Measuring moral hazard using insurance panel data

Juha Tolvanen

November 14, 2015
Latest version

Abstract

This paper describes two new methods for testing and measuring moral hazard that are applicable to a variety of insurance settings where panel data is available. The methods are able to measure moral hazard separately from adverse selection. The first method uses the variation in risk over different objects with different levels of insurance owned by the same person to isolate the average effect that insurance has on precautionary effort. For the second method, I show that many insurance contracts imply different incentives for precautionary effort for both the beginning and end of the contract period. These differences can be used to test moral hazard. I demonstrate the methods using a detailed insurance data set of virtual space ship insurance from a large virtual world called EVE Online. Because of moral hazard, the daily hazard rates for insured ships are sixfold compared to uninsured ships. Moral hazard is especially high at the beginning of each contract with newly insured ships having thirteen times higher daily hazard rates compared to ships with only 15 days left in their contract.

1 Introduction

Moral hazard and adverse selection can cause severe market failures in insurance markets.[1] Both phenomena can lead to inefficient levels of coverage and prices or even market breakdown. Being able to measure the two phenomena separately would help to direct resources more optimally to either alleviate moral hazard by investing in monitoring technologies or to reduce adverse selection by more accurately screening the insured population. As was already noted by Chiappori and Salanié (2000), separately estimating the two is difficult, if not impossible, using only cross-sectional data on accidents and coverage decisions. In this paper I introduce two new methods for separately measuring moral hazard with insurance panel data by using either variation in risk between differently insured items owned by the same agent or dynamic variation in risk over the contract period.

My first method considers situations where an agent purchases different levels of insurance for a number of similar items. Because all these items will share the agent's

[1] See, for example, Pauly, 1968 and Holmström, 1979 on moral hazard and Rothschild and Stiglitz, 1976 on adverse selection.

risk type, any variation in risk between them should be attributed to moral hazard. For example, consider a person who buys two televisions of the same make and buys product insurance or an additional warranty for one but not the other. Since both TVs are identical and owned by the same agent, differences in their accident risks cannot be due to adverse selection. If the insured television is still more likely to break down or get stolen, this difference in risk must be a result of differences in precautionary effort. Hence, estimating a model akin to a fixed effects model with agent fixed effects will allow the estimation of the average effect an increase in insurance coverage has on the decision maker's precautionary effort.

The key identification assumption guaranteeing the separation of moral hazard from adverse selection is that, from the agent's perspective, the insurable items are fairly similar. As I show in the theory section of the paper, the items need not be completely identical. For the empirical strategy to work, it is enough that the marginal cost and benefit from precautionary effort are the same for the two items when equal amounts of effort are spent on protecting them. The method works especially well with new items for which first insurance decisions are made at the time of purchase, because prior to owning the items the agent is unlikely to have any private information about their riskiness.

Furthermore, there is a wealth of situations that can be mapped to the theoretical model even if the insurable items seem initially different. For example, consider a situation where a person owns two otherwise identical cars but knows that one of them has some problems with its brakes and insures only that car. If, in absence of insurance, the agent repaired the brakes, then failing to repair them when the car is insured should still be classified as moral hazard. On the other hand, if the person did not touch the brakes even when the car is uninsured, then the broken brakes of the insured car should be classified as car-specific adverse selection. In the latter case my method will still provide useful information by separating the decision maker-specific risk type from the item-specific combined effect of adverse selection and moral hazard. The separation is valuable, because it helps in evaluating the relative benefit of screening agents versus screening the property they own.

My second method identifies moral hazard using the variation in risk over the contract period. It utilizes the following two common features of insurance contracts: first, many of them offer protection for a fixed period time; second, accidents that result in the loss of the insured property also terminate the insurance contract and only the payout of the contract is awarded to the insured agent. Specifically, many common insurance contracts do not refund any fraction of the price of the contract if it is terminated before the contract period ends. One example of contracts that have this structure are warranties and additional product insurance policies. If a TV with a warranty breaks down beyond repair, the owner will get a refund and the warranty will then terminate, as there is no TV to be covered by the warranty. Similarly, car insurance policies in the United Kingdom often have a structure where no premiums are returned upon termination of the contract, if the insured person has made a claim during the period of insurance.

This contract structure implies that, in case of an accident, the insured agent will not only lose her deductible but also the intrinsic value of the remaining insurance coverage.

Consequently, the agent faces different incentives for precautionary effort depending on how much time is left in the insurance contract. The value of the contract will be higher, the longer the remaining coverage period. Therefore, the incentives for precautionary effort should be higher in the beginning of the insurance contract and decrease towards its end. As long as agents' risk types are static over time, adverse selection will not produce a similar dynamic pattern.

I also show how the two methods can be combined, if there is data that simultaneously fulfills the prerequisites of both of the methods. The combined method boils down to estimating the variation in risk over items owned by the same person with different amounts of time left in their insurance contracts. The logic above implies that at any given point in time the item with more time left in its contract should be less risky than any item owned by the same agent that is close to the end of its contract period. This combined test is hence a version of the second test that is robust to seasonality and heterogeneity in risk and tastes, because comparisons are made separately within each time period and between different objects owned by the same agent.

On the econometric side, I use versions of the Cox proportional hazards model (Cox, 1972) to construct my tests for moral hazard. The model has multiple benefits over commonly used probit and logit models used in earlier cross-sectional studies. First, it allows for highly flexible agent-specific base line hazard rates that let me capture dynamically varying heterogeneity in both tastes and risk types. Second, the model helps me to paint a more detailed picture of the agent's effort choices over time and flexibly estimate dynamic changes in moral hazard. For example, the model easily allows for time-varying covariates such as the time left in the insurance contract or the age of an insured item.

To illustrate the application of my methods I use data from a large multiplayer online game called EVE Online, where the players have the option to insure their virtual space ship against destruction. The game offers an ideal, almost laboratory-like environment for applying the methods for three reasons: First, the incentives in the game are clearly defined and, for example, the value of a ship or an insurance contract can be measured in real monetary terms. Second, the insurance contracts in the game are very similar to many real world insurance contracts and they satisfy all of the identifying assumptions required by my methods. Last, the server log data allows me to accurately measure all of the key variables without error and with a high dynamic frequency.

Using my first method that measures the average difference in risk between insured and uninsured items owned by the same person over the whole contract period, reveals that moral hazard leads to an insured ship being about 6 times more likely to get destroyed each day compared to an uninsured ship. When I then focus on the detailed evolution of risk over the insurance contract, an interesting pattern emerges: I show that a newly insured ship is about 13 times more likely to get destroyed each day compared to an identical ship with only 15 days left in its contract when both ships are owned simultaneously by the same agent. Although moral hazard causes insured ships to be more risky over the whole contract period, this empirical pattern of decreasing dynamic risk contradicts the theoretical prediction outlined above for how optimally behaving agents should be reducing their precautionary effort towards the end of the contract period. Last, I show that experience or learning is a key factor in the magnitude of moral hazard with the daily effect of moral hazard on risk being almost five times as big for experienced players than for new-comers.

It is noteworthy that, even though my model and econometric methods are framed in terms of insurance, their logic applies potentially to other contracts such as employment contracts. For example, the insights of my dynamic model could be used to measure moral hazard in temporary project contracts with hard deadlines. If poor output can lead to an early termination of the contract and if at least some wage payments are made over the course of the contract, then the incentive structure for the employee is identical to the one given in my insurance model: When the project is close to its end, by shirking the employee loses only the value of the remaining few wage payments. However, early in the project the value of the potential stream of payments is much higher. Consequently, effort should be lessening towards the end of the contract. The result is robust to even reputation effects, as long as the reputational cost of a premature termination is not lower in the beginning of the contract. These ideas are somewhat similar to what was presented in Gibbons and Murphy (1992).

On a more general level, this paper highlights that despite our imperfect understanding of optimal dynamic contracting with both adverse selection and moral hazard, theoretical modeling of existing contract structures can still be used to derive predictions of optimal responses to these contracts that are testable using panel data. The econometric methods used in this paper to measure dynamic variation in risk can be flexibly applied to other contracting situations with completely different incentive structures.

The literature on measuring moral hazard separately from adverse selection is fairly new.[2] The papers that come thematically closest to mine are the ones by Abbring, Chiappori and Pinquet (2003) and Dionne, Michaud and Dahchour (2013).[3] Both papers use the bonus malus feature of French car insurance that increases the premium for all future years after an accident and reduces the premium after a year without an accident. The authors argue that moral hazard would imply that there should be a discontinuous increase in effort after an accident. The authors find little evidence of moral hazard in their data. Dionne, Michaud and Dachour (2013) look at a panel data over a longer period and are able to control for learning about one's risk type that could confound the results if bad drivers start exerting more effort after learning about their risk type from previous accidents. Their parametric model finds statistically significant evidence for moral hazard among drivers with less than 15 years of driving experience but does not find any evidence of moral hazard for drivers with more than 15 years of driving experience. It is noteworthy that a method that bases its identification on previous accidents may yield biased results. If an accident leads to the car being damaged or the driver being injured or traumatized, the agent is likely to drive her car much less than normally. Consequently, the likelihood of future accidents is lower during the months following an accident for reasons that have little to do with the agent's effort choices. The empirical pattern is much like what is generated by the incentives from the bonus malus system.

Another strain of literature estimating moral hazard uses discontinuities or kinks

[2]There is a large body of empirical literature measuring the combined effect of moral hazard and adverse selection. For recent reviews, see Einav et al. (2010) and Cohen and Siegelman (2010).

[3]See also Abbring, Heckman, Chiappori and Pinquet (2003).

in the incentive structure of offered contracts. Dobbie and Skiba (2013) consider the payday lending market where moral hazard would imply that borrowers with larger loans have a smaller opportunity cost of defaulting, while adverse selection implies that borrowers with an ex-ante higher risk of defaulting know they are likely to default and hence select a loan with a higher value. Dobbie and Skiba notice that offered loan sizes change discontinuously with respect to a borrower's pay. Assuming that borrowers around the discontinuity are on average fairly similar, the authors are able to measure how much increased loan size affects the likelihood of a given borrower paying back the loan and thus get a measure of moral hazard. Einav, Finkelstein and Schrimpf (forthcoming) is conceptually fairly similar.[4] It uses discontinuities in the co-insurance rate of Medicare Part D prescription drug coverage to estimate the effect insurance has on people's drug purchases.

My paper is unique in the way it uses variation in risk within a person, either between separate items owned by the same agent or over time for a given item, to factor out adverse selection and identify moral hazard by keeping the agent's risk type constant. Adams, Einav and Levin (2009) share a similarity with this idea, as they also use a special feature of their data to hold the agent's risk type constant while varying the incentives for effort.[5] It uses data from a large auto sales company that lends money to its customers who have a hard time obtaining credit elsewhere. The forms of moral hazard and adverse selection are conceptually the same as in Dobbie and Skiba (2013). Their data contains two sources of variation in loan size which allow them to separate moral hazard from adverse selection: the choice of the down payment and the variation in the pricing of the purchased car which they claim to be orthogonal to the borrower's risk type conditional on his/her down payment choice. The first source of variation allows the authors to pin down the contract choice and hence control for any private information that the customer has about their credit worthiness. As long as the remaining variation in loan size that comes from the price of the offered car is not related to the customer's risk type in a way not captured by the down payment choice, then this variation will not be influenced by adverse selection and can be used to measure the effect that loan size has on moral hazard.

I also illustrate how virtual worlds can be used as living laboratories where economists can test their theories under fairly idealized incentive structures while still retaining extremely large sample sizes. The possibility of using virtual worlds as a testing ground for theories from social sciences is a fairly novel idea (see Bainbridge, 2007). These environments usually include many of the institutions found in normal societies, while the total complexity of social interaction is considerably reduced. This creates laboratory-like environments with study populations that are a thousand fold to what researchers usually can afford to hire for their experiments. Furthermore, if the researcher has access to the log data of the virtual world, he or she can accurately observe practically everything that happens in the world. Last, observing log data allows the researcher to run studies where the population under study is unaware of the fact that they are being observed. Such studies can avoid the usual threats to internal validity due to reactivity

[4] See also Einav, Finkelstein, Kluender and Schrimpf (2015).

[5] See also Einav, Jenkins and Levin (2012).

of the subjects to being in an experiment such as the Hawthorne effect.[6]

Economists have slowly become aware of the research potential of data gathered from the virtual worlds. Initial studies were mostly explorative trying to document what was happening in these online hang-outs (see, for example, Castronova 2001). The more recent papers have used virtual worlds to study social interaction in trust games (Fiedler et al., 2011 and Fiedler and Haruvy, 2009), online labor markets (Horton et al., 2011), the market for avatars (Castronova, 2004) and the general potential of virtual worlds for experiments (Chesney et al. 2009). All of these papers are experiments and not observational studies that would directly use the server log data the way I do here. For a previous study that uses log data, see Golde (2008) on the efficiency of thick markets.

The rest of the paper is organized as follows. The next section presents my theoretical models and formalizes my two identification strategies. Section 3 describes my virtual world data. Section 4 presents the results from applying my methods to the data. Section 5 investigates the effect of experience on moral hazard and Section 6 concludes the study.

2 Theoretical model and hypotheses

This section starts by introducing a simple static model wherein each agent owns two insurable objects and which provides my identificaton method with the agent “fixed effects”. I then add dynamic effort choice to the model but drop the second insurable item to illustrate the idea behind my dynamic identification strategy. Last, I combine the two models by assuming dynamic effort choice with two insurable items. This produces a method that requires the most from the data but that controls for the richest variety of alternative explanations. All of the proofs for this section can be found in the Appendix.

2.1 A static model with two insurable units

In short, the model in this section is as follows: A population of heterogeneous, risk averse agents each owns two insurable items. The heterogeneity can be in terms of risk types, preferences or effort costs. They first choose which of the items to insure (if any) and then select a separate level of precautionary effort for both items. Finally, some items are randomly involved in an accident that destroys them resulting in a monetary loss for the agent. The probability of an accident is inversely related to how much precautionary effort was taken to protect it. If the item was insured, some of the loss is covered by the insurance. Conditional on the agent’s type and effort choices the accidents happen independently of each other. I show that in a model like this, for a given person owning both an insured and an uninsured item, the person will always exercise more precautionary effort to protect the uninsured item. This difference can then be used to test moral hazard.

[6]See, for example, Adair (1984) for a more detailed description of the Hawthorne effect and a review of empirical literature documenting it.

Formally, I assume that every agent $j \in \mathbb{N}$ owns two items indexed by $i = 1, 2$, both of which can be insured against a monetary loss of L for a premium C to get a pay-out P in case there is an accident. I assume that $L - P > 0$, i.e. that there is a deductible in the insurance contract. Let $I_j \in \{0, 1\}^2$ be the agent's choice of insurance contracts where $I_{ij} = 1$ is interpreted to mean that item i is insured.

Suppose that agent j's monetary endowment is equal to ω_j. I assume that ω_j is independently and identically distributed according to some probability distribution F_ω for all $j \in \mathbb{N}$. Furthermore, I allow heterogeneity in terms of a random variable $\varepsilon_j \in \mathbb{R}$ that is independently and identically distributed according to some F_ε. This ε_j captures arbitrary variation in tastes, risk types and risk-aversion in ways that will become clear below.

The consumer's von Neumann-Morgenstern utility from consuming wealth c when his/her heterogeneity term is ε is given by $u(c, \varepsilon)$, where

$$u \colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$$

I assume that u is continuous everywhere and twice continuously differentiable with respect to c. Furthermore, I require that $\partial_c u(c, \varepsilon) > 0$, $\partial_c^2 u(c, \varepsilon) < 0$ for all $c \in \mathbb{R}_+$ and $\varepsilon \in \mathbb{R}$. Notice that u can depend arbitrarily on ε.

Each agent chooses precautionary effort levels (x_{1j}, x_{2j}) that determine the probability of an accident independently for each item. Let $D_j \in \{0, 1\}^2$ be a binary random vector where $D_{ij} = 1$ is interpreted as an accident that happened to item $i = 1, 2$ owned by agent j. I assume that

$$\mathbb{P}(D_{ij} = 1 \mid \varepsilon_j, x_{1j}, x_{2j}) = p(x_{ij}, \varepsilon_j),$$

where $p \colon \mathbb{R}_{++} \times \mathbb{R} \to [0, 1]$ is a mapping that translates effort choices and risk/taste types to accident probabilities. Notice that I am assuming that given the same precautionary effort for both items owned by the same person, their accident probabilities will also be given by the same mapping. I require that conditional on ε_j and (x_{1j}, x_{2j}), D_{1j} is independent of D_{2j}. The assumption of conditional independence combined with the assumption that the accident probability for item 1 does not depend on the effort choice for item 2 precludes settings where investing in precautionary effort may cause spillover effects to the accident probability of the other item. For example, investing in fire alarms for your garage will not only protect the building but also the car parked in it and any other buildings near-by. This type of interdependence is not allowed by the model.

Furthermore, D_j is assumed to be independent of D_k for $j \neq k$. I also assume that $p(\cdot, \varepsilon_j)$ is twice continuously differentiable, decreasing and convex for every ε_j and that the mapping $(x_{1j}, x_{2j}) \mapsto p(x_{1j}, \varepsilon_j) p(x_{2j}, \varepsilon_j)$ is strictly convex. The latter assumption guarantees the concavity of the agent's optimization problem.

The utility cost of effort is measured by a function $E \colon \mathbb{R}_+^2 \times \mathbb{R} \to \mathbb{R}_+$. Here, $E(x_{1j}, x_{2j}, \varepsilon_j)$ is interpreted as the utility cost from effort levels (x_{1j}, x_{2j}) for person j. This function is also allowed to depend arbitrarily on ε. The key identifying assumption for this section is that the marginal cost of effort for the two items is equal at equal effort levels:

Assumption 1 (Diagonal symmmetry)**.** *The cost function satisfies:* $\partial_1 E(x, x, \varepsilon_j) = \partial_2 E(x, x, \varepsilon_j)$ *for all* ε_j *and all* $x \in \mathbb{R}_+$.

This is clearly more general than assuming that E is symmetric with respect to x_{1j} and x_{2j}. Hence, the two items need not be completely identical, as long as the marginal return to effort is equal for the two items at equal effort levels. I also require that E is twice continuously differentiable, increasing and convex in its first two arguments for every ε_j and that $\partial_{12}E(x_1, x_2, \varepsilon_j) \geq 0$. To guarantee that first-order conditions are satisfied at some (x_{1j}, x_{2j}) I assume that for all x_{-i}, $\lim_{x_i \to 0} \partial_i E(x_1, x_2) = 0$ and $\lim_{x_i \to \infty} \partial_i E(x_1, x_2) = \infty$.

The utility maximization problem for agent j with a risk/taste type ε_j is given by

$$\begin{aligned}
&\max_{I_j, x_{1j}, x_{2j}} p(x_{1j}, \varepsilon) p(x_{2j}, \varepsilon_j) u(\omega_j - 2L + (I_{1j} + I_{2j})(P - C), \varepsilon_j) \\
&+ p(x_{1j}, \varepsilon_j)(1 - p(x_{2j}, \varepsilon_j)) u(\omega_j - L - (I_{1j} + I_{2j})C + I_{1j}P, \varepsilon_j) \\
&+ p(x_{2j}, \varepsilon_j)(1 - p(x_{1j}, \varepsilon_j)) u(\omega_j - L - (I_{1j} + I_{2j})C + I_{2j}P, \varepsilon_j) \\
&+ (1 - p(x_{1j}, \varepsilon_j))(1 - p(x_{2j}, \varepsilon_j)) u(\omega_j - (I_{1j} + I_{2j})C, \varepsilon_j) \\
&- E(x_{1j}, x_{2j}, \varepsilon_j).
\end{aligned} \tag{2.1}$$

The first result of this section shows that, if an agent finds herself with both or none of her items insured, then the accident probabilities for both of the items will be the same. The point of this result is to highlight that even extreme asymmetries in the cost function that are off the diagonal will not translate into differences in effort choices when both items have the same insurance status.[7]

Proposition 1. *The following equalities hold:*

1. $\mathbb{E}[p(x_{1j}, \varepsilon_j) \mid I_{1j} = 0, I_{2j} = 0] = \mathbb{E}[p(x_{2j}, \varepsilon_j) \mid I_{1j} = 0, I_{2j} = 0]$
2. $\mathbb{E}[p(x_{1j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 1] = \mathbb{E}[p(x_{2j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 1]$.

By potentially re-indexing the items, it is without loss of generality to assume that if the agent insures only one item, she insures item number 1. The key identification result from this model is given in the proposition below:

Proposition 2. *Optimal effort choices imply that*

$$\mathbb{E}[p(x_{1j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 0] - \mathbb{E}[p(x_{2j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 0] > 0. \tag{2.2}$$

That is, if it is known that a person has insured one of her items but not the other, then one would expect a higher accident probability for the insured item. The result is highly intuitive: ceteris paribus, an agent should be taking better care of her uninsured items. It is easy to also see that given the assumptions on identical items, if I shut down the choice of effort and consider a model with only heterogeneity in risk type, tastes

[7]Notice that I allow for selection into insurance based on differing marginal costs of effort between the items off the diagonal. However, I classify any resulting differences in risk between the insured and uninsured items as moral hazard, because when both items are uninsured the items would have the same accident probability.

and wealth, then the accident probabilities for each insurable object owned by each person will be the same and hence the inequality above does not hold. Consequently, the moment condition in Proposition 2 can be used to identify moral hazard separately from adverse selection.

The key assumptions are that at least some agents own multiple *nearly identical* objects and purchase varying degrees of coverage for them. When this is satisfied the test accounts for a large variety of possible sources of heterogeneity ranging from differences in wealth through differences in risk aversion all the way to differences in risk types.

As argued by Chiappori and Salanié (2000), testing whether the difference

$$E[p(x_{ij}, \varepsilon_j) \mid I_{ij} = 1] - \mathbb{E}[p(x_{ik}, \varepsilon_k) \mid I_{ik} = 0] \tag{2.3}$$

is positive yields a test of the combined effect of adverse selection and moral hazard. Insured items are more risky, because more risky agents buy more insurance and insurance reduces incentives for precautionary effort. It is tempting to argue that one gets a measure for adverse selection by taking the difference between the combined effect of moral hazard and adverse selection from 2.3 and the pure effect of moral hazard from 2.2. However, this intuition is not strictly correct. The problem is that it may be optimal for agents who insure one of their items to shift effort from the insured item to the uninsured item. If this "substitution" effect is strong enough it is possible to generate situations where the agents who buy insurance have a higher risk type but their uninsured item ends up being less risky than the items of agents who do not buy insurance. If this happens, it is possible that the combined effect of moral hazard and adverse selection from 2.3 is in fact lower than the effect of moral hazard from 2.2.

The idea is illustrated by the following numerical example. Consider agents for whom $u(x) = x^{\frac{9}{10}}$, $E(x, y) = (x + y)^2$ and $p(x, \varepsilon) = \frac{1}{1+x} + \varepsilon$. Let $\omega_j = 15$ for all agents and let $L = 7$, $P = 1.2$ and $C = 1$. In Figure 1 I plot the agents' expected utility given optimal effort choices as a function of their risk type ε and depending on their choice of insurance. One line shows the value of the maximization problem, if they insure one unit and the other shows what they get, if they leave both items uninsured. The dashed lines show two risk types of whom the left one never wants to insure her items while the one on the right chooses to insure one item.[8]

In Figure 2 I plot the optimal accident probabilities for items under different insurance regimes as a function of the agent's risk type. The highest line shows the accident probability of an insured item when the other unit is uninsured. The middle line shows the accident probabilities of the items of an agent who left both uninsured and finally, the lowest line depicts the accident probability of the uninsured unit of a person who insured one of her items. As can be seen from the picture, with this parametric specification the agent always chooses lower risk for the uninsured item when she owns one insured item than what she would choose if both items were uninsured. In other words, she "transfers" risk from the uninsured item to the insured one.

Consider then a scenario where most of the population is of the lower risk type with $\varepsilon = 0.14$ (the first dashed line) and a very small fraction has risk type $\varepsilon = 0.15$. It

[8]It can be shown that neither of the types wants to insure both items.

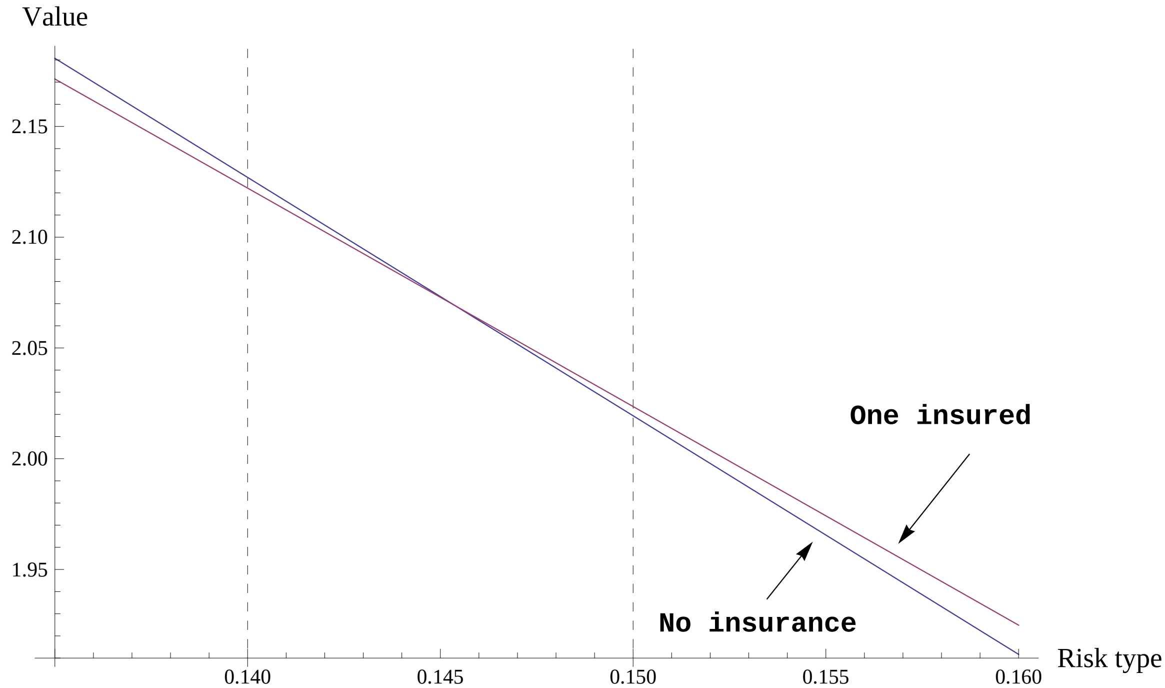

Figure 1: Value of the maximization problem as a function of risk type

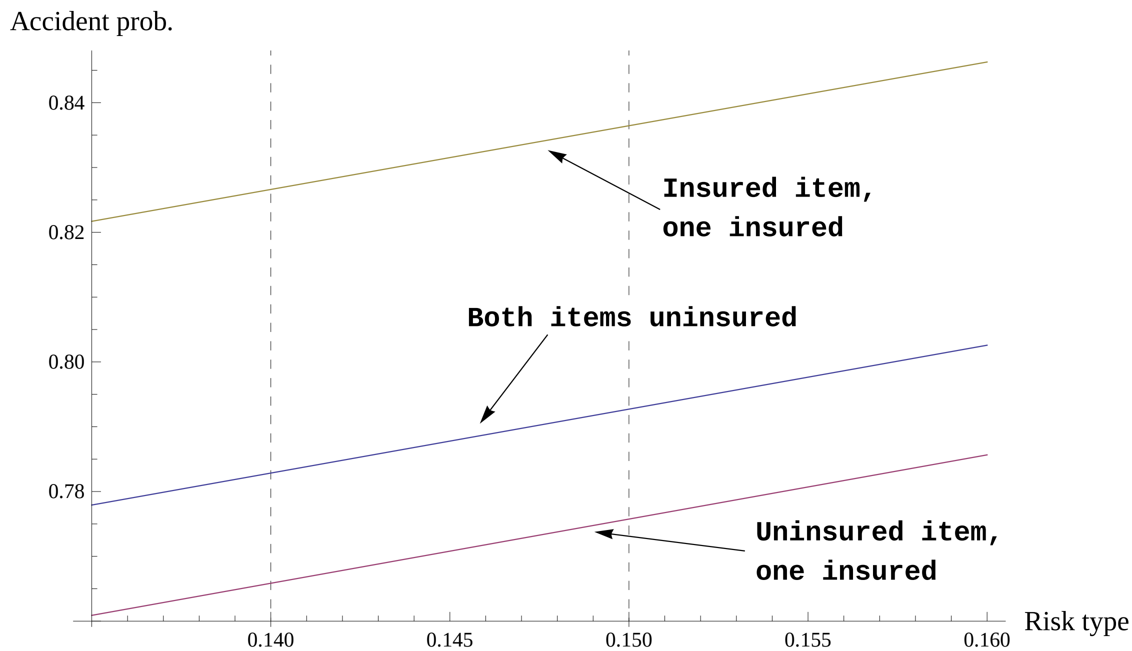

Figure 2: Accident probabilities as a function of risk type

is easy to check that neither of these types wants to insure both items. Now, the first picture shows that the lower risk type never insures any of her items while the higher risk type prefers to insure one unit. If I estimate the ratio of accident probabilities between insured and uninsured items owned by the same person, I will be estimating the ratio between the highest and the lowest line at $\varepsilon = 0.15$, because these are the only types in the population who own both an insured and an uninsured item. The ratio of these accident probabilities is about $0.836/0.776 \approx 1.08$. However, when estimating 2.3 most of the uninsured observations come from the population for whom $\varepsilon = 0.14$ as they were an overwhelming majority in the population. Consequently, the compound estimate will end up comparing the middle line at $\varepsilon = 0.14$ to the highest line at $\varepsilon = 1.50$. This results in a hazard ratio of about $0.836/0.783 \approx 1.07 < 1.08$. Consequently, a too hasty application of this technique would lead me to claim that there is advantageous selection in the population even though in reality the converse is true.

The problem with looking at the residual positive correlation in the cross section that is not explained by my fixed effects method for moral hazard is, that it always contains some element of comparing apples with oranges: To measure adverse selection, I would like to compare the risks of either insured or uninsured items of two risk types conditional on them facing the same incentives for precautionary effort. However, when one of the types is insuring one item and the other is insuring none, not only the insured items, but also the uninsured items have different returns to protective effort. As I never observe the two risk types making the same insurance choices, the incentives for precautionary effort will also vary between both items owned by the two types. Hence, without fixing functional forms for the utility function and the way effort is transformed into accident risk, it is hard to clean the estimate of adverse selection of the effect that insurance has on the effort choice for the uninsured item when the person's other item is insured.

2.2 A dynamic model with a single insurable unit

I will next add dynamic choice of effort into the model but assume that each person owns only a single insurable unit. I also assume that the insurance contracts have the form described in the introduction: they are taken for l units of time or until the next accident happens. The offered length l is set exogenously by the insurance provider. I show that in the beginning of the contract the net utility loss from an accident is higher than in the end because the longer the remaining insurance coverage period the more valuable it is and hence the more it hurts to lose it. The section gives a set of assumptions under which items with more time left in their contracts are also cross-sectionally less risky than items whose contract is about to end. My goal in this section is to first formally illustrate the dynamic moral hazard outlined above and then discuss some of the restrictions of trying to measure this effect across different agents. The assumptions of this section are fairly demanding but most of them are relaxed in the following section where I discuss my model with dynamic effort choice for two insurable items.

I assume that time is discrete and the timing of events in each period, $t = 1, 2, \ldots$, is as follows: In the beginning of the period the agent observes how much time $I'_{jt} \in \{0, 1, \ldots, l-1\}$ she has left in her insurance contract. Here $l \in \{1, 2, \ldots\}$ is the length of the offered contract and it is set by the insurance company. After observing I'_{jt} the agent decides how much precautionary effort x_{jt} she would like to exercise and whether she wants to renew her contract by setting $I_{jt} = l$ by paying C or just keep whatever time she has in her contract (if any) by setting $I_{jt} = I'_{jt}$ for free. After the agent's choice the accident risk is realized. If I let D_{jt} be an indicator of an accident for person j in period t, then person j's monetary endowment at the end of period t is given by

$$\omega - D_{jt}(L - \mathbb{1}\{I_{jt} > 0\}P) - \mathbb{1}\{I_{jt} = l\}C,$$

where $\mathbb{1}A$ is an indicator of event A. Just as before, ω is the person's per period monetary income, L is the size of the loss, P is the insurance payout and C is the cost of renewing the insurance contract. The agent then enjoys her utility for period t given by

$$u(\omega - D_{jt}(L - \mathbb{1}\{I_{jt} > 0\}P) - \mathbb{1}\{I_{jt} = l\}C) - E(x_{jt}, \varepsilon_j)$$

Here $u(\cdot)$ is just as in section 2.1 and $E(\cdot, \varepsilon_j)$ is a twice continuously differentiable, increasing and convex cost of effort. For this subsection, I assume that there is only heterogeneity in risk types and effort costs, and this heterogeneity is captured by $\varepsilon_j \in \mathbb{R}$ with higher values interpreted as higher risk types. I require that

$$\partial_x \partial_\varepsilon E(x, \varepsilon) \geq 0.$$

This single-crossing condition guarantees that, true to my interpretation of ε, increasing risk type does not imply that effort becomes relatively cheaper. Otherwise, agents with higher ε might optimally choose lower accident probabilities and hence, calling them higher risk types would be misleading. I also assume that $\lim_{x\to 0} \partial_x E(x, \varepsilon_j) = 0$ and $\lim_{x\to\infty} \partial_x E(x, \varepsilon_j) = \infty$. Given an effort level x, agent j's accident probability is $p(x, \varepsilon_j)$ where $p(\cdot, \varepsilon_j)$ is twice continuously differentiable, decreasing and convex for every ε_j. Furthermore, I assume that $\partial_\varepsilon p(x, \varepsilon) > 0$, and $\partial_x \partial_\varepsilon p(x, \varepsilon) \geq 0$. The interpretation for the first is that higher risk types have higher accident probabilities at every effort level and the second implies that increases in effort do not buy larger decreases in risk for higher risk types. The latter assumption complements the assumption that $\partial_x \partial_\varepsilon E(x, \varepsilon) \geq 0$ in making sure that, ceteris paribus, higher risk types will not have lower realized risk. Last, I assume that every agent discounts the future by a common discount factor $\delta \in (0, 1)$.

A central identification assumption in this section is that an accident resets the time left in the insurance contract to zero. Formally, I assume that the time left in the insurance contract at the beginning of period $t+1$ is given by $I'_{t+1j} = (1 - D_{jt})(\max\{0, I_{tj} - 1\})$. In other words, either an accident resets the time left to 0 or, in the absence of an accident, the person starts her next period with one unit of time less left in her contract. Of course, if her contract has already run out of time, she will start the period with zero units of time in her contract.

I assume for simplicity that there are no endogenous saving decisions or heterogeneity in wealth and that the agent's per period monetary income is fixed.[9] Without assuming constant absolute risk aversion, heterogeneity or fluctuations in unobserved wealth translate to potential heterogeneity in risk preferences that considerably complicate the problem.

Let $h_t = \{x_{js}, I_{js}, D_{js}\}_{s=0}^{t-1}$ be the t length history of accidents and insurance and effort choices before period t and let $\mathcal{H}$ be the set of all of these histories. The initial

[9] I allow for arbitrary saving decisions in my most general model below.

levels (x_{j0}, I_{j0}, D_{j0}) are assumed exogenous.[10] If the agent starts at $t = 0$ with I' units of insurance, then the agent's problem is to choose functions $x_j : \mathcal{H} \to \mathbb{R}_+$ and $I_j : \mathcal{H} \to \{0, 1, 2, \ldots, l\}$ to solve

$$V_j(I') := \max_{x_j, I_j} \mathbb{E}\Bigg[\sum_{t=0}^{\infty} \Big\{u(\omega - D_{jt+1}(L - \mathbb{1}\{I_j(h_t) > 0\}P) - \mathbb{1}\{I_j(h_t) = l\}C)$$
$$-E(x_j(h_t), \varepsilon_j)\Big\}\delta^t \mid x_j, I_j, \varepsilon_j\Bigg], \tag{2.4}$$

where the expectation is over the realization of histories and where for a history h_t, $I_j(h_t)$ is constrained to the set $\{\max\{0, (1 - D_{jt-1})I_j(h_{t-1}) - 1\}, l\}$. This is a Markov decision process where the time left in the insurance contract when entering the period, I'_t, can be taken as the state. Arguments given in, for example Bertsekas (1995), can be used to show that a maximizing policy can be found from the set of policies that depend only on I'_t. A solution to (2.4) is then obtained by solving the following Bellman equation:

$$V(I', \varepsilon_j) =$$
$$\max_{x_j(I'), I_j(I')} \Big\{p(x_j(I'), \varepsilon_j)\,[u(\omega - L + \mathbb{1}\{I_j(I') > 0\}P - \mathbb{1}\{I_j(I') = l\}C) + \delta V_j(0)]$$
$$+ (1 - p(x_j(I'), \varepsilon_j))\,[u(\omega - \mathbb{1}\{I_j(I') = l\}C) + \delta_j V_j(\max\{0, I_j(I') - 1\})]$$
$$- E(x_j(I'), \varepsilon_j)\Big\}, \tag{2.5}$$

subject to $I_j(I') \in \{l, I'\}$.

The first key observation is that for any individual the probability of an accident increases as the end of the insurance contract becomes closer. The formal version of this idea is given in the following lemma:

Lemma 1. *If $I(I') > I(I'') > 0$, then $p(x_j(I'), \varepsilon_j) < p(x_j(I''), \varepsilon_j)$ for all ε_j.*

The full proof of this lemma is in the Appendix. Here I highlight the key step behind the proof. I consider first the harder case, where $I(I') < l$. For $\hat{I} \in \{I', I''\}$ the first-order condition with respect to x is:

$$p'(x, \varepsilon_j)\Big(u(\omega - L + P) - u(\omega) + \delta_j V(0, \varepsilon_j) - \delta_j V(I(\hat{I}) - 1, \varepsilon_j)\Big) = E'(x, \varepsilon_j).$$

This can be interpreted as a standard balancing of marginal cost and marginal benefit of effort. On the left-hand side the benefit comes from two sources: First, higher effort (x) decreases the probability of an accident (p). Increasing x moves probability mass at rate p' from the event where the flow utility is $u(\omega - L + P)$ to the event where it is $u(\omega) > u(\omega - L + P)$. It also increases the likelihood of the continuation utility being $V(I(\hat{I}) - 1, \varepsilon_j)$ rather than $V(0, \varepsilon_j)$. Because the agent is risk averse, she finds

[10] Of these only the level of I_{j0} will affect the solution to the agent's maximization problem and none of them change my results.

insurance valuable and the more periods that remain in the contract, the better. In other words, V is increasing in its first argument. I show this intuitive result formally in the Appendix. But this then implies that $V(I(I')-1, \varepsilon_j) > V(I(I'')-1, \varepsilon_j) \geq V(0, \varepsilon_j)$. Thus, the marginal benefit for effort is higher when the insurance choice is $I(I')$ than when it is $I(I'')$ implying higher effort choices and hence lower accident probabilities. When $I(I') = l$ the effect is even stronger, since the contemporaneous utility loss from an accident is $u(\omega - L + P - C) - u(\omega - C) < u(\omega - L + P) - u(\omega)$ where the inequality follows from concavity of u. The income effect of the momentary wealth loss due to the insurance payment makes the agent more sensitive to the absolute monetary loss from the accident. Of course, allowing saving and borrowing would smooth some of this effect over several periods. However, this observation is otherwise very general and holds with much richer assumptions about the individuals.

Now, consider contracts that are fully separating in the sense that all of the agents who buy the insurance contract share the same risk type ε. Then all of the agents buying insurance will be identical and consequently will also choose the same level of precautionary effort when faced with the same amount of time left in their contracts. Lemma 10 then implies that on expectation an agent with less time left in her contract will be more risky than an agent who is in the beginning of her contract. Formally, this idea is given by the following corollary:

Corollary 1. *Assume that the contract is such that if $I_j > 0$ and $I_k > 0$ then $\varepsilon_j = \varepsilon_k =: \hat{\varepsilon}$. In that case $I(I', \varepsilon) > I(I'', \varepsilon) > 0$ implies that*

$$p(x(I'), \hat{\varepsilon}) = \mathbb{E}[p(x, \varepsilon) \mid I(I', \varepsilon)] < \mathbb{E}[p(x, \varepsilon) \mid I(I'', \varepsilon)] = p(x(I''), \hat{\varepsilon}).$$

This is a testable implication with my data.

Consider then a model without effort choice and hence without moral hazard where each agent solves:

$$\begin{aligned} V_j(I') = & \\ & \max_{I_j(I')} \Big\{ p(\varepsilon_j) \left[u(\omega - L + \mathbb{1}\{I_j(I') > 0\}P - \mathbb{1}\{I_j(I') = l\}C) + \delta_j V_j(0) \right] \\ & + (1 - p(\varepsilon_j)) \left[u(\omega - \mathbb{1}\{I_j(I') = l\}C) + \delta_j V_j(\max\{0, I_j(I') - 1\}) \right] \Big\} \end{aligned} \tag{2.6}$$

Here, observing $I_{jt} > I_{kt} > 0$ implies that person j has on average been a shorter time without an accident than person k. In other words, observing person k with less time in her contract than what j has implies that she is on average *less risky* than j because she is known to have seen at least $l - I_{kt}$ periods without an accident, while for person j the same lower bound is only $l - I_{jt}$. This insight can be used to prove the following result:

Proposition 3. *In the model without moral hazard given in (2.6), if $I_{jt} > I_{kt}$, then $\mathbb{E}[D_{jt} \mid I_{jt}] \geq \mathbb{E}[D_{kt} \mid I_{kt}]$.*

This inequality holds as an equality only if there is almost surely no heterogeneity of risk types in the population buying the contract – in other words, if the contract is separating in the sense described above. This is my second testable result. To summarize, whenever the contract is non-separating and there is both moral hazard and adverse selection, the econometrician may observe new contracts being riskier than old contracts or vice versa. If old contracts are riskier than new ones, that can be taken as proof of moral hazard while new contracts being riskier than old ones implies that there is a possibility of adverse selection. If the hazard rate is constant over the span of the insurance contract, the test is inconclusive, either the contract is separating in the above sense[11] and there is no moral hazard, or the heterogeneity in risk types completely offsets the effect of moral hazard. I would like to emphasize that heterogeneity of risk types in the pool of insuring agents does not strictly speaking imply adverse selection, as the selection to the insurance pool may be based on factors such as tastes that may be completely orthogonal to the person's risk type. This could yield the same distribution of risk types in both insuring and non-insuring populations with any differences in risk between the two populations being completely explained by moral hazard. This ambiguity is fully resolved by the richer model in the next section.

A notable restriction of the model is that it does not allow for risk to vary over time. Even very uniform seasonal variation in risk may cause problems, if it results in different risk types buying insurance at different times. Patterns where riskier types buy insurance when seasonal risk is lower can potentially lead to observing periods where agents with more time left in their contracts are less risky even if the agents cannot affect their riskiness. Controlling for seasonality with something like time fixed effects can mitigate the problem as long as the effect of the seasonality is relatively homogeneous for different risk types. However, this problem is not present in my third identification strategy presented in the next subsection.

Second, the model above does not allow for heterogeneity in risk preferences or wealth. Allowing for heterogeneity in wealth is clearly almost synonymous to allowing for variations in risk preferences unless I assume constant absolute risk aversion. Variation in risk preferences, on the other hand, is problematic because it does not yield clear predictions of how it will translate into variations in accident probabilities. This is the old dichotomy between self-insurance and self-protection (see Ehrlich and Becker, 1972, Dionne and Eeckhoudt, 1984 and Briys and Schleisinger, 1990). While more risk averse agents are always willing to accept higher reductions to their welfare in the good state in exchange for an increase in welfare in the bad state (self-insurance), the same is not true for paying a cost in both states for reducing the probability of the bad state (self-protection). In fact, the authors above present examples where more risk averse agents exhibit more risky behavior. The intuition behind the result is that self-protection, i.e. increasing effort to reduce the accident probability, makes the agent worse off in both the good and the bad state. In contrast, the mean preserving contractions that are always preferred by a risk averse agent have the feature that they always increase the payoff in the bad state.

The result from Proposition 3 clearly still holds if I allow for variation in risk preferences, because it assumes no variation in effort. If I assume enough on the utilities to guarantee that more risk averse agents choose higher levels of effort, and if there

[11] Notice that this contains the case where there is no adverse selection.

are multiple risk preference types in the pool of insured agents, then it seems that this would increase the chance of observing average risk decreasing over the span of the contract, as the agents with higher risk aversion choose more effort in each period and hence are more likely to survive longer. This effect still has to offset each individual's decreasing effort choices over the span of the contract. The model in the following section goes around this ambiguity by comparing multiple items owned by the same person with different amounts of time left in their contracts. This allows me to factor out the effects that risk type and preferences have on survivability.

As I alluded to earlier in the introduction, this section's incentive structure is not unique to insurance contracts. Many labor market contracts, especially in service and construction work, have a structure where a contract lasts for a set amount of time, payments happen periodically within that time frame and poor outcomes may lead to an early termination of the contract. For instance, payments for construction work occur over the span of the project and the contractor or an individual worker can be fired any time during the contract because of shoddy workmanship. It is easy to modify the model above to show that these contracts will potentially lead to similarly decreasing incentives for effort over the span of the contract. As long as the agent is relatively patient compared to the value of the periodic payments, terminating the contract early will lead to a higher loss in expected value to the agent than terminating it late, because more installments are lost with an early termination. This idea is fairly similar with the one given in Gibbbons and Murphy (1992).

2.3 Dynamic model with two insurable items

Next, I will combine the two models above to get an identification strategy that is most demanding on the data but at the same time is the most robust to alternative explanations such as seasonality. Each agent again owns two identical insurable objects. Either of these can be insured at any point in time with a separate contract for l units of time. In the case of an accident, the insurance offers a fixed payout of P after which the contract on that item ends, no matter how much time was left in the contract. Just as in the model above, I assume that new contracts purchased while an old one is still running only renew the length of the contract but do not affect the coverage in any other way. The main result of this section is that if a person owns two items that are both insured but one of them has less time left in its contract than the other, then the item with less time in its contract is on average more risky. Here, the fact that I observe multiple items in different states from the same person allows me to factor out all heterogeneity and seasonality - even when these affect the agent's effort choices.

Let $x_j^s = (x_{j1}^s, x_{j2}^s) \in \mathbb{R}_+^2$ be the vector of effort choices of the agent in period s. Similarly, $I_j^s = (I_{1j}^s, I_{2j}^s) \in \{0, 1, \ldots, l\}^2$ is the time left in the insurance contracts in period s after new insurance purchases have been made. Denote by $D_j^s = (D_{1j}^s, D_{2j}^s) \in \{0, 1\}^2$ the indicators of accidents for both items in period s. Now, in the beginning of period t the agent recalls the history $h_j^t = \{x_j^s, I_j^s, D_j^s\}_{s=0}^{t-1}$. The timing of events in each period is the same as in the previous section with the modification that whenever the agent is making a choice, she is simultaneously choosing effort or insurance levels for both of the two items. Similarly, she learns whether each item had an accident simultaneously. Define $\mathcal{H}$ as the set of all histories. I assume that conditional on the history of effort choices, D^s_{ij} is independent of D^s_{-ij} and of D^r_{kj} for all $r < s$ and all k. I also assume that $\mathbb{P}(D^t_{ij} = 1 \mid x^t_{1j}, x^t_{2j}, \varepsilon^t_j, t) = p(x^t_{ij}, \varepsilon^t_j, t)$ and that the agent is fully aware of this mapping. In particular, she is able to perfectly forecast all possible future accident probabilities. I continue assuming that $p(\cdot, \varepsilon^t_j, t)$ is continuously twice differentiable, decreasing and strictly convex and that $(x_1, x_2) \mapsto p(x_1, \varepsilon^t_j, t)p(x_2, \varepsilon^t_j, t)$ is also strictly convex for all ε^t_j and t.

I allow the agent's risk/taste type ε^t_j and her periodic monetary endowment ω^t_j to vary arbitrarily over time. For simplicity, I assume that the agent knows the full paths of these processes. Notice that this assumption includes the case where the agent earns a deterministic stream of income and is potentially committed to a predetermined plan for saving and borrowing decisions. However, I am not allowing the wealth distribution to depend on the history of accidents, exercised effort or insurance choices. Relaxing this assumption is unlikely to completely change the results in this section. Nevertheless, introducing, for example, rational saving decisions into the model can dampen some of the dynamic results provided below through consumption smoothing.

The agent's problem is then to choose policy functions $x_j : \mathcal{H} \to \mathbb{R}^2_+$ and $I_j : \mathcal{H} \to \{0, 1, 2, \ldots, l\}^2$ to maximize the expectation

$$\mathbb{E}\Bigg[\sum_{t=0}^{\infty} u\left(\omega^t_j + \sum_{i=1}^{2} D^{t+1}_{ij}(\mathbb{1}\{I_{ij}(h^t_j) > 0\}P - L) - \mathbb{1}\{I_{ij}(h^t_j) = l\}C, \varepsilon^t_j\right)\delta^t_j \\ -E(x_{1j}(h^t_j), x_{2j}(h^t_j), \varepsilon_j)\delta^t_j \mid (\varepsilon^t_j)_{t\in\mathbb{N}}, (\omega^t_j)_{t\in\mathbb{N}}, x_j, I_j\Bigg], \tag{2.7}$$

where $\delta_j \in (0, 1)$ is the discount factor and the maximization is subject to

$$I_{ij}(h^t_j) \in \left\{l, \max\left\{0, (1 - D^{t-1}_{ij})(I_{ij}(h^{t-1}_j) - 1)\right\}\right\}.$$

In other words, the choice of insurance for each period must be made from the set $\{0, l\}$, if the object was in an accident in the previous period. Otherwise, one can either just keep the existing contract (with one unit less time left) or alternatively buy a new contract with l units of time. I assume that u and E are just as in Section 2.1, except that I need E to be fully symmetric:

Assumption 2 (Symmetric effort cost)**.** *The cost function satisfies $E(x, y) = E(y, x)$ for all $x, y \in \mathbb{R}_+$.*

Results from Bertsekas (1995) can again be applied to show that the problem has a recursive structure with respect to the state (t, I'_{j1}, I'_{j2}) where I'_{jk} is the time left in the insurance contract $k \in \{1, 2\}$ owned by agent j when entering period t before any choices are made. I let the accident probabilities, monetary endowments and risk/taste type depend on the time period. I therefore have to add it as a state.

The next proposition argues that the idea of the changing option value embedded in the contract carries over from the previous model to choices made by a single person on how to distribute effort over two insured items with different amounts of time left in their contracts. Specifically, a person who owns two insured items is going to take better care of the item that has more time left in its contract.

Proposition 4. *With dynamic moral hazard, the following inequalities hold for all $t \in \mathbb{N}$:*

1. *If $0 < I_1' < I_2'$, then*

$$\begin{aligned} & \mathbb{E}[p(x_{1j}(t, I_1', I_2'), \varepsilon_j^t, t) \mid I_1', I_2', I_{ij}(t, I_1', I_2') = I_i', i = 1, 2] \\ \geq \ & \mathbb{E}[p(x_{2j}(t, I_1', I_2'), \varepsilon_j^t, t) \mid I_1', I_2', I_{ij}(t, I_1', I_2') = I_i', i = 1, 2]. \end{aligned} \tag{2.8}$$

2. *If $0 < I_1'$, then*

$$\begin{aligned} & \mathbb{E}[p(x_{1j}(t, I_1', I_2'), \varepsilon_j^t, t) \mid I_1', I_{1j}(t, I_1', I_2') = I_1', I_{2j} = l] \\ \geq \ & \mathbb{E}[p(x_{2j}(t, I_1', I_2'), \varepsilon_j^t, t) \mid I_1', I_2', I_{1j}(t, I_1', I_2') = I_1', I_{2j} = l]. \end{aligned} \tag{2.9}$$

3. *Both inequalities are strict, if an agent never re-insures an item that still has time left in its contract.*

The third point does not necessarily always hold, because buying insurance "too early" can serve as a way to smooth out consumption. For example, a very high accident risk tomorrow implies lower expected utility. Buying insurance today lets the agent shift some of the higher welfare today forward in time and smooth expected consumption over periods. However, if the agent is allowed to make rational borrowing and saving decisions, smoothing with front loading insurance decisions is unlikely to be optimal, and hence the third point is likely to hold in most applications. Its validity can also be easily checked from the data.

The proof of Proposition 4 is somewhat more involved than the proof of the result in the previous section but the intuition behind it is the same: More time left in the contract of one of the items implies that the marginal return for effort is higher for that item, as an accident will result not only in losing the item but also in the termination of the longer remaining insurance coverage.

In the absence of moral hazard the two sides of the inequality should clearly be equal, because the two items share the agent's risk type. Since I am comparing the risk of two items owned by the same person in the same time period, heterogeneity of risk types combined with survivorship is not dampening the effect like it was in the previous section. Hence, even when both phenomena are present, rationality of the agents would imply that the inequalities above should hold in the population. Notice also that the model allows for arbitrary seasonal variation in risk across individuals, as well as variation in wealth and risk preferences. This flexibility is a result of the ability to compare the risk across two items owned by the same person at any given time.

3 Data

I will start this section by introducing my data source and describing its insurance market. After that I will present some descriptive statistics of my data.

3.1 EVE Online and its spaceship insurance

My data source, EVE Online, is a science fiction multiplayer game played by over 300,000 players worldwide in a single virtual world. The game is run by an Iceland based game company called CCP and is part of the massively multiplayer online game industry that was estimated to generate about $10 billion in revenue in 2014, which was about 15% of the total video game market value. During the data collection period, to play EVE a player had to pay a subscription fee between $14 and $20 per month depending on the method of payment and how many months were purchased at once.

A player in EVE is a spaceship pilot in a complex society consisting of all of the game's subscribers. The game itself does not give players any clear goals that would lead to somehow completing the game. Hence, players are left to their own devices at deciding what activities they find worth doing. For many players, important measures of how well they are doing in the game are such things as economic wealth, their status in the game's social community, the fraction of the virtual universe that their group of players, called a corporation, controls or how well they are doing in spaceship combat against other players.

All of these activities require some economic resources and hence some involvement in the game's player-run economy. The economy consists of over 5,000 spatially distinct markets where players trade thousands of intermediate and final products manufactured by players who are often highly specialized in certain manufacturing activities. Common tasks in the economy involve mining asteroids for raw materials, manufacturing goods for markets, freighting products from where they are cheap to where they are expensive, running errands for richer players or non-player characters, or highway robbery, i.e. threatening a mining vessel or a freighter with a war-ship in the hope that the owner will rather pay some protection money than see hours of his or her work blown to bits. The game has even seen multiple player-run banks as well as some bank runs that resembled their real-life counterparts.

Entering many of the areas in the game runs a high risk of getting one's ship attacked by space pirates or members of a hostile corporation who have a stake on the area and its economic wealth. Because large fractions of players' wealth consists often of the ships they own, the threat of violence creates a demand for ship insurance. As much of the time in the game goes into gathering means to purchase space-ships and to tuning them with parts and weapons sold in the market, losing the ship in combat may mean losing many days of work. The value of a given ship can even be measured in real currency using the unofficial floating exchange rate between the in-game currency ISK and euros. At the time of the data collection a player could buy 6 months of play time from the game company for 105 euros. The purchased play time could then be turned into six single month play-time extensions that could be sold in the in-game markets for in-game currency. At the time of data collection, a play-time extension cost around 480,000,000 ISK. Hence, a euro bought about 27,400,000 ISK. A Drake, one of the most popular combat ships that will appear multiple times later in this article, cost approximately 55,000,000 ISK. Using the implied exchange rate above, this is equivalent to approximately 2 euros of real money. The most expensive ships, Titans, cost around 4,950,000,000 ISK or around 180 euros.

EVE's spaceship insurance is provided by the game company and it offers seven levels of insurance for every ship type. The contracts differ by the percentage of the full payout covered by the insurance ranging from the free 40% insurance to the highest 100% insurance. To determine the full payout, the developers of the game periodically calculate the market prices of the raw materials that go into producing each ship. This cost is usually considerably less than the market value of the ship, because the formula does not take into account the profits of the ship builder and all the market intermediaries nor the materials wasted in the production process. The different insurance contracts are then based solely on this cost and what is called the ship type group. Ship type groups are categories of ships that broadly share the same purpose (mining barge, freighter, different combat roles, etc.) and size. Every ship type group has a different percentage of the material costs covered by the full payout. For instance, for the Battleship group all of the material costs are covered while for the Heavy Interdictors only 70% of the material costs get covered. Since the payout of the insurance is based on material costs and not on the market price, even the 100% coverage contract for Battleships entails a de facto deductible. Also the premiums of different insurance contracts are based on the ship type group level full payout. Adding 10 percentage points more coverage increases the premium by 5 percentage points of the full payout. Thus starting from the free 40% insurance, the contract that yields the maximum payout costs 30% of the full payout. All of the contracts cover the ship for 12 weeks. If a ship is destroyed before 12 weeks have passed, the insured agent is immediately awarded the predetermined payout and the contract ends. Hence the structure of the contracts corresponds to the structure in my model.

Although the goal of the insurance system is primarily to make the game more entertaining and not to maximize profits, it still adheres to both of the supply side predictions described by Chiappori and Salanié (2000, p. 58) which should prevail in an insurance market with adverse selection: "(1) observationally equivalent agents are faced with menus of contracts, among which they are free to choose and (2) within the menu, contracts with more comprehensive coverage are sold at a higher premium". Furthermore, Chiappori and Salanié point out that complex non-linear pricing of insurance contracts may cause a spurious correlation between coverage and accident probabilities, if the characteristics on which the pricing is based are not properly controlled for. In EVE Online, the pricing categories are extremely simple (ship types) and hence easily controlled in an econometric model. In this sense the setup has the transparency, clarity and simplicity of a laboratory experiment. More complex contracts may also be hard for the prospective insurees to comprehend and, as was demonstrated by Fang, Keane and Silverman (2008), cognitive ability may become a more important factor than risk type in deciding how much insurance people will buy. This is unlikely to happen with EVE's very simple contracts.

Another benefit from using log data from an online game is that it is high frequency data that is extremely detailed, extremely reliable and the potential number of observations is very large.[12] Furthermore, because the payout is automatic and no player action

[12]In particular, the only things not collected automatically by the servers are the player demographics, age and gender, which the players fill in when they subscribe to the game. These are not visible to other players and hence there should be little incentives to misreport.

is required for reporting accidents, I observe all accidents and not only the ones that get reported. As is commonly pointed out, people usually do not report claims that are less than or barely exceed their deductible. Thus tests based on claims as a proxy for risk and the deductible as a measure of coverage may often yield spurious correlation between coverage and the probability of an accident, as people with higher deductibles simply report fewer claims even though their number of accidents may be higher (see e.g. Cohen and Siegelman 2010). Since all accidents are automatically reported, I am able to circumvent this problem. I also observe all of the objects that could have been insured, not only those that end up in an insurance company's books. Consequently, my data is able to capture the selection into the insurance pool, not only selection effects between different levels of coverage offered by a given company. Finally, as I noted in the introduction, compared to experimental studies, my empirical results are considerably less susceptible to psychological reactivity effects, such as the Hawthorne effect, as the players do not know that they are being studied.

Some people have voiced a concern over the generalizability of even qualitative empirical results from a computer game. Especially, people have worried about whether the players' actions are driven by the same incentives as in the real world, because the players in a game might be just "playing" and hence might act irrationally even when risking the loss of property that has real monetary value. This may be a potential explanation for when empirical results show players acting irrationally or not adhering to the proposed theory. However, whenever players appear to be acting rationally, the results should be generalizable to any populations acting at least as rationally as the player population when faced with similar incentives.

The dog-eat-dog environment of EVE strongly rewards rational actions and harshly punishes blatant irrationality, as those who are the most able to utilize all available information are much more likely to gather most wealth, build the best space-ships and stations and hence get the upper hand in the game's fierce competition. For instance, the traders in the EVE markets often use special automated Excel spreadsheets with macros that take in large amounts of market data and use it to suggest optimal trade routes and goods that yield the largest profits. There are also multiple sites on the internet where speculators can follow the development of market prices of different commodities in all spatial markets.[13] Hence, at least on the surface, it seems that EVE players respond to the incentives of the virtual economy very much like people in the real economy.

3.2 Some descriptive statistics

My data consists of a random sample of 60,022 EVE players. For these players I have the basic, self-reported real life demographics, gender and age. One player may own multiple characters or avatars which are their embodiments in the game. I have a comprehensive set of in-game demographic variables for each of the 99,730 characters owned by my 60,022 players, including attributes such as the character gender, monetary wealth, number of ships owned, the market value of these ships, the distribution of skill

[13] Good examples are `http://www.eve-markets.net` and `http://eve-central.com`.

points[14] and the total number of times the character lost a space ship. Last, I have a record of all the ships that were owned by my players and in working condition some time between May 22, 2011 and July 18, 2011 (meaning 1,873,343 ships). However, due to server capacity constraints the game company does not save event logs for the cheapest of ships. Consequently, my sample contains full records that detail how, where and when a ship got destroyed for 759,865 ships excluding all of the ships meant primarily for the new players. The total monetary value of the observed ships is about 75 times as large as the total monetary value of the excluded ships.[15] Hence, in terms of monetary value, my sample contains arguably the more interesting half of the stock of ships. Furthermore, the cheaper ships are mostly used by inexperienced, new players and hence excluding them is likely to yield a sample of players that better understand both the insurance system and the risks that come with their actions. I observe these ships for nearly three months or until they get destroyed starting from May 22, 2011 with the last observations being on July 18, 2011. My data on the insurance spells for these ships goes even further. The insurance contracts are sold in two month spells and the first contract start date in my data is February 25, 2006.

Table 1: Descriptive statistics about players and their ship ownership

	Variable	Mean	Median	Std. Dev.
Player attributes	Sex (1=Male)	0.96	1	0.20
	Age	31.52	30	9.31
	Hrs. played	1737	880	2294
	Hrs. played July 1-18	30.60	10.46	45.63
	Characters	1.66	2	0.45
Ship ownership	Active ships	31.21	20	40.27
	Insured ships	1.54	0	3.16

N=60,022

Tables 1 and 2 report some descriptive statistics of the players in the sample. Table 1 contains normal player level statistics. However, as players may have multiple charac-

[14] Players use skill points to develop skills which make them more proficient in different areas of the game.

[15] The value of the stock of cheap, excluded ships in my sample, using the implied exchange rate above, was about €32,000. The total value of the included ships was about €2,400,000.

Table 2: Descriptive statistics about characters and their ships

	Variable	Mean	Median	Std. Dev.
Character attributes	Sex (1=Male)	0.68	1	0.47
	Monetary wealth (million ISK)	490	13	3300
	Lifetime income (billion ISK)	6.8	0.41	60
	Skill points (million)	17	5.0	24
	Security status	1.3	0.29	2.3
	Lowest sec. system visited	0.09	0.35	0.58
Ship variables	Owned ships	36.79	11	390.37
	Value of ships (billion ISK)	1.6	0.11	5.1
	Active ships	19	9	33
	Ships lost	1.2	0	10
	Others' ships destroyed	0.82	0	7.4

N=99730

ters, some statistics such as the fraction of male characters in the population are more naturally reported at the character level. These variables are collected in Table 2.

Some of the variables may require a small explanation: The Hrs. played and the Hrs. played in June 2011 variables both report the total number of hours a player stayed logged on to the game in total and in July 2011, respectively.

The Active ships and Insured ships show the mean, median and standard deviation for the number of ships owned by a player or a character. I call a ship active, if it is not in a storage container. For any market-related purposes such as freighting or selling the ship, it needs to be packaged into a container before it can be put on sale or shipped on a freighter. Packaged ships cannot be used before unpacking. Therefore, an unpacked ship is a good indicator of the ship being currently in use and not just stored to be sold in the market. In Table 2 Ships lost and Others' ships destroyed show the same statistics for how many ships a player lost and how many other players' ships she destroyed in combat during the data collection period.

Skill points are something that players accumulate automatically over time in the game and what the players can then invest to develop the efficiency of their characters. Skill points can be used for a wide variety of purposes ranging from reducing waste in manufacturing and the costs incurred from selling objects to higher combat efficiency. Skill points accumulate automatically whenever the player has paid her monthly subscription fee. Hence, the total number of skill points can be viewed as the age of the character. Furthermore, it reflects the level of proficiency of the character.

Last, there are the Security status and Lowest sec. system visited variables. The first one is lower the more crimes the player has committed in the space policed by the in-game police, Concord. The lower this rating, the more there are star systems where the galactic police will automatically attack the player. Furthermore, players with security status -5 or less are considered outlaws and can be attacked by other players without the Concord intervening or having the attackers security status lowered. The Lowest sec. system visited is the security rating of the star system with the lowest security rating visited by the player during the data collection period. The lower this number, the more lawless and perilous the system is and the easier it is to get away with crime in that system. Hence, visits to systems with a low security rating are indicative of high risk taking.

A striking feature of these tables is the sex distribution: only 4% of the whole player population are female. This is possibly even more extreme than in many other online role-playing games. As my aim is to showcase my methods and test the existence of moral hazard, not so much to give accurate point estimates with external validity, the skewed gender distribution is not a big problem. Another noteworthy fact is that only 68% of the characters are male. This implies that a large fraction of the male players chose to play with a female character. At the time of the data collection, the character gender in the game affected only the small picture of the avatar visible to other players.

The mean and median player ages are slightly lower than the ones found in many real world countries.[16] However, they are higher than what one might expect from a computer game. This reflects the fact that EVE Online is a relatively complex game directed to more mature audiences.

Many of the remaining variables in both of the tables show a common feature where the mean is much higher than the median. As most of these variables are measures of wealth, success and time invested in the game, this is compatible with the commonly shared preconception that EVE's laissez-faire, dog-eat-dog world is a very hard place for newcomers and success requires a lot of playing. At the same time, a small elite controls large quantities of capital with which they are able to generate large income flows. These players may daily sell dozens of ships that might cost over a hundred times a newcomer's monthly in-game income. In general, taking part in the more economically

[16]For instance, the median age in the USA in 2010 was 35.3 years (Howden and Meyer 2011)

lucrative activities in EVE demands considerable amounts of wealth as these activities often happen in the lawless 0.0 space where operations usually require well equipped, expensive and high quality ships. Furthermore, the risk of losing one's ship in combat is relatively high in 0.0 and consequently players operating there must have enough liquid property to be able to overcome occasional losses.[17] This picture is reflected by the fact that a median player had not lost any ships or destroyed any opposing players' ships, and on average both of these numbers were close to one. However, a player one standard deviation away from the mean lost over 10 ships and destroyed over 7 other players' ships.

The descriptive statistics of both of the security status variables reinforce this interpretation. Security status is accumulated by destroying pirates in low security space and running errands for computerized agents which often require destroying pirates or traveling to low security space. Wealthier players in their more expensive ships have a much better chance to successfully pursue both of these activities. Similarly, due to their superior equipment wealthier players have a much higher chance to survive trips to low security space, and replacing any ships lost there does not imply personal bankruptcy. Thus it is not surprising to find that the mean security status is higher than the median and that the median character stayed in a much safer space than the mean.

The insurance status of the ships in my data is fairly polarized between the two extremes. This can be seen in Table 3. About 91% of the ships are insured only with the free 0.4 insurance level and out of the people who paid for their insurance (categories 0.5-1.0) over 85% have the highest insurance status. Consequently, I will reclassify all ships in categories 0.5-1.0 as insured and any ship in the free 0.4 category as uninsured. All results presented in this paper are robust to different potential classifications.

Table 3: Distribution of assembled ships by insurance level

Insurance level	Frequency	Percent
0.4	693,459	91.26
0.5	4,128	0.54
0.6	1,375	0.18
0.7	1,256	0.17
0.8	1,552	0.20
0.9	1,632	0.21
1.0	56,463	7.43
Total	759,865	100

Since two out of the three of my identification strategies use variation in risk between ships of the same type, owned by the same person but have different amounts of

[17] Although there are some player-run banks the financial markets in EVE are arguably far from perfect. Hence, most players are heavily credit constrained and cannot borrow the assets required to operate in 0.0.

insurance coverage, it is worthwhile to document that such individuals exist. In total, there are 11,272 players who own at least 2 ships of the same type out of which at least one is uninsured and at least one is insured. Table 4 in the appendix breaks this down by ship type over the ten most popular ship types.

4 Econometric model and results

This section is divided into three subsections where each one of the subsections discusses the results from applying each one of the three identification strategies to the EVE Online data. Each subsection first describes the econometric model used and then displays and discusses the estimation results.

4.1 Difference in risk between insured and uninsured ships owned by the same person

My first identification strategy measures the average difference in risk between differently insured items owned by the same agent. As was formally argued in Section 2.1, controlling for the agent's identity controls for agent-specific adverse selection. Any remaining positive correlation between the insurance coverage and accident risk between items owned by the same agent is caused by different choices of precautionary effort. I will measure this remaining correlation using a highly flexible semiparametric econometric model.

I assume that the probability that ship j of type k owned by person i survives without an accident for more than t units of time is given by

$$S_{ji}(t;k) := \exp\left[-\int_0^t \lambda_{ik}(s)\exp(\beta Z_{ji}(s))\,\mathrm{d}s\right],$$

where $Z_{ji}(s)$ equals one if ship j is insured at time s, and zero otherwise. The function $\lambda_{ik}(\cdot)$ is a person-ship-type specific baseline hazard rate which can vary arbitrarily over time. The parameter β measures the effect that insurance status has on accident risk because of the agent's effort choices. This is a standard Cox proportional hazards model (Cox 1972 and 1975) with time varying covariates stratified at the person-ship type level. The parameter β can be consistently estimated using partial maximum likelihood without estimating the potentially infinite dimensional λ_{ik} (see Ridder and Tunali, 1999 or Kalbfleisch and Prentice, 2002).

Proposition 2 implies that if there is moral hazard, β will be positive. If both insured and uninsured ships of the same type, owned by the same person have the same hazard rate, it will still be captured by $\lambda_{ik}(t)$. This is true even if this common hazard rate varies over time. Consequently, if there is no moral hazard, β will be zero.

My identification strategy requires that I compare multiple *simultaneous* effort choices made by the same agent. In other words, I need to compare items owned by the same person with different insurance statuses at the same *calendar* time. Hence the origin of the time axis in the econometric model should not vary between the items owned by the same person. My empirical models will therefore use the same calendar date as their origin.

The model allows me to control for multiple sources of possible dynamic variation in the agent's risk. First, it allows for times when the person was logged out of the game and hence had a zero baseline hazard rate. The same holds true for episodes of high intensity gaming such as weekends or private vacation days. Figure 3 presents the frequency of ship destruction events as a function of time. One bar of the histogram represents a day during my observation period. As can be seen from the figure, there is a strong weekly cycle in destruction events. Furthermore, there are three days with abnormally few destruction events that are likely due to server down time. My model accommodates not only these features but any heterogeneity in gaming habits. It also allows the player's baseline hazard rate to change when an agent becomes more experienced with a specific ship type or just generally learns more about the environment.[18] All of these sources of variation have clear analogues in most real-life insurance settings. The flexibility of the model is therefore valuable also outside the game.

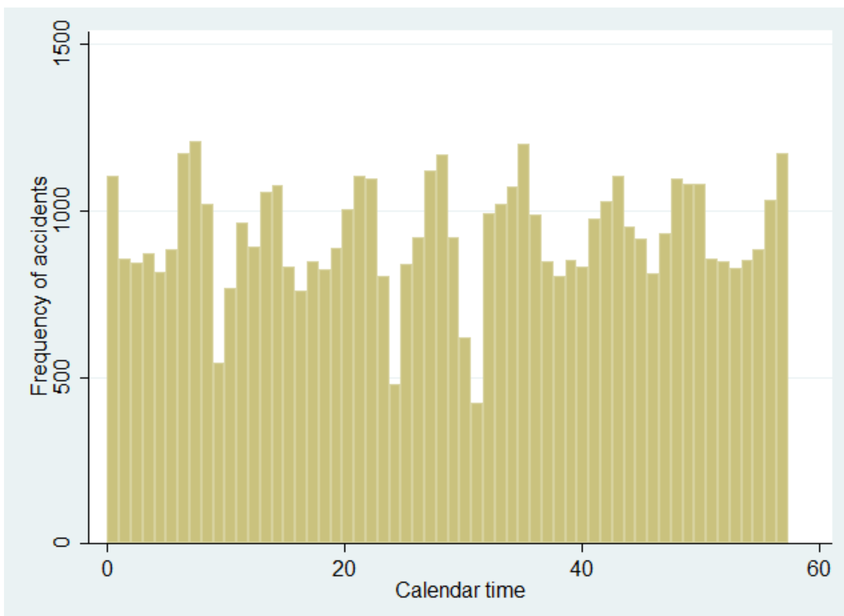

Figure 3: The number of destruction events over the observation period.

The model does not account for variation in risk over the life cycle of an item. The most prominent reason why this type of variation might matter in most settings is depreciation. If old or depreciated items are both riskier and insured more often, the model above would confound moral hazard with item-specific adverse selection. Many real world insurance contracts account for depreciation. In these settings stratifying the model at the contract and agent level will alleviate the problem.

In EVE, a ship can get partially damaged in combat. The more damaged the ship, the easier it will get destroyed. Damage can be repaired for a low cost but the insurance does not cover this cost. It is also likely that older ships are on average more

[18]I discuss the effect of experience on moral hazard in Section 5.

damaged than newer ships. This depreciation may generate a potential source of a bias, if damaged old ships are insured more than new ships. However, in EVE depreciation is unlikely to be a big factor, since players can fit their ships with very cheap repair units that repair any damage for free.

In some settings, depreciation can be approximated by age. As long as the effect of the age of an item on the hazard rate is uniform across individuals I can control for it by appending the model with a set of age fixed effects. Formally, the survival probability is then given by

$$S_{ji}(t;k) := \exp\left[-\int_0^t \lambda_{ik}(s)\exp(\beta Z_{ji}(s) + \sum_{h=1}^{H} \delta_{A_h}\mathbb{1}\{a_j(s) \in A_h\})\,\mathrm{d}s\right],$$

where $a_j(s)$ is the age of ship j at time s, the set $A_h \in \{A_1, \ldots, A_H\}$ is an arbitrary partition of possible ship ages and δ_{A_h} is the effect that belonging to age group A_h has on the ship's hazard rate.

Table 4 presents estimation results from different specifications of the model. Rather than reporting the estimate for β which is harder to interpret, it gives the estimates for $\exp(\beta)$. This corresponds to the ratio of the hazard rates between an insured and an uninsured observation from the same strata with otherwise identical covariates. Formally, for example in the model with age dummies, for any t,

$$\exp(\beta) = \frac{\lambda_{ik}(t)\exp(\beta + \sum_{h=1}^{H} \delta_{A_h}\mathbb{1}\{a_j(s) \in A_h\})}{\lambda_{ik}(t)\exp(\sum_{h=1}^{H} \delta_{A_h}\mathbb{1}\{a_j(s) \in A_h\})}.$$

Consequently, $\exp(\beta)$ measures the relative change in the hazard rate due to the item having insurance. Below its estimates I also report the 95% confidence interval for this hazard ratio. In the model with ship age controls, I add dummies for the first 7 days of a given ship, as well as a dummy for ships that are older than 180 days.[19]

The non-stratified estimates measure the compound effect of moral hazard and adverse selection given in (2.3). These estimates allow for seasonality in the common baseline hazard rate shared by the population. The first row of the table shows that this combined effect implies that insured ships are on average eight times as likely to get destroyed each day compared to their uninsured counterparts. Estimates from different markets are not highly comparable due to differences in the fraction of the risk carried by the agent under different contracts, opportunity cost of effort and losses from poor outcomes. Nevertheless, eightfold *daily* accident probability of insured items compared to uninsured items is much higher than the estimates found in the earlier literature. For example, Dobbie and Skiba (2013) find that agents who choose a *two-week* payday loan that is \$50 larger have a 28-44% higher default risk. For car insurance, Cohen (2005) shows that having a low deductible is associated with about a 4% higher probability of having at least one accident during the *year* in which the policyholder took out the insurance. The high combined effect of moral hazard and adverse selection in EVE's

[19]This is approximately the median age of a ship in my sample both at the end of the data collection period and at the time of destruction for destroyed ships. I experimented with other partitions of the age distribution without considerable effect on the results.

Table 4: Estimates of the insurance parameter from stratified Cox regressions

	Strata	Hazard ratio No age dummies	Hazard ratio Age dummies
Pooled Sample	Ship type	8.02	4.38
N=759,865		(7.68, 8.37)	(4.23, 4.54)
	Ship type,	5.83	6.16
	User ID	(5.32, 6.40)	(5.64, 6.73)
	Ship type,	5.72	6.09
	Character ID	(5.22, 6.27)	(5.58, 6.66)
Drakes	None	6.56	4.04
N=34,039		(6.02, 7.14)	(3.73, 4.37)
	User ID	6.44	6.19
		(5.14, 8.08)	(5.06, 7.56)
	Character ID	6.38	6.13
		(5.09, 7.80)	(5.02, 7.49)
Badger IIs	None	7.04	3.39
N=22,248		(5.15, 9.63)	(2.51, 4.58)
	User ID	4.28	9.25
		(1.42, 12.89)	(1.58, 54.14)
	Character ID	4.28	10.00
		(1.42, 12.89)	(1.55, 64.43)

Note: 95% confidence intervals in parenthesis. Implicit standard errors clustered at the user level. Time recorded in days.

spaceship insurance is not that surprising given the nature of the game and its insurance contracts. One of the design goals of the game's insurance contracts is to make highly risky modes of gameplay, such as spaceship combat and exploring the high risk space, more appealing to the players.

When baseline hazard rates are allowed to vary from player to player, I get the effect of moral hazard on hazard rates. These estimates are given on the second row for the pooled sample. An insured ship is almost 6 times more likely to get destroyed each day compared to an uninsured, identical ship owned by the same individual. The result indicates a high level of moral hazard in EVE's insurance market. The third row of the table allows the baseline hazard rate to vary also between different characters owned by the same user. In principle, it could be possible that the player played differently with different characters. Some of this effect could be incorrectly captured in my agent-level estimate of moral hazard. However, the effect of moral hazard from this regression is commensurate with the one that stratifies only at the user level. This indicates that players' risk taking behavior does not vary considerably between characters.

The level of moral hazard is likely to vary by the purpose of the insurable item. When an agent is driving a sports car she may behave differently than when driving a truck. The rows 4-9 of the table run the same model but restrict the pool of observations only to Drakes or Badger Mark IIs. Drake is the most popular ship in my sample and it is mostly used for fighting other players or hostile ships controlled by the game server. Badger Mark II, on the other hand, is the most popular mainly non-combat ship.[20] Looking at these separate estimates I can get some indication of how the purpose of the ship affects the effect of moral hazard. From rows 5 and 6 one can see that the effect of moral hazard on the risk of Drakes is somewhat higher than the average effect in the population of all ships. However, the Badger's point estimate is somewhat smaller (rows 8-9). This seems natural, since the peaceful purpose of the Badger implies that the benefits from taking risks with it are likely to be smaller. However, the standard errors for these estimates are so high that the population average is inside the 95% confidence interval for both ships.

The last column shows what happens to the estimates if I add the ship age dummies into the model. In general, if there is persistent heterogeneity in the baseline hazard rates of the items owned by the same person, then an average item with a lower risk type has survived longer and is consequently older than an average item of a higher risk type. Therefore, the item's age would be negatively correlated with its risk type. If there is item-specific adverse selection, adding item age as a control should therefore reduce the estimate of the effect of insurance on hazard rates in the model stratified at the agent level.[21] In this sense, the last column serves as a test of my identifying assumption that ships of the same type owned by the same person have the same risk type. As can be seen from Table 4, adding these dummies has a negligible effect on the estimates stratifying at the user or character level. This finding supports my claim that ships of the same type are fairly identical in EVE. Adding the time dummies for Badgers doubles the estimate on the effect of moral hazard on insured ships' hazard rates. However, the increase in the estimate can be completely due to the increased inaccuracy in the estimates. Furthermore, as argued above, in case of ship-specific adverse selection, controlling for ship age should decrease, not increase, the estimate for the effect of insurance. The reason for the decreased accuracy is the relatively small number of players (704) that own both an insured and an uninsured Badger Mark II.

Comparing the first and the second row of Table 4, I find that the hazard ratio between insured and uninsured items is at least 40% higher than what would be implied by only moral hazard in a fully homogeneous population. As was pointed out in the end of Section 2.1 this way of measuring adverse selection is not completely accurate. When I add age dummies to the model stratified only at the ship type level, the hazard ratio falls from 8.02 to 4.38. This is even lower than what is obtained in the model stratified also at the user level. Since adding age dummies controls for at least some of the adverse

[20]Badger Mark II is one of the best freight carriers in the game.

[21]If there can be multiple accidents per insurable object, time since the previous accident serves the same purpose.

selection, the relatively low hazard ratio could be indicative of the “substitution effect” outlined in Section 2.1, dominating any remaining adverse selection.

I assess the goodness of fit of the model and especially the appropriateness of the proportional hazards assumption in Section E of the appendix. The model seems to fit the data exceptionally well.

4.2 The evolution of risk over the insurance contract in the cross section

I now turn to the identification strategy outlined in Section 2.2. The goal is to estimate how the time left in the insurance contract affects the item’s hazard rate. If the assumptions from Section 2.2 hold, moral hazard should increase the hazard rate towards the end of the contract. Conversely, heterogeneity of risk types in the insurance pool will have the opposite effect of decreasing the average hazard rate over the span of the contract.

For the purpose of evaluating whether ships are more likely to get destroyed in the beginning or in the end of the insurance contract, I estimate the following Cox proportional hazards model:

$$S_j(t;k) := \exp\left[-\int_0^t \lambda_k(s)\exp\left(\sum_{m=1}^{M}\gamma_m \mathbb{1}\left\{I_j(s)\in\left(\frac{l(m-1)}{M},\frac{lm}{M}\right]\right\}\right)\,\mathrm{d}s\right].$$

Here j identifies the ship, t is the time period, k indicates the ship’s type, $\lambda_k(s)$ is a ship type specific, possibly time varying baseline hazard rate, and γ_m is the effect of having more than $\frac{l(m-1)}{M}$ but no more than $\frac{lm}{M}$ days left in the insurance contract. In other words, I split each insurance contract into M parts of equal length and add dummies for each of the parts. Notice that the econometric model is not specific to the EVE’s insurance setting. It can be used to flexibly estimate the evolution of risk over the span of a contract in a multitude of other insurance settings.

In contrast to the model in the previous section, I do not allow the baseline hazard rate to depend on the person owning the ship. This restriction is applied simply to demonstrate the method from section 2.2. I will relax it in the next section.

Figure 4 plots the results from this estimation when the insurance contract is split into episodes of 4 days. Each estimate is plotted with its 95% confidence interval. Notice that, for a given ship, time in the figure runs from right to left. As can be seen from the figure, the first days of the insurance contract are a period of extremely high risk with hazard rates nearly 50 times as high as what they are for uninsured ships. The hazard ratio falls rapidly during the first few days. Once about 10 days have passed from the purchase of the contract, the hazard rate of an average ship is only about 5 times higher than the hazard rate of an uninsured ship. To get a clearer idea of what happens after 10 days from the purchase of the contract, I plot the same graph excluding the first 8 days of the contract. The result can be seen in Figure 5. The risk is fairly monotonously decreasing also over the remaining 76 days of insurance.

If one is willing to accept the theoretical model from Section 2.2, the results above imply that heterogeneity of risk types and potentially adverse selection is dominating

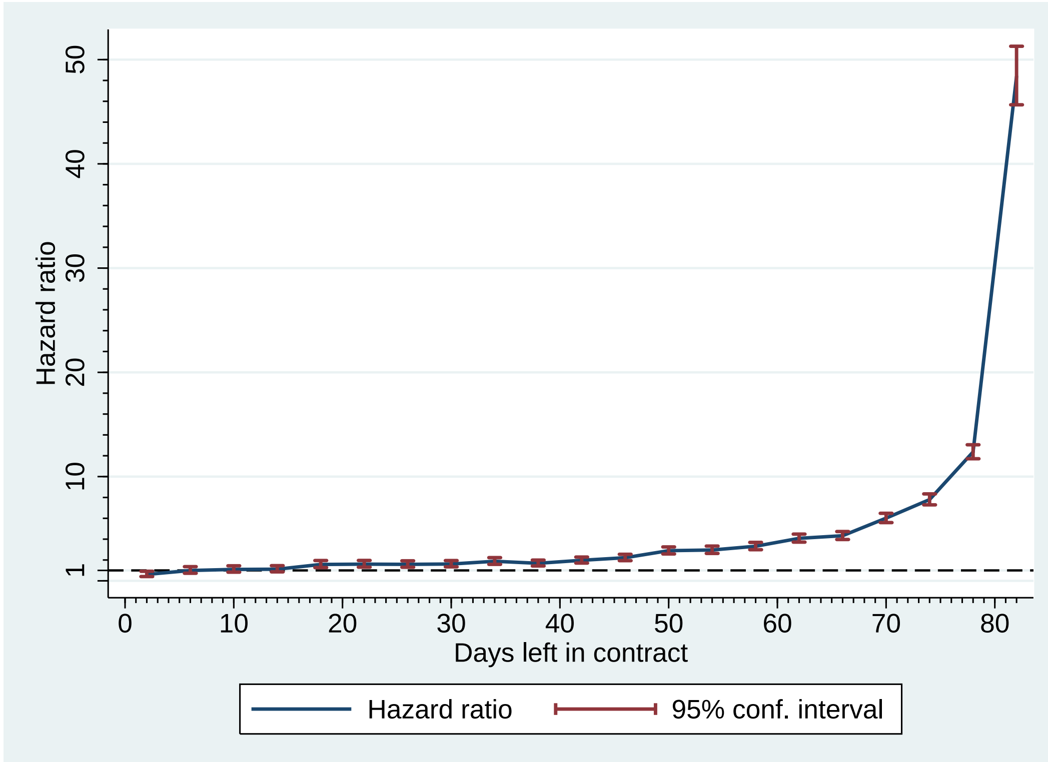

Figure 4: Hazard ratio as a function of time left in the insurance contract. 95% confidence intervals calculated using standard errors clustered at the player level.

moral hazard in EVE's spaceship insurance market. The hazard ratio is not increasing even at the end of the contract. Hence, the dynamic effect that the contract structure has on players' incentives seems relatively weak. However, the next section illustrates that interpreting the decreasing risk over the contract period as adverse selection could be erroneous. It is likely caused mostly by behavioral patterns or transaction costs not present in the theoretical model. I discuss these alternative explanations for the observed empirical pattern in the following section.

The results are robust to splitting the contract into even shorter time intervals as well as looking only at agents who own only one ship of a given kind. A specification where a person owns only a single ship corresponds better, strictly speaking, to my theoretical model in Section 2.2.

4.3 Test for dynamic moral hazard using multiple ships owned by the same agent

As was outlined in Section 2.3, here I will estimate the difference in hazard rates for two items, conditional on them being owned by the same person and them having different amounts of time left in their insurance contracts. To this end, I need to only slightly alter the model from the previous section by allowing the baseline hazard rate to be person specific. In other words, the survival probability for ship j of type k, owned by

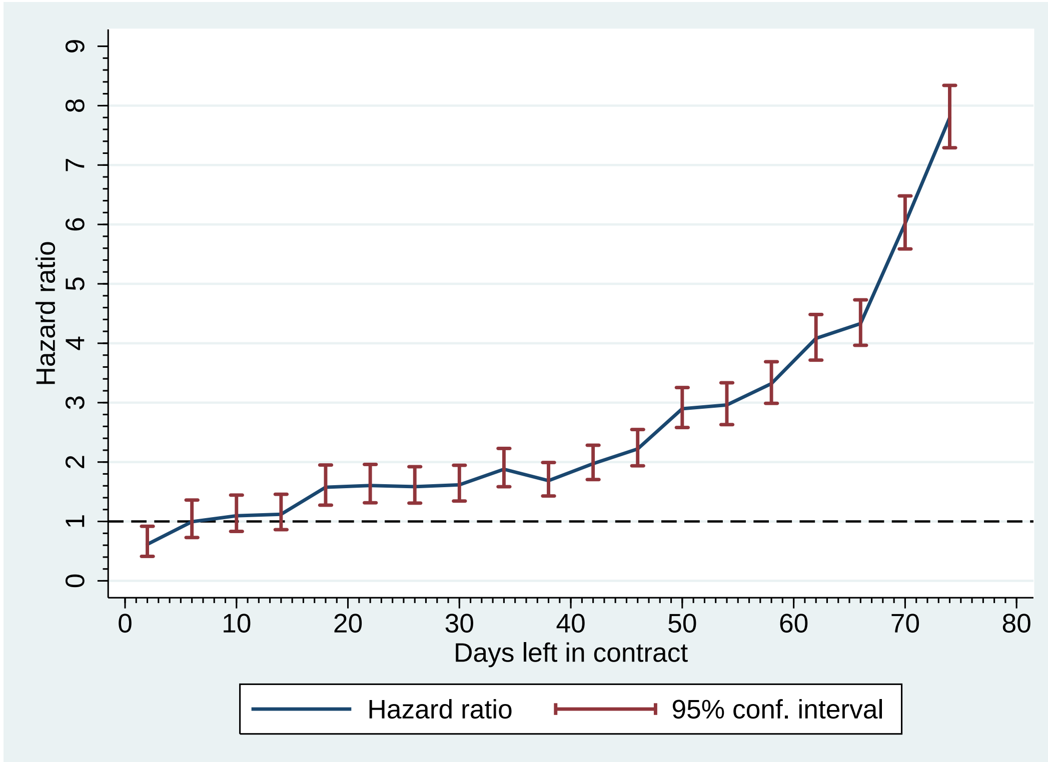

Figure 5: Hazard ratio as a function of time left in the insurance contract excluding first 8 days of the contract. 95% confidence intervals calculated using standard errors clustered at the player level.

person i is given by:

$$S_{ij}(t;k) := \exp\left[-\int_0^t \lambda_{ik}(s)\exp\left(\sum_{m=1}^{M}\gamma_m \mathbb{1}\left\{I_{ij}(s) \in \left(\frac{l(m-1)}{M}, \frac{lm}{M}\right]\right\}\right) \mathrm{d}s\right].$$

The only difference compared to the previous section is that the baseline hazard rate λ_{ik} is allowed to depend arbitrarily on the person in addition to the ship's type. The coefficients γ_m are identified using the variation in risk between ships of the same type owned by the same person that have different amount of time left in their contracts at a given point in time. As long as ships of the same type and owned by the same person share a risk type at each point in time,[22] the same amount of effort from the user should result in equal hazard rates for the two ships. Consequently, any observed differences in the hazard rates should be attributed to differences in precautionary effort. If the players find longer remaining contracts more valuable, I should observe ships with less time left in their contracts being more hazardous than ships closer to the beginning of their contract. In other words, I expect the pictures from the previous section to be reversed.

The results from estimating the model are presented in Figures 6 and 7. The striking first finding is that the decreasing overall hazard ratio from the previous section is

[22] This risk type can be arbitrarily time varying.

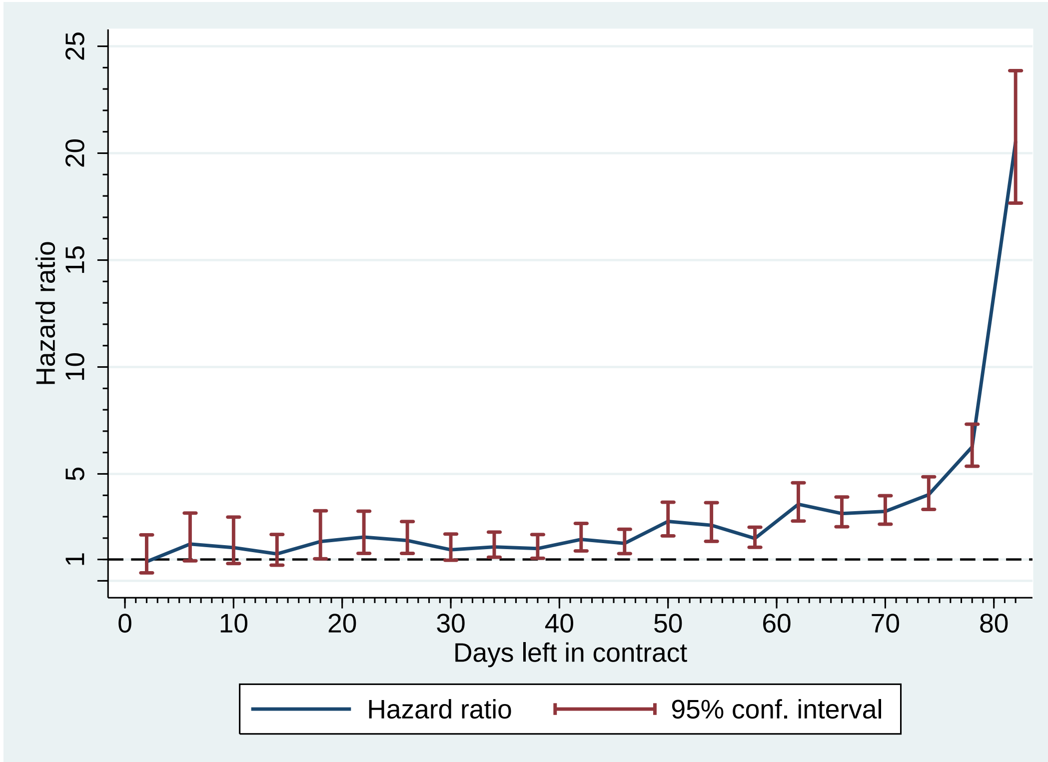

Figure 6: Hazard ratio as a function of time left in the insurance contract. Model includes a user specific baseline hazard rate. 95% confidence intervals calculated using standard errors clustered at the player level.

mostly preserved. Computing the estimates using only insured ships does not change the results.

Rationalizing the decreasing hazard ratio requires an explanation for why an agent might use the recently insured item more than the item that has only a single day left in its contract. The benefit from allocating effort this way has to be larger than the expected value of the remaining long period of future insurance protection enjoyed by the newly insured ship. One possible explanation is that the items are being used in different locations and the environmental risk or the benefit from risk taking in these locations varies over time. For example, consider an agent who owns two cars, one in Chicago and the other in New York. Assume that the car in Chicago has only liability insurance while the vehicle in New York has more comprehensive coverage. If Chicago is hit by a snow storm, using the car there becomes more risky for a short while. The agent has two options: she can either drive the car with more coverage from New York to Chicago or buy more coverage for the car in Chicago. The second choice might be more appealing if the opportunity cost of the long drive is higher than the price of additional coverage. If this is true, we might observe a similar pattern to that presented in Figure 6. Hence, even if the insurable items are otherwise identical, the combination of highly localized use, dynamically and spatially varying risk and high cost of moving items between locations can pose a threat to my dynamic identification strategy.

Ships in EVE are also used in different locations and the reward for taking risky

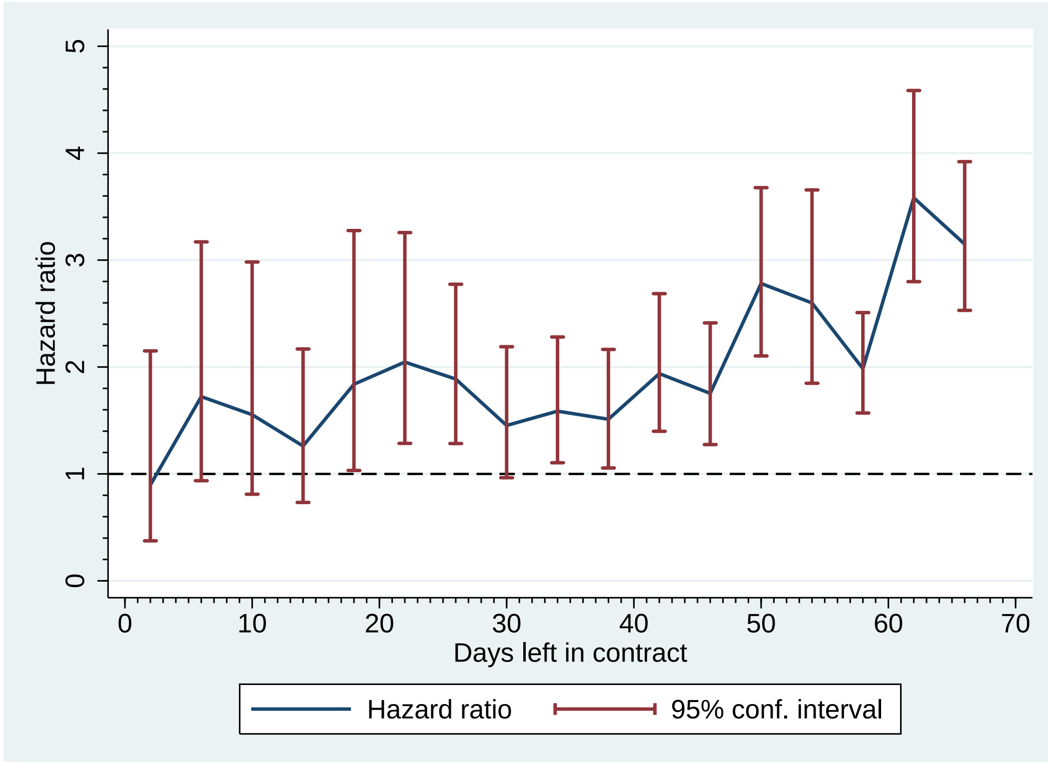

Figure 7: Hazard ratio as a function of time left in the insurance contract excluding first 8 days of the contract. Model includes a user specific baseline hazard rate. 95% confidence intervals calculated using standard errors clustered at the player level.

actions is likely to vary between different locations and over time. The explanation still requires that most players find driving the existing insured ship to the new location a dreadful task whose cost overweighs the expected cost of lost insurance protection. The shocks to the riskiness of different locations need to be relatively persistent to be able to explain why a ship that has been insured for over 30 days still has a hazard rate that is over twofold compared to a ship at the end of its insurance contract.

Another plausible explanation is more behavioral: the ship that got recently insured has been the target of the user's attention. The longer it has passed from the purchase of the previous insurance contract the more likely it is that the user has forgotten that the ship is available for use. In other words, the insurance purchase acts as a proxy for attention paid to the ship. This story could also generate the monotonous decreasing relationship between hazard ratios and time since the previous insurance purchase.

Just like before, adding dummies for ship age should partially control for ship-specific heterogeneity in the insurance pool. Heterogeneity would imply reductions in hazard ratios over the whole span of the insurance contract. In Figures 8 and 9 I present the results from the estimation with the same age dummies as in Section 4.1. The effect of the first days of insurance on the hazard ratio drops from about 20 to 13. However, compared to the model without the age dummies, the remainder of the graph stays at a considerably higher level. For example, for the last 40 days of the contract the hazard ratio with age dummies is consistently above 2 compared to the earlier range between

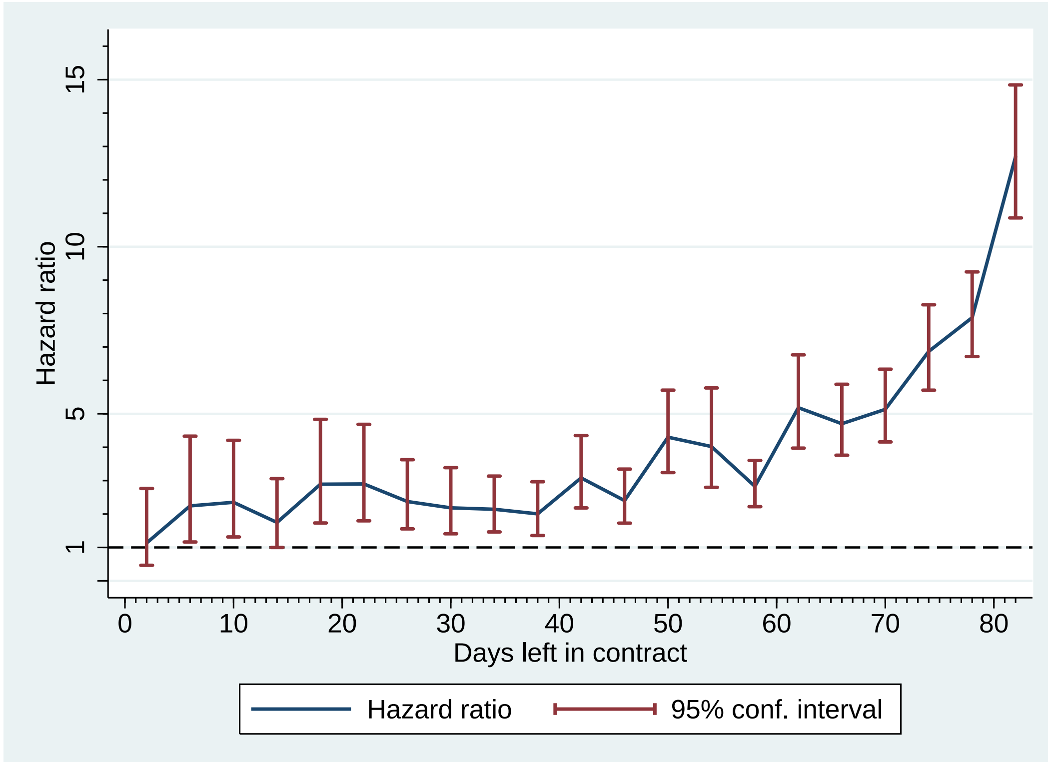

Figure 8: Hazard ratio as a function of time left in the insurance contract. Model includes a user specific baseline hazard rate and ship age dummies. 95% confidence intervals calculated using standard errors clustered at the player level.

1.5 and 2 in the model without age dummies.[23] Because adding the dummies reduces the hazard ratio only in the very beginning of the contract, the presence of ship-specific adverse selection is unlikely.

The age dummies themselves are statistically significant. The first day after a ship is being unpacked is especially risky with a hazard ratio of 15 while ships older than 180 days are fairly safe with a hazard ratio of about 0.38. This pattern is consistent with the attention based story, because the act of unpacking a ship is a likely indicator of an intention to use it. Attention would also explain why adding age dummies had a negligible effect on the stratified estimates in section 4.1 while the age dummies themselves are statistically significant: Old ships are less risky only because they are more likely to be forgotten. Once an agent insures an old ship, she is fully aware of its existence. Consequently, she will treat it just like a recently unpacked, insured ship.

Both, Figure 7 and 9, show a considerable increase in risk when about 20 days of the contract is left. According to the point estimate from the model with the ship age dummies, when a ship has 24 days left in its contract its hazard rate is about 45% higher than the hazard rate of a ship that has 40 days left in its contract. In the model without the ship age dummies this gap is about 36%. For Drakes the spike in risk

[23]Alternative specifications with differently paced age dummies produced qualitatively very similar results.

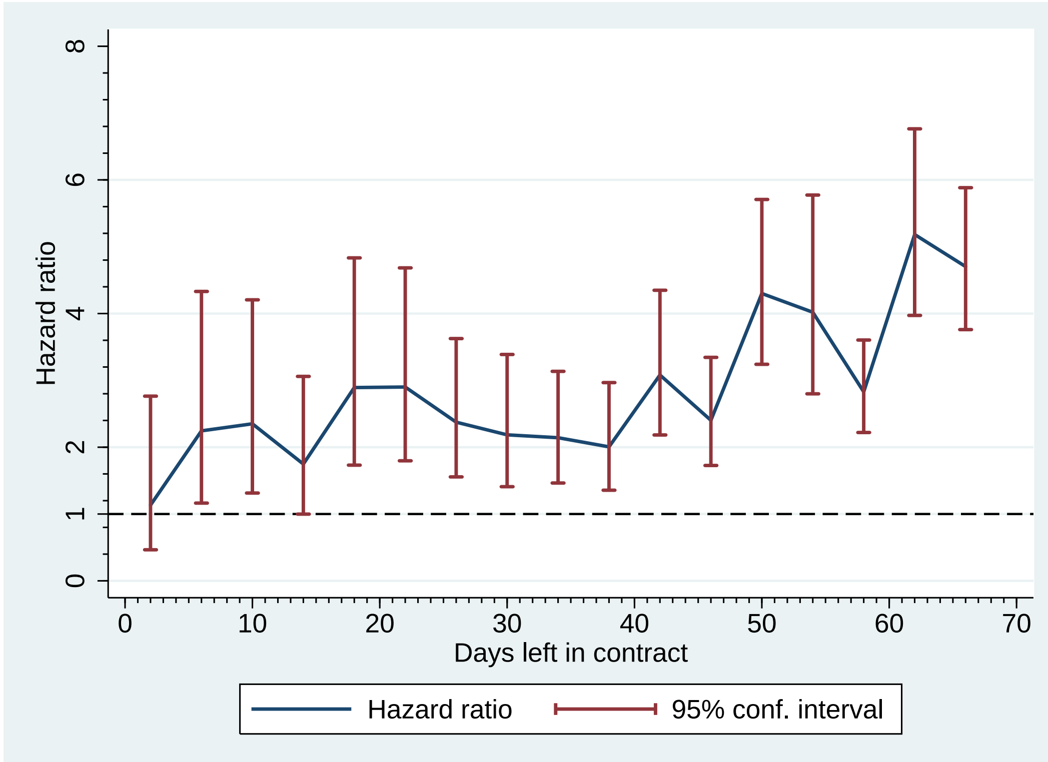

Figure 9: Hazard ratio as a function of time left in the insurance contract excluding first 8 days of the contract. Model includes a user specific baseline hazard rate and ship age dummies. 95% confidence intervals calculated using standard errors clustered at the player level.

for ships with 20-24 days left in their contracts is even more pronounced. Looking at Figures 10 and 11, one can see that the hazard ratio for these days is over 5 times as high as the hazard ratio for the ships with 44-48 days left in their contract. This effect could be taken as partial evidence for dynamic moral hazard increasing as the contract comes close to its end. However, the accuracy of these estimates is so poor that I cannot completely rule out the gap being purely a result of random variation. Moreover, the fall in risk over the last 20 days of the contract is difficult to reconcile with the theory.

5 Experience and moral hazard

There are at least three reasons why agents' level of experience can be an important factor in insurance markets. First, more experienced agents may better understand what type of behavior is likely to lead to an accident and choose safer actions. Second, learning one's risk type potentially increases the information asymmetry between the insurance provider and the agent. Hence, more experienced policy holders may be more adversely selected. Finally, better understanding of the relationship between available actions and risk helps agents to choose more optimal precautionary effort when faced with different levels of insurance coverage. More experienced agents may hence exhibit

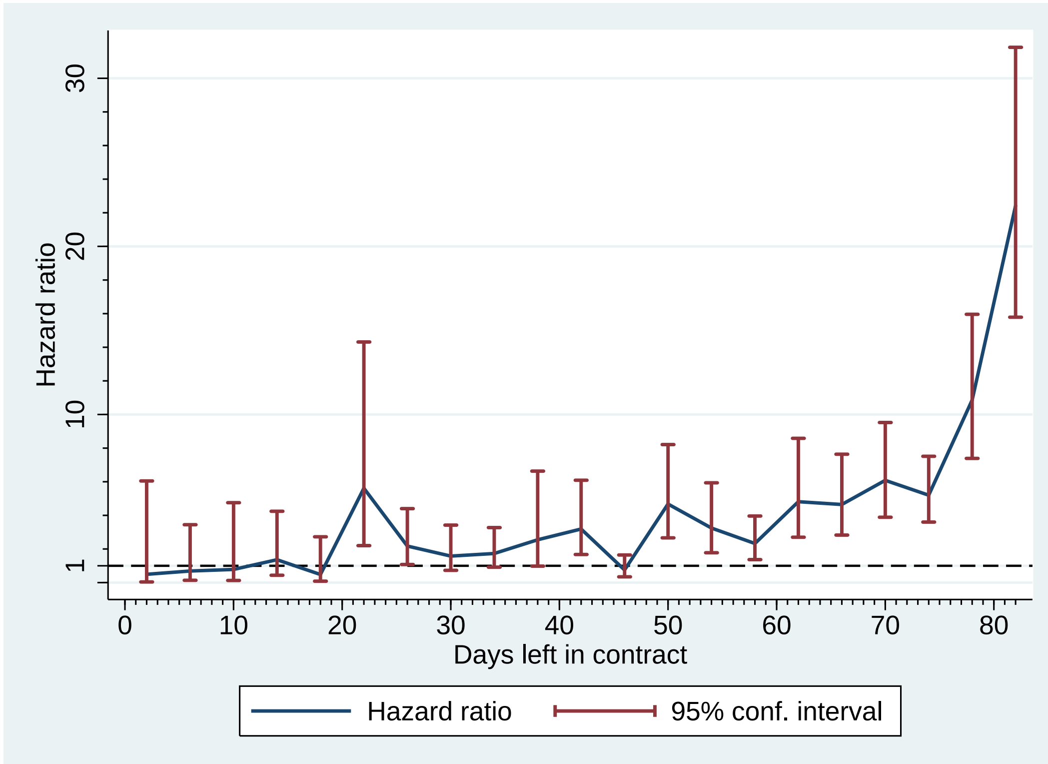

Figure 10: Drakes' hazard ratios as a function of time left in the insurance contract. Model includes a user specific baseline hazard rate and ship age dummies. 95% confidence intervals calculated using standard errors clustered at the player level.

more moral hazard if they are better able to reduce their effort costs by choosing risky actions when insured and exercise more relevant precautionary effort when uninsured. Consequently, offering separate contracts for experienced and inexperienced agents can be an important way of increasing efficiency in the insurance market. Furthermore, insights on the channels through which experience affects risk can help target monitoring and screening effort to the right subpopulation.

I will measure player's experience by the number of days she has spent logged into the game. This is analogous to many real world studies that measure experience, for example, by using the time since a driver got her driver's license. My measure is potentially more accurate, because being logged into the game is a strong indicator of actively using a ship. Most activities in EVE involve the use of ships and hence actively playing the game is almost synonymous with using a ship. In comparison, two persons who got their licenses at the same time are likely to have considerably different amounts of driving experience.

To establish the fact that players become better at protecting their property over time, I compared the ship destruction rates of the bottom and top quartiles of players sorted by their time spent in the game. As expected, new players seem to get their spacecraft destroyed much more frequently than the old-timers: 11.6% of the ships owned by the least experienced 25% of the sample got destroyed, while only 5.4% of the ships owned by the most experienced 25% faced the same fate. In other words, the

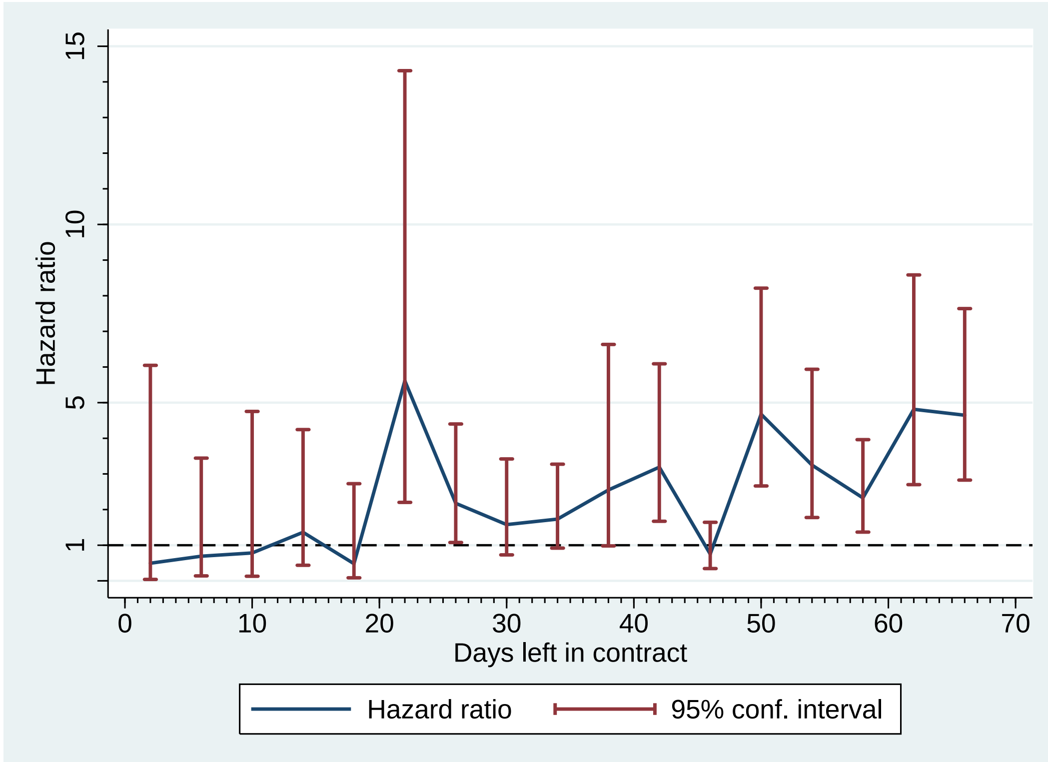

Figure 11: Drake's hazard ratios as a function of time left in the insurance contract excluding first 8 days of the contract. Model includes a user specific baseline hazard rate and ship age dummies. 95% confidence intervals calculated using standard errors clustered at the player level.

ship of a new player was almost twice as likely to get destroyed than that of a seasoned veteran.

Next, I estimated the model from Section 4.1 separately for these two quartiles. The results are presented in Table 5. The first thing to notice is that the combined effect of moral hazard and adverse selection is over 3 times as high for the more experienced players than for the new-comers. Even though the new players are very likely to get their ships destroyed, the experienced players' ships are much more likely to be insured *when* they get destroyed. The direction of the effect is in line with what one would expect if it takes time for the players to learn their risk type or the mapping from effort to accident probabilities. The magnitude of the effect, however, is striking.

The results on the row where the estimates are stratified at the user level show that the effect of *moral hazard* on the ship's risk is almost five times as high for the veterans compared to the new players. The impact of experience on moral hazard is slightly smaller when I look only at Drakes. However, due to the relatively small number of Drakes in the sample, the estimates become inaccurate when stratified at the user level. For the players in EVE, learning and experience seem to play a key role in figuring out the optimal amount of precautionary effort and how insurance factors into it. If a similar finding is true in other insurance contexts, having higher deductibles or coinsurance rates for young people might be neither profit maximizing nor socially

Table 5: Estimates of the insurance parameter from stratified Cox regressions for the most and least experienced players

Strata	Included observations	Hazard ratio Bottom 25%	Hazard ratio Top 25%
Ship type	All	3.53	10.82
		(3.30, 3.78)	(10.04, 11.66)
		N=74,519	N=367,710
Ship type,	All	2.19	9.94
User ID		(1.76, 2.73)	(8.40, 11.76)
		N=74,519	N=367,710
None	Drakes	3.00	11.09
		(2.34, 3.83)	(9.32, 13.20)
		N=2,798	N=16,042
User ID	Drakes	4.08	12.26
		(1.68, 9.96)	(7.54, 19.95)
		N=2,798	N=16,042

Note: 95% confidence intervals in parenthesis. Implicit standard errors clustered at the user level. Time recorded in days.

optimal. If information asymmetries are less of a concern with young agents, a better practice would have them paying higher premiums but face lower coinsurance rates than what is offered to more experienced agents.

I next apply the model from Section 4.3 separately to the bottom and top 25% of players ranked by the hours they have spent logged into the game. To gain enough within person variation in both insurance contracts and accidents to accurately estimate the model, I split the insurance contracts into 12-day periods instead of the previous 4-day periods. The results are presented in Figures 12 and 13.

The first interesting pattern from Figure 12 is that the insured ships of the least experienced players are statistically significantly more risky than uninsured ships only in the beginning of the insurance contract. After 12 days of the contract has passed, insured ships have hazard rates that are indistinguishable from uninsured ships. The same is not true for experienced players for whom the hazard ratio between insured and uninsured ships is clearly above one for the whole contract period. Hence, the inexperienced players seem to be planning their risky activities only for a very short horizon.

As can be seen in Figure 13, the spike in risk in the beginning of the contract is even higher for the experienced players. The daily hazard rate of a ship that was insured less than 12 days ago is almost 25 times as high as the hazard rate of an uninsured ship and about 8 times as high as the hazard rate of a ship that is close to the end of its insurance contract. This is consistent with experienced players better knowing when

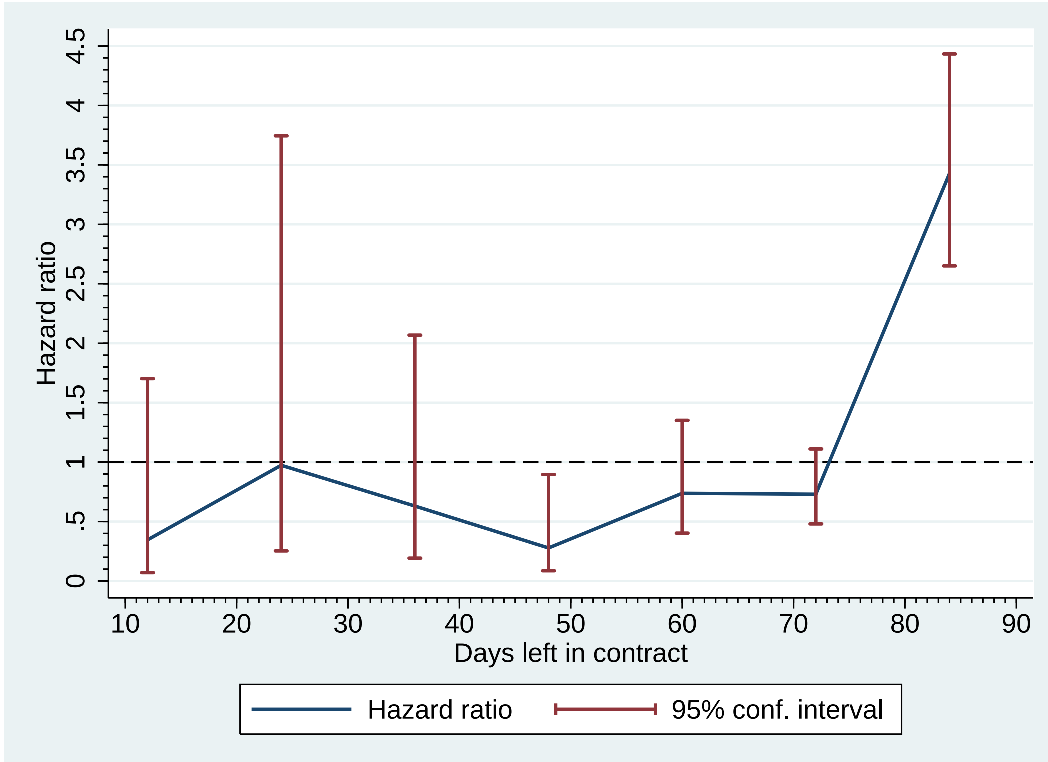

Figure 12: Hazard ratio as a function of time left in the insurance contract for the *least* experienced quartile of players. Model includes a user specific baseline hazard rate. 95% confidence intervals calculated using standard errors clustered at the player level.

they need insurance. Notice also that after the first 12 days, the hazard ratio settles to a relatively stable level between 2 and 3. Hence, it seems that these players do not find dynamic optimization of their precautionary effort worthwhile but just choose to use their insured ships equally more over the remaining 72 days of the contract. In contrast to the inexperienced agents, these players exhibit strong moral hazard over the whole span of the insurance contract but do not seem to adjust their effort choices after the first 12-24 days.

The results have interesting implications for contract design, if similar patterns hold even in real world insurance. If purchase of insurance is a strong predictor of increased risk for the short term but not for the long horizon, and if this effect is due to effort choices, insurance contracts should have higher deductibles for accidents that happen early in the contract. Furthermore, if there are inexperienced agents who exhibit moral hazard only in the beginning of the contract, then these agents could be potentially nearly fully insured for accidents that happen later in the contract.

6 Conclusions

Better measurement of moral hazard and how people respond to incentives that change over time may translate into more efficient contracts. Interesting dynamic patterns, such

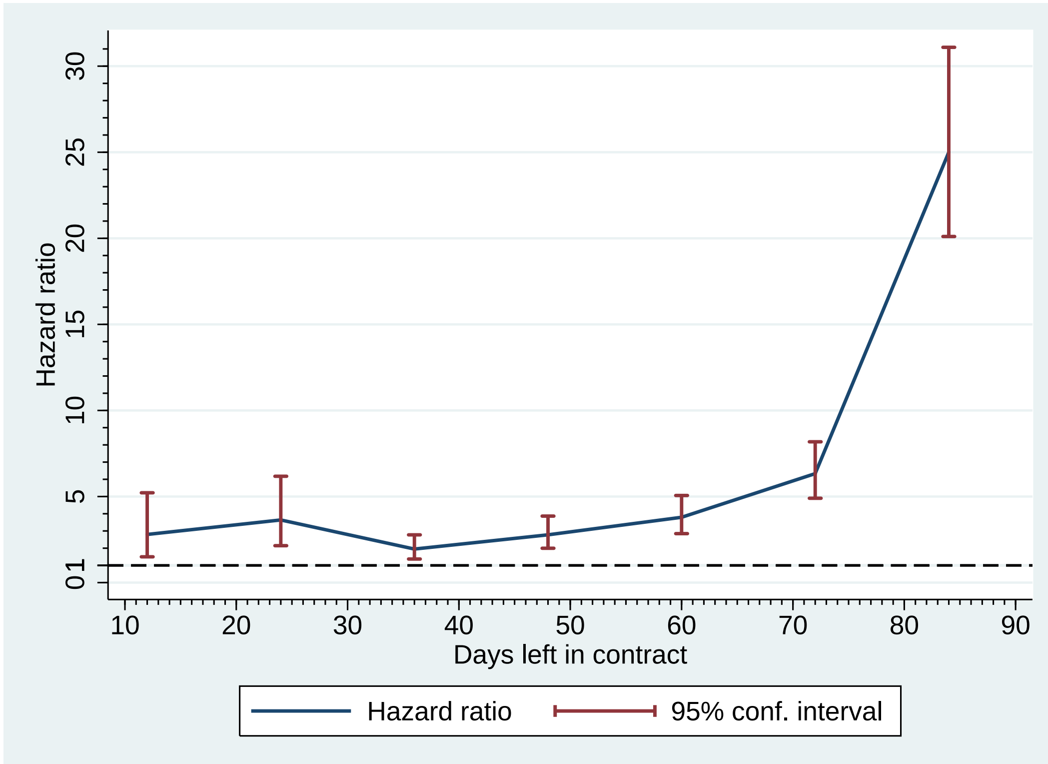

Figure 13: Hazard ratio as a function of time left in the insurance contract for the *most* experienced quartile. Model includes a user specific baseline hazard rate. 95% confidence intervals calculated using standard errors clustered at the player level.

as the high initial risk in the beginning of the contract, as presented in this paper, could potentially be used to generate slack in the more binding constraints of the contract design problem. Hence, a better understanding of these dynamic patterns could lead to contracts that are able to reduce information rents and yield socially more desirable outcomes. This paper has introduced methods that hopefully help us to achieve these goals.

More specifically, it would interesting to learn whether the front-loading of risk to the beginning of the contract period is a more ubiquitous phenomenon or whether it is special to the insurance contracts in EVE. Similarly, documenting the effect of experience on moral hazard in other insurance contracts can have high social returns. If inexperienced contracting agents exhibit less moral hazard also in other institutional environments, it could be beneficial for both society and the insurance companies to sell young or otherwise inexperienced agents contracts that cover a higher fraction of all losses.

Applying my methods to short-term labor contracting is another interesting avenue of future research. Measuring how output and the risk of premature termination of the contract varies over the span of a temporary work contract and how the dynamic variation in output is affected by the timing of wage payments can be especially important if the role of temporary and flexible work continues to grow in modern societies.

Last, my econometric method for measuring the changes in hazard rates over the course of the contract is not specific to EVE's institutional setup and could be applied to measure dynamic variation in risk even when the intertemporal incentives for risk taking are completely different from the ones induced by EVE's insurance contracts. Tolvanen (in progress) uses a similar method to identify dynamic incentives for moral hazard in Nordic car insurance. In some of the insurance company's contracts the owner of a new car is guaranteed an identical replacement vehicle during the first 3 years of the car in case the original car is totaled. Because the car still depreciates during the first 3 years of the car, the contract structure should imply decreasing effort over the first 3 years. Furthermore, since after the first 3 years the depreciation is taken discontinuously into account when valuing the totaled car, there should be a discontinuous increase in effort after the third year. This identification strategy, even though being different on the surface shares a strong similarity with the dynamic identification strategy in this paper and similar econometric modeling can be used to measure the dynamic changes in risk induced by the structure of these contracts.

A appendix

The following well-known lemma on decreasing differences of a concave function is useful throughout the proofs:

Lemma 2. *Let $u\colon \mathbb{R} \to \mathbb{R}$ be differentiable, increasing and strictly concave. If $z > 0$, then*

$$\frac{\mathrm{d}\left(u(x) - u(x-z)\right)}{\mathrm{d}x} < 0$$

for all $x \in \mathbb{R}$.

Proof. By direct calculation,

$$\frac{\mathrm{d}\left(u(x) - u(x-z)\right)}{\mathrm{d}x} = u'(x) - u'(x-z).$$

Since u is strictly concave, u' is decreasing. This in turn implies that $u'(x-z) > u'(x)$ proving the lemma. □

A.1 Proofs for section 2.1

For the proofs in this section, it is useful to define the utility value from any given effort levels and insurance choices as:

$$\begin{aligned} V(x_{1j}, x_{2j}; I_j, \varepsilon_j) \;:=\;& p(x_{1j}, \varepsilon_j)p(x_{2j}, \varepsilon_j)u(\omega_j - 2L + (I_{1j} + I_{2j})(P - C), \varepsilon_j) \\ &+p(x_{1j}, \varepsilon_j)(1 - p(x_{2j}, \varepsilon_j))u(\omega_j - L - (I_{1j} + I_{2j})C + I_{1j}P, \varepsilon_j) \\ &+p(x_{2j}, \varepsilon_j)(1 - p(x_{1j}, \varepsilon_j))u(\omega_j - L - (I_{1j} + I_{2j})C + I_{2j}P, \varepsilon_j) \\ &+(1 - p(x_{1j}, \varepsilon_j))(1 - p(x_{2j}, \varepsilon_j))u(\omega_j - (I_{1j} + I_{2j})C, \varepsilon_j) \\ &-E(x_{1j}, x_{2j}, \varepsilon_j), \end{aligned}$$

I will first show that for any I_j there exists levels of effort that satisfy the first-order conditions of the agent's optimization problem:

Lemma 3. *For any I_j and ε_j there exists (x_{1j}, x_{2j}) such that $\nabla V(x_{1j}, x_{2j}; I_j, \varepsilon_j) = 0$.*

Proof. Fix any I_j and ε_j. For this proof I will drop the dependence on both ε_j and j since the result clearly does not depend on them. The two first-order conditions are now given by:

$$\begin{aligned}
&p'(x_1)p(x_2)u(\omega - 2L + (I_1 + I_2)(P - C)) \\
&+p'(x_1)(1 - p(x_2))u(\omega - L - (I_1 + I_2)C + I_1 P) \\
&-p'(x_1)p(x_2)u(\omega - L - (I_1 + I_2)C + I_2 P) \\
&-p'(x_1)(1 - p(x_2))u(\omega - (I_1 + I_2)C) \\
&-\partial_1 E(x_1, x_2) = 0
\end{aligned} \tag{A.1}$$

and

$$\begin{aligned}
&p'(x_2)p(x_1)u(\omega - 2L + (I_1 + I_2)(P - C)) \\
&-p'(x_2)p(x_1))u(\omega - L - (I_1 + I_2)C + I_1 P) \\
&+p'(x_2)(1 - p(x_1))u(\omega - L - (I_1 + I_2)C + I_2 P) \\
&-p'(x_2)(1 - p(x_1))u(\omega - (I_1 + I_2)C) \\
&-\partial_2 E(x_1, x_2) = 0.
\end{aligned} \tag{A.2}$$

Fix any x_2 and write equation (A.1) as

$$p'(x_1)\left(p(x_2)(a - c) + (1 - p(x_2))(b - d))\right) - \partial_1 E(x_1, x_2) = 0 \tag{A.3}$$

where

$$\begin{aligned}
a &:= u(\omega - 2L + (I_1 + I_2)(P - C)) \\
b &:= u(\omega - L - (I_1 + I_2)C + I_1 P) \\
c &:= u(\omega - L - (I_1 + I_2)C + I_2 P) \\
d &:= u(\omega - (I_1 + I_2)C)
\end{aligned}$$

Now since u is increasing and $L > P$ both $a - c$ and $b - d$ are strictly negative. Furthermore, since p is decreasing and strictly convex, both $\lim_{x\to 0} p'(x)$ and $\lim_{x\to\infty} p'(x)$ exist in $[-\infty, 0]$ and $\lim_{x\to 0} p'(x) < 0$ and $\lim_{x\to\infty} p'(x) > -\infty$. Consequently, $\lim_{x_1\to 0} E(x_1, x_2) = 0$ implies then that

$$\lim_{x_1\to 0} p'(x_1)\left(p(x_2)(a - c) + (1 - p(x_2))(b - d)\right) - \partial_1 E(x_1, x_2) > 0.$$

Similarly, $\lim_{x_1\to\infty} \partial_1 E(x_1, x_2) = \infty$ implies that

$$\lim_{x_1\to\infty} p'(x_1)\left(p(x_2)(a - c) + (1 - p(x_2))(b - d)\right) - \partial_1 E(x_1, x_2) = -\infty.$$

But then the intermediate value theorem implies that for each x_2 there exists $x_1(x_2)$ such that

$$p'(x_1(x_2))\left(p(x_2)(a - c) + (1 - p(x_2))(b - d)\right) - \partial_1 E(x_1(x_2), x_2) = 0.$$

Plugging this back in to (A.2), it will be enought to show that

$$p'(x_2)\left(p(x_1(x_2))(a-b)+(1-p(x_1(x_2)))(c-d)\right)-\partial_2 E(x_1(x_2),x_2)=0$$

has a solution. Just as above, $a-b$ and $c-d$ are both strictly negative. This together with the assumption that $\partial_{12}E(x_1,x_2)>0$ implies that

$$\begin{aligned}&\lim_{x_2\to\infty} p'(x_2)\left(p(x_1(x_2))(a-b)+(1-p(x_1(x_2)))(c-d)\right)-\partial_2 E(x_1(x_2),x_2)\\ \leq\ &\lim_{x_2\to\infty} p'(x_2)\left(p(x_1(x_2))(a-b)+(1-p(x_1(x_2)))(c-d)\right)-\partial_2 E(0,x_2)=-\infty\end{aligned}$$

Now if I consider $x_2=0$, then (A.2) becomes

$$p'(0)\left(p(x_1(0))(a-b)+(1-p(x_1(0)))(c-d)\right)-\partial_2 E(x_1(0),0)>0.$$

Applying the intermediate value theorem again implies that there exists x_2 such that

$$p'(x_2)\left(p(x_1(x_2))(a-b)+(1-p(x_1(x_2)))(c-d)\right)-\partial_2 E(x_1(x_2),x_2)=0$$

proving the lemma. □

The next lemma implies that for any choice of insurance, a choice of effort levels where the gradient of V vanishes yields maximal utility for that level of insurance:

Lemma 4. *V is strictly concave in (x_{1j},x_{2j}) .*

Proof. Fix any I_j. I will again drop the dependence on j and ε_j in this proof. Furthermore, I will make the notation somewhat clearer by writting:

$$\begin{aligned}W(x_1,x_2;I) &= p(x_1)p(x_2)a+p(x_1)(1-p(x_2))b+p(x_2)(1-p(x_1))c\\ &\quad +(1-p(x_1))(1-p(x_2))d,\end{aligned}$$

where again

$$\begin{aligned}&a:=u(\omega-2L+(I_1+I_2)(P-C))\\ &b:=u(\omega-L-(I_1+I_2)C+I_1P))\\ &c:=u(\omega-L-(I_1+I_2)C+I_2P)\\ &d:=u(\omega-(I_1+I_2)C)\end{aligned}$$

Since E is convex and the sum of two concave functions is concave, it is enough to show that W is concave. Now,

$$\frac{\partial^2}{\partial x_1^2}W(x_1,x_2) = p''(x_1)\left((a-c)\,p(x_2)+(b-d)\,(1-p(x_2))\right)<0,$$

where the negativity follows, because u is increasing and p is convex. Similarly,

$$\frac{\partial^2}{\partial x_2^2}W(x_1,x_2) = p''(x_2)\left((a-b)\,p(x_1)+(c-d)\,(1-p(x_1))\right)<0.$$

Thus it is enough to show that the determinant of the Hessian of W, $|H_W|$, is strictly positive. Calculating this Hessian requires the cross-derivative:

$$\frac{\partial^2}{\partial x_1 \partial x_2} W(x_1, x_2) = (a - b - c + d) p'(x_1) p'(x_2).$$

Now a straightforward calculation yields,

$$\begin{aligned}
|H_W| &= \frac{\partial^2}{\partial x_1^2} W(x_1, x_2) \frac{\partial^2}{\partial x_2^2} W(x_1, x_2) - \left(\frac{\partial^2}{\partial x_1 \partial x_2} W(x_1, x_2) \right)^2 \\
&= p''(x_1) p''(x_2) \left((p(x_2)(a - b - c + d) + b - d \right) \left(p(x_1)(a - b - c + d) + c - d \right) \\
&\quad - \left((a - b - c + d) \, p'(x_1) p'(x_2) \right)^2 \qquad \text{(A.4)}
\end{aligned}$$

Now, since u is concave, $a - b - c + d$ is easily verified negative. Furthermore, since $b - d < 0$ and $c - d < 0$, it follows that

$$\begin{aligned}
& \left((p(x_2)(a - b - c + d) + b - d \right) \left(p(x_1)(a - b - c + d) + c - d \right) \\
= \; & |(p(x_2)(a - b - c + d) + b - d| \, |p(x_1)(a - b - c + d) + c - d| \\
> \; & p(x_1) p(x_2) (a - b - c + d)^2
\end{aligned}$$

Now, I can plugg this back into (A.4) and use the fact that $p'' > 0$ to obtain

$$|H_W| > \left(p''(x_1) p''(x_2) p(x_1) p(x_2) - p'(x_1)^2 p'(x_2)^2 \right) (a - b - c + d)^2 \qquad \text{(A.5)}$$

Now the first term is simply the determinant of the Hessian of the mapping $(x_1, x_2) \mapsto p(x_1)p(x_2)$ which was assumed to be strictly convex. Hence, that term is strictly positive implying that $|H_W|$ must be strictly positive. □

Together Lemmas 3 and 4 and the fact that I_j is selected from a finite set imply that the agent's maximization problem has a unique solution. Next I will show that an insured item always has a higher accident risk than an uninsured item owned by the same person.

Lemma 5. *Take any person j with $(\varepsilon_j, \omega_j)$ such that $I_j = (0, 0)$ or $I_j = (1, 1)$. Then for this person the optimal choice of effort (x_{1j}^*, x_{2j}^*) satisfies: $p(x_{1j}^*, \varepsilon_j) = p(x_{2j}^*, \varepsilon_j)$.*

Proof. For the proof of this lemma I may again drop the dependence on j and ε_j. The proofs for the two cases are identical and I will only consider the slightly more complex case where $I_j = (1, 1)$. The case where $I_j = (0, 0)$ follows by dropping both C and P from the proof below. I will abuse the notation slightly and write

$$\begin{aligned}
V_2(x_1, x_2) &= p(x_1) p(x_2) u(\omega - 2L + 2P - 2C) \\
&\quad + p(x_1)(1 - p(x_2)) u(\omega - L - 2C + P) \\
&\quad + p(x_2)(1 - p(x_1)) u(\omega - L - 2C + P) \\
&\quad + (1 - p(x_1))(1 - p(x_2)) u(\omega - 2C) \\
&\quad - E(x_1, x_2),
\end{aligned}$$

where the subscript 2 stands for the two insured units. With this notation, it is enough to prove that if (x_1^*, x_2^*) maximizes V_2, then $x_1^* = x_2^*$. The first-order conditions for this maximization problem are given by

$$x_1\colon \quad \partial_1 V_2(x_1^*, x_2^*) = p'(x_1^*)p(x_2^*)A - p'(x_1^*)\left(u(\omega - 2C) - u(\omega - L - 2C + P)\right) - \partial_1 E(x_1^*, x_2^*) = 0 \tag{A.6}$$

$$x_2\colon \quad \partial_2 V_2(x_1^*, x_2^*) = p'(x_2^*)p(x_1^*)A - p'(x_2^*)\left(u(\omega - 2C) - u(\omega - L - 2C + P)\right) - \partial_2 E(x_1^*, x_2^*) = 0 \tag{A.7}$$

where

$$A\colon = u(\omega - 2C) - 2u(\omega - L - 2C + P) + u(\omega - 2C - 2L + 2P) < 0,$$

because u is concave.

Notice that $\lim_{x\to 0} \partial_1 V_2(x, x) > 0$, because $\lim_{x\to 0} \partial_1 E(x, x) = 0$, $p'(x) < 0$, $A < 0$, and $u(\omega - 2C) - u(\omega - L - 2C + P > 0)$, by the fact that u is increasing. On the other hand, $\lim_{x\to\infty} \partial_1 V_2(x, x) = -\infty$, because $\lim_{x\to\infty} E(x, x) = \infty$. Consequently, there exists an $\hat{x} \in \mathbb{R}_+$ such that $\partial_1 V_2(\hat{x}, \hat{x}) = 0$. Moreover,

$$\partial_2 V_2(\hat{x}, \hat{x}) = \partial_1 V_2(\hat{x}, \hat{x}) + \partial_1 E(\hat{x}, \hat{x}) - \partial_2 E(\hat{x}, \hat{x}) = 0,$$

as, by assumption $\partial_1 E(\hat{x}, \hat{x}) - \partial_2 E(\hat{x}, \hat{x})$. Therefore, $\hat{x}$ satisfies both of the first-order conditions. As, by the previous lemma, V_2 is strictly concave, $(\hat{x}, \hat{x})$ must be the unique maximizer of V_2. □

Then the proof of Proposition 1 follows trivially from the law of iterated expectations:

Proof of Proposition 1. Let $k \in \{0, 1\}$. The law of iterated expectations together with the previous lemma imply:

$$\begin{aligned}
& \mathbb{E}[p(x_{1j}, \varepsilon_j) \mid I_{1j} = k, I_{2j} = k] \\
= \; & \mathbb{E}\left[\mathbb{E}\left[p(x_{1j}, \varepsilon_j) \mid I_{1j} = k, I_{2j} = k, \varepsilon_j, \omega_j\right] \mid I_{1j} = k, I_{2j} = k\right] \\
= \; & \mathbb{E}\left[\mathbb{E}\left[p(x_{2j}, \varepsilon_j) \mid I_{1j} = k, I_{2j} = k, \varepsilon_j, \omega_j\right] \mid I_{1j} = k, I_{2j} = k\right] \\
= \; & \mathbb{E}[p(x_{2j}, \varepsilon_j) \mid I_{1j} = k, I_{2j} = k].
\end{aligned}$$

□

Lemma 6. *Take any person j with $(\varepsilon_j, \omega_j)$ such that $I_j = (1, 0)$. Then for this person the optimal choice of effort (x_{1j}^*, x_{2j}^*) satisfies: $p(x_{1j}^*, \varepsilon_j) > p(x_{2j}^*, \varepsilon_j)$.*

Proof. I continue suppressing the dependence on j and ε_j. I will abuse the notation slightly and write

$$\begin{aligned}
V_1(x_1, x_2) \; = \; & p(x_1)p(x_2)u(\omega - 2L + P - C) \\
& + p(x_1)(1 - p(x_2))u(\omega - L - C + P) \\
& + p(x_2)(1 - p(x_1))u(\omega - L - C) \\
& + (1 - p(x_1))(1 - p(x_2))u(\omega - C) \\
& - E(x_1, x_2).
\end{aligned}$$

Now, it is enough to prove that if (x_1^*, x_2^*) maximizes V, then $x_1^* < x_2^*$. The first-order conditions for this maximization problem are given by

$$
\begin{aligned}
x_1\colon \quad & \partial_1 V_1(x_1^*, x_2^*) = p'(x_1^*)p(x_2^*)B - p'(x_1^*)\left(u(\omega - C) - u(\omega - L - C + P)\right) \\
& -\partial_1 E(x_1^*, x_2^*) = 0 \qquad \text{(A.8)} \\
x_2\colon \quad & \partial_2 V_1(x_1^*, x_2^*) = p'(x_2^*)p(x_1^*)B - p'(x_2^*)\left(u(\omega - C) - u(\omega - L - C)\right) \\
& -\partial_2 E(x_1^*, x_2^*) = 0 \qquad \text{(A.9)}
\end{aligned}
$$

where

$$B\colon = u(\omega - C) - u(\omega - L - C + P) - u(\omega - C - L) + u(\omega - C - 2L + P),$$

which is negative, since u is concave. For any x_2 define

$$\hat{x}_1(x_2) := \arg\max_{x_1} V_1(x_1, x_2).$$

Now, $\hat{x}_1(x_2)$ solves $\partial_1 V_1(x_1, x_2) = 0$ for any given x_2. Thus, by the implicit function theorem,

$$
\begin{aligned}
\frac{d\hat{x}_1(x_2)}{dx_2} &= -\frac{\partial_{12} V_1(\hat{x}_1(x_2), x_2)}{\partial_{11} V_1(\hat{x}_1(x_2), x_2)} \\
&= -\frac{p'(\hat{x}_1)p'(x_2)B - \partial_{12}E(\hat{x}_1, x_2)}{p''(\hat{x}_1)p(x_2)B - p''(\hat{x}_1)\left(u(\omega - C) - u(\omega - L - C + P)\right) - \partial_{11}E(\hat{x}_1, x_2)},
\end{aligned}
$$

where $\hat{x}_1$ stands for $\hat{x}_1(x_2)$. Because p' and B are negative and $\partial_{12}E$ is positive, the numerator is clearly negative. Furthermore, since the convexity of E and p imply that $\partial_{11}E$ and p'' are positive, the denominator must be negative. Thus I have that $\frac{d\hat{x}_1(x_2)}{dx_2} < 0$. This in turn implies that the graph of $\hat{x}_1(x_2)$ has one and only one intersection with the graph of $f(x_2) = x_2$. Hence, $\hat{x}_1(x_2)$ has a unique fixed point, $\hat{x}_2$.

Define next an auxiliary function

$$F(x_1, x_2) := \partial_1 V_1(x_1, x_2) - \partial_2 V_1(x_1, x_2).$$

The first-order conditions imply that

$$F(\hat{x}_1(x_2^*), x_2^*) = F(x_1^*, x_2^*) = 0. \qquad \text{(A.10)}$$

On the other hand, $\partial_1 E(\hat{x}_2, \hat{x}_2) = \partial_2 E(\hat{x}_2, \hat{x}_2)$ implies that

$$F(\hat{x}_1(\hat{x}_2), \hat{x}_2) = F(\hat{x}_2, \hat{x}_2) = p'(\hat{x}_2)\left(u(\omega - L - C + P) - u(\omega - L - C)\right) < 0. \qquad \text{(A.11)}$$

This implies that the lemma holds, if I can show that $F(\hat{x}_1(x_2), x_2)$ is strictly increasing. To complete the proof, notice that

$$
\begin{aligned}
\frac{dF(\hat{x}_1(x_2), x_2)}{dx_2} &= \left(\partial_{11} V_1(\hat{x}_1(x_2), x_2) - \partial_{12} V_1(\hat{x}_1(x_2), x_2)\right) \frac{d\hat{x}_1(x_2)}{dx_2} \\
&\quad + \partial_{12} V_1(\hat{x}_1(x_2), x_2) - \partial_{22} V_1(\hat{x}_1(x_2), x_2) \\
&= -\left(\partial_{11} V_1(\hat{x}_1(x_2), x_2) - \partial_{12} V_1(\hat{x}_1(x_2), x_2)\right) \frac{\partial_{12} V_1(\hat{x}_1(x_2), x_2)}{\partial_{11} V_1(\hat{x}_1(x_2), x_2)} \\
&\quad + \partial_{12} V_1(\hat{x}_1(x_2), x_2) - \partial_{22} V_1(\hat{x}_1(x_2), x_2) \\
&= \frac{\partial_{12} V_1(\hat{x}_1(x_2), x_2)^2}{\partial_{11} V_1(\hat{x}_1(x_2), x_2)} - \partial_{22} V_1(\hat{x}_1(x_2), x_2) \qquad \text{(A.12)}
\end{aligned}
$$

Since V_1 is striclty concave, $\partial_{11}V_1 < 0$. Thus (A.12) is strictly positive, if and only if

$$\partial_{11}V_1(\hat{x}_1(x_2), x_2)\partial_{22}V_1(\hat{x}_1(x_2), x_2) - \partial_{12}V_1(\hat{x}_1(x_2), x_2)^2 > 0.$$

But this holds always, since V is strictly concave and hence its Hessian must have a strictly positive determinant. □

Proof of Proposition 2. After Lemma 6, the proof for this proposition follows easily from the law of total probability:

$$\begin{aligned} & \mathbb{E}[p(x_{1j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 0] \\ = \; & \mathbb{E}\left[\mathbb{E}\left[p(x_{1j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 0, \varepsilon_j, \omega_j\right] \mid I_{1j} = 1, I_{2j} = 0\right] \\ > \; & \mathbb{E}\left[\mathbb{E}\left[p(x_{2j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 0, \varepsilon_j, \omega_j\right] \mid I_{1j} = 1, I_{2j} = 0\right] \\ = \; & \mathbb{E}[p(x_{2j}, \varepsilon_j) \mid I_{1j} = 1, I_{2j} = 0]. \end{aligned}$$

□

B Proofs for Section 2.2

I will start by showing that the agent's maximization problem has a recursive structure with respect to $I_t' = \max\{0, (1 - D_t)I_t - 1\}$. For this end I will need some auxiliary definitions. Let

$$U_j(x, I, D) = u(\omega - D(L - \mathbb{1}\{I\}P) - \mathbb{1}\{I = l\}C) - E(x, \varepsilon_j).$$

Then the agent's maximization problem in this notation is

$$\sup_{(x_j, I_j) \in C} \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t U(x_j(h_t), I_j(h_t), D_{t+1}) \mid x_j, I_j, h_0\right],$$

where

$$C := \{(x_j, I_j) \colon \mathcal{H} \mapsto \mathbb{R}_+ \times \{1, \dots, l\}, I_j(h_t) \in \{l, \max\{0, (1 - D_{jt-1})I_{jt-1} - 1\}\} \,\forall h_t \in \mathcal{H}\}.$$

Notice that since

$$\lim_{x \to \infty} U_j(x, I, D) = -\infty$$

and $U_j(0, 0, 1) > -\infty$, there exists $M > 0$ such that choosing the controls from

$$\hat{C} := \{(x_j, I_j) \colon \mathcal{H} \mapsto [0, M] \times \{1, \dots, l\}, I_j(h_t) \in \{l, \max\{0, (1 - D_{jt-1})I_{jt-1} - 1\}\} \,\forall h_t \in \mathcal{H}\}$$

will yield the same supremum as choosing them from C. If one endows $\{0, 1, \dots, l\}$ with its discreet topology, then $\hat{C}$ is clearly compact in the product topology by the Tychonoff's Theorem.

For a given policy x_j, I_j, let $W(x_j, I_j, h_t)$ be the continuation utility from history h_s onwards:

$$W(x_j, I_j, h_s) := \mathbb{E}\left[\sum_{t=s}^{\infty} \delta^{t-s} U(x_j(h_t), I_j(h_t), D_{t+1}) \mid x_j, I_j, h_s\right]$$

Since both p and E are continuous in x, one can also show that $W(\cdot,\cdot,h_0)$ is continuous for each h_0. Consequently, $\max_{(x_j,I_j)\in C} W(x_j,I_j,h_0)$ is well defined.

Denote an arbitrary outcome in period t by $\theta_{jt} = (x_{jt}, I_{jt}, D_{jt})$. Then I can write an arbitrary history $h_t = (\theta_{jt})_{s=0}^{t-1}$. For any $s < u$ I will denote the partial history $(\theta_{jr})_{r=s}^{u-1}$ of the full history $h_t = (\theta_r)_{r=0}^{t-1}$ by h_s^u. Furthermore, for any two histories $h_t = (\theta_{jr})_{r=0}^{t-1}$ and $h'_s = (\theta'_{jr})_{r=0}^{s-1}$ and for any $0 \geq a < b \leq t$ and $0 < c < d \leq s$ I will define the combined history

$$h_a^b \otimes h_c^{\prime d} := (h_a^b, h_c^{\prime d}) = (\theta_{ja}, \theta_{ja+1}, \ldots, \theta_{jb-1}, \theta'_{jc}, \theta'_{jc+1}, \ldots, \theta'_{jd-1}).$$

The following lemma shows that for any two histories that share the same amount of insurance in the beginning of the period, the supremum of the continuation payoffs must be the same.

Lemma 7. *Let* h_t *and* h'_s *be such that* $\max\{0,(1-D_{jt-1})I_{t-1}-1\} = \max\{0,(1-D'_{js-1})I'_{s-1}-1\}$. *Then*

$$\max_{(x_j,I_j)\in C} W(x_j,I_j,h_t) = \max_{(x_j,I_j)\in C} W(x_j,I_j,h'_s).$$

Proof. Let (x_j, I_j) be an arbitrary policy that satisfies $I_j(h_t) \in \{l, \max\{0,(1-D_{jt-1})I_{t-1}\}\}$ for every h_t. Partition $\mathcal{H}$ by setting

$$\mathcal{H} \mid h'_s = \left\{h_r \colon \exists h_q \in \mathcal{H} \text{ such that } \ h_r = h_0^{\prime s} \otimes h_0^q\right\}$$

Then define (x'_j, I'_j) by setting $(x'_j(h_r), I'_j(h_r)) = (x_j(h_r), I_j(h_r))$ if $h_r \notin \mathcal{H} \mid h'_s$. If $h_r = h_0^{\prime s} \otimes h_0^q$, then set

$$(x'_j(h_r), I'_j(h_r)) = (x_j(h_0^t \otimes h_0^q), I_j(h_0^t \otimes h_0^q)).$$

Notice that this policy is trivially an element of C, because (x_j, I_j) is an element in it *and* the same choices are available in C after histories h'_s and h_t by assumption. Notice then that for any $(x, I, D) \in \mathbb{R}_+ \times \{0,\ldots,l\} \times \{0,1\}$

$$\mathbb{P}\left[(x_{js}, I_{js}, D_{js}) = (x,I,D) \mid x'_j, I'_j, h'_s\right] = \mathbb{P}\left[(x_{jt}, I_{jt}, D_{jt}) = (x,I,D) \mid x_j, I_j, h_t\right],$$

since both are equal to $Dp(x) + (1-D)(1-p(x))$, if $x = x_j(h_t)$ and $I = I_j(h_t)$, and zero otherwise. Then an easy induction step can be used to establish that more generally, for any $h_q \in \mathcal{H}$

$$\mathbb{P}\left[h_0^{\prime s} \otimes h_0^q \mid x'_j, I'_j, h'_s\right] = \mathbb{P}\left[h_0^t \otimes h_0^q \mid x_j, I_j, h_t\right].$$

Since the utility stream U depends in any arbitrary period r only on θ_r, we must have that

$$\begin{aligned} W(x_j, I_j, h_t) &= \mathbb{E}\left[\sum_{r=t}^{\infty} \delta^{r-t} U(x_j(h_r), I_j(h_r), D_{r+1}) \mid x_j, I_j, h_t\right] \\ &= \mathbb{E}\left[\sum_{r=s}^{\infty} \delta^{r-s} U(x'_j(h_r), I'_j(h_r), D_{r+1}) \mid x'_j, I'_j, h'_s\right] \\ &= W(x'_j, I'_j, h'_s) \leq \max_{(\hat{x}_j, \hat{I}_j)\in C} W(\hat{x}_j, \hat{I}_j, h'_s), \end{aligned} \tag{B.1}$$

Because (x_j, I_j) was arbitrary element of C, it must be the case that

$$\max_{(\hat{x}_j,\hat{I}_j)\in C} W(\hat{x}_j, \hat{I}_j, h_t) \leq \max_{(\hat{x}_j,\hat{I}_j)\in C} W(\hat{x}_j, \hat{I}_j, h'_s).$$

As h'_s and h_t were arbitrary histories sharing the same amount of time left in the contract, one can switch the roles of h'_s and h_t in the argument above to obtain

$$\max_{(\hat{x}_j,\hat{I}_j)\in C} W(\hat{x}_j, \hat{I}_j, h_t) \geq \max_{(\hat{x}_j,\hat{I}_j)\in C} W(\hat{x}_j, \hat{I}_j, h'_s),$$

which proves the lemma. □

Applying this lemma to the case of initial history yields the following corollary:

Corollary 2. *If h_0 and h'_t are such that*

$$\max\{0, (1 - D_{j0})I_{j0} - 1\} = \max\{0, (1 - D'_{jt})I'_{jt} - 1\},$$

then $V_j(h_0) = \max_{(x_j,I_j)\in C} W(x_j, I_j, h'_t)$.

Furthermore, with slight abuse of notation, I can now write $V(I'_0)$ instead of $V(h_0)$ where $I'_0 = \max\{0, (1 - D_0)I_0 - 1\}$.

Then one can apply standard verification arguments such as those given for Propositions 2.2 and 2.3 in Bertsekas (1995) to argue that the optimal value function is given by 2.5 and that optimal policies can be chosen to depend only on $I'_t = \max\{0, (1-D_t)I_t-1\}$, if such policies exist that solve 2.5. It is easy to see that the maximization problem 2.5 is concave and maximizing functions $x_j(I')$ and $I_j(I')$ exist. I will prove these details for the more general case with 2 items in the following section. The next lemma shows that more time left in the insurance contract is valuable:

Lemma 8. *$V(I', \varepsilon) \geq V(I'', \varepsilon)$ for any $I' > I''$ and $\varepsilon \in \mathbb{R}$.*

Proof. This proof is a simpler version of the proof of Lemma 12. Since the result does not depend on ε, I will opt for clearer notation and write all functions as if they did not depend on ε. Assume first that the optimal choice of insurance level in state I'' is l. Then,

$$\begin{aligned} V(I') &= p(x(I'))\left[u\left(\omega - L + \mathbb{1}\{I(I') > 0\}P - \mathbb{1}\{I(I') = l\}C\right) + \delta V(0)\right] \\ &\quad + (1 - p(x(I')))\left[u(\omega - \mathbb{1}\{I(I') = l\}C) + \delta V(\max\{0, I(I') - 1\})\right] - E(x(I')) \\ &\geq p(x(I''))\left[u\left(\omega - L + P - C\right) + \delta V(0)\right] \\ &\quad + (1 - p(x(I'')))\left[u(\omega - C) + \delta V(l - 1)\right] - E(x(I'')), \\ &= V(I'') \end{aligned} \tag{B.2}$$

where the inequality follows since $x(I'')$ and l are available also when the state is I'. I will show the case where $I(I'') = I''$ using induction. Assume first that $I'' = 0$. Then

$$\begin{aligned} V(0) &= p(x(0))\left[u(\omega - L) + \delta V(0)\right] \\ &\quad + (1 - p(x(0)))\left[u(\omega) + \delta V(0)\right] - E(x(0)) \\ &< p(x(0))\left[u(\omega - L + P) + \delta V(0)\right] \\ &\quad + (1 - p(x(0)))\left[u(\omega) + \delta V(0)\right] - E(x(0)) \\ &\leq V(1), \end{aligned} \tag{B.3}$$

where the first inequality follows, since u is strictly increasing and $p(x(0)) > 0$. The second follows, because both $x(0)$ and $I(1) = 1$ are available policies when the state is 1. Now, for the induction step, assume that $V(k) \geq V(k-1)$ for $1 \leq k < l-1$. It is enough to show that if $I(k) = k$, then $V(k+1) \geq V(k)$. Using estimates like those above I get that

$$\begin{aligned} V(k) &= p(x(k))\,[u(\omega - L + P) + \delta V(0)] \\ &\quad +(1 - p(x(k)))\,[u(\omega) + \delta V(k-1)] - E(x(k)) \\ &\leq p(x(k))\,[u(\omega - L + P) + \delta V(0)] \\ &\quad +(1 - p(x(k)))\,[u(\omega) + \delta V(k)] - E(x(k)) \\ &\leq V(k+1), \end{aligned}$$

where the first inequality follows from the induction assumption and the second from the fact that $x(k)$ and $I(k+1) = k+1$ are available in the state $k+1$. This completes the induction step.

Notice also that $V(I', \varepsilon) > V(I'', \varepsilon)$, if I can show that it is never optimal to buy a new contract when there is still time left in the old one. The result follows, because in that case (B.2) will be a strict inequality. Since (B.3) is strict, the whole induction argument can be argued using strict inequalities. The next lemma shows that it is indeed true that renewing contracts when there is still time left is never optimal. □

Lemma 9. *If $I' > 0$, then $I(I') = I'$.*

Proof. Define the following set of numbers:

$$\begin{aligned} W_l &= \max_x \big\{p(x)\,(u(\omega - L + P - C) + \delta V(0)) \\ &\quad +(1 - p(x))\,(u(\omega - C) + \delta V(l-1)) - E(x)\big\} \end{aligned} \tag{B.4}$$

and

$$\begin{aligned} W_I &= \max_x \big\{p(x)\,(u(\omega - L + P) + \delta V(0)) \\ &\quad +(1 - p(x))\,(u(\omega) + \delta V(I-1)) - E(x)\big\} \end{aligned} \tag{B.5}$$

where $I \in \{1, \ldots, l-1\}$. It is enough to show that $W_I > W_l$ for all $I \in \{1, \ldots, l-1\}$. Let $x(l)$ and $x(I)$ be maximizers for the problems (B.4) and (B.5), respectively. The proof is by induction over k, for W_{l-k}. Assume first that $k = 1$, $l - 1 > 0$ and assume by way of contradiction that $W_l \geq W_{l-1}$. This directly implies that $V(l-1) = W_l$. On the other hand,

$$\begin{aligned} V(l-1) &\geq W_{l-1} \\ &\geq p(x(l))u(\omega - L + P) + (1 - p(x(l)))u(\omega) \\ &\quad +p(x(l))\delta V(0) + (1 - p(x(l)))\delta W_l - E(x(l)) \\ &> p(x(l))u(\omega - L + P - C) + (1 - p(x(l)))u(\omega - C) \\ &\quad +p(x(l))\delta V(0) + (1 - p(x(l)))\delta V(l-1) - E(x(l)) \\ &= W_l, \end{aligned}$$

which clearly contradicts $V(l-1) = W_l$. The second inequality follows by noticing that $x(l)$ is available also when the state is $l-1$ and the agent is always free to re-insure in the state $l-2$ which will yield her W_l. The third inequality results from u being strictly increasing.

Assume next that $W_{l-k} > W_l$ for some $k \in \{1, \ldots, l-1\}$ and $l-k-1 > 0$. Assume again by way of contradiction that $W_l \geq W_{l-k-1}$. In other words, $V(l-k-1) = W_l$. Now, arguments identical to those above yield that

$$\begin{aligned} V(l-k-1) &\geq W_{l-k-1} \\ &\geq p(x(l-k))u(\omega - L + P) + (1 - p(x(l-k)))u(\omega) \\ &\quad + p(x(l-k))\delta V(0) + (1 - p(x(l-k)))\delta W_l - E(x(l-k)) \\ &= p(x(l-k))u(\omega - L + P) + (1 - p(x(l-k)))u(\omega) \\ &\quad + p(x(l-k))\delta V(0) + (1 - p(x(l-k)))\delta V(l-k-1) - E(x(l-k)) \\ &= W_{l-k} > W_l, \end{aligned}$$

where the second inequality follows, since $x(l-k)$ is available when the state is $l-k-1$ and the agent can re-insure in state $l-k-2$ if she gets that far without an accident. The last inequality follows from the induction assumption. The result contradicts $V(l-k-1) = W_l$ and hence completes the proof. □

By the arguments given at the end of the proof for Lemma 8, I now get the following corollary:

Corollary 3. *$V(I', \varepsilon) > V(I'', \varepsilon)$ for any $I' > I''$ and $\varepsilon \in \mathbb{R}$.*

With these results I have the ingredients for the key result of the section:

Lemma 10. *If $I(I') > I(I'') > 0$, then $p(x_j(I'), \varepsilon_j) < p(x_j(I''), \varepsilon_j)$ for all ε_j.*

Proof. Let first $I(I') = I' < l$. Then the optimal $x(I')$ solves the first-order condition:

$$\begin{aligned} \frac{E'(x(I'))}{p'(x(I'))} &= (u(\omega - L + P) - u(\omega) + \delta V(0) - \delta V(I' - 1)) \\ &< (u(\omega - L + P) - u(\omega) + \delta V(0) - \delta V(I'' - 1)) \\ &= \frac{E'(x(I''))}{p'(x(I''))}, \end{aligned}$$

where the inequality follows from the corollary above and the last equality from the first-order condition for $x(I'')$. Notice then that

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{E'(x)}{p'(x)} = \frac{E''(x)p'(x) - p''(x)E'(x)}{p'(x)^2} < 0$$

which implies that $x(I'') < x(I')$ proving the first part.

Assume then that $I(I') = l$. The first-order condition for $x(I')$ now solves:

$$\frac{E'(x(I'))}{p'(x(I'))} = (u(\omega - L + P - C) - u(\omega - C) + \delta V(0) - \delta V(I' - 1)).$$

Hence, the same proof as above will work as long as I can show that

$$u(\omega - L + P - C) - u(\omega - C) < u(\omega - L + P) - u(\omega).$$

Notice then that this inequality can be written as

$$u(\omega' + P - L) - u(\omega') < u(\omega + P - L) - u(\omega),$$

where

$$\omega' := \omega - C < \omega.$$

But then the inequality follows the concavity of u by Lemma 2, as $P - L < 0$. □

The previous lemma implies the trivial result that if only one risk type buys insurance, then more time left in the contract implies lower expected risk:

Corollary 4. *Assume that the contract is such that if $I_j > 0$ and $I_k > 0$ then $\varepsilon_j = \varepsilon_k =: \hat{\varepsilon}$. In that case $I(I', \varepsilon) > I(I'', \varepsilon) > 0$ implies that*

$$p(x(I'), \hat{\varepsilon}) = \mathbb{E}[p(x, \varepsilon) \mid I(I', \varepsilon)] < \mathbb{E}[p(x, \varepsilon) \mid I(I'', \varepsilon)] = p(x(I''), \hat{\varepsilon}).$$

The next result shows that heterogeneity of risk types in the pool of insured has the opposite effect, because less risky agents are more likely to survive longer into the contract causing a selection effect.

Lemma 11. *Consider the model without moral hazard. In other words, assume that $p(x, \varepsilon) = p(x', \varepsilon) \equiv p(\varepsilon)$ for all x. Then, if $I > I' > 0$,*

$$\mathbb{E}[p(\varepsilon) \mid I] \geq \mathbb{E}[p(\varepsilon) \mid I'].$$

Proof. Let $F_\varepsilon(e \mid I)$ be the cumulative distribution function of ε conditional on I and let $I > 0$. I will show that $F_\varepsilon(e \mid I + 1)$ first-order stochastically dominates $F_\varepsilon(e \mid I)$. The result then follows from the fact that $p(\varepsilon)$ is increasing.

For notational convenience I will reintroduce the time period into the notation. In what follows $I = I_t$ is the time left in the insurance contract at the beginning of time period t and $D_t \in \{0, 1\}$, where $D_t = 1$ is interpreted as the item getting destroyed in the end of period t.

Now,

$$\begin{aligned}
F_\varepsilon(e \mid I_t = I) &= \mathbb{P}(\varepsilon \leq e \mid I_{t-1} = I + 1, D_{t-1} = 0) \\
&= \frac{\mathbb{P}(\varepsilon \leq e, D_{t-1} = 0 \mid I_{t-1} = I + 1)}{\mathbb{P}(D_{t-1} = 0 \mid I_{t-1} = I + 1)} \\
&= \frac{\int_{-\infty}^{e} 1 - p(z) \, \mathrm{d}F_\varepsilon(z \mid I_{t-1} = I + 1)}{\int_{-\infty}^{\infty} 1 - p(z) \, \mathrm{d}F_\varepsilon(z \mid I_{t-1} = I + 1)} \\
&= F_\varepsilon(e \mid I_{t-1} = I + 1) \frac{1 - \frac{1}{F_\varepsilon(e \mid I_{t-1} = I+1)} \int_{-\infty}^{e} p(z) \, \mathrm{d}F_\varepsilon(z \mid I_{t-1} = I + 1)}{1 - \mathbb{E}\left[p(\varepsilon) \mid I_{t-1} = I + 1\right]} \\
&= F_\varepsilon(e \mid I_{t-1} = I + 1) \frac{1 - \mathbb{E}\left[p(\varepsilon) \mid \varepsilon \leq e, I_{t-1} = I + 1\right]}{1 - \mathbb{E}\left[p(\varepsilon) \mid I_{t-1} = I + 1\right]} \\
&\geq F_\varepsilon(e \mid I_{t-1} = I + 1),
\end{aligned} \tag{B.6}$$

where the inequality follows since, p is increasing and hence

$$\mathbb{E}\left[p(\varepsilon) \mid \varepsilon \leq e, I_{t-1} = I+1\right] \leq \mathbb{E}\left[p(\varepsilon) \mid I_{t-1} = I+1\right].$$

□

C Proofs for Section 2.3

It is straightforward to adapt the proofs from the section above to show that given a person with some e_j and ω_j, then policies that depend only on (t, I_1', I_2'), where $I_i' = \max\{0, (1-D_i^t)I_i^t - 1\}$, the time left in insurance policy i when entering period t, can always do at least as well as fully history dependent policies. This follows from two facts: First, the available controls after any two histories that share (t, I_1', I_2') is the same. Second, given a triplet (t, I_1', I_2'), the probability measure over the tomorrow's history h^{t+1} depends only on today's choice of controls. Hence, for any two histories that share the same (t, I_1', I_2'), one can use an argument identical to the one above to show that the continuation value from those two histories must be the same. Just as above one can then apply the results from Bertsekas (1995) to show that the optimal value of the problem satisfies the Bellman equation:

$$\begin{aligned}
& V(t, I_1', I_2') \\
= \quad & \sup_{x_1, x_2, \hat{I}_1, \hat{I}_2} \Bigg\{ \\
& p(x_1,t)p(x_2,t)\Big(u(\omega - 2L + (\mathbb{1}\{\hat{I}_1 > 0\} + \mathbb{1}\{\hat{I}_2 > 0\})P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, 0)\Big) \\
+ \quad & p(x_1,t)(1-p(x_2,t))\Big(u(\omega - L + \mathbb{1}\{\hat{I}_1 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, \max\{\hat{I}_2 - 1, 0\})\Big) \\
+ \quad & (1-p(x_1,t))p(x_2,t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, 0)\Big) \\
+ \quad & (1-p(x_1,t))(1-p(x_2,t))\Big(u(\omega - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, \max\{\hat{I}_2 - 1, 0\})\Big) - E(x_1, x_2)\Bigg\}
\end{aligned} \tag{C.1}$$

It is also straightforward to apply standard results, like those in Bertsekas (1995), to show that this functional equation has a unique solution in the space of bounded functions and that solution is the actual maximal value function. I will next show that there exists a policy that achieves the supremum.

I will start by showing the following result which shows that the agent always prefers more insurance to less insurance:

Lemma 12. *The following inequalities hold:*

1. $V(t, I_1', I_2') \geq V(t, I_1'', I_2')$ *for all* $I_1' > I_1''$
2. $V(t, I_1', I_2') \geq V(t, I_1', I_2'')$ *for all* $I_2' > I_2''$.

Proof. By the symmetry of the problem, it is clearly enough to prove just the first inequality. Since the set from which $(\hat{I}_1, \hat{I}_2)$ is chosen is finite, I can find $(\hat{I}_1, \hat{I}_2)$ and a sequence $(x_{1n}'', x_{2n}'')_{n\in\mathbb{N}}$ such that

$$\begin{aligned}
& p(x_{1n}'', t)p(x_{2n}'', t)\Big(u(\omega - 2L + (\mathbb{1}\{\hat{I}_1 > 0\} + \mathbb{1}\{\hat{I}_2 > 0\})P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, 0)\Big) \\
+ \quad & p(x_{1n}'', t)(1 - p(x_{2n}'', t))\Big(u(\omega - L + \mathbb{1}\{\hat{I}_1 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, \max\{\hat{I}_2 - 1, 0\})\Big) \\
+ \quad & (1 - p(x_{1n}'', t))p(x_{2n}'', t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, 0)\Big) \\
+ \quad & (1 - p(x_{1n}'', t))(1 - p(x_{2n}'', t))\Big(u(\omega - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, \max\{\hat{I}_2 - 1, 0\})\Big) - E(x_{1n}'', x_{2n}'') \underset{n\to\infty}{\longrightarrow} V(t, I_1'', I_2'). \qquad \text{(C.2)}
\end{aligned}$$

Consider first the case when $\hat{I}_1 = l$ achieves the supremum when the state is (t, I_1'', I_2').

Then I have:

$$
\begin{aligned}
& V(t, I_1', I_2') \\
= & \sup_{x_1, x_2, \hat{I}_1, \hat{I}_2} \Bigg\{ \\
& p(x_1,t)p(x_2,t)\Big(u(\omega - 2L + (\mathbb{1}\{\hat{I}_1 > 0\} + \mathbb{1}\{\hat{I}_2 > 0\})P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, 0)\Big) \\
+ & \ p(x_1,t)(1-p(x_2,t))\Big(u(\omega - L + \mathbb{1}\{\hat{I}_1 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, \max\{\hat{I}_2 - 1, 0\})\Big) \\
+ & \ (1-p(x_1,t))p(x_2,t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, 0)\Big) \\
+ & \ (1-p(x_1,t))(1-p(x_2,t))\Big(u(\omega - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, \max\{\hat{I}_2 - 1, 0\})\Big) - E(x_1, x_2)\Bigg\} \\
\geq & \ p(x''_{1n},t)p(x''_{2n},t)\Big(u(\omega - 2L + (1 + \mathbb{1}\{\hat{I}_2 > 0\})P - (1 + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, 0)\Big) \\
+ & \ p(x''_{1n},t)(1-p(x''_{2n},t))\Big(u(\omega - L + P - (1 + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, \max\{\hat{I}_2 - 1, 0\})\Big) \\
+ & \ (1-p(x''_{1n},t))p(x''_{2n},t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - (1 + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, l-1, 0)\Big) \\
+ & \ (1-p(x''_{1n},t))(1-p(x''_{2n},t))\Big(u(\omega - (1 + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, l-1, \max\{\hat{I}_2 - 1, 0\})\Big) - E(x''_{1n}, x''_{2n}) \qquad \text{(C.3)}
\end{aligned}
$$

for every $n \in \mathbb{N}$. The inequality follows, since $(x''_{1n}, x''_{2n}, l, \hat{I}_2)$ is also available when the state is (t, I_1', I_2'). Now the right-hand side converges to $V(t, I_1'', I_2')$ as $n \to \infty$ and hence the claim must hold when in the state (t, I_1'', I_2') the supremum is achieved with $\hat{I}_1 = l$.

To cover the case when the supremum is achieved at $\hat{I}_1 = I_1''$ when the state is (t, I_2'', I_2') I proceed by induction to show that $V(t, I_1'' + 1, I_2') \geq V(t, I_1'', I_1')$. This together with the result above clearly implies the whole lemma. Assume first that

$I''_1 = 0$ and that the supremum is attained at $\hat{I}_1 = 0$ when the state is (t, I''_1, I'_2), i.e.

$$
\begin{aligned}
V(t,0,I'_2) \; \underset{n\to\infty}{\longleftarrow} \;& p(x''_{1n},t)p(x''_{2n},t)\Big(u(\omega - 2L + \mathbb{1}\{\hat{I}_2 > 0\}P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,0)\Big) \\
+ \;& p(x''_{1n},t)(1-p(x''_{2n},t))\Big(u(\omega - L - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) \\
+ \;& (1-p(x''_{1n},t))p(x''_{2n},t)\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,0)\Big) \\
+ \;& (1-p(x''_{1n},t))(1-p(x''_{2n},t))\Big(u(\omega - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) - E(x''_{1n}, x''_{2n}) \\
\leq \;& p(x''_{1n},t)p(x''_{2n},t)\Big(u(\omega - 2L + (1+\mathbb{1}\{\hat{I}_2 > 0\})P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,0)\Big) \\
+ \;& p(x''_{1n},t)(1-p(x''_{2n},t))\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) \\
+ \;& (1-p(x''_{1n},t))p(x''_{2n},t)\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,0)\Big) \\
+ \;& (1-p(x''_{1n},t))(1-p(x''_{2n},t))\Big(u(\omega - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) - E(x''_{1n}, x''_{2n}),
\end{aligned}
$$

where the inequality follows, because u is increasing. The right-hand side is exactly the value from choosing policy $(x''_{1n}, x''_{2n}, I'_1 - 1, \hat{I}_2)$ when the state is $(t, I'_1 = I''_1 + 1, I'_2) = (t, 1, I'_2)$. Since this policy is a valid choice in that state, the value of that policy must be at most $V(t, 1, I'_2)$. Now by the convergence on the first row, I know that for every

$\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that

$$\begin{aligned} V(t,0,I_2') \leq & \; p(x_{1n}'',t)p(x_{2n}'',t)\Big(u(\omega - 2L + (1+\mathbb{1}\{\hat{I}_2 > 0\})P - \mathbb{1}\{\hat{I}_2 = l\}C) \\ & +\delta V(t+1,0,0)\Big) \\ + & \; p(x_{1n}'',t)(1-p(x_{2n}'',t))\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\ & +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) \\ + & \; (1-p(x_{1n}'',t))p(x_{2n}'',t)\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\ & +\delta V(t+1,0,0)\Big) \\ + & \; (1-p(x_{1n}'',t))(1-p(x_{2n}'',t))\Big(u(\omega - \mathbb{1}\{\hat{I}_2 = l\}C) \\ & +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) - E(x_{1n}'',x_{2n}'') + \varepsilon \\ \leq & \; V(t,1,I_2') + \varepsilon. \end{aligned}$$

Taking $\varepsilon \to 0$ yields the first step of the induction.

Next, suppose that for any $t \in \mathbb{N}$ and $k \in \{1,\ldots,l-1\}$, $V(t,k,I_2') \geq V(t,k-1,I_2')$. I want to show that $V(t,k+1,I_2') \geq V(t,k,I_2')$. If $V(t,k,I_2')$ is a limit of a sequence of effort levels when, $\hat{I}_1 = l$ then our claim is proven by the first part of the lemma. Hence assume that optimal insurance choice for state (t,k,I_2') is k. Then for every $\varepsilon > 0$ there exists $x_{1\varepsilon}$ and $x_{2\varepsilon}$ such that

$$\begin{aligned}
V(t,k,I_2') \leq\ & p(x_{1\varepsilon},t)p(x_{2\varepsilon},t)\Big(u(\omega - 2L + (1+\mathbb{1}\{\hat{I}_2 > 0\})P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,0)\Big) \\
+\ & p(x_{1\varepsilon},t)(1-p(x_{2\varepsilon},t))\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) \\
+\ & (1-p(x_{1\varepsilon},t))p(x_{2\varepsilon},t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,k-1,0)\Big) \\
+\ & (1-p(x_{1\varepsilon},t))(1-p(x_{2\varepsilon},t))\Big(u(\omega - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,k-1,\max\{\hat{I}_2 - 1, 0\})\Big) - E(x_{1\varepsilon},x_{2\varepsilon}) + \varepsilon \\
\leq\ & p(x_{1\varepsilon},t)p(x_{2\varepsilon},t)\Big(u(\omega - 2L + (1+\mathbb{1}\{\hat{I}_2 > 0\})P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,0)\Big) \\
+\ & p(x_{1\varepsilon},t)(1-p(x_{2\varepsilon},t))\Big(u(\omega - L + P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,0,\max\{\hat{I}_2 - 1, 0\})\Big) \\
+\ & (1-p(x_{1\varepsilon},t))p(x_{2\varepsilon},t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,k,0)\Big) \\
+\ & (1-p(x_{1\varepsilon},t))(1-p(x_{2\varepsilon},t))\Big(u(\omega - \mathbb{1}\{\hat{I}_2 = l\}C) \\
& +\delta V(t+1,k,\max\{\hat{I}_2 - 1, 0\})\Big) - E(x_{1\varepsilon},x_{2\varepsilon}) + \varepsilon \qquad \text{(C.4)}
\end{aligned}$$

where the inequality follows from the induction assumption. Now the right-hand side is ε plus the value from choosing $(x_{1\varepsilon}, x_{2\varepsilon}, k+1, \hat{I}_2)$ when the state is $(t, k+1, I_2')$. Since this option is available in that state, I get that $V(t,k,I_2') \leq V(t,k+1,I_2') + \varepsilon$ for all $\varepsilon > 0$. Taking the limit as $\varepsilon \to 0$ completes the induction. □

Given the value function V, define the value of choosing policy $(x_1, x_2, \hat{I}_1, \hat{I}_2)$ in period t when the insurance status is (I_1', I_2') as

$$\begin{aligned}
& W(x_1, x_2, \hat{I}_1, \hat{I}_2; t, I_1', I_2') \\
:= \; & p(x_1,t)p(x_2,t)\Big(u(\omega - 2L + (\mathbb{1}\{\hat{I}_1 > 0\} + \mathbb{1}\{\hat{I}_2 > 0\})P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& + \delta V(t+1, 0, 0)\Big) \\
+ \; & p(x_1,t)(1 - p(x_2,t))\Big(u(\omega - L + \mathbb{1}\{\hat{I}_1 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& + \delta V(t+1, 0, \max\{\hat{I}_2 - 1, 0\})\Big) \\
+ \; & (1 - p(x_1,t))p(x_2,t)\Big(u(\omega - L + \mathbb{1}\{\hat{I}_2 > 0\}P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& + \delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, 0)\Big) \\
+ \; & (1 - p(x_1,t))(1 - p(x_2,t))\Big(u(\omega - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& + \delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, \max\{\hat{I}_2 - 1, 0\})\Big) - E(x_1, x_2) \qquad \text{(C.5)}
\end{aligned}$$

In what follows I will suppress the dependence of W on (t, I_1', I_2'). The next two lemmas show that there exists an optimal effort choice vector (x_1, x_2) and that effort choice is given by the first-order condition over (x_1, x_2) at some insurance choice $(\hat{I}_1, \hat{I}_2)$.

Lemma 13. *Given any choice of insurance levels $(\hat{I}_1, \hat{I}_2)$ (and any state (t, I_1', I_2')), the mapping $(x_1, x_2) \mapsto W(x_1, x_2, \hat{I}_1, \hat{I}_2)$ is strictly concave.*

Proof. Define $\hat{W} : (x_1, x_2) \mapsto W(x_1, x_2, \hat{I}_1, \hat{I}_2) + E(x_1, x_2)$. Because E is strictly convex, it is enough to show that $\hat{W}$ is concave. I will do this by proving that $\partial_{11}\hat{W}(x_1, x_2) < 0$ and

$$\partial_{11}\hat{W}(x_1, x_2)\partial_{22}\hat{W}(x_1, x_2) - \partial_{12}\hat{W}(x_1, x_2)^2 > 0.$$

Since the notation for the general case is nearly illegible, I will prove only the case where $\hat{I}_1 = l$ and $0 < \hat{I}_2 < l$. The argument for the other cases is identical. Notice first that

$$\partial_{11}\hat{W}(x_1, x_2) = p''(x_1, t)\Big[p(x_2,t)A + B + p(x_2,t)C_1 + (1 - p(x_2,t))D_1\Big],$$

where I write $p'(x,t)$ instead of $\partial_x p(x,t)$ and where

$$A := u(\omega - 2L + 2P - C) - 2u(\omega - L + P - C) + u(\omega - C),$$

$$B := u(\omega - L + P - C) - u(\omega - C),$$

$$C_1 := \delta\left[V(t+1, 0, 0) - V(t+1, l-1, 0)\right] \qquad \text{(C.6)}$$

and

$$D_1 := \delta\Big[V(t+1, 0, \hat{I}_2 - 1) - V(t+1, l-1, \hat{I}_2 - 1)\Big]. \qquad \text{(C.7)}$$

All of these terms are negative: the negativity of A follows from concavity of u, B is negative because u is strictly increasing and C_1 and D_1 are negative by the previous lemma. Hence, $\partial_{11}\hat{W}$ is negative.

Next, notice that

$$\partial_{22}\hat{W}(x_1,x_2) = p''(x_2,t)\Big[p(x_1,t)A + B + p(x_1,t)C_2 + (1-p(x_1,t))D_2\Big],$$

where A and B were defined above and

$$C_2 := \delta\left[V(t+1,0,0) - V(t+1,0,\hat{I}_2-1)\right] \tag{C.8}$$

and

$$D_2 := \delta\left[V(t+1,l-1,0) - V(t+1,l-1,\hat{I}_2-1)\right] \tag{C.9}$$

which are also all negative. Hence, $\partial_{22}\hat{W}(x_1,x_2) < 0$. It is also easy to verify that $E := C_1 - D_1 = C_2 - D_2$. Hence, I have:

$$\begin{aligned} & \partial_{11}\hat{W}(x_1,x_2)\partial_{22}\hat{W}(x_1,x_2) \\ = \; & p''(x_1,t)p''(x_2,t)\left[p(x_2,t)(A+E)+B+D_1\right] \\ & \times\left[p(x_1,t)(A+E)+B+D_2\right] \\ = \; & p''(x_1,t)p''(x_2,t)\big|p(x_2,t)(A+E)+B+D_1\big| \\ & \times\big|p(x_1,t)(A+E)+B+D_2\big|. \end{aligned} \tag{C.10}$$

I already argued that $p(x_i,t)(A+E)+B+D_i < 0$ and that $B, D_i < 0$ for $i = 1,2$. This then implies that

$$(C.10) > p''(x_1,t)p''(x_2,t)p(x_1,t)p(x_2,t)\left(A+E\right)^2. \tag{C.11}$$

Next, consider the cross partial derivative:

$$\begin{aligned} & \partial_{12}W(x_1,x_2) \\ = \; & p'(x_1,t)p'(x_2,t)\Bigg[u(\omega-2L+2P-C) - 2u(\omega-L+P-C) + u(\omega-C) \\ & +\delta\Big(V(t+1,0,0) - V(t+1,0,\hat{I}_2-1) \\ & -V(t+1,l-1,0) + V(t+1,l-1,\hat{I}_2-1)\Big)\Bigg] \\ = \; & p'(x_1,t)p'(x_2,t)\left(A+E\right) \end{aligned} \tag{C.12}$$

Now, if I combine (C.11) and (C.12), I get that

$$\begin{aligned} & \partial_{11}W(x_1,x_2)\partial_{22}W(x_1,x_2) - \partial_{12}W(x_1,x_2)^2 \\ > \; & (A+E)^2\Big(p''(x_1,t)p''(x_2,t)p(x_1,t)p(x_2,t) - \left(p'(x_1,t)p'(x_2,t)\right)^2\Big) > 0, \end{aligned}$$

where the inequality follows from the maintained assumption that $(x_1,x_2) \mapsto p(x_1,t)p(x_2,t)$ is convex. □

Lemma 14. *For each* $(\hat{I}_1, \hat{I}_2)$ *(and* (t, I'_1, I'_2)*) there exists a unique* (x_1, x_2) *that maximizes* $W(x_1, x_2, \hat{I}_1, \hat{I}_2)$*. Furthermore, this maximizer is the unique solution to*

$$\nabla_{(x_1,x_2)} W(x_1, x_2, \hat{I}_1, \hat{I}_2) = 0 \tag{C.13}$$

Proof. For any $M > 0$, let $S_M = \{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 \leq M, x_2 \leq M\}$. I start by showing that there exists $M > 0$ such that

$$\sup_{(x_1,x_2)\in\mathbb{R}^2_+} W(x_1, x_2, \hat{I}_1, \hat{I}_2) = \sup_{(x_1,x_2)\in S_M} W(x_1, x_2, \hat{I}_1, \hat{I}_2). \tag{C.14}$$

This together with the compactness of S_M and continuity of W guarantees the existence of a maximum. To prove the equality, notice that

$$\begin{aligned}
& W(x_1, x_2, \hat{I}_1, \hat{I}_2) \leq u(\omega - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, \max\{\hat{I}_2 - 1, 0\}) - E(x_1, x_2) \\
\leq\ & u(\omega - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, \max\{\hat{I}_1 - 1, 0\}, \max\{\hat{I}_2 - 1, 0\}) - E(0, x_2) \underset{x_2\to\infty}{\to} -\infty,
\end{aligned} \tag{C.15}$$

where the first inequality follows by Lemma 12 and because u is increasing, the second inequality follows because E is increasing in both of its arguments and the limit uses the fact that V is bounded and E is increasing and strictly convex and hence $\lim_{x_2\to\infty} E(x_1, x_2) = \infty$ for all $x_1 \in \mathbb{R}_+$. Then there exists $M_2 > 0$ such that $W(x_1, x_2, \hat{I}_1, \hat{I}_2) < 0$ for any $x_1 \in \mathbb{R}_+$ and $x_2 \geq M_2$. Choosing any such x_2 is clearly suboptimal, since choosing $(0, 0)$ instead yields:

$$\begin{aligned}
W(0, 0, \hat{I}_1, \hat{I}_2) \ \geq\ & u(\omega - 2L + (\mathbb{1}\{\hat{I}_1 > 0\} + \mathbb{1}\{\hat{I}_2 > 0\})P - (\mathbb{1}\{\hat{I}_1 = l\} + \mathbb{1}\{\hat{I}_2 = l\})C) \\
& +\delta V(t+1, 0, 0) - E(0, 0) > -\infty.
\end{aligned}$$

A symmetrical argument shows that choosing a policy pair with $x_1 \geq M_1$ for some large enough M_1 is also suboptimal. Setting $M = \max\{M_1, M_2\}$ yields the equality (C.14).

Potential maximizers in the square S_M must be either on the boundary of S_M or in the interior. If I can rule out any maximizers on the boundary, then the strict concavity of W from Lemma 13 guarantees that there can be only one maximizer in the interior. Furthermore, the system of equations (C.13) is a necessary condition for an interior optimum and hence ruling out maximizers on the boundary of S_M will prove the lemma.

Notice first that by the arguments above, there exists M large enough such that, $W(x_1, x_2) < -E(0, 0)$ on the set

$$\{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 = M, x_2 \leq M\} \cup \{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 \leq M, x_2 = M\}.$$

Just like above, choosing $(0, 0)$ yields a better result and hence I am left with only the lower part of the boundary given by

$$\{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 = 0, x_2 < M\} \cup \{(x_1, x_2) \in \mathbb{R}^2_+ : x_2 = 0, x_1 < M\}.$$

By symmetry, it is enough to show that it is never optimal to set $x_1 = 0$. Because $\partial_1 E(0, x_2) = 0$, it is enough to show that $\partial_1 \hat{W}(0, x_2) > 0$, where $\hat{W}$ was defined in the proof of the previous lemma, since this would imply that $\partial_1 W(0, x_2, \hat{I}_1, \hat{I}_2) > 0$ which in turn is enough to show that setting $x_1 = 0$ is not optimal for any level of x_2. Again, the proof follows by checking all of the possible configurations of $(\hat{I}_1, \hat{I}_2)$. The steps behind all of the cases are the same and I again consider only the one where $(\hat{I}_1, \hat{I}_2) = (l, I_2')$ for some $I_2' > 0$, then it is easy to see that,

$$\partial_1 \hat{W}(x_1, x_2) = p'(x_1, t)\Big[p(x_2, t)A + B + p(x_2, t)C_1 + (1 - p(x_2, t))D_1\Big],$$

where A, B, C_1, D_1 are just as in the proof of Lemma 13. There I also argued that the term in square brackets is negative. Now the result simply follows from the assumption that p is strictly decreasing, so $\partial_1 \hat{W}(0, x_2) > 0$. □

The following result shows that at individual level effort levels are ordered the same way as is required by Proposition 4

Lemma 15. *If $0 < I_1' < I_2'$, and t, ε_j and ω_j are such that $I_{ij}(t, I_1', I_2') = I_i'$ for $i = 1, 2$ then*

$$p(x_1(t, I_1', I_2'), \varepsilon_j, t) \geq p(x_2(t, I_1', I_2'), \varepsilon_j, t)$$

The same inequality holds also, if $0 < I_1'$, t, ε_j and ω_j are such that $I_{1j}(t, I_1', I_2') = I_1'$ and $I_{2j}(t, I_1', I_2') = l$ are optimal without further restrictions on I_2'.

Proof. The proof is relatively similar to that of Lemma 6. Again, after fixing j or ε_j so that the prequisites of the lemma are met, the result does not depend on these variables and hence I can drop this dependence in favor of notational clarity. I will also write $p'(x, t)$ instead of $\partial_x p(x, t)$. Consider first an agent for whom the assumptions of the first part hold. Now since $I_{1j}(t, I_1', I_2') = I_1'$ and $I_{2j}(t, I_1', I_2') = I_2'$ are optimal it is enough to look at the first-order conditions on effort choices given these levels of insurance. Define the value from choosing effort level (x_1, x_2) as

$$\begin{aligned} W(x_1, x_2) \;:=\; & p(x_1, t)p(x_2, t)\left(u\left(\omega - 2L + P\right) + \delta V\left(t + 1, 0, 0\right)\right) \\ +\; & p(x_1, t)(1 - p(x_2, t))\left(u\left(\omega - L + P\right) + \delta V\left(t + 1, 0, I_2' - 1\right)\right) \\ +\; & (1 - p(x_1, t))p(x_2, t)\left(u\left(\omega - L + P\right) + \delta V\left(t + 1, I_1' - 1, 0\right)\right) \\ +\; & (1 - p(x_1, t))(1 - p(x_2, t))\left(u\left(\omega\right) + \delta V\left(t + 1, I_1' - 1, I_2' - 1\right)\right) \\ -\; & E(x_1, x_2) \end{aligned} \tag{C.16}$$

The first-order conditions of the agent's maximization problem are then given by:

$$\begin{aligned} \partial_1 W(x_1, x_2) \;=\; & p'(x_1, t)p(x_2, t)\left(u\left(\omega - 2L + 2P\right) + \delta V\left(t + 1, 0, 0\right)\right) \\ +\; & p'(x_1, t)(1 - p(x_2, t))\left(u\left(\omega - L + P\right) + \delta V\left(t + 1, 0, I_2' - 1\right)\right) \\ -\; & p'(x_1, t)p(x_2, t)\left(u\left(\omega - L + P\right) + \delta V\left(t + 1, I_1' - 1, 0\right)\right) \\ -\; & p'(x_1, t)(1 - p(x_2, t))\left(u\left(\omega\right) + \delta V\left(t + 1, I_1' - 1, I_2' - 1\right)\right) \\ -\; & \partial_1 E(x_1, x_2) = 0 \end{aligned} \tag{C.17}$$

and

$$
\begin{aligned}
\partial_2 W(x_1,x_2) &= p'(x_2,t)p(x_1,t)\left(u\left(\omega-2L+2P\right)+\delta V\left(t+1,0,0\right)\right)\\
&- p'(x_2,t)p(x_1,t)\left(u\left(\omega-L+P\right)+\delta V\left(t+1,0,I_2'-1\right)\right)\\
&+ p'(x_2,t)(1-p(x_1,t))\left(u\left(\omega-L+P\right)+\delta V\left(t+1,I_1'-1,0\right)\right)\\
&- p'(x_2,t)(1-p(x_1,t))\left(u\left(\omega\right)+\delta V\left(t+1,I_1'-1,I_2'-1\right)\right)\\
&- \partial_2 E(x_1,x_2)=0 \qquad \text{(C.18)}
\end{aligned}
$$

Now, define

$$
\begin{aligned}
A' &:= u(\omega-2L+P)-2u(\omega-L+P)+u(\omega)\\
&\quad +\delta V(t+1,0,0)-\delta V(t+1,0,I_2'-1)\\
&\quad -\delta V(t+1,I_1'-1,0)+\delta V(t+1,I_1'-1,I_2'-1).
\end{aligned}
$$

I will prove the lemma in two parts: first when $A'\leq 0$ and then when $A'>0$.

Assume that $A'\leq 0$. Just like in Lemma 6, equation (C.17) implicitly defines a function $\hat{x}_1(x_2)$, a level of effort for taking care of item 1 that maximizes the agents utility given that item 2 is protected at effort level x_2. I next want to show that there exists $\hat{x}\in\mathbb{R}$ such that $\partial_1 W(\hat{x},\hat{x})=0$. To see that this holds, notice first that

$$
\begin{aligned}
\partial_1 W(x,x) &= p'(x,t)\Bigg\{p(x,t)\Big[u\left(\omega-2L+P\right)-2u(\omega-L+P)+u(\omega)\\
&+ \delta V\left(t+1,0,0\right)-\delta V\left(t+1,0,I_2'-1\right)\\
&- \delta V\left(t+1,I_1'-1,0\right)+\delta V\left(t+1,I_1'-1,I_2'-1\right)\Big]\\
&+ u(\omega-L+P)-u(\omega)+\delta V(t+1,0,I_2'-1)-\delta V(t+1,I_1'-1,I_2'-1)\Bigg\}\\
&- \partial_1 E(x,x). \qquad \text{(C.19)}
\end{aligned}
$$

Denote the term in the curly brackets by T. Then,

$$
\begin{aligned}
T &\leq p(x,t)\Big[u\left(\omega-2L+P\right)-2u(\omega-L+P)+u(\omega)\\
&+ \delta V\left(t+1,0,0\right)-\delta V\left(t+1,0,I_2'-1\right)\Big]\\
&+ u(\omega-L+P)-u(\omega)<0, \qquad \text{(C.20)}
\end{aligned}
$$

where the first inequality holds, since $V(t+1,0,I_2'-1)-V(t+1,I_1',I_2')\leq 0$ by Lemma 12 and because $p(x,t)\leq 1$. The second follows because the sum of the first three terms is negative by concavity of u and the sum of the V terms is non-positive by Lemma 12. Inequality (C.20) then implies that

$$\partial_1 W(x,x)>\partial_1 E(x,x),$$

because $p'(x)<0$. Taking a limit as $x\to 0$ I get that

$$\lim_{x\to 0}\partial_1 W(x,x)>\lim_{x\to 0}\partial_1 E(x,x)=0.$$

Consider then the limit as $x \to \infty$. By assumption,

$$\lim_{x\to\infty} \partial_1 E(x,x) = \infty \text{ and } \lim_{x\to\infty} p'(x) = 0.$$

Because T is always finite, I get that

$$\lim_{x\to\infty} \partial_1 W(x,x) = 0 - \lim_{x\to\infty} \partial_1 E(x,x) = -\infty.$$

Since, $\partial_1 W(x,x)$ is continuous, there must exist $\hat{x} \in \mathbb{R}_+$ such that $\partial_1 W(\hat{x},\hat{x}) = 0$. In other words, $\hat{x}_1(\hat{x}) = \hat{x}$.

It also turns out that $\hat{x}$ is unique. This clearly follows, if I can show that $g(x) := \partial_1 W(x,x)$ is strictly monotonous. Differentiate g to get

$$g'(x) = \partial_{11} W(x,x) + \partial_{12} W(x,x). \tag{C.21}$$

Since W is concave, I also have:

$$\begin{aligned} & \partial_{11} W(x,x) \partial_{22} W(x,x) - \partial_{12} W(x,x)^2 > 0 \\ \Leftrightarrow\ & \partial_{11} W(x,x)^2 > \partial_{12} W(x,x)^2 \\ \Rightarrow\ & |\partial_{11} W(x,x)| > |\partial_{12} W(x,x)| \geq \partial_{12} W(x,x), \end{aligned} \tag{C.22}$$

where the second line follows, since E is symmetric. Note further that

$$\partial_{11} W(x,x) = p''(x,t)T - \partial_{11} E(x,x) < 0,$$

because $p'' > 0$, $T < 0$ and $\partial_{11} E > 0$. Hence,

$$|\partial_{11} W(x,x)| = -\partial_{11} W(x,x).$$

Plugging this back into (C.22) and reorganizing terms yields:

$$g'(x) = \partial_{11} W(x,x) + \partial_{12} W(x,x) < 0.$$

Hence g is strictly decreasing implying the uniqueness of $\hat{x}$.

The rest the proof for the part where $A' \leq 0$ follows an idea similar to the proof of Lemma 6. Define

$$\begin{aligned} F(x_1,x_2) \ &:= \ \partial_1 W(x_1,x_2) - \partial_2 W(x_1,x_2) \\ &= \ A'\left(p'(x_1,t)p(x_2,t) - p'(x_2,t)p(x_1,t)\right) \\ &+ \ \left(p'(x_1,t) - p'(x_2,t)\right)\left(u(\omega - L + P) - u(\omega) - \delta V(t+1, I_1' - 1, I_2' - 1)\right) \\ &+ \ p'(x_1,t)\delta V(t+1, 0, I_2' - 1) - p'(x_2,t)\delta V(t+1, I_1' - 1, 0) \\ &- \ \partial_1 E(x_1,x_2) + \partial_2 E(x_1,x_2) \end{aligned} \tag{C.23}$$

If (x_1^*, x_2^*) is the optimal effort level (which is guaranteed to be unique by the strict concavity of W), then the first-order conditions imply that $F(x_1^*, x_2^*) = 0$. Notice then that

$$F(\hat{x},\hat{x}) = p'(\hat{x},t)\delta V(t+1, 0, I_2' - 1) - p'(\hat{x},t)\delta V(t+1, I_1' - 1, 0) \leq 0 \tag{C.24}$$

because $p'(\hat{x}, t) < 0$.

Now an identical argument to the one given in Lemma 6 shows that the concavity of W implies that $F(\hat{x}_1(x_2), x_2)$ is strictly increasing in x_2. This together with (C.24) implies that $x_2^* \geq \hat{x}$ for any $\hat{x}$ that solves $\hat{x}_1(\hat{x}) = \hat{x}$. Now, suppose I was able to show that $\hat{x}_1'(x_2) < 1$. Then it would follow that

$$\begin{aligned} x_1^* = \hat{x}_1(x_2^*) = \int_{\hat{x}}^{x_2^*} \hat{x}_1'(z)\,\mathrm{d}z + \hat{x}_1(\hat{x}) \\ \leq \int_{\hat{x}}^{x_2^*} 1\,\mathrm{d}z + \hat{x}_1(\hat{x}) = x_2^* - \hat{x} + \hat{x}_1(\hat{x}) = x_2^*, \end{aligned} \tag{C.25}$$

where the second equality follows from the fundamental theorem of calculus and the last one is implied by $\hat{x}$ being a fixed point of $\hat{x}_1$. The inequality $x_1^* \leq x_2^*$ together with $p' < 0$ would clearly imply the claim of the lemma. Hence, the proof of the first part will be complete, if I can show that $\hat{x}_1'(x_2) < 1$.

Now, by the implicit function theorem,

$$\begin{aligned} \frac{d\hat{x}_1(x_2)}{dx_2} &= -\frac{\partial_{12} W(\hat{x}_1(x_2), x_2)}{\partial_{11} W(\hat{x}_1(x_2), x_2)} \\ &= -\frac{p'(x_1, t)p'(x_2, t)A' - \partial_{12} E(x_1, x_2)}{\partial_{11} W(\hat{x}_1(x_2), x_2)}. \end{aligned}$$

The numerator is clearly negative, because $p' < 0$, $A' \leq 0$ and $\partial_{12} E \geq 0$. Furthermore, I already showed in Lemma (13) that $\partial_{11} W < 0$. Consequently, $\frac{d\hat{x}_1(x_2)}{dx_2} \leq 0 < 1$ proving the first part.

Assume then that $A' > 0$ and assume by way of contradiction that $x_1^* > x_2^*$. By the first-order conditions, I must still have $F(x_1^*, x_2^*) = 0$. Notice first that, since E is convex and symmetric, it is also Schur convex and hence[24]

$$\partial_2 E(x_1^*, x_2^*) \leq \partial_1 E(x_1^*, x_2^*).$$

This together with Lemma 12 implies that

$$\begin{aligned} F(x_1^*, x_2^*) \;\leq\;& A' \left(p'(x_1^*, t)p(x_2^*, t) - p'(x_2^*, t)p(x_1^*, t)\right) \\ +\;& \left(p'(x_1^*, t) - p'(x_2^*, t)\right)\left(u(\omega - L + P) - u(\omega)\right) \\ +\;& \left(p'(x_1^*, t) - p'(x_2^*, t)\right)\left(\delta V(t+1, I_1' - 1, 0) - \delta V(t+1, I_1' - 1, I_2' - 1)\right). \end{aligned}$$

Now, $p'(x_1^*, t)p(x_2^*, t) \leq p'(x_1^*, t)p(x_1^*, t)$, because p is decreasing. Now, since $A' > 0$, I also get that

$$\begin{aligned} & A' \left(p'(x_1^*, t)p(x_2^*, t) - p'(x_1^*, t)p(x_1^*, t)\right) \\ \leq\;& A' p(x_1^*, t)(p'(x_1^*, t) - p'(x_2^*, t)) \\ \leq\;& A'(p'(x_1^*, t) - p'(x_2^*, t)), \end{aligned}$$

[24]See, for example, Theorem C on p. 259 in Roberts and Varberg (1973).

where the last inequality follows, since the the whole term is positive and the left-hand side is the right-hand side multiplied by $p(x_1, t) \leq 1$. Plugging this back into the inequality I had for $F(x_1^*, x_2^*)$ yields:

$$\begin{aligned} F(x_1^*, x_2^*) &\leq (p'(x_1^*, t) - p'(x_2^*, t)) \Big(A' + u(\omega - L + P) - u(\omega) \\ &\quad + \delta V(t+1, I_1' - 1, 0) - \delta V(t+1, I_1' - 1, I_2' - 1) \Big) \\ &= (p'(x_1^*, t) - p'(x_2^*, t)) \Big(u(\omega - 2L + P) - u(\omega - L + P) \\ &\quad + \delta V(t+1, 0, 0) - \delta V(t+1, 0, I_2' - 1) \Big) < 0, \end{aligned} \tag{C.26}$$

by the definition of A' and since $p'(x_1^*, t) > p'(x_2, t)$ as p' is increasing. This clearly contradicts $F(x_1^*, x_2^*) = 0$.

Identical arguments hold for the case when $\hat{I}_{1j} = I_1'$ and $\hat{I}_{2j} = l$ by simply replacing $u(x)$ with $u(x - C)$ for all values of x and I_2' by l everywhere in the argument above. □

Corollary 5. *If it is never optimal to buy a new contract for an item that still has time left in its contract and if* $0 < I_1' < I_2'$, *and* t, ε_j *and* ω_j *are such that* $\hat{I}_{i,j} = I_i'$ *for* $i = 1, 2$ *then*

$$p(x_1(t, I_1', I_2'), \varepsilon_j, t) > p(x_2(t, I_1', I_2'), \varepsilon_j, t)$$

The same inequality holds also, if $0 < I_1'$, t, ε_j *and* ω_j *are such that* $\hat{I}_{1j} = I_1'$ *and* $\hat{I}_{2j} = l$ *are optimal.*

Proof. It is easy to check first that the uniqueness of the optimal policy and the additional assumption of renewing contracts before the previous contract ends being suboptimal imply that Lemma 12 holds with a strict inequality. This in turn allows replacing the relevant inequalities in the proof above by strict inequalities. □

D Supporting tables

Table 6: Ownership of uninsured and insured ships by ship type

Ship type	Ships	Fraction insured	Owners	Owners w/ insured & uninsured
Drake	34,039	0.20	16,970	2,363
Badger Mk. II	22,248	0.03	11,396	705
Hurricane	21,634	0.23	10,358	1,608
Catalyst	16,909	0.12	10,448	577
Bestower	16,638	0.04	9,008	150
Hulk	16,084	0.06	12,266	179
Noctis	15,946	0.07	13,239	157
Thrasher	15,781	0.12	8,683	518
Iteron Mk. V	15,731	0.02	7,993	92
Cormorant	15,464	0.10	10,041	406

The first column of the table documents the total number of ships, the second column shows the fraction of those ships that are insured, the third column contains the number of players who own a ship of that type and the last column is the most important one, containing the number of players who own both an uninsured and an insured ship of the given type. As can be seen from the table, there is considerable variation in how large a fraction of ships of a given type is insured. These differences are likely to be driven by the ship's use. All of the ship types of which more than 10% are insured are combat ships. Since this is the riskiest activity in the game, the result is not surprising. The rest of the ships in the table are ship types with non-combative intended purpose such as mining, salvaging of destroyed ships or freighting. A peaceful main purpose translates also to a low number of individuals that own both insured and uninsured ships of that type. Hence the moral hazard estimates that use variation in risk between ships of a given type owned by the same individual are mostly driven by the dedicated combat ships.

E Proportional hazards assumption and goodness of fit

Let T_i be the destruction time of ship i and let $Z_i(t) \in \{0,1\}$ denote the insurance status of ship i. To assess the fit of the model I can first non-parametrically estimate the probability:

$$S_{0i}(t) := \mathbb{P}\left[T_i \geq t \mid Z_i(s) = 0 \; \forall s \leq t\right].$$

A standard way of estimating this is using the so called Kaplan-Meier estimator for the subsample where the insurance status is zero. Call this estimator $\hat{S}_0(t)$. The Kaplan-Meier estimator is a straightforward maximum likelihood estimator of the cumulative distribution function of accidents. For more details see Kalbfleisch and Prentice (2002). Now take any Borel measurable insurance process $z_i : [0,\infty) \to \{0,1\}$. If I assume the accident risk in the population to be given by the Cox proportional hazards model

$$\mathbb{P}\left[T \geq t \mid Z_i(s) = z_i(s) \; \forall s \leq t\right] = \exp\left[-\int_0^t \lambda(s)\exp(\beta z_i(s))\,\mathrm{d}s\right] =: S_{cox,i}(t),$$

then for $z_i \equiv 1$, I get that the survival probability is given by

$$S_{cox,1}(t) := \exp\left[-\int_0^t \lambda(s)\,\mathrm{d}s\right]^{\exp(\beta)} \equiv S_{0i}(t)^{\exp(\beta)}.$$

This suggests an estimator of $S_{cox,1}(t)$ by setting $\hat{S}_{cox,1}(t) = \hat{S}_{0i}(t)^{\exp(\hat{\beta})}$, where $\hat{\beta}$ is the standard partial maximum likelihood estimator of β. On the other hand, I can apply the Kaplan-Meyer estimator to the the population of insured ships to directly obtain an estimator of

$$\mathbb{P}\left[T_i \geq t \mid Z_i(s) = 1 \; \forall s \leq t\right] =: S_{1i}(t).$$

If I denote this other estimator by $\hat{S}_1(t)$, then the appropriateness of the Cox model can be assessed by evaluating how well the graphs of the estimated $\hat{S}_1$ and $\hat{S}_{cox,1}$ agree.

This method does not allow for many strata, as one would need to draw a separate line for each of them. Arguably, adding strata to the model can only improve the fit and hence the pictures presented below can be thought of as lower bounds for the fit of the models used, especially, in Section 4.1.

Figure 14 plots the survival curves treating all ship types equal. Even with this very crude choice the fit of the model seems extremely good. The model slightly overestimates the likelihood of survival of insured ships. The figure also illustrates the magnitude of the effect that insurance has on survival probabilities either because of adverse selection or moral hazard with insured ships having much lower survival probabilities throughout the sample.

I next apply the technique to the subpopulation of Battlecruisers, a group of combat ship types that includes also the Drake, as well as only Drakes. The results are shown in Figures 15 and 16. As can be seen from the pictures, the fit becomes the better the fewer ship types are included. Notice also that the model in Figure 16 is exactly the same model as on the fourth row of Table 4. The fit of that model is almost perfect in comparison to the completely non-parametric Kaplan-Meier estimates with only slight overestimation of survival probabilities in the middle of the analysis period.

References

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

[2] AbbringJ, Chiappori P-A, Pinquet J. 2003. Moral Hazard and Dynamic Insurance Data. Journal of the European Economic Association 1(4), 767-820.

[3] Adair J. 1984. The Hawthorne Effect: A Reconsideration of the Methodological Artifact. Journal of Applied Psychology 69(2), 334-345.

[4] Bainbridge W. 2007. The Scientific Research Potential of Virtual Worlds. Science 317, 472-476.

[5] Bertsekas D. P. 1995. Dynamic Programming and Optimal Control. Athena Scientific. Belmont.

[6] Castronova E. 2001. Virtual Worlds: A First-Hand Account of Market and Society on the Cyberian Frontier. CESifo Working Paper No. 618.

[7] Castronova E. 2004. The Price of Bodies: A Hedonic Pricing Model of Avatar Attributes in a Synthetic World. Kyklos 57(2), 173-196.

[8] Chesney T, Chuah S-H, Hoffmann R. 2009. Virtual world experimentation: An exploratory study. Journal of Economic Behavior & Organization 72(1), 618-635.

[9] Chiappori P-A, Salanié B. 2000. Testing for asymmetric information in insurance markets. J. Polit. Econ. 108(1), 56-78.

[10] Cohen A. 2005. Asymmetric Information and Learning in the Automobile Insurance Market, The Review of Economics and Statistics, 87(2), 197-207.

[11] Cohen A, Siegelman P. 2010. Testing for Adverse Selection in Insurance Markets. The Journal of Risk and Insurance (77)1, 39-84.

[12] Cox D. 1972. Regression Models and Life Tables (with discussion). J. Roy. Statist. Soc. Ser. B 34(2), 187-220.

[13] Cox D. 1975. Partial Likelihood. Biometrika 62(2), 269-276.

[14] Dionne G, Michaud P-C, Dahchour M. 2013. Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France. Journal of the European Economic Association (11)4, 897-917.

[15] Einav L, Cohen A. 2007. Estimating Risk Preferences from Deductible Choice. American Economic Review 97(3), 745-788.

[16] Einav L, Finkelstein A, Kluender R, Schrimpf P. 2015. Beyond Statistics: The Economic Content of Risk Scores. `http://www.stanford.edu/~leinav/BeyondStat.pdf`, retrieved 2015-04-08.

[17] Einav L, Finkelstein A, Levin J. 2010. Beyond Testing: Emprical Models of Insurance Markets. Annual Review of Economics 2, 311-336.

[18] Einav L, Finkelstein A, Ryan S, Schrimpf P, Cullen M. 2013. Selection on moral hazard in health insurance. American Economic Review 103(1).

[19] Einav L, Finkelstein A, Schrimpf P. Forthcoming. The Response of Drug Expenditure to Non-Linear Contract Design: Evidence from Medicare Part D. Quarterly Journal of Economics.

[20] Fang H, Keane M, Silverman D. 2008. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. Journal of Political Economy (116)2, 303-350.

[21] Fiedler M, Haruvy E. 2009. The Lab Versus the Virtual Lab and Virtual Field – An Experimental Investigation of Trust Games with Communication. Journal of Economic Behavior & Organization 72, 716-724.

[22] Fiedler M, Haruvy E, Li S. 2011. Social Distance in a Virtual World Experiment. Games and Economic Behavior 72(2), 400-426.

[23] Finkelstein A, McGarry K. 2006. Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. Am. Econ. Rev. 96(4), 938-958.

[24] Gibbons R, Murphy K. 1992. Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. Journal of Political Economy 100(3), 468-505.

[25] Golde S. 2008. The Economics of Market Thickness. PhD Thesis. Stanford University.

[26] Holmström B. 1979. Moral Hazard and Observability. The Bell Journal of Economics 10(1), 74-91.

[27] Horton J, Rand D, Zeckhauser R. 2011. The Online Laboratory: Conducting Experiments in a Real Labor Market. Experimental Economics 14, 399-425.

[28] Howden L, Meyer J. 2011. Age and Sex Composition: 2010. 2010 Census Briefs. U.S. Census Bureau. `http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf`, retrieved 2012-07-23.

[29] Kalbfleisch J, Prentice R. 2002. The Statistical Analysis of Failure Time Data. Wiley, New Jersey.

[30] Neyman J, Scott E L. 1948. Consistent estimation from partially consistent observations. Econometrica 16, 1-32.

[31] Pauly M. 1968. The Economics of Moral Hazard: Comment. The American Economic Review 58(3), 531-537.

[32] Ridder G, Tunali İ. 1999. Stratified partial likelihood estimation. Journal of Econometrics 92, 193-232.

[33] Roberts A. W., Vargerg D. 1973. Convex Functions. Academic Press. New York.

[34] Rothschild M, Stiglitz J. 1976. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. The Quarterly Journal of Economics 90(4), 629-649.

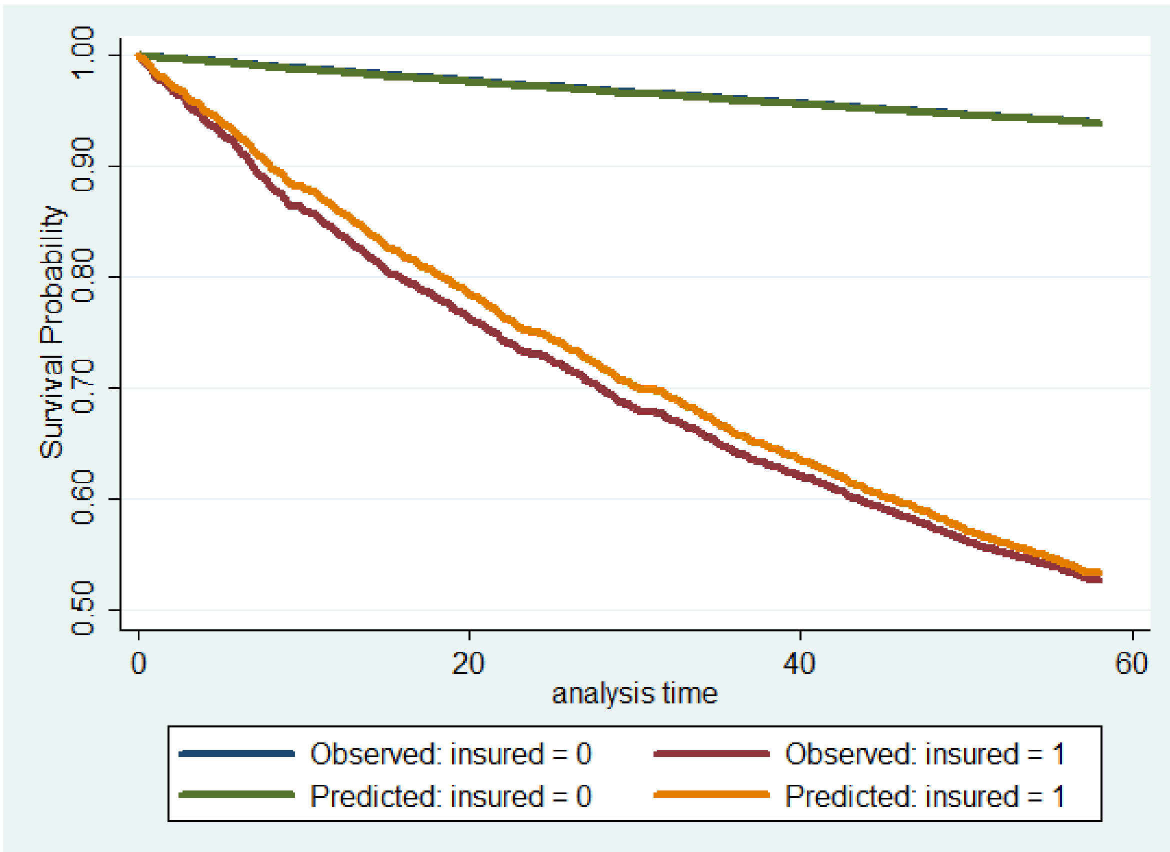

Figure 14: Kaplan-Meier and Cox proportional hazards survival curves for insured and uninsured ships.

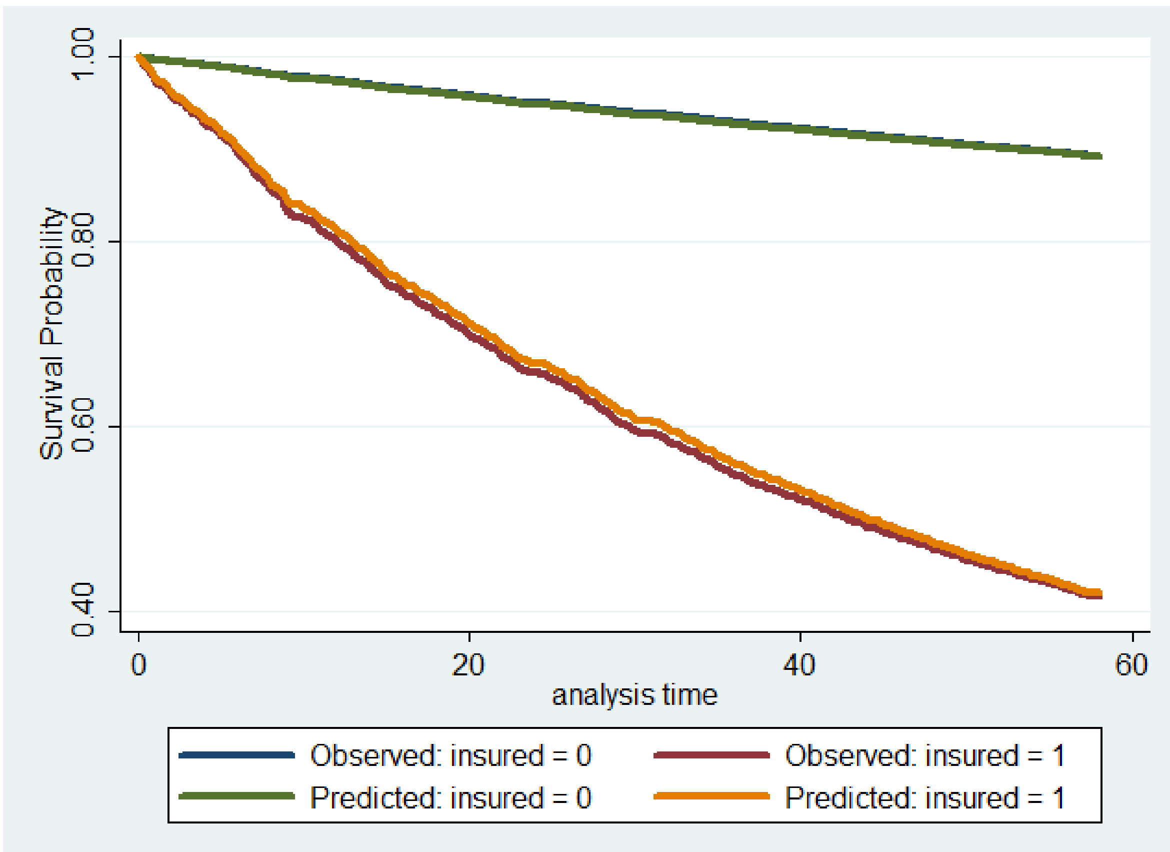

Figure 15: Kaplan-Meier and Cox proportional hazards survival curves for insured and uninsured Battlecruisers.

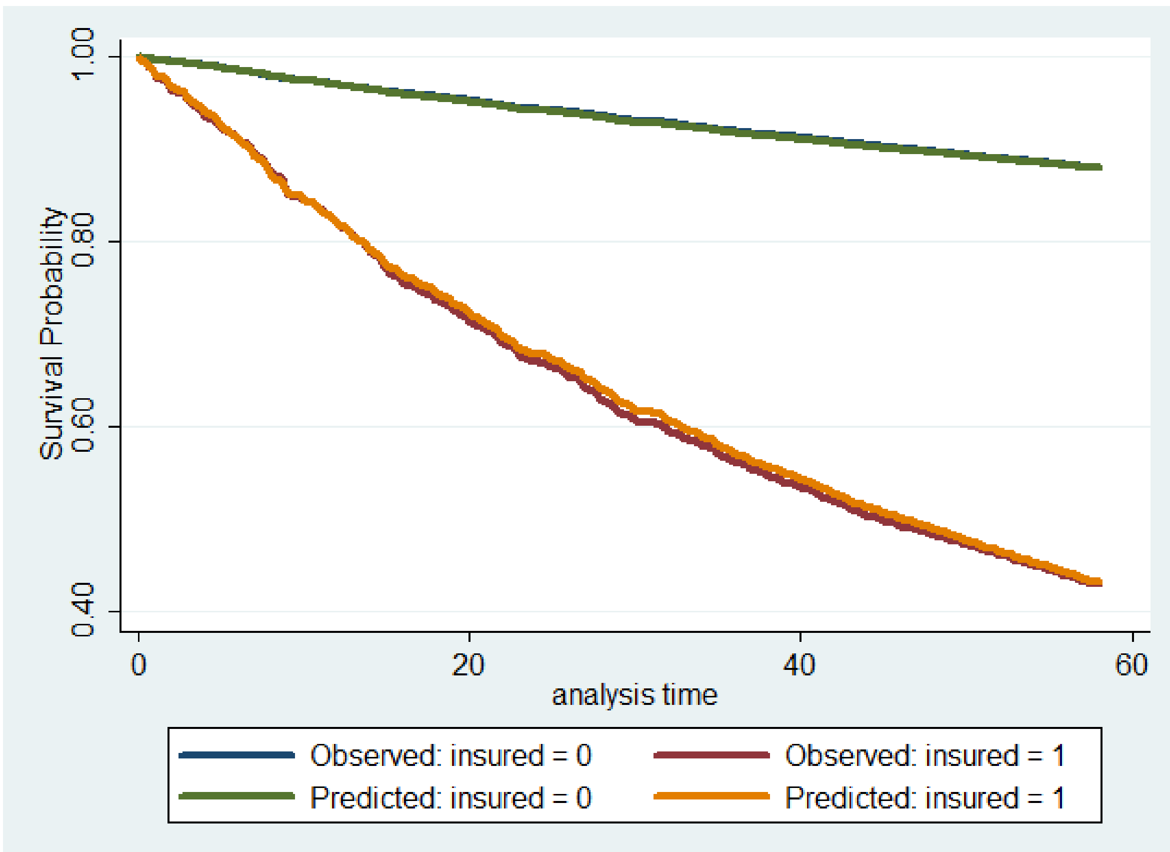

Figure 16: Kaplan-Meier and Cox proportional hazards survival curves for insured and uninsured Drakes.