implementation of the firm switchers design is that individuals are somewhat inert, and decide to reoptimize their UI choices only at specific times, like, for instance, when they switch firm.

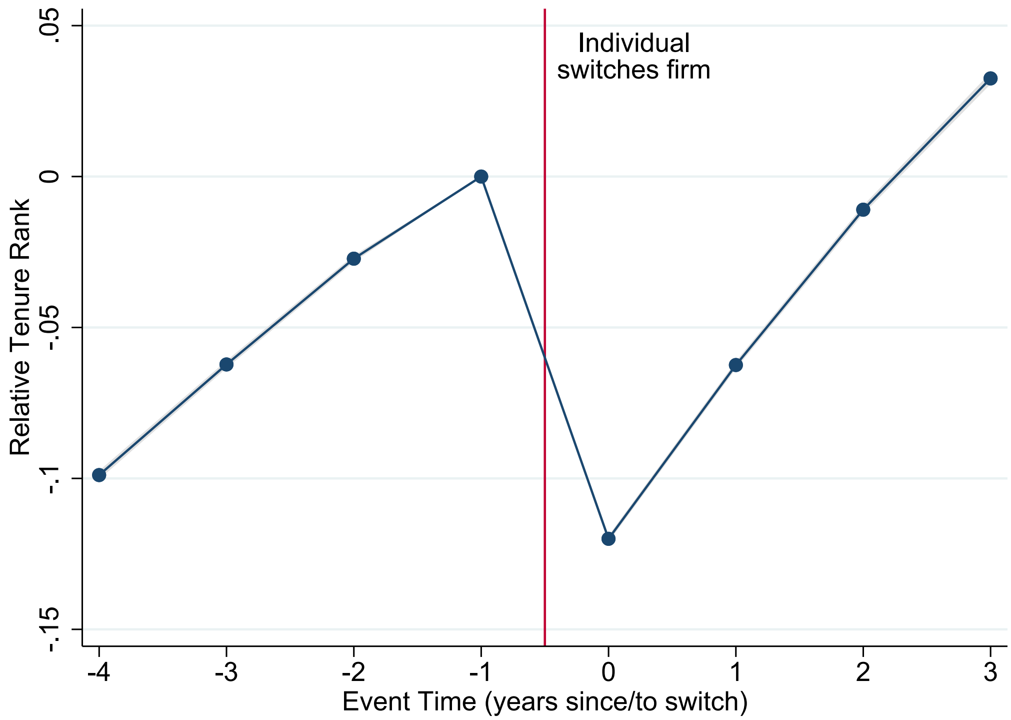
To mitigate the concern that the surge in UI coverage at the time of the switch is the result of the specific timing of UI choices and not a response to the change in underlying risk, we exploit additional variation in risk in the switchers design coming from changes in underlying firm layoff risk. While all switchers experience an increase in their displacement probability due to the decline in their tenure ranking, the effect of a switch on individual displacement probability exhibits large differences according to whether their destination firm is much riskier ("positive shock") or a lot less risky ("negative shock") than their origin firm. We therefore split the population of switchers according to their rank in the distribution of $\Delta_{j,j'}\bar{\pi}_{-i} = \bar{\pi}_{-i,j'} - \bar{\pi}_{-i,j}$, the change in their underlying average firm layoff risk when moving from firm $j$ to firm $j'$. In Figure C.4 panel B, we contrast individuals in the bottom decile of $\Delta_{j,j'}\bar{\pi}_{-i}$ (large negative shock, i.e., individuals who experience a large negative decline in their firm layoff risk, going from a high risk to a low risk firm), and individuals in the top decile of $\Delta_{j,j'}\bar{\pi}_{-i}$ (large positive shock, i.e., individuals who experience a large increase in their firm layoff risk going from a low risk to a high risk firm). The Figure confirms that individuals experiencing a large positive shock in their firm layoff risk exhibit a significantly larger increase, of about 2 percentage point, in their displacement probability at the time of the switch, relative to individuals experiencing a large negative shock.

In panel B of Figure C.5, we now compare the evolution of insurance choices around firm switch for individuals experiencing large positive vs large negative shocks. We run event study specification (26) with $D_i$, a dummy for buying the comprehensive UI coverage as an outcome, separately for the sample of individuals experiencing large positive shocks and for the sample of individuals experiencing large negative shocks. The graph indicates that the increase in the probability to buy UI around firm switch is significantly larger (by about 1.5 percentage point) among individuals moving to significantly more risky firms relative to those moving to less risky firms. We also report on the graph the estimated coefficient $\beta_{2SLS} = .57$ (.08) of the two-stage model:

$$\begin{aligned} D_{i,n} &= \kappa_i + \beta_{2SLS} \cdot \pi_{i,n} + \textstyle\sum_k \delta_k \cdot \mathbb{1}[n=k] + \mathbf{Z'_{i,n}}\boldsymbol{\alpha_1} + \epsilon_{i,n} \\ \pi_{i,n} &= \nu_i + \zeta \cdot \mathbb{1}[n \geq 0] \cdot \Delta\bar{\pi}_{-i,j} + \mathbf{Z'_{i,n}}\boldsymbol{\alpha_2} + \eta_{i,n} \end{aligned} \tag{28}$$

This model uses firm switch interacted with the change in average firm level layoff risk $\Delta\bar{\pi}_{-i,j}$ as risk shifter for individual displacement probability. This model estimated on the sample of all workers who ever experience a firm switch between 2002 and 2006 and who have more than 1 year of tenure in the origin firm. The results suggest that the probability to buy the comprehensive coverage is strongly correlated with average firm layoff risk, even after controlling for individual unobserved heterogeneity with this switcher design strategy.

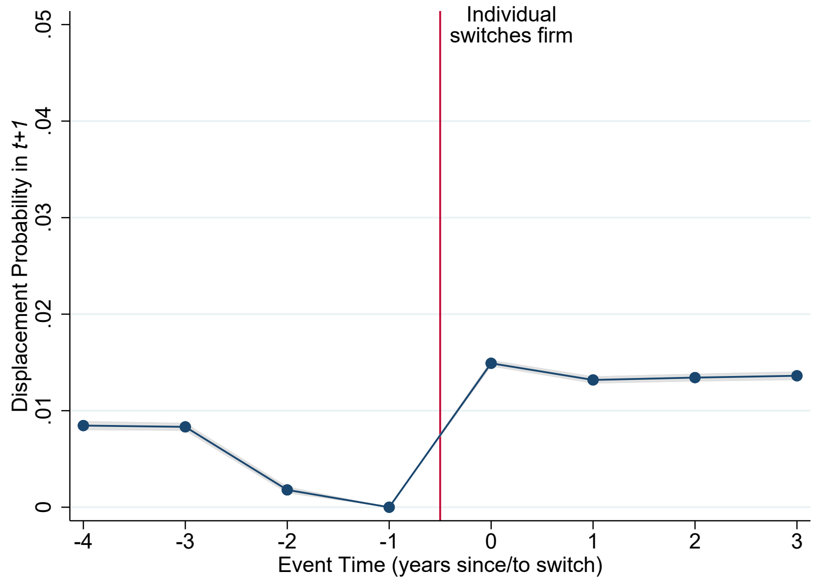
Figure C.3: Switchers Design: Relative Tenure Ranking as a Function of Event Time



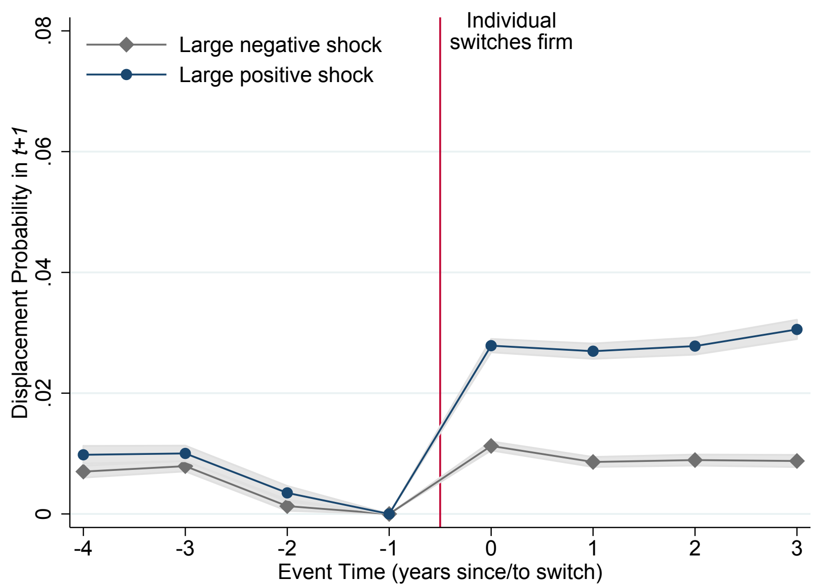

**Notes:** The Figure focuses on "firm switchers", i.e. individuals moving from having a labor contract with firm $j$ to having a contract with firm $k$, without any recorded non-employment spell between these two contracts. We focus on individuals with more than 1 year of tenure in the origin firm. In this Figure we show that switchers experience a variation in their layoff risk coming from underlying variation in their relative tenure ranking. Relative tenure ranking affects displacement probability due to the strict enforcement of the Last-In-First-Out (LIFO) principle in Swedish labor laws. To follow the rules pertaining to the application of LIFO, relative tenure ranking is defined within each establishment times occupation group using the RAMS employer-employee data since 1985. The chart displays estimates of the event study specification (26) using relative tenure ranking as an outcome. The graph shows that relative tenure ranking drops abruptly at the time of the firm switch. Panel A of Figure C.4 shows that this drop in tenure ranking translates in a significant increase in displacement risk.

Figure C.4: Firm Switchers - Displacement Rate in $t+1$ as a Function of Time to/since Firm Switch

A. All Switchers



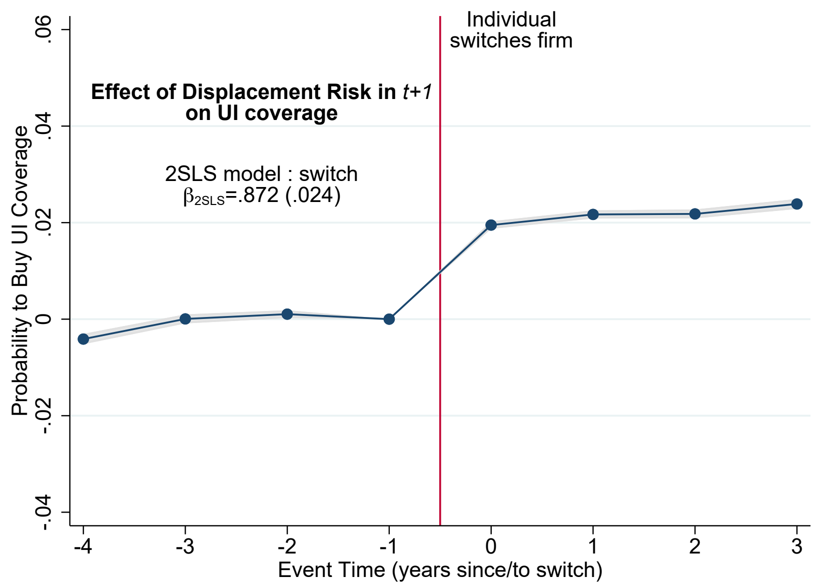

B. Switchers Experiencing Large Positive Firm Layoff Risk Shock
vs Large Negative Firm Layoff Risk Shock



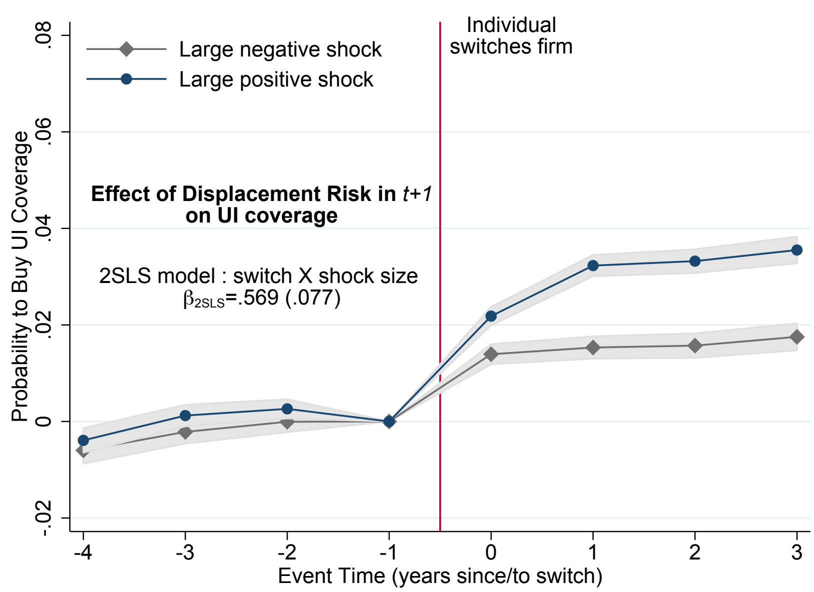

**Notes:** The Figure focuses on "firm switchers", i.e. individuals moving from having a labor contract with firm $j$ to having a contract with firm $k$, without any recorded non-employment spell between these two contracts. We focus on individuals with more than 1 year of tenure in the origin firm. Switchers experience a variation in their layoff risk coming from underlying variation in both risk shifters: their tenure ranking changes, and so does their underlying firm layoff risk. In panel A, we display estimates of the event study specification (26) using displacement risk in $t+1$ as an outcome. The graph shows that the displacement risk increases sharply and significantly at the time of the firm switch. In panel B, we split the population of switchers according to their rank in the distribution of $\Delta_{j,j'}\pi_{-i} = \pi_{-i,j'} - \pi_{-i,j}$, the change in their underlying firm risk when moving from firm $j$ to firm $j'$. We focus on individuals in the bottom decile of $\Delta_{j,j'}\pi_{-i}$ (large negative shock, i.e., individuals going from a high risk to a low risk firm), and individuals in the top decile of $\Delta_{j,j'}\pi_{-i}$ (large positive shock).

Figure C.5: Firm Switchers - UI Coverage Choices as a Function of Time to/since Firm Switch

A. All Switchers




B. Switchers Experiencing Large Positive Firm Layoff Risk Shock vs Large Negative Firm Layoff Risk Shock




**Notes:** The Figure focuses on "firm switchers". In panel A, we display estimates of the event study specification (26) using UI coverage $V$ as an outcome. The Figure shows that the probability of buying the comprehensive coverage increases sharply at the time of the firm switch. In panel B, we split the population of switchers according to their rank in the distribution of $\Delta_{j,j'}\pi_{-i} = \pi_{-i,j'} - \pi_{-i,j}$, the change in their underlying firm risk when moving from firm $j$ to firm $j'$, as in Figure C.4 panel B. The graph indicates that the increase in the probability to buy UI around firm switch is significantly larger among individuals moving to significantly more risky firms relative to those moving to less risky firms. On both panels, we display the coefficient from a two-stage least square fixed-effect specification similar to (24) where we use firm switch (and firm switch interacted with shock size) as risk shifter $Z$ for individual displacement probability.

## C.3 Risk Shifter & UI Choices II: Layoff Notifications and LIFO

The previous section suggests a strong correlation between firm layoff risk and UI choices, indicative of the presence of significant adverse selection. As explained above though, firm layoff risk may be correlated with willingness-to-pay for UI, either through unobserved individual specific heterogeneity ($\kappa_i$) or unobserved firm specific heterogeneity ($\rho_j$). The firm switcher design above deals with individual specific heterogeneity ($\kappa_i$), but may still pick up selection on firm level heterogeneity $\rho_j$ if firm heterogeneity is correlated with $\Delta\bar{\pi}_{-i,j}$.

We now show how layoff notifications and the application of the Last-In-First-Out (LIFO) principle enables to identify the effect of predictable risk on UI choices controlling jointly for firm level heterogeneity $\rho_j$ and individual level heterogeneity $\kappa_i$. We leverage the fact that layoff notifications and LIFO creates variation in layoff risk both within firm and across individuals over time.

In section 2.1, we described the institutional details of the layoff notification system and its interaction with the LIFO rule. We also explained and demonstrated in Appendix A, that layoff notifications signal a significant change in a firm layoff risk. In particular, we reported in Figure A.2 that the displacement probability of workers increases sharply and significantly around the first layoff notification event in the history of the firm. We also showed in Figure A.1 panel B that the effect of a layoff notification on displacement probability is strongly heterogenous depending on the relative tenure ranking of workers. Workers with relative tenure ranking below .5 have a much higher probability of being laid-off following a layoff notification than workers with relative tenure ranking above .5.

We now show how UI choices correlate with this variation in risk stemming from the interaction between a notification event and relative tenure ranking. We follow the same event study empirical approach as in section Appendix A around the event of a layoff notification. We define event year $n$ as the year to/since the firm emits its first layoff notification.

Our sample is the panel of workers who are employed in the firm at the date this layoff notification is emitted to the PES. All these workers constitute our treatment group. We follow, as in Appendix A a matching strategy and create a control panel of workers. To do this, we use nearest-neighbor matching to select a set of firms that are similar, along a set of observable characteristics, to the firms emitting a layoff notification, but never emit a layoff notification.[51] We allocate to the matched firm in the control group a placebo event date equal to the layoff notification date of her nearest-neighbor in the treated group of firms. We then select workers that are in the control firm at the time of the placebo event date to create our matched panel of control individuals.

In Figure C.6 we split the sample by tenure ranking at the time the layoff notification is emitted and report the evolution of the average fraction of individuals buying the supplemental coverage in our treatment group and in the matched control group.[52] Panel A of Figure C.6 reports the

[51]The covariates used for matching are the number of employees, the 4 digit sector codes of the firm, the average earnings and average years of education of workers in the firm.

[52]For control workers we use their tenure ranking at the time of the placebo layoff notification.

evolution of the fraction buying the supplemental UI coverage for workers with relative tenure ranking below 50% in year $n = 0$. The graph shows that UI coverage increases significantly among the treated group, starting one year before the layoff notification is sent, which suggests the existence of some degree of private information among workers regarding the timing of the layoff notification. In panel B, we report the evolution of UI choices for the sample of workers with relative tenure ranking above 50% in year $n = 0$. The graph displays no sign of variation in the fraction of individuals buying the comprehensive coverage around the notification event.

On both panels, we also report estimates $\hat{\beta}$ of the reduced form specification:

$$D_{i,n} = \kappa_i + \rho_j + \beta \cdot \mathbb{1}[n \geq 0] \cdot \mathbb{1}[T = 1] + \theta \cdot \mathbb{1}[n \geq 0] + \mathbf{Z'_{i,n}}\alpha_{\mathbf{1}} + \epsilon_{i,n} \tag{29}$$

as well as estimates $\hat{\beta}_{2SLS}$ from the following two-stage specification:

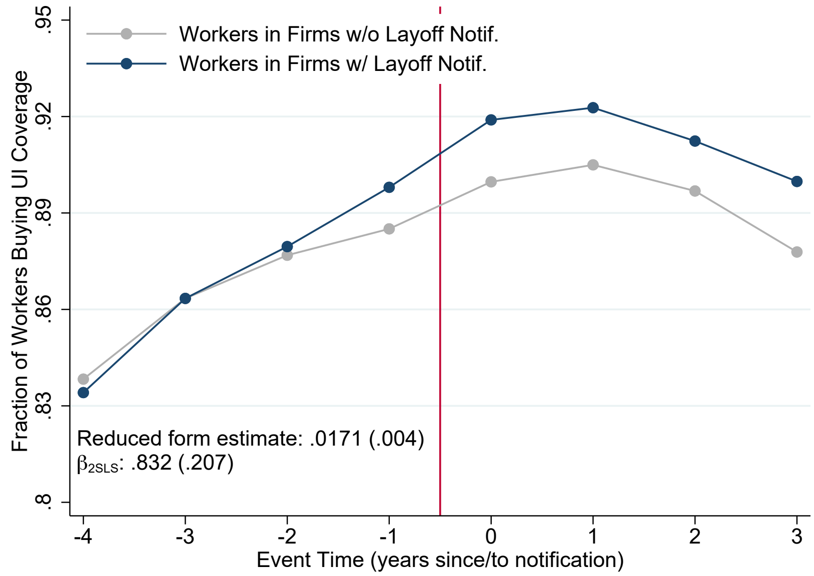
$$\begin{aligned} D_{i,n} &= \kappa_i + \rho_j + \beta_{2SLS} \cdot \pi_{i,n} + \textstyle\sum_k \delta_k \cdot \mathbb{1}[n = k] + \mathbf{Z'_{i,n}}\alpha_{\mathbf{1}} + \epsilon_{i,n} \\ \pi_{i,n} &= \nu_i + \gamma_j + \zeta \cdot \mathbb{1}[n \geq 0] \cdot \mathbb{1}[T = 1] + \theta \cdot \mathbb{1}[n \geq 0] + \mathbf{Z'_{i,n}}\alpha_{\mathbf{2}} + \eta_{i,n} \end{aligned} \tag{30}$$

The above two-stage model specification uses variation in risk stemming from being in a firm having emitted a layoff notification, and controls for both individual fixed effects ($\kappa_i$) and firm fixed effects ($\rho_j$). The comparison between the estimates for the low vs high tenure ranking sample further exploits the additional layer of variation in displacement risk coming from the interaction between a notification event and relative tenure ranking. Results show that individuals with low tenure ranking strongly respond to the variation in risk arising from a layoff notification and are significantly more likely to buy the comprehensive coverage as a result: $\beta_{2SLS} = .84$ (.21). To the contrary, the UI choices of individuals with high tenure ranking do not significantly respond to a layoff notification.

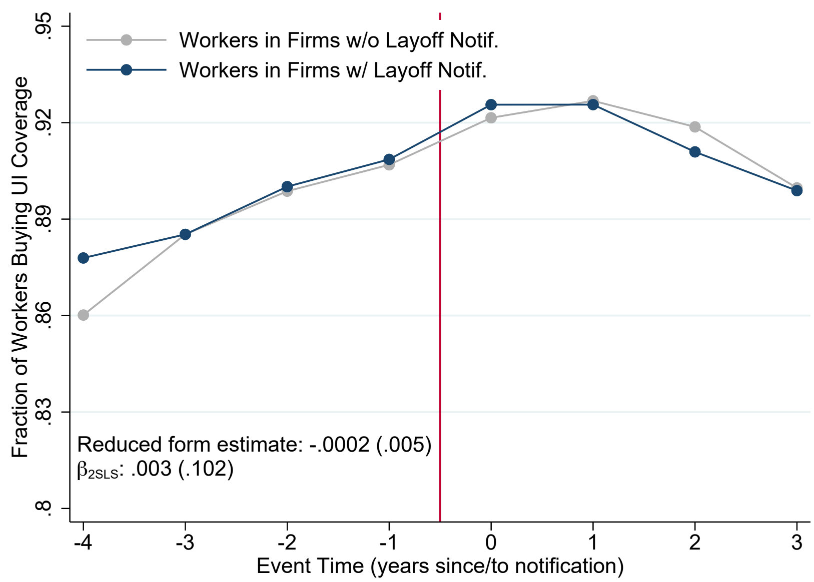
**Summary of evidence** Taken together, the evidence from this appendix strongly suggests that UI choices do significantly respond to the various sources of variations in individuals' predictable unemployment risk. The different strategies clearly differ in terms of the way they control for underlying selection on unobserved heterogeneity into the comprehensive coverage, as well as in terms of the population of compliers. Yet, we systematically find a strong positive and significant relationship between the probability to buy the comprehensive coverage and the observable risk shifters entering our predicted risk model. This overall confirms that the strong correlation between predictable risk and UI choices documented in section 4.2 does capture the presence of significant adverse selection into the comprehensive coverage.

Figure C.6: Layoff Notification

A. Workers with Relative Tenure Ranking < .5 at Event Time 0




B. Workers with Relative Tenure Ranking ≥ .5 at Event Time 0




**Notes:** The Figure uses layoff notification events interacted with relative tenure ranking as a source of variation in displacement risk to investigate how UI coverage choices react to variations in underlying risk. Panel A reports the evolution of UI coverage around the time of the first layoff notification for the panel of workers in the treated group and for workers in our placebo (control) group, restricting the sample to workers with relative tenure ranking below 50% in year $n = 0$. The Figure shows that UI coverage increases significantly among the treated group, starting one year before the layoff notification is sent, which suggests the existence of some degree of private information among workers regarding the timing of the layoff notification. In panel B, we report similar estimates but for the sample of workers with relative tenure ranking above 50% in year $n = 0$. The graph displays no sign of variations in individuals insurance coverage among the event. On both panels, we display the estimated coefficient $\beta_{2SLS}$ of our two-stage least square model using the layoff event interacted with tenure and a dummy for being in the treatment group as a risk shifter $Z$.

# Appendix D Price Variation: Additional Material

In this appendix we present additional results using the 2007 price variation to identify adverse selection:

(i) we present further non-parametric evidence of adverse selection using additional risk outcomes

(ii) we show how adverse selection would survive the inclusion of many unused demographic observables in the Swedish UI policy

(iii) we provide evidence showing that our ranking of willingness-to-pay for the comprehensive coverage correlates strongly with proxies for the value of unemployment insurance and for risk preferences.

(iv) we address potential concerns, such as inertia, to the validity of our ranking of individuals by willingness-to-pay.

**Alternative risk outcomes** In our baseline analysis of the 2007 reform in section 5, we use total number of days unemployed in 2008 as our main outcome. Here, we show that our estimates of adverse selection are again robust to using alternative risk outcomes. We look at the displacement probability in t+1, t+2,... up to t+5. To control for observables $Z$, we model the probability of displacement as a probit:

$$E(\pi|Z) = \Phi(Z'\beta + \sum_j \alpha_j \cdot \mathbb{1}[D = j]) \tag{31}$$

where $\Phi(.)$ is the standard normal c.d.f.

In Figure D.1 we report the correlation between willingness-to-pay in 2007 and displacement outcomes in t+1, t+2,... up to t+5. We report for each year the semi elasticity

$$\text{Semi}^{t+k}_{M(\mathbf{p})} = \frac{E(\pi_{t+k}|Z, D = M) - E(\pi_{t+k}|Z, D = 0)}{E(\pi_{t+k}|Z, D = 0)}$$

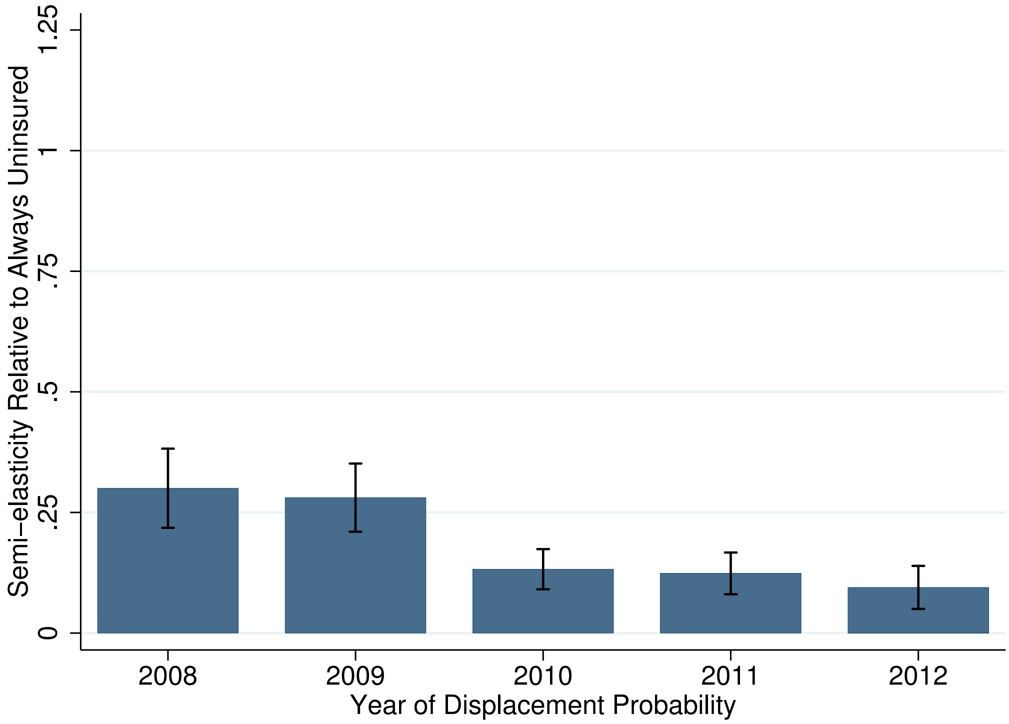
of the displacement probability in $t+k$ for the marginals $D = M$ relative to the individuals in the basic coverage throughout $D = 0$. The figure reveals that the correlation between unemployment risk and willingness-to-pay decreases rapidly as we consider later years, but remains statistically significant up to five years.

**Role of Unpriced Heterogeneity** The 2007 price reform allows us to investigate how much of the risk-based selection is driven by selection on specific unpriced observables correlated with risk.

We do so by sequentially including in specification (16) a set of controls $X$: dummies for age, gender, marital status, education (four categories), industry (1-digit code) and and wealth level (quartiles). We then report for each specification the semi-elasticity $\text{Semi}^{X}_{M(\mathbf{p})} = \frac{E(\pi|Z,X,D=M)-E(\pi|Z,X,D=0)}{E(\pi|Z,X,D=0)}$.

Interestingly, the semi-elasticity increases compared to our baseline when including age as a control. Age is therefore a driver of advantageous selection into UI. Adding rich sets of controls

Figure D.1: Price Variation: Using Displacement Probability as a Risk Outcome



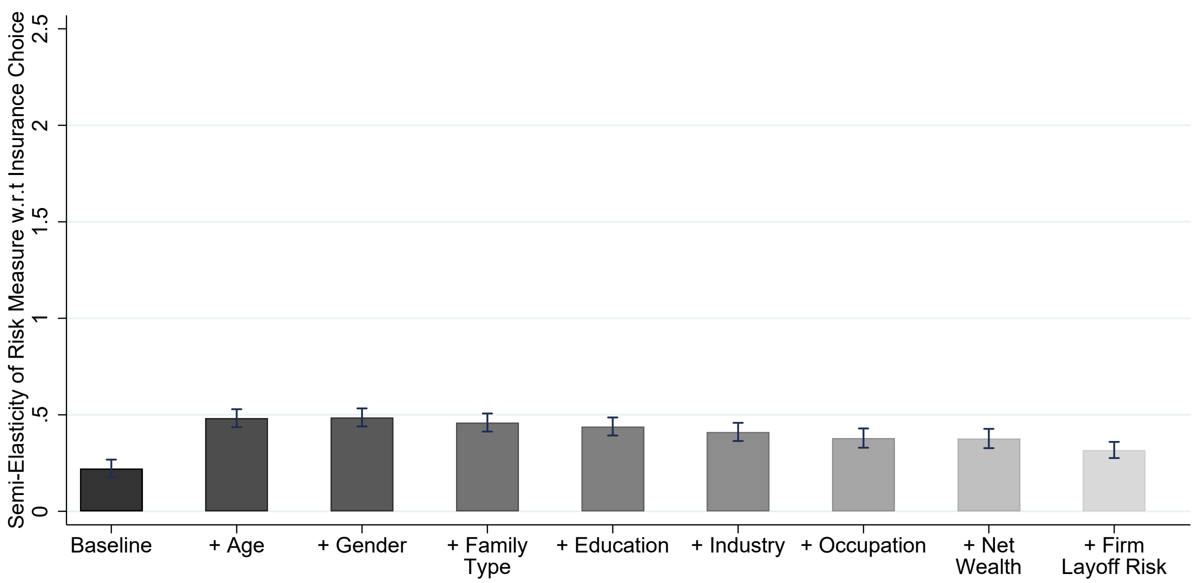

**Notes:** The Figure reports the correlation between willingness-to-pay in 2007 and realized displacement outcomes in 2008, 2009,.. up to 2012. We report for each year, the semi-elasticity $\text{Semi}^{t+k}_{M(\mathbf{p})}$ of the displacement probability in year $t+k$ for the marginals $M$ relative to the individuals from group 0.

for education, industry, occupation and wealth decreases the estimated correlation only slightly, indicating that there is little risk-related selection along these margins.

Overall, this suggests that demographic characteristics, and age especially, provide advantageous selection on average, such that if contracts were differentiated along these observable dimensions, adverse selection into comprehensive coverage would actually be more severe.

Furthermore, controlling for these unpriced observables does not exhaust risk-based selection in the supplemental UI coverage. In other words, even if the supplemental coverage policy were to price this rich set of observable characteristics, a significant amount of adverse selection would remain.

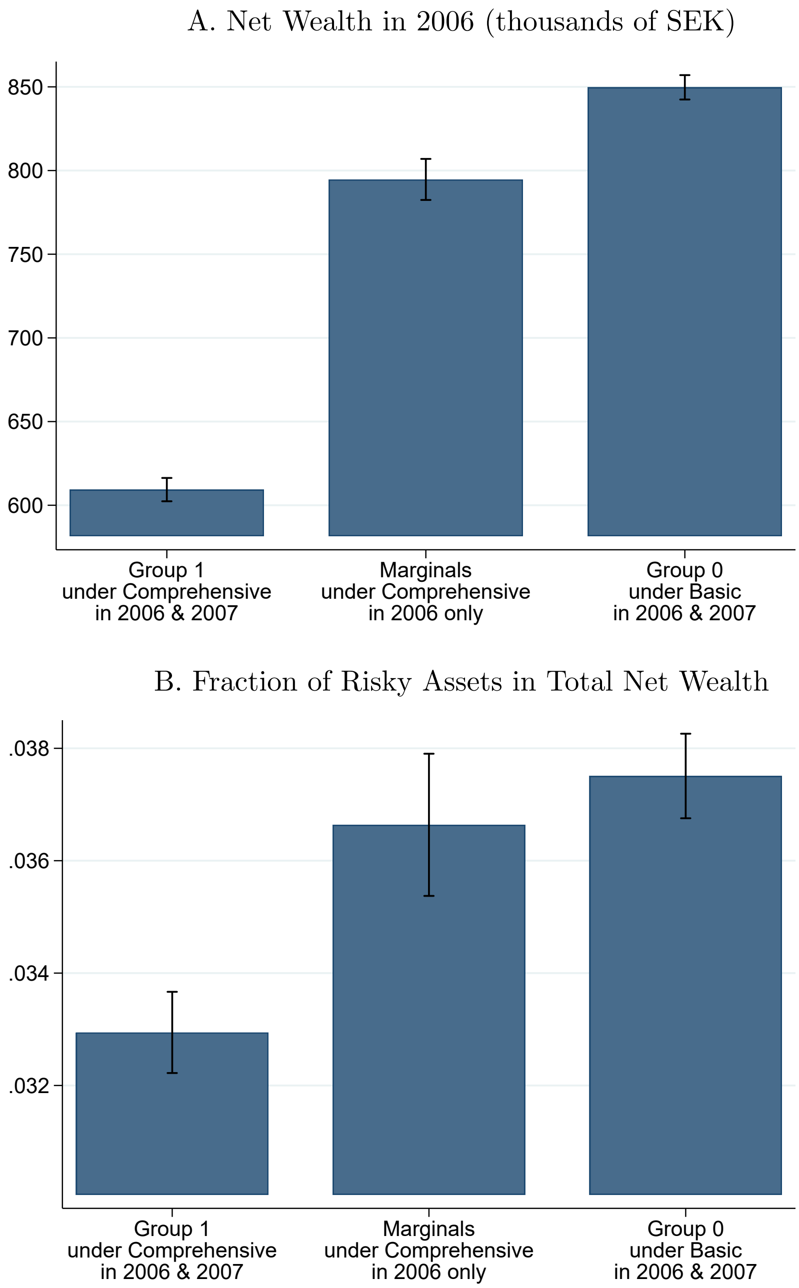
Figure D.2: Price Variation: Role of Unpriced Heterogeneity




**Notes:** The Figure explores to what extent estimated adverse selection using the 2007 price variation is driven by selection on observable characteristics that are unpriced in the Swedish UI system. We report the semi-elasticity $\text{Semi}^X_{M(\mathbf{p})}$ of the number of days spent unemployed in 2008 for the marginals $M$ relative to the individuals from group 0. We start with the baseline estimate only controlling for the characteristics affecting the actual UI policy, and show how the semi-elasticity evolves as we add sequentially more characteristics to the vector of controls $X$. We start by adding demographic controls (age, then gender, and marital status), then controls for skills and other labor market characteristics (controls for education (four categories), industry (1-digit code), occupation (1-digit code) and wealth level (quartiles).

**Selection on preferences and value of UI** The 2007 price reform also allows to investigate patterns of selection along dimensions other than risk. In Figure D.3, we examine how characteristics that determine the value of unemployment insurance and proxy for risk preferences correlate with willingness-to-pay for insurance revealed by the 2007 price variation. Panel A correlates the level of individuals' net wealth in 2006 in thousands of SEK with their willingness-to-pay controlling for age. Individuals with larger net wealth have more means to smooth consumption in case of displacement, and as a result, should value extra coverage less. The graph indeed confirms the presence of a clear monotonic relationship between net wealth and willingness-to-pay: individuals from group 0 have significantly larger net wealth than the marginals $M$, who have significantly more net wealth than the individuals from group 1. In panel B, we probe into the potential amount of selection based on risk-preferences. To proxy for risk aversion, we use the fraction of total net wealth invested in risky assets (stocks). The graph shows that the individuals in comprehensive coverage have a significantly larger fraction of risky assets in their portfolio than the marginals and the individuals in basic coverage, conditional on net wealth. This evidence is in line with more risk-averse individuals valuing the extra coverage more.

Figure D.3: Price Variation: Selection on Preferences

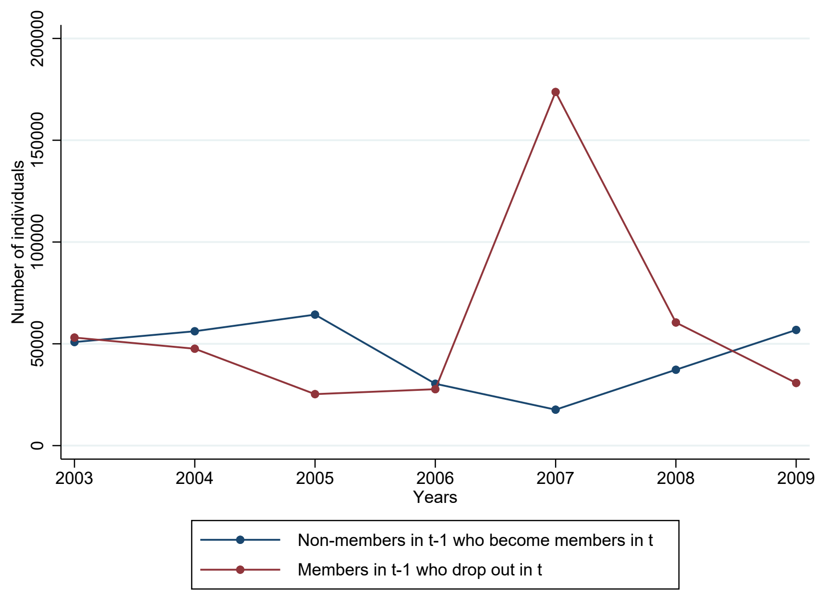



**Notes:** The Figure uses the 2007 price reform to rank individuals according to their willingness-to-pay for the supplemental coverage $\mathbf{u}$, and uses this ranking to correlate $\mathbf{u}$ with proxies for the value of unemployment insurance and risk preferences. In both panels, individuals are ranked by decreasing order of $\mathbf{u}$. Group 1 on the left are individuals who are insured with the comprehensive coverage both in 2006 and 2007 and have the highest level of $\mathbf{u}$. The middle group corresponds to the marginals ($M(\mathbf{p})$): individuals who were insured with the comprehensive coverage in 2006 but switch out in 2007 when the premium increases. They have a lower level of $\mathbf{u}$ than group 1, but a higher level of $\mathbf{u}$ than the last group on the right (0), of individuals who neither buy the supplemental coverage in 2006, nor in 2007. Using this ranking, we correlate in panel A willingness-to-pay with the level of net wealth in 2006. Individuals with higher net wealth have better means to smooth consumption in case of displacement and should have a lower valuation of additional unemployment insurance. We winsorize net wealth and eliminate the bottom and top percentile of the distribution. In panel B, we proxy for risk aversion using the fraction of total net wealth invested in risky assets (stocks). In both panels we report the average outcome of each group controlling for our baseline vector of characteristics $Z$ plus a cubic polynomial for age, and a cubic for net wealth in panel B.

**Robustness** Our partition of the population in terms of willingness-to-pay implicitly assumes that $\mathbf{u}$ is constant over time, or to be more precise that the ranking of individuals' willingness-to-pay is the same in 2006 and 2007. In practice $\mathbf{u}$ may change over time, due for instance to idiosyncratic shocks to risk, or preferences, thus creating a flow of individuals switching out of the comprehensive plan, even absent price changes. Appendix Figure D.4 provides evidence that the flow of individuals who switch out of the supplemental coverage was in fact very small prior to the 2007 price reform, but experienced a sudden surge in 2007. This alleviates the concern that our ranking of individuals by willingness-to-pay is confounded by underlying changes in individuals' preferences or risks.

We also note that our partition of the population ignores a negligible fourth group of individuals, who were not buying the comprehensive plan in 2006, but switched in the comprehensive plan in 2007. The size of this group is seven times smaller than the group of individuals switching out of the comprehensive plan in 2007. The ranking of this fourth group in terms of willingness-to-pay is also ambiguous, as one would need to include idiosyncratic shocks to $\mathbf{u}$ to account for the fact that these individuals switched in the comprehensive coverage in 2007 despite the increase in prices $\mathbf{p}$. We display in Appendix Figure D.4 the evolution of the flow of individuals not buying the comprehensive plan in $t-1$ but switching in the comprehensive plan in $t$. The graph shows that this flow of individuals was small prior to 2007, and equivalent in size to the flow of individuals switching out, hence the stability in the fraction of individuals insured. The flow of individuals switching in seems to decrease with the 2007 reform, but only slightly. The average unemployment risk of the workers switching into the comprehensive plan was the highest among the four groups throughout this period.

Figure D.4: The 2007 Price Reform: Flows of Individuals Switching in and Switching out of the Comprehensive Coverage over Time




**Notes:** The Figure reports the evolution of the absolute flows of individuals "switching in" and "switching out" of the comprehensive coverage over time. The sample is restricted to individuals were meeting the work eligibility requirement. Individuals who switch in are individuals who were not buying the comprehensive coverage in year $t-1$ but are buying in year $t$ (blue curve). Individuals who switch out are individuals who were buying the comprehensive coverage in year $t-1$ but are no longer buying in year $t$ (red curve). The Figure shows a large and sudden increase in the flow of individuals switching out and a decrease in the flow of individuals switching in, following the large increase in the the premia paid for the supplemental coverage in 2007.

**Inertia** Inertia is a potentially important behavioral friction, which has been shown to be extremely relevant in other social insurance contexts, such as health insurance. We investigate here the role of inertia and how it affects adverse selection identified in the context of the 2007 price variation. In line with the existing the literature (e.g., Handel (2013)), we use job switchers to proxy for differential exposure to inertia. New employees in a firm arguably face a more active choice environment than existing employees. The former have to reoptimize many choices, while the latter remain in a more passive choice environment. In practice, Figure C.7 above confirms that switching job is indeed associated with a significant change in insurance choices.

In Figure D.5 below, we start by looking at how the price reform of 2007 affected insurance choices for individuals in active choice environments (job switchers) relative to individuals in passive choice environments (job stayers). We find that the 2007 price reform immediately decreased the demand for the comprehensive coverage, in similar proportions, in both the active and the more passive choice environment. But we do find a larger response one year after (in 2008) for job switchers than for non-switchers, which suggests the presence of inertia. Overall, though, the graph suggest that inertia plays a relatively limited role in our setting: individuals in passive choice environments reacted strongly to the reform, and their long run demand response is quite similar to that of individuals facing more active choice environments.

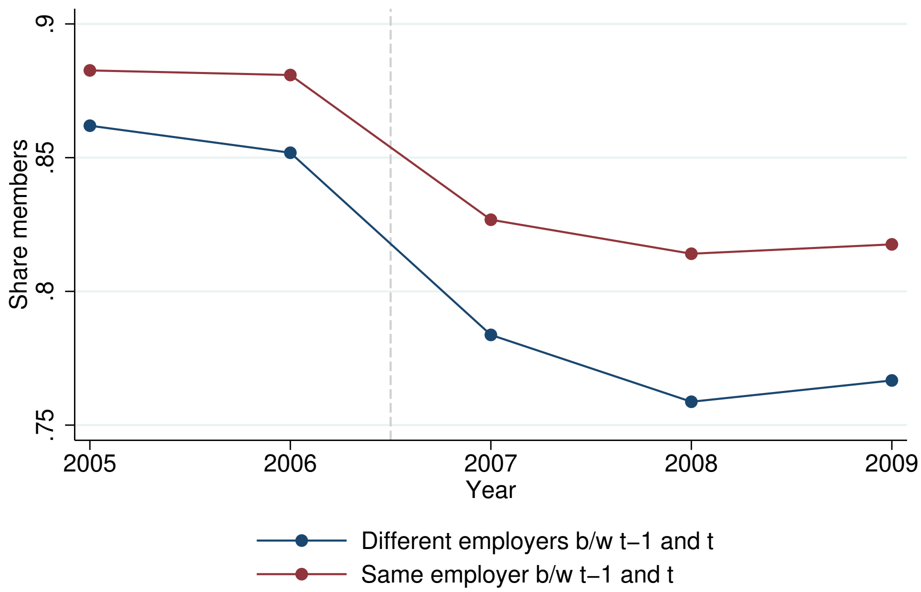
In Figure D.6, we further investigate whether the adverse selection created by these demand responses is different for individuals in active vs passive choice environments. We report the semi-elasticity of the predicted risk $\hat{\pi}_j, j \in \{0, 1\}$ of marginals versus individuals always in the basic coverage, splitting the sample by active vs passive choice environment in 2007:

$$\text{Semi}^{\hat{\pi}_j}_{M(\mathbf{p})} = \frac{E(\hat{\pi}_j|Z, D = M) - E(\hat{\pi}_j|Z, D = 0)}{E(\hat{\pi}_j|Z, D = 0)}$$

where $Z$ is a vector of characteristics affecting the contract space. We find that the adverse selection identified by the 2007 price reform is slightly larger for predicted risk in the basic coverage for workers observed in active compared to workers observed in passive choice environment. But we do not find any significant difference in adverse selection for risk in the comprehensive coverage.

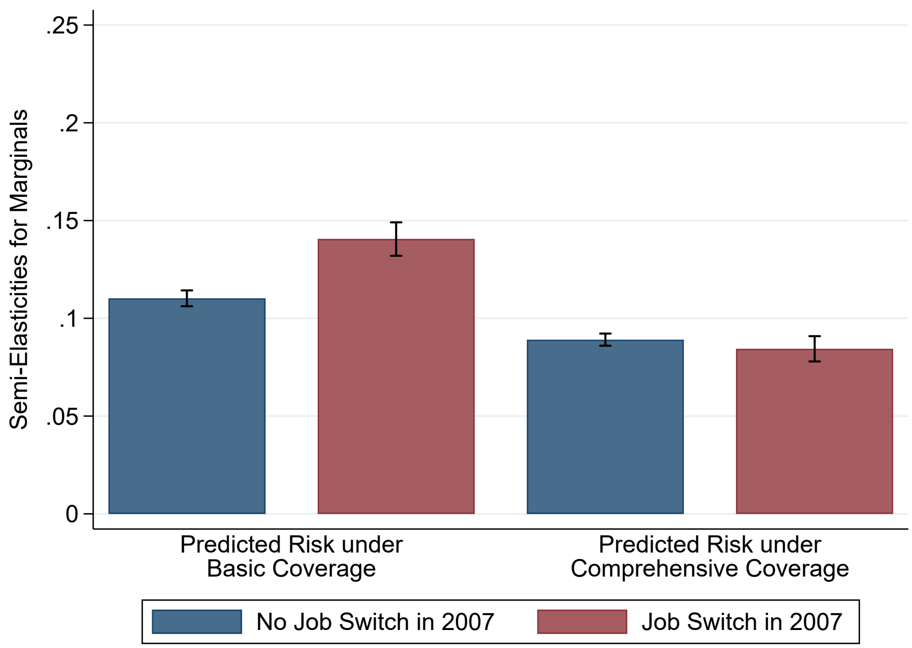
Put together, this evidence suggests that inertia does not seem to critically affect our estimates of the demand and marginal cost curves. It is worth noting though that the relatively modest inertia we find is likely due to the fact that the 2007 reform was large, and salient.

Figure D.5: Inertia: Fraction of Workers Buying the Comprehensive Coverage around the 2007 Reform by Job Switching Status




**Notes:** The Figure reports the evolution of the fraction of individuals buying the comprehensive coverage around the 2007 by job switching status. In line with the existing the literature (e.g. Handel (2013)), we use job switchers to proxy for differential exposure to inertia. New employees in a firm face a more active choice environment than existing employees. The former have to reoptimize many choices, while the latter remain in a more passive choice environment.

Figure D.6: Inertia & Adverse Selection: Relative Risk of the Marginals Compared to Individuals in the Basic Coverage by Job Switching Status




**Notes:** The Figure reports the estimated adverse selection created by the 2007 price reform for two sets of workers who are differently exposed to inertia. The red bars refer to individuals who switch job in 2007. These individuals are facing an active choice environment in 2007, at the moment of the price change. The blue bars refer to individuals who stay with their employers. These individuals are facing a passive choice environment in 2007. For both groups of workers, we report the semi-elasticity of the predicted risk $\hat{\pi}_j, j \in \{0, 1\}$ of marginals versus individuals always in the basic coverage:

$$\text{Semi}^{\hat{\pi}_j}_{M(\mathbf{p})} = \frac{E(\hat{\pi}_j|Z, D = M) - E(\hat{\pi}_j|Z, D = 0)}{E(\hat{\pi}_j|Z, D = 0)}$$

where $Z$ is a vector of characteristics affecting the contract space.

# Appendix E Benefit Variation: Additional Material

In this appendix we provide additional material regarding our RKD estimation of the effect of benefit variation on insurance choices and risk-based selection:

(i) we present results assessing the validity of our RK design.

(ii) we present results assessing the sensitivity of our RKD estimates.

Table E.2 provides the summary statistics for the sample used for the RKD analysis.

## E.1 Assessing Validity of the RK Design

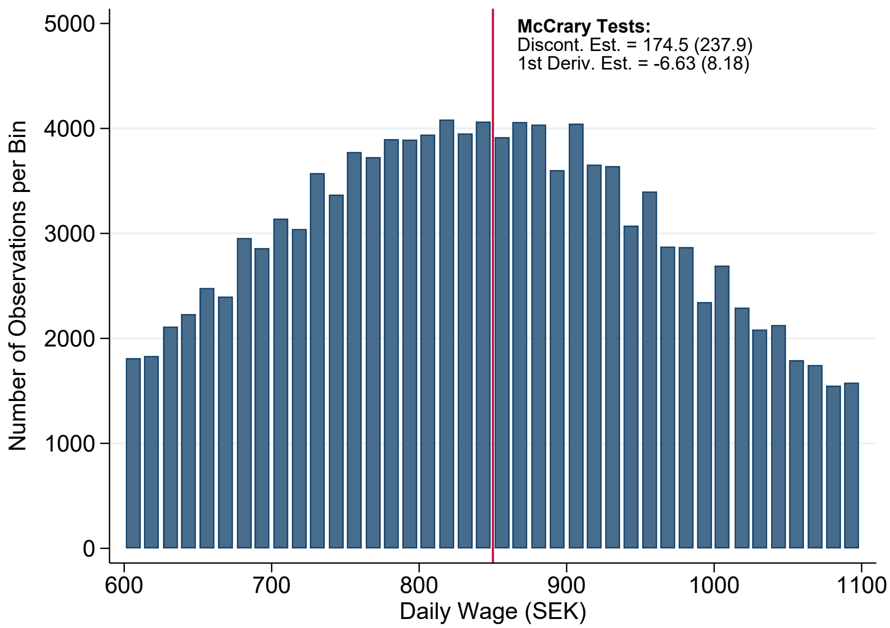
The key identifying assumption of the RK design is the existence of a smooth relationship at the threshold $w = 850SEK$ between the assignment variable and any pre-determined characteristics affecting the demand for insurance. To assess the credibility of this assumption, we conduct two types of analysis [see also Kolsrud et al. [2018]].

**Smoothness of the distribution of the assignment variable at the kink** First, we focus on the probability density function of the assignment variable, to detect manipulation or lack of smoothness around the kink that could indicate the presence of selection. Figure E.1 shows that the pdf of daily wage does not exhibit a discontinuity nor lack of smoothness at the kink, which is confirmed by the results of formal McCrary tests.

**Covariate Tests** Second, we investigate the presence of potential selection along observable characteristics around the kink. For this purpose, instead of looking at each characteristics in isolation, we aggregate them in a covariate index. The index is a linear combination of a vector of characteristics $X$ that correlate with our outcomes of interest for the RKD, which includes age, gender, level of education, region, family type and industry. The coefficients in the linear combination are obtained from a regression of the outcome variable on these covariates. In Figure E.2, we display the relationship between this covariate index and the assignment variable for our three outcomes of interest: the choice of coverage, and the predicted risk under basic and comprehensive coverage. The relationship between the index and daily wage appears smooth around the 850SEK threshold. Yet, formal tests of non-linearity suggest the presence of a significant (although economically small) kink at the threshold for insurance choice. But for predicted risk, we do not find any significant non-linearity in the covariate index at the kink.

Figure E.1: Regression Kink Design: Testing for Manipulation of Assignment Variable

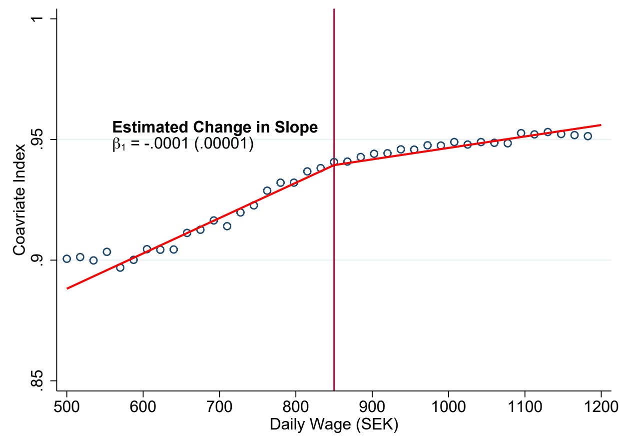
Pdf of Assignment Variable



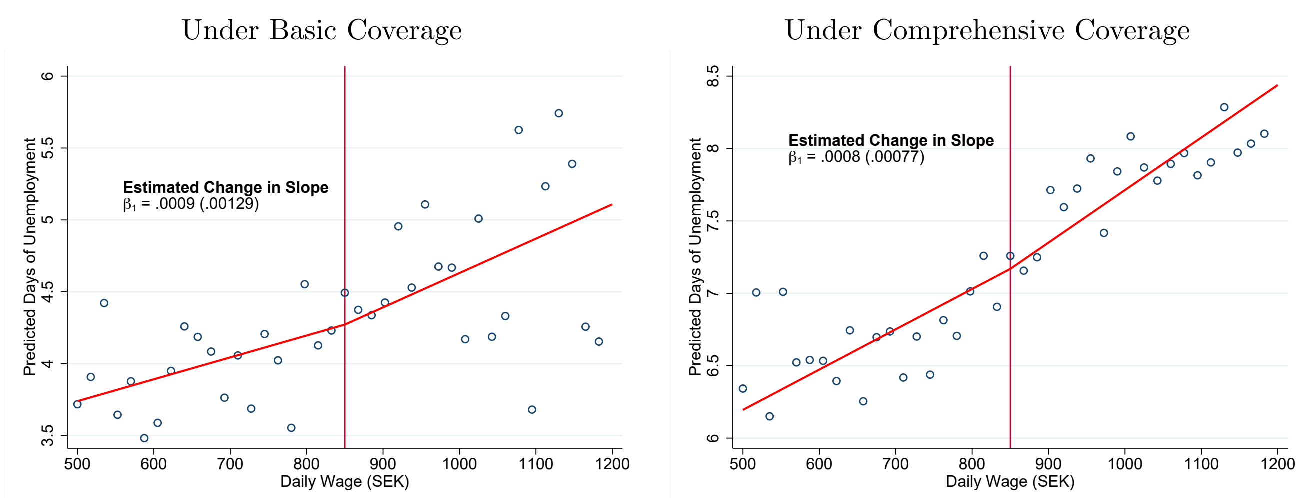

**Notes:** The panel displays the probability density function of daily wage. We also report on the graph formal McCrary tests for the existence of a discontinuity and of lack of smoothness of the pdf at the 850SEK threshold.

Figure E.2: Regression Kink Design: Smoothness of Distribution of Observables Characteristics

### A. Covariate Index vs Assignment Variable: Insurance Choices




### B. Covariate Index vs Assignment Variable: Predicted Risk




**Notes:** The Figure investigates the presence of potential selection along observable characteristics around the kink. For this purpose, we aggregate observable characteristics into a covariate index. The index is a linear combination of a vector of characteristics $X$ that correlate with the outcome, and which includes age, gender, level of education, region, family type and industry. The coefficients in the linear combination are obtained from a regression of the outcome variable on these covariates. Panel A displays the relationship between the assignment variable and the covariate index for the probability to buy the comprehensive coverage. Panel B displays the corresponding graph for the covariate indexes of predicted risk under basic and under comprehensive coverage. We also report on each graph formal tests of non-linearity, i.e. the coefficients $\beta_1$ obtained from a specification similar to (18).

### E.2 Assessing Sensitivity of the RKD estimates

**Sensitivity to bandwidth choice** Our baseline RK results use a bandwidth of 350SEK for the daily wage. We start by investigating how sensitive our results are to different bandwidth choices. In Figure E.3, we plot for our three outcomes the value of the RK estimate and its 95% confidence interval for various values of the bandwidth. The graph shows that estimates are stable across bandwidth size. We also computed the optimal bandwidth from Calonico, Cattaneo, & Titiunik (2014), and find 358SEK for the predicted risk and 175SEK for insurance choice.

**Sensitivity to inclusion of controls** We next investigate how sensitive our results are to the inclusion of the set of controls $X$. In table E.1, we report in panel A column (1) the estimate of the change in slope $\beta_1$ from specification (18), where we do not include the vector $X$. In column (2), we add controls for age, gender and family types. In column (3), we also add controls for education, region of residence, and industry. We find that the results are stable across these specifications. We then replicate this analysis for predicted risk. In panel B, we focus on predicted risk under basic coverage, and in panel C on predicted risk under comprehensive coverage. Each column reports the estimate $\beta_1$ from specification (19), and we vary across columns the set of controls included in the residualization procedure

$$E(\hat{\pi}_j|Z,X) = (1 - f(0|Z,X)) \exp(Z'\gamma_Z + X'\gamma_X)$$

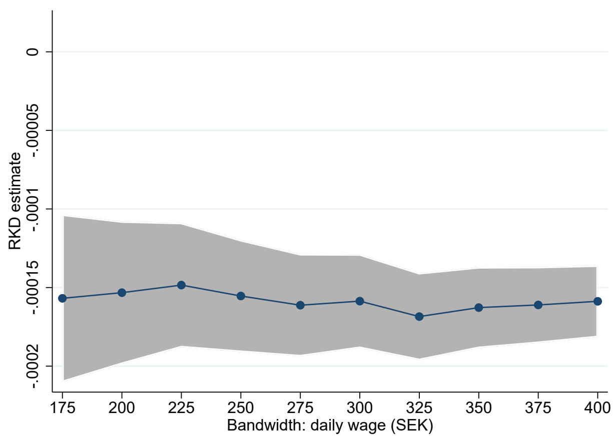
We find that results are also stable to the inclusion of controls.

**Inference** Finally, we explore the robustness of our inference approach to non-linearities in the relationship between the assignment variable and the outcome. We implement a permutation test and compare the coefficient estimate at the true kink to those at "placebo" kinks placed away from the true kink.

In Figure E.4, we report the probability density function of the estimated change in slope $\beta_1$ for 1000 placebo kinks outside the 750SEK-950SEK range. Panel A shows the distribution of placebo RK estimates, using specification (18), for the probability of buying the comprehensive coverage. The estimated coefficient at the true kink lies markedly below all the placebo estimates, indicating that our estimates are unlikely to pick up some non-linearity in the relationship between daily wage and insurance choice. In Panel B we report the distribution of placebo RK estimates for the predicted risk under basic and comprehensive using specification (19). In both cases, we find that the vast majority of placebo estimates is negative, so that if anything, there is non-linearity in the opposite direction than the one detected at the true kink. The probability to find a placebo estimate larger than the estimate at the true kink is, in both cases, inferior to 5%.

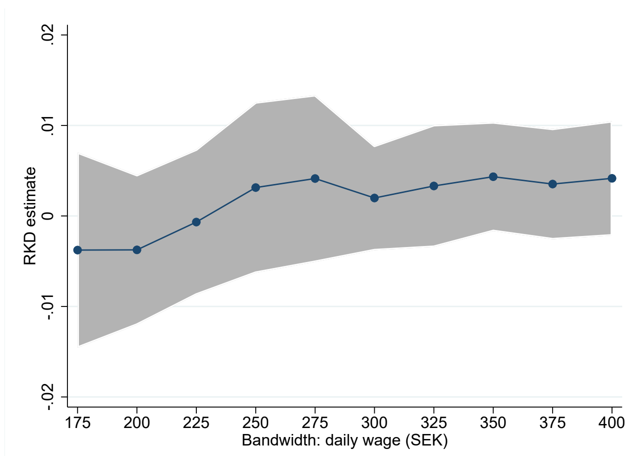
Figure E.3: Regression Kink Design: Sensitivity to Bandwidth Choice

### A. RKD Estimates of Insurance Choice Response by Bandwidth

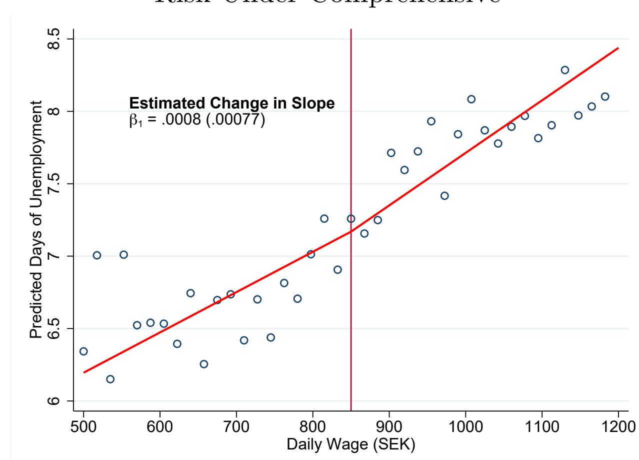



### B. RKD Estimates of Predictable Risk by Bandwidth

Risk Under Basic




Risk Under Comprehensive




**Notes:** The Figure investigates the sensitivity of our estimates to the choice of bandwidth for the RK estimation. Our baseline bandwidth is 350. Panel A plots the value of the RK estimate and its 95% confidence interval for various values of the bandwidth. The graph shows that estimates are stable across bandwidth size. Panel B plots similar graphs for the predicted risk under basic and under comprehensive coverage.

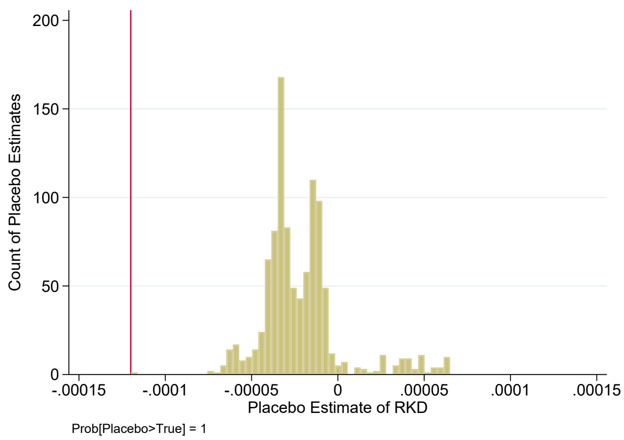
Table E.1: Regression Kink Design: Sensitivity to Inclusion of Controls

| | **A. Probability to Buy Comprehensive** | | | **B. Risk Under Comprehensive** | | | **C. Risk Under Basic** | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| $\beta_1$ | - .016 | - .013 | - .012 | .307 | .359 | .279 | .646 | .520 | .204 |
| | (.001) | (.001) | (.001) | (.094) | (.093) | (.105) | (.467) | (.330) | (.144) |
| N | 110,123 | 110,123 | 110,123 | 89,576 | 89,576 | 89,576 | 3,998 | 3,998 | 3,998 |
| Baseline | × | × | × | × | × | × | × | × | × |
| Age, gender family type | | × | × | | × | × | | × | × |
| Education, region industry | | | × | | | × | | | × |

**Notes:** The baseline controls refer to the vector $Z$ of characteristics affecting premia. It consists in a dummy for union membership, a dummy for eligibility and year fixed effects.

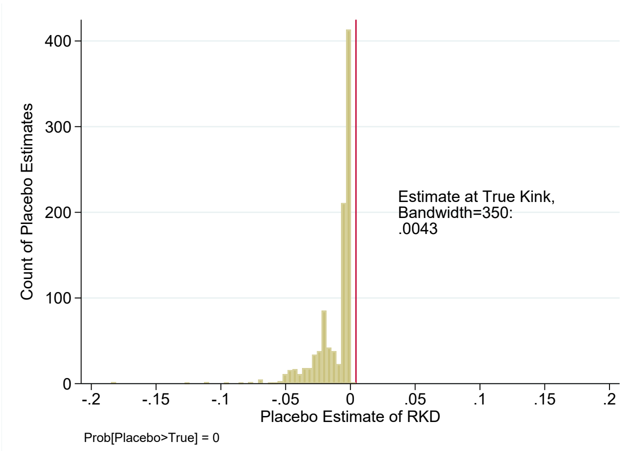
Figure E.4: Regression Kink Design: Permutation-Based Inference

### A. Distribution of Placebo Estimates of Insurance Choice Response

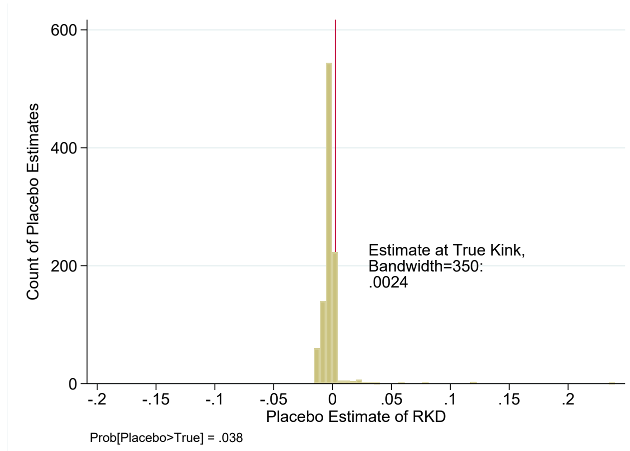



### B. Distribution of Placebo Estimates of Predicted Risk

Risk Under Basic




Risk Under Comprehensive




**Notes:** The Figure reports the probability density function of the estimated change in slope $\beta_1$ for 1000 placebo kinks outside the 750SEK-950SEK range. Panel A shows the distribution of placebo RK estimates for the probability of buying the comprehensive coverage. Panel B reports the distribution of placebo RK estimates for the predicted risk under basic and comprehensive. We also report on all three graphs the probability to find a placebo estimate larger than the estimate at the true kink.

Table E.2: Summary Statistics - RKD Population

| | Mean | P10 | P50 | P90 |
|---|---|---|---|---|
| | **I. Unemployment** | | | |
| Days unemployed | 180.07 | 21 | 145 | 365 |
| Predicted days unemployed under Basic | 4.81 | 2.33 | 3.19 | 5.87 |
| Predicted days unemployed under Comprehensive | 8.15 | 3.88 | 5.26 | 11.5 |
| Unemployment spell duration (days) | 410 | 91 | 301 | 910 |
| Fraction receiving layoff notification | .04 | - | - | - |
| | **II. Union and UI Fund Membership** | | | |
| Union membership | .78 | - | - | - |
| UI fund membership | .96 | - | - | - |
| Switch from coverage 0 to 1 | .04 | - | - | - |
| Switch from coverage 1 to 0 | .01 | - | - | - |
| | **III. Demographics** | | | |
| Age | 37.95 | 23 | 37 | 55 |
| Years of education | 12.18 | 10 | 12 | 16 |
| Fraction men | .56 | - | - | - |
| Fraction married | .33 | - | - | - |
| | **IV. Income and Wealth SEK 2003(K)** | | | |
| Gross earnings | 127.6 | 0 | 127 | 246 |
| Net wealth | 153.7 | -182 | 0 | 644 |
| Bank holdings | 28.9 | 0 | 0 | 72 |
| N | 140,777 | | | |

**Notes:** The Table provides summary statistics for the RKD sample. See Table 1 for our main sample of interest.

# Appendix F Welfare Analysis: Additional Material

This appendix provides further detail and derivations for the theoretical analysis in Section 3 and supporting material for the empirical implementation in Section 6.

## F.1 Conceptual Framework: Further Details

To evaluate the design of the social insurance system allowing for choice, we characterized the welfare impact of small changes in prices and coverages with two plan options, $\{(b_j, p_j)\}_{j=1,0}$. This relies on the social welfare function being concave and differentiable so that we can characterize the optimal contract using first-order conditions. We also assumed that workers' preferences are quasi-linear in prices so that an individual's risk $\pi_j$, conditional on plan choice $j$, does not depend on prices and neither does the ranking of individuals' valuations $\mathbf{u}_i$. This is a standard assumption in the insurance literature, but the implications from relaxing it are evident from the analysis of coverage changes.

Regarding the timing of the model, we stick closely to the structure of the Swedish UI system where individuals become eligible to receive the supplemental benefits when they have been contributing for one year to the comprehensive coverage, and can opt in and out of the comprehensive plan at any time. As a consequence, the value $u_j^t$ of plan $j$ in year $t$ depends on unemployment risk $\pi^{t+1}$, the expected number of days spent unemployed in year $t+1$. With this in mind, we have dropped time subscripts, but $u\left(\pi\right)$ always refers to $u^t\left(\pi^{t+1}\right)$, unless otherwise specified.

Our framework is highly stylized, but allows for multi-dimensional heterogeneity and endogenous actions. Following a recent tradition in the social insurance literature [see Chetty and Finkelstein [2013]], we choose not to explicitly model the underlying heterogeneity and actions. The key micro-foundations for our analysis are the resulting plan valuations and costs and how they change with the plans' prices and coverage levels. For example, in a setting with expected utility and binary unemployment risk, the value of a plan to a worker equals

$$u_i\left(b_j, p_j\right) = \max_{a'} \pi\left(a'|\theta_i\right) \tilde{u}\left(b_j - p_j, a'|\mu_i\right) + \left(1 - \pi\left(a'|\theta_i\right)\right) \tilde{u}\left(w_i - p_j, a'|\mu_i\right), \tag{32}$$

while the insurer's cost equals $c_i\left(b_j, p_j\right) = \pi(a'|\theta_i)b_j$. The value and cost are interdependent through the risk parameter $\theta$ and the effort choice $a$, which in turn depends on the preference parameter $\mu$ that also affects the valuation. In general, a worker's expected utility depends both on the probability of job loss and the time spent unemployed. The government's expected cost would depend only on the expected number of days spent unemployed when the benefit profile is flat.

**Sorting Effect** The fiscal externality in both Propositions 1 and 2 depends on how many individuals change in response to the policy (as captured by the demand elasticity) and the cost characteristics of those who switch. We develop here formally why the cost characteristics of the switchers in response to a change in coverage are different than for a change in price under multi-dimensional heterogeneity.

We first re-write the social welfare function, ranking individuals based on their utility gain of the comprehensive relative to the basic plan,

$$
\begin{aligned}
\mathcal{W} &\equiv \int_{\mathbf{u}_i \geq 0} \omega\left(u_i\left(b_1, p_1\right)\right) di + \int_{\mathbf{u}_i < 0} \omega\left(u_i\left(b_0, p_0\right)\right) di + \lambda\left\{F_1\left[p_1 - E_1\left(\pi_1\right) b_1\right] + F_0\left[p_0 - E_0\left(\pi_0\right) b_0\right]\right\}, \\
&= \int_{\mathbf{u} \geq 0} E(\omega\left(u\left(b_1, p_1\right)\right) | \mathbf{u}) dG\left(\mathbf{u}\right) + \int_{\mathbf{u} < 0} E(\omega\left(u\left(b_0, p_0\right)\right) | \mathbf{u}) dG\left(\mathbf{u}\right) \\
&\quad + \lambda\left\{\left(1 - G\left(0\right)\right)\left[p_1 - E_1\left(\pi_1\right) b_1\right] + G\left(0\right)\left[p_0 - E_0\left(\pi_0\right) b_0\right]\right\}.
\end{aligned}
$$

Here $G(\cdot)$ is the distribution of $\mathbf{u} = u_1 - u_0$, which depends on the plan characteristics, with $G(0) = F_0$ and $1 - G(0) = F_1$. Following the derivation in Veiga and Weyl [2016] and Handel et al. [2019], we then find

$$
\frac{\partial}{\partial x_j}\left[\left(1 - G\left(0\right)\right) E_1\left(z_1\right)\right] = E\left(z_1 \frac{\frac{\partial \mathbf{u}}{\partial x_j}}{E\left(\frac{\partial \mathbf{u}}{\partial x_j} | \mathbf{u} = 0\right)} | \mathbf{u} = 0\right) \frac{\partial\left(1 - G\left(0\right)\right)}{\partial x_j}. \tag{33}
$$

assuming no direct effect of the policy variable $x_j$ on the outcome $z_1$.

The argument proceeds as following. First, using iterated expectations and introducing notation $\mathbf{u}' \equiv \frac{\partial \mathbf{u}}{\partial x_j}$, we can write

$$
\begin{aligned}
\frac{\partial}{\partial x_j}\left[\left(1 - G\left(0\right)\right) E_1\left(z_1\right)\right] &= \frac{\partial}{\partial x_j}\left[\int_{\mathbf{u} \geq 0} E\left(z_1 | \mathbf{u}\right) dG\left(\mathbf{u}\right)\right] \\
&= \frac{\partial}{\partial x_j}\left[\int_{\mathbf{u} \geq 0} \int E\left(z_1 | \mathbf{u}, \mathbf{u}'\right) f\left(\mathbf{u}' | \mathbf{u}\right) g\left(\mathbf{u}\right) d\mathbf{u}' d\mathbf{u}\right] \\
&= \int \frac{\partial}{\partial x_j}\left[\int_{\mathbf{u}_\varepsilon \geq -\mathbf{u}' \times [x_j - x_\varepsilon]} E\left(z_1 | \mathbf{u}_\varepsilon, \mathbf{u}'\right) g^{\varepsilon}\left(\mathbf{u}_\varepsilon | \mathbf{u}'\right) d\mathbf{u}_\varepsilon\right] f\left(\mathbf{u}'\right) d\mathbf{u}'
\end{aligned}
$$

The last equality follows from (1) using $f\left(\mathbf{u}' | \mathbf{u}\right) g\left(\mathbf{u}\right) = g\left(\mathbf{u} | \mathbf{u}'\right) f\left(\mathbf{u}'\right)$, (2) approximating $\mathbf{u}\left(b_j\right) \cong \mathbf{u}\left(b_\varepsilon\right) + \mathbf{u}' \times \left[b_j - b_\varepsilon\right]$**,** and substituting the variable in the integral $\mathbf{u}\left(b_j\right) (\equiv \mathbf{u})$ by $\mathbf{u}\left(b_\varepsilon\right) (\equiv \mathbf{u}_\varepsilon)$, where $d\mathbf{u} = d\mathbf{u}_\varepsilon$, conditional on $\mathbf{u}'$.

We can then apply Leibniz rule and find after re-substituting,

$$
\begin{aligned}
\frac{\partial}{\partial x_j}\left[\int_{\mathbf{u} \geq 0} E\left(z_1 | \mathbf{u}\right) dG\left(\mathbf{u}\right)\right] &= \int \left[E\left(z_1 \mathbf{u}' | \mathbf{u} = 0, \mathbf{u}'\right) f\left(\mathbf{u}' | \mathbf{u} = 0\right) d\mathbf{u}'\right] g\left(0\right) \\
&= E\left(z_1 \frac{\partial \mathbf{u}}{\partial x_j} | \mathbf{u} = 0\right) g\left(0\right). \\
&= E\left(z_1 | \mathbf{u} = 0\right) E\left(\frac{\partial \mathbf{u}}{\partial x_j} | \mathbf{u} = 0\right) g\left(0\right) + cov\left(z_1, \frac{\partial \mathbf{u}}{\partial x_j} | \mathbf{u} = 0\right)
\end{aligned}
$$

Note also that the effect on the share of individuals buying comprehensive coverage equals

$$\frac{\partial\left[1-G\left(0\right)\right]}{\partial x_j}=\frac{\partial}{\partial x_j}\left[\int_{\mathbf{u}\geq 0}dG\left(\mathbf{u}\right)\right]=E\left(\frac{\partial\mathbf{u}}{\partial x_j}|\mathbf{u}=0\right)g\left(0\right).$$

Taken together, we can then indeed write

$$\frac{\partial}{\partial x_j}\left[\left(1-G\left(0\right)\right)E_1\left(z_1\right)\right]=E\left(z_1\frac{\frac{\partial\mathbf{u}}{\partial x_j}}{E\left(\frac{\partial\mathbf{u}}{\partial x_j}|\mathbf{u}=0\right)}|\mathbf{u}=0\right)\frac{\partial\left(1-G\right)}{\partial x_j}.$$

Similarly, we can find

$$\frac{\partial}{\partial x_j}\left[G\left(0\right)E_0\left(z_0\right)\right]=E\left(z_0\frac{\frac{\partial\mathbf{u}}{\partial x_j}}{E\left(\frac{\partial\mathbf{u}}{\partial x_j}|\mathbf{u}=0\right)}|\mathbf{u}=0\right)\frac{\partial G}{\partial x_j}$$

Applying this now to the difference in average risk for switchers in response to a coverage and price change, we find

$$E_{M(b_j)}(\pi_k)-E_{M(\mathbf{p})}(\pi_k)=\frac{cov\left(\pi_k,\frac{\partial u_j}{\partial b_j}|\mathbf{u}=\mathbf{0}\right)}{E\left(\frac{\partial u_j}{\partial b_j}|\mathbf{u}=\mathbf{0}\right)}.$$

Individuals with higher unemployment risk tend to value extra coverage more. However, with heterogeneity in risk aversion (in addition to risk heterogeneity), the correlation between risks and the marginal value of coverage among the marginal buyers can become negative [Ericson et al. [Forthcoming]]. Relatedly, the risk-based selection into supplemental coverage may well worsen as the basic coverage level increases. The reason is that the variation in willingness-to-pay coming from heterogeneity in risk aversion, which would mute the risk-based selection, decreases as the basic coverage becomes more generous [see Ericson et al. [Forthcoming]].

## F.2 Proof of Proposition 1

We denote again by $\mathbf{p}$ and $\mathbf{c}$, the difference in prices and costs between the two plans, e.g. $\mathbf{p}=p_1-p_0$.

*Proof.* Differentiating the social welfare with respect to $p_0$, we get:

$$\begin{aligned}\frac{\partial\mathcal{W}}{\partial p_0}=&\,G\left(0\right)E_0\left(\frac{\partial\omega_0}{\partial p_0}\right)\\&+\lambda\left\{G\left(0\right)-\frac{\partial G\left(0\right)}{\partial p_0}\mathbf{p}-\frac{\partial\left[\left(1-G\left(0\right)\right)E_1\left(c_1\right)+G\left(0\right)E_0\left(c_0\right)\right]}{\partial p_0}\right\}.\end{aligned}$$

Here, we are invoking the envelope theorem for the resorting of marginal individuals, $u_i\left(b_1,p_1\right)=$

$u_i(b_0, p_0)$. We can rewrite this expression as

$$\frac{\partial \mathcal{W}}{\partial p_0} = G(0) E_0\left(\frac{\partial \omega_0}{\partial p_0}\right) + \lambda\left\{G(0) - (\mathbf{p} - E_M(\mathbf{c}))\frac{\partial G(0)}{\partial p_0}\right\}$$

using the result in equation (33),

$$\begin{aligned}\frac{\partial\left[(1-G(0))E_1(c_1) + G(0)E_0(c_0)\right]}{\partial p_0} &= E_M(c_1)\frac{\partial\left[1-G(0)\right]}{\partial p_0} + E_M(c_0)\frac{\partial G(0)}{\partial p_0} \\ &= -E_M(\mathbf{c})\frac{\partial G(0)}{\partial p_0}.\end{aligned}$$

Note that the quasi-linear assumption implies that prices do not cause any moral hazard response, $E_0\left(\frac{\partial c_0}{\partial p_0}\right) = E_0\left(b_0\frac{\partial \pi_0}{\partial p_0}\right) = 0$. Hence, the only impact on the budget constraint is the re-sorting response. This response itself also simplifies due to the quasi-linearity assumption. Since $\frac{\partial \mathbf{u}}{\partial p_0}$ is constant, it depends on the demand response $\frac{\partial G(0)}{\partial p_0}$ multiplied by the cost of providing the supplemental coverage to workers at the margin, which simplifies to the unweighted average among the marginals, $E_M(\mathbf{c}) = E(\mathbf{c}|\mathbf{u}=0)$. Using $\frac{\partial G(0)}{\partial p_0} = -\frac{\partial G(0)}{\partial \mathbf{p}}$, we then find that the FOC with respect to the price of basic coverage, $\frac{\partial \mathcal{W}}{\partial p_0} = 0$, is equivalent to

$$-E_0\left(\frac{\partial \omega_0}{\partial p_0}\right) = \lambda\left\{1 + AS_{\mathbf{p}}\frac{\partial \ln G(0)}{\partial \mathbf{p}}\right\}.$$

In a similar way, we can get the FOC for the price of the comprehensive plan

$$-E_1\left(\frac{\partial \omega_1}{\partial p_1}\right) = \lambda\left\{1 + AS_{\mathbf{p}}\frac{\partial \ln(1-G(0))}{\partial \mathbf{p}}\right\}$$

Putting the two FOCs together, we get the expression in the Proposition. □

## F.3 Proof of Proposition 2

*Proof.* We consider the welfare impact of an increase in $b_0$, for given prices and coverage $b_1$. The impact of a change in $b_0$ on the government's budget depends both on the change in selection into both plans and the direct effect from increasing the coverage,

$$\frac{\partial}{\partial b_0}\left[\int_{\mathbf{u}\geq 0}\left[p_1 - E(c_1|\mathbf{u})\right]dG(\mathbf{u}) + \int_{\mathbf{u}<0}\left[p_0 - E(c_0|\mathbf{u})\right]dG(\mathbf{u})\right]$$

$$\begin{aligned} &= \left[\mathbf{p} - E\left(\mathbf{c}\times\frac{\frac{\partial \mathbf{u}}{\partial b_0}}{E\left(\frac{\partial \mathbf{u}}{\partial b_0}|\mathbf{u}=0\right)}|\mathbf{u}=0\right)\right]\frac{\partial(1-G(0))}{\partial b_0} - G(0)\frac{\partial E_0 c_0}{\partial b_0} \\ &\equiv -\left[\mathbf{p} - E_{M(b_0)}(\mathbf{c})\right]\frac{\partial G(0)}{\partial b_0} - G(0)\frac{\partial E_0 c_0}{\partial b_0}\end{aligned}$$

By analogy to the subsidy change, we decompose the change in cost from providing coverage due to the change in selection as the demand effect $\frac{\partial G}{\partial b_0}$ multiplied by the fiscal externality $\mathbf{p} - E_{M(b_0)}(\mathbf{c})$, caused by the switching of individuals who respond to the coverage change. This fiscal externality differs from the fiscal externality of the subsidy as different individuals respond to a change in coverage depending on their marginal value of basic coverage $\frac{\partial \mathbf{u}}{\partial b_0}$, explaining the weights put on the costs of the different marginals with $\mathbf{u} = 0$. This is discussed in detail in Appendix F.1. In addition to the selection response, an increase in coverage of the basic plan affects the government's expenditures directly, but also indirectly through a moral hazard response. That is,

$$\frac{\partial E_0 c_0}{\partial b_0} = E_0(\pi_0) + E_0\left(\frac{\partial \pi_0}{\partial b_0}\right) b_0 = E_0(\pi_0)\left[1 + FE_{b_0}^{MH}\right].$$

Invoking now the envelope condition for the individuals at the margin (i.e., $\mathbf{u} = 0$), we find

$$d\mathcal{W} = G(0)\, E_0\left(\frac{\partial \omega_0}{\partial b_0}\right) - \lambda G(0)\, E_0(\pi_0)\left[1 + FE_{b_0}^{MH}\right] - \lambda\left[\mathbf{p} - E_{M(b_0)}(\mathbf{c})\right]\frac{\partial G}{\partial b_0},$$

where

$$E_0\left(\frac{\partial \omega_0}{\partial b_0}\right) = \frac{1}{G(0)}\int_{\mathbf{u}<\mathbf{0}} E\left(\omega'(u_0)\,\frac{\partial u(b_0, p_0)}{\partial b_0}|\mathbf{u}\right) dG(\mathbf{u})$$

Hence, at an (interior) optimum, we need $d\mathcal{W}$ =0 and thus

$$E_0\left(\frac{\partial \omega_0}{\partial b_0}\right)/E_0(\pi_0) = \lambda\left\{1 + FE_{b_0}^{MH} + \left[\mathbf{p} - E_{M(b_0)}(\mathbf{c})\right]\frac{\partial \ln G}{\partial b_0}/E_0(\pi_0)\right\}.$$

In a similar way, we can get the FOC for the coverage of the comprehensive plan,

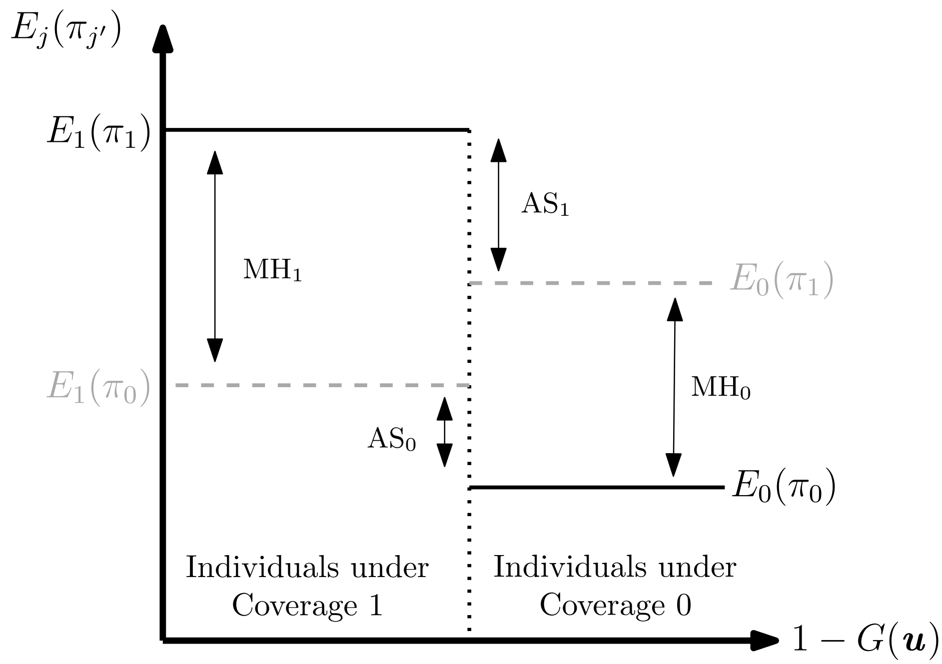
$$E_1\left(\frac{\partial \omega_1}{\partial b_1}\right)/E_1(\pi_1) = \lambda\left\{1 + FE_{b_1}^{MH} - \left[\mathbf{p} - E_{M(b_1)}(\mathbf{c})\right]\frac{\partial \ln(1-G)}{\partial b_1}/E_1(\pi_1)\right\}.$$

Putting the two FOCs together, we get the expression in the Proposition.

□

## F.4 Graphical Representation of AS vs. MH

Figure F.1: DECOMPOSITION OF PCT STATISTIC




**Notes:** The Figure illustrates the decomposition of the positive correlation test (PCT) statistic $E_1(\pi_1) - E_0(\pi_0)$ studied in Section 3. Workers opt for the comprehensive plan 1 when $\mathbf{u} \geq 0$ and for the basic plan 0 otherwise.$E_j(\pi_{j'})$ denotes the average unemployment risk for workers who opt for coverage $j$ when under plan $j'$. There are two complementary ways to quantify the respective roles of adverse selection and moral hazard due to the fact that the measurement of the differences in risk due to adverse selection is plan-dependent, while the measurement of the differences in risk due to moral hazard is group-dependent. A first decomposition consists of adverse selection in the comprehensive plan ($AS_1$) plus moral hazard for the group selecting basic coverage ($MH_0$). A second decomposition consists of moral hazard for the group selecting comprehensive coverage ($MH_1$) and adverse selection in the basic plan ($AS_0$). Relating this to the textbook analysis of selection and treatment effects, the moral hazard response can be interpreted as the treatment effect on risk from providing supplementary coverage. This treatment effect can be different for workers who select into treatment compared to those who do not. The difference in treatments effects between the two groups depends on the difference in risks under the comprehensive and basic coverage respectively.

## F.5 Policy Implications: Supporting Material

**Risk and Cost Curves** Using the price variation, we can infer how the risk under basic and comprehensive coverage changes with willingness-to-pay for the supplemental coverage. In Section 6, we convert the resulting *marginal* risk curves into marginal cost curves like in Einav et al. [2010b], accounting for the coverage levels. The risk curves are displayed in Figure F.2. Just like for the cost curves, we position the three groups of individuals $i = 1, M, 0$ on the x-axis according to their willingness-to-pay. Individuals who choose the basic coverage (0) both in 2006 and 2007 are on the right hand side, while individuals who always buy the comprehensive coverage (1) are on the left-hand side of the graph. The marginals correspond to the group in between, and their share is given in Figure 4, where we see the fraction of individuals buying dropping from 86 to 78% with the 2007 price increase. We plot with green dots the observed average realized risk, measured by the number of days spent unemployed in year 2008, for the three groups. Note that all risk measures in Figure F.2 are conditional on the vector of characteristics $Z$, and normalized to the average risk under basic coverage of individuals observed under basic coverage $E_0(\pi_0)$. The difference in realized risk (green dots) between the marginals and group 0 is therefore equivalent to our semi-elasticity estimate $\text{Semi}_{M(\mathbf{p})}^{Baseline}$.

We then plot with blue triangles the average predicted risk under basic coverage and with red triangles the average predicted risk under comprehensive coverage for all three groups. The blue triangles identify the marginal risk curve in the basic coverage, while the red triangles plot the marginal risk curve in the comprehensive coverage. These two curves provide all the information necessary to compute the fiscal externality term $FE_{\mathbf{p}}^{AS}$ that determines the optimal price level in Proposition 1. Moreover, the difference between the red and blue triangles for each group identifies the moral hazard of moving these individuals from basic to comprehensive coverage.

**Consumption Smoothing Gains** When considering to further differentiate the coverage levels following Proposition 2, the fiscal externalities need to be traded off against the relative consumption smoothing gains from extra coverages: $\frac{E_1\left(\frac{\partial\omega_1}{\partial b_1}\right)/E_1(\pi_1)}{E_0\left(\frac{\partial\omega_0}{\partial b_0}\right)/E_0(\pi_0)}$. Given our estimates of the fiscal externalities, further differentiation would therefore be socially valuable if the marginal utility of an increase in benefits for individuals under the comprehensive coverage is worth at least 27% more than the marginal utility of an increase in benefits for individuals under the basic coverage.

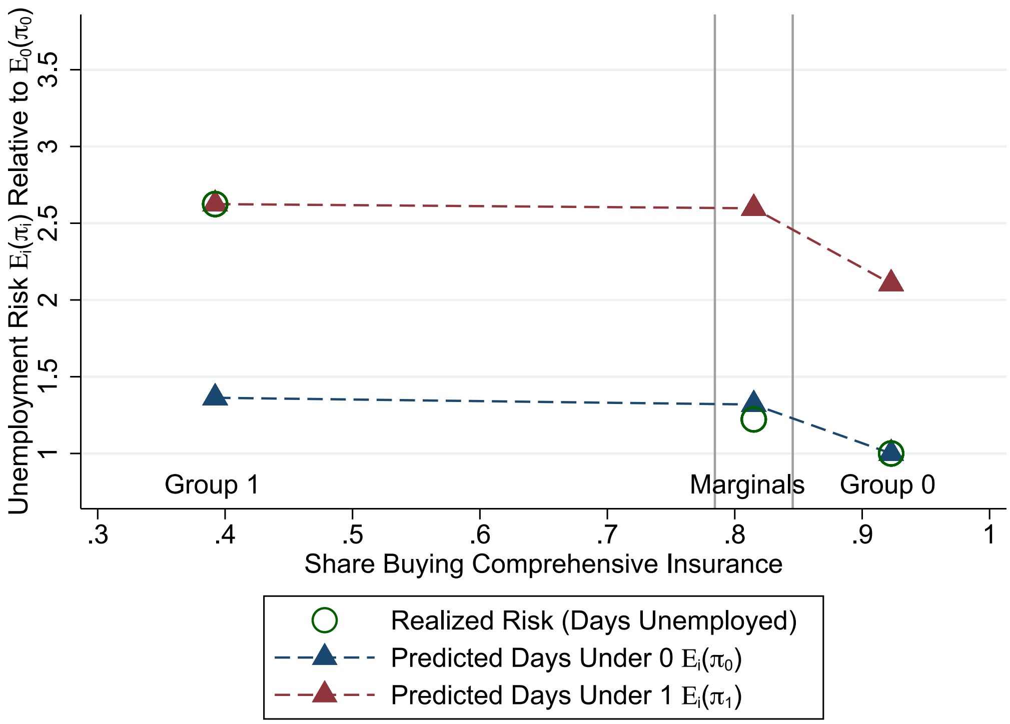
To estimate the magnitude of the relative consumption smoothing gains, we need to know about the two basic forces that underlie them. The first is heterogeneity in the value of insurance, conditional on risk: people who buy the comprehensive coverage may do so because they value an additional kroner of consumption when unemployed more. Landais and Spinnewijn [Forthcoming] find in the Swedish context that the mark-up workers under comprehensive coverage are willing to pay for the supplemental coverage is 160% larger compared to workers under basic coverage, conditional on their risk: $E_1(\overline{MRS}_{b_0,b_1}) = 2.6 \cdot E_0(\overline{MRS}_{b_0,b_1})$.[53] This number underlines that there is indeed significant heterogeneity in willingness-to-pay for insurance conditional on risk, which is a strong force pushing for coverage differentiation.

The second force is diminishing marginal utility of consumption. The marginal utility of further coverage should be estimated at coverage level $b_1$ for individuals buying the comprehensive coverage and at coverage level $b_0$ for individuals in the basic plan. Diminishing marginal utility of consumption makes the value of an additional kroner lower when evaluated at $b_1$ than at $b_0$.

Using a Taylor approximation, we can provide a back-of-the envelope calculation of the relative consumption smoothing gains from extra coverage. We can indeed write that: $E_j\left(\frac{\partial\omega_j}{\partial b_j}\right)/E_j\left(\pi_j\left(b_j\right)\right) \equiv$

[53] The mark-up $\overline{MRS}_{b_0,b_1}$ is defined as the average marginal rate of substitution between consumption when employed and unemployed evaluated between coverage level $b_0$ and $b_1$. These estimates come from the model of Table 3 column (2), for which they find $E_1\left(\overline{MRS}_{b_0,b_1}\right) = 2.92$, $E_0\left(\overline{MRS}_{b_0,b_1}\right) = 1.12$

Figure F.2: Price Variation: Decomposition and Marginal Risk Curves




**Notes:** The Figure uses estimates of Figure 5 to trace out the marginal risk curves under basic coverage and comprehensive coverage using the sample of all eligible workers observed in 2006 and 2007. We start by positioning the three groups of individuals $j = 1, M, 0$ on the x-axis according to their willingness-to-pay. Individuals who choose the basic coverage (0) both in 2006 and 2007 are on the right hand side, while individuals who always buy the comprehensive coverage (1) are on the left-hand side of the graph. The marginals correspond to the group in between, and their share is given in Figure 4, where we see the fraction of individuals buying dropping from 86 to 78% with the 2007 price increase. We plot with green dots the observed average realized risk, measured by the number of days spent unemployed in year $t+1$, for the three groups. All risk measures are conditional on the vector of characteristics $Z$, and normalized to the average risk under basic coverage of individuals observed under basic coverage $E_0(\pi_0)$. The difference in realized risk (green dots) between the marginals and group 0 is therefore equivalent to our semi-elasticity estimate $\text{Semi}_{M(\mathbf{p})}^{Baseline}$. We then plot with blue triangles the average predicted risk under basic coverage and with red triangles the average predicted risk under comprehensive coverage for all three groups. The blue triangles identify the marginal risk curve in the basic coverage, while the red triangles plot the marginal risk curve in the comprehensive coverage. The difference between the red and blue triangles for each group identifies the moral hazard of moving these individuals from coverage 0 to coverage 1. These two curves provide all the information necessary to compute the fiscal externality term $FE_{\mathbf{p}}^{AS}$ that determines the optimal price level in Proposition 1.

$E_j\left(u'\left(c_u\left(b_j\right)\right)/u'\left(c_e\left(b_j\right)\right)\right) \cong 1+\gamma\times E_j\left(\frac{\Delta c}{c}\left(b_j\right)\right)$, where $c_u$ and $c_e$ denote the consumption levels when unemployed and employed, and $\gamma = \frac{u''(c)c}{u'(c)}$ is the parameter of relative risk aversion. Similarly, using a linear approximation for the mark-up, we have that: $E_j\left(\overline{MRS}_{b_0,b_1}\right) \cong 1+\gamma\times E_j\left(\frac{\Delta c}{c}\left(b_0\right)+\frac{\Delta c}{c}\left(b_1\right)\right)/2$. This allows to link the average mark-up $\overline{MRS}_{b_0,b_1}$ estimated in Landais and Spinnewijn [Forthcoming], to the marginal mark-ups evaluated at $b_0$ and $b_1$ respectively. In practice, note that we unfortunately cannot observe the counterfactual consumption wedge between employment and unemployment $\left(\frac{\Delta c}{c}\left(b_j\right)\right)$ for workers when put under a different plan $j$. However, we can easily provide bounds for these counterfactual drops. Assuming the consumption drop were to double when changing from comprehensive to basic coverage (i.e. $E_j(\frac{\Delta c}{c}(b_1)) = E_j(\frac{\Delta c}{c}(b_0))/2$), we get a value of 1.97 for the left-hand side in Proposition 2. Note that Landais and Spinnewijn [Forthcoming] find that the difference in consumption wedges for workers under comprehensive and basic coverage is actually quite small.[54] Assuming that the consumption drop doubles when changing from comprehensive to basic coverage is therefore a conservative upper bound. In other words, estimates from Landais and Spinnewijn [Forthcoming] suggest that the value of extra coverage is likely much more than 27% larger for workers under comprehensive coverage compared to workers under basic coverage, even when evaluated at $b_1$ and $b_0$ respectively. It follows that further differentiation in coverage levels would probably enhance welfare.

[54] They indeed find that $E_0\frac{\Delta c}{c}(b_0) = -.178(.25)$ and $E_1\frac{\Delta c}{c}(b_1) = -.138(.036)$.