Asymmetric Information in Secondary Insurance Markets: Evidence from the Life Settlement Market

Daniel Bauer, Jochen Russ, Nan Zhu*

Abstract

We use data from a large US life expectancy provider to test for asymmetric information in the secondary life insurance—or *life settlement*—market. We compare realized lifetimes for a subsample of settled policies relative to all (settled and non-settled) policies, and find a positive settlement-survival correlation indicating the existence of informational asymmetry between policyholders and investors. Estimates of the "excess hazard" associated with settling show the effect is temporary and wears off over approximately eight years. This indicates individuals in our sample possess private information with regards to their near-term survival prospects and make use of it, which has economic consequences for this market and beyond.

*Bauer (corresponding author): Department of Economics, Finance, and Legal Studies, University of Alabama, Tuscaloosa, AL 35487, dbauer@cba.ua.edu. Phone: +1-(205)-348-8486; Fax: +1-(205)-348-0590. Russ: Institut für Finanz- und Aktuarwissenschaften and Ulm University, Lise-Meitner-Straße 14, 89081 Ulm, Germany, j.russ@ifa-ulm.de. Zhu: Department of Risk Management, Pennsylvania State University, University Park, PA 16802, nanzhu@psu.edu. We are grateful for support from Fasano Associates, especially to Mike Fasano. Moreover, we thank Lauren Cohen, Liran Einav, Ken Froot, Daniel Gottlieb, Lu Mao, Stephen Shore, Petra Steinorth, as well as seminar participants at the 2015 ASSA meetings, the 2014 ARIA meeting, DAGStat2014, 2014 ESEM, the 2014 NBER insurance workshop, the 2016 WEAI conference, Georgia State University, LUISS Guido Carli, Pennsylvania State University, St. John's University, Towson University, the University of Alabama, the University of Cincinnati, the University of Connecticut, the University of Georgia, the University of Illinois Urbana-Champaign, the University of Iowa, the University of Nebraska-Lincoln, and the University of St. Thomas. Bauer and Zhu also gratefully acknowledge financial support from the Society of Actuaries under a CAE Research Grant.

1 Introduction

Asymmetric information in insurance markets is an important and intensive area of research.[1] This paper makes two primary contributions to the existing body of knowledge. First, we provide evidence for asymmetric information in the secondary life insurance market—the market for so-called *life settlements*—between policyholders and investors. To the best of our knowledge, this is the first empirical study of informational frictions in a secondary personal insurance market.[2] This complements research from primary insurance markets, where the decision problem is different in nature but the underlying risk is the same. Second, we are able to characterize the evolution of the informational friction over time. While there is a significant impact immediately after selling the insurance coverage, the effect dissipates over approximately eight years. This suggests that the policyholders in our sample are competent in evaluating their own *relative* survival prospects over the near future, in a situation where they are prompted with relevant information and where there are significant monetary consequences to their decision. This complements research from the behavioral literature suggesting that individuals fare poorly at appraising their own mortality.

Within a life settlement, a policyholder sells—or *settles*—her life-contingent insurance payments for a lump sum to a life settlement (LS) company, where the offered price depends on an individualized estimation of her survival probabilities by a third party *life expectancy* (LE) *provider*. Clearly, ceteris paribus, an LS company will pay more for a life insurance policy with shorter estimated life expectancy since, on average, survival-contingent premiums have to be paid for a shorter period whereas the death benefit is disbursed sooner. The company profits from a short realized lifespan relative to the estimate. The policyholder, on the other hand, benefits from a life expectancy estimate that is (too) short—whereas she may walk away from the transaction if the estimate notably overstates her true life expectancy. This wedge creates the possibility of asymmetric information between the policyholder and the life settlement company influencing the transactions.

We use the dataset of a large US LE provider to test for this informational asymmetry. Leaning on the literature that studies asymmetric information in primary insurance markets, we derive a test that hinges on the correlation between selling insurance coverage and (ex-post) risk. We find that individuals selling their policy live significantly longer (relative to their LE estimate)

[1] While seminal theoretical contributions have emphasized the importance of informational frictions since the 1960s (Arrow, 1963; Akerlof, 1970; Rothschild and Stiglitz, 1976; Stiglitz and Weiss, 1981), the corresponding empirical literature has flourished only relatively recently (Puelz and Snow, 1994; Cawley and Philipson, 1999; Chiappori and Salanié, 2000; Dionne et al., 2001; Cardon and Hendel, 2001; Finkelstein and Poterba, 2004; Finkelstein and McGarry, 2006; Cohen and Einav, 2007; Einav et al., 2010b, among others).

[2] Our findings are in line with a recent industry study by Granieri and Heck (2014) that postdates earlier drafts of our paper. More precisely, based on simple comparisons of survival curves for different populations, the authors conclude that within the life settlement market "insureds use the proprietary knowledge of their own health to select against the investor."

than those retaining the insurance coverage, providing evidence that individuals possess private information regarding their mortality prospects. It is important to distinguish our result from the notion that individuals wishing to sell their insurance coverage, as a group, live longer, e.g. because they are wealthier per se[3] or because the absence of dependents requiring protection implies the availability of resources to spend on their own care. Rather, what we find is that among those seeking out the opportunity to sell their policy, those deciding to pull the trigger will on average live longer, conditional on all observables. The identification then relies on the idea that for two individuals with the same observable characteristics, the quoted price will be more attractive to the one (privately) expecting a longer life, ceteris paribus. Example calculations for an average 75-year old male policyholder suggest that the effect amounts to roughly half a year of additional expected lifetime, relative to a life expectancy of a little over ten years, although this result is sensitive to underlying assumptions.

To analyze the pattern of the deviation between the two groups, we derive non-parametric estimates of the *excess hazard* (or excess mortality) for policyholders choosing to settle. These show that the difference in the hazard is most pronounced immediately after settling the policy but wears off over the course of roughly eight years. Survival regressions confirm this observation: When including a time trend interacted with the settlement dummy, the model fit improves markedly and the effect becomes stronger at settlement but weakens over time, zeroing after the same approximately eight-year time frame. Thus, while there is a large asymmetry immediately after selling the insurance coverage, the influence of the factors leading to the difference in mortality dissipates over time. This structure is in line with adverse selection with regards to the policyholder's initial condition as the origin of the asymmetry, but not with other explanations such as permanent changes in behavior, confounding attributes, or information gains due to the transaction process. A key conclusion is that individuals participating in the life settlement market appear competent in evaluating their propensity to survive in the near future.

While the basic empirical approach and the basic results are straightforward, there are a number of aspects in the identification process that warrant further investigation. In particular, our comparison group includes both non-settled and (unknown) settled cases, which complicates identifying the qualitative and the quantitative impact of asymmetric information. With regards to the quantitative impact, we derive "correction" formulae that give estimates for more relevant settled- vs. non-settled comparison given certain assumptions.[4] With regards to the qualitative impact, we are able to stave off pressing concerns on sample selection and omitted variables through a

[3]Face values in life settlement transactions are relatively large (3.92 million USD on average in our sample), indicating that market participants are generally relatively wealthy.

[4]We note that this situation of a mixed comparison sample is similar to what individual investors in the LS market face, since they know what policies they purchased and bid on, but do not generally know what happened to the policies that they did not submit a (successful) bid on. Hence, our econometric approach is relevant to them as well.

combination of theoretical and empirical robustness considerations.

More precisely, although our sample of settled policies covers a significant portion of the entire market, one concern is that the sample of *known* settled policies differs from the (unknown) set of *all* settled policies in some relevant and systematic way. Since we are controlling for all observables used by the LE provider, these would need to be additional characteristics, such as details on the policies or a second LE estimate from another provider. We have policy face value available for a fraction of the sample, and robustness analyses that include it as a covariate reveal the same significant patterns. Moreover, we show theoretically that additional information on the individuals' mortality, e.g. a second LE estimate, will lead to a bias *against* our results *if* the proportion of settlements is increasing in mortality—which is true in the data. Additional robustness analyses include considering modified samples that exclude early deaths in the comparison group and repeating the analysis using the latest observation date for a policy in our LE database to address concerns on (post) selection bias based on survival experience. Again our findings are robust.

Related Literature and Organization of the Paper

Our paper relates to the large literature on asymmetric information in insurance markets (see Footnote 1 for a list of references). In this context, several contributions highlight the merits of insurance data for testing theoretical predictions (Cohen and Siegelman, 2010; Chiappori and Salanié, 2013), although heterogeneity along multiple dimensions may impede establishing or characterizing informational asymmetries (Finkelstein and McGarry, 2006; Cohen and Einav, 2007; Cutler et al., 2008; Fang et al., 2008). We contribute by carrying out tests in a secondary insurance market, which offers the same benefits of insurance data but considers a different decision problem—namely selling rather than purchasing insurance coverage. To our knowledge, this aspect has not been explored thus far.

Our results are of immediate interest and have implications for the life settlement market, for instance in view of pricing the transactions (Zhu and Bauer, 2013) and regarding equilibrium implications (Daily et al., 2008; Fang and Kung, 2017). In addition, our findings corroborate empirical results from the primary life insurance market that policyholders, or at least a subset of policyholders, possess superior information regarding their mortality prospects (He, 2009; Wu and Gan, 2013). We complement these studies in that we are able to provide insights on the characteristics of the informational advantage.

More broadly, our results provide positive evidence on individuals' ability to make financial decisions that depend on their mortality prospects. This contrasts research from the behavioral literature comparing individual forecasts of *absolute* life expectancies to actuarial estimates, which suggests that individuals fare poorly at appraising their own mortality prospects (Elder, 2013;

Payne et al., 2013; and references therein). Our results indicate that individuals participating in the life settlement market are competent in evaluating their *relative* life expectancy, when prompted with relevant information on population mortality. This may be the more material task in situations where there are significant monetary consequences and when appropriate "default" choices that are suitable for average individuals are provided, such as retirement planning.

In what follows, we first provide background information on life settlements and the possible relevance of asymmetric information in this market. We then describe our dataset and our basic empirical approach. The next sections present our analysis of the time trend of the informational asymmetry and a variety of robustness analyses. Section 6 discusses the impact and the origin of the informational friction, and the final section concludes. An online appendix collects details on derivations and supplemental results.

2 Life Settlements and Asymmetric Information

Within a life settlement transaction, a policyholder offers her life insurance contract, typically via a broker, to an LS company. Based on individual LE reports (typically two) from established LE providers, the company then prepares an offer. If the offer is accepted, the policy—and, particularly, all life-contingent insurance benefits and premiums—will be transferred to the company, who then holds it in its own portfolio or on behalf of capital market investors. Emerging from so-called *viatical settlements* with terminally ill insureds in the 1980s, a typical life settlement transaction involves senior policyholders with a below average life expectancy. According to recent industry figures, in 2016 the total market volume amounted to approximately USD 100 billion in face value, which is less than one half percent of the total US life insurance market (Roland, 2016).

As indicated in the Introduction, an LS company will pay more for a policy with shorter life expectancy, ceteris paribus, and profits from a relatively short realized lifespan. The policyholder, on the other hand, gains from a short life expectancy estimate relative to her true (privately) expected lifespan. This creates the possibility for asymmetric information affecting the transactions. To illustrate, we consider a simple one-period model. We assume that at time zero, the policyholder is endowed with a one-period term-life insurance policy that pays \$1 at time one in case of death before time one and nothing in case of survival thereafter. The probability for dying (mortality probability) before time one is $\mathbb{P}(\tau < 1) = q$, where τ is the time of death.[5]

Suppose the policyholder is offered a life settlement at price π. For simplicity, we assume she assesses her settlement decision $\Delta = \boldsymbol{1}_{\{\text{policyholder settles}\}}$ by comparing the settlement price to the

[5]While the model is very simple, it illustrates the basic points and it facilitates the discussion of robustness of our results in Section 5. In particular, we provide an extension that delivers less obvious implications in Appendix A.3.

present value of her contract (the risk-free rate is set to zero):

$$\Delta = 1 \Leftrightarrow \pi > q - \psi, \tag{1}$$

and ψ characterizes the policyholder's proclivity for settling.[6] The latter may originate from risk-averse policyholder preferences with different bequest motives, liquidity constraints, etc. Here, we simply use ψ to capture deviations from a value-maximizing behavior, under which the market may unravel due to a "lemons problem" as in Akerlof (1970). The key assumption is that the policyholder is more likely to settle when offered a higher price.

Thus, from the policyholder's perspective, the question of whether or not to settle the policy based on Equation (1) is deterministic. However, this will not be the case from the perspective of the LS company offering to purchase the policy since it will have imperfect information with respect to q and/or ψ.[7] More precisely, assume that the policyholder has private information on the mortality probability q but the LS company solely observes the expected value, $\mathbb{E}[q]$, conditional on various observable characteristics such as age, medical impairments, etc. Then we obtain for the mortality probability conditional on the observation that the policyholder settles her policy:

$$\mathbb{P}(\tau < 1|\Delta = 1) = \mathbb{E}\left[q|\Delta = 1\right] = \mathbb{E}\left[q\left|q < \pi + \psi\right.\right] \leq \mathbb{E}[q] = \mathbb{P}(\tau < 1). \tag{2}$$

Hence, if there exists private information on q, we will observe a *negative* relationship between settling and dying.

Note that we can alternatively represent the result in (2) as:

$$\mathbb{E}[\,\boldsymbol{1}_{\{\tau<1\}}\,\Delta] - \mathbb{E}[\,\boldsymbol{1}_{\{\tau<1\}}]\,\mathbb{E}[\Delta] \leq 0 \Leftrightarrow \text{Corr}\left(\Delta,\ \boldsymbol{1}_{\{\tau<1\}}\right) \leq 0 \Leftrightarrow \text{Corr}\left(\Delta,\ \boldsymbol{1}_{\{\tau\geq1\}}\right) \geq 0. \tag{3}$$

Therefore, this is simply a version of the well-known *correlation test* for the presence of asymmetric information that examines whether (ex-post) risk and insurance coverage are positively related (Chiappori and Salanié, 2000, 2013). However, since we are considering secondary market transactions, the mechanism is reversed: A policyholder will be more inclined to settle—i.e., sell—her policy if she is a low risk from the insurer's perspective—i.e., if she has a low probability of dying. The intuition for this result is straightforward: If the policyholder has private insights on her lifetime distribution, she will gladly agree to beneficial offers from her perspective while she will

[6]We do not consider partial settlement. While private information may affect the contract choice in theory, the possibility of owning multiple policies, the non-exclusivity of the contractual relationship, and the presence of different sources of uncertainty (q and ψ) may hinder screening. Importantly, partial settlements are not common in the marketplace.

[7]Of course, such an informational asymmetry may affect the pricing of the transaction, i.e. the choice of π. We refer to Zhu and Bauer (2013) for a corresponding analysis. Here, we focus on the implications when the settlement price is given.

walk away from bad offers; hence, a pool of settled policies, on average, will display relatively longer life expectancies than the entire population of policyholders that considered selling their policy, controlling for observables.

Asymmetric information with respect to ψ alone, e.g. arising from heterogeneous preferences or liquidity constraints, cannot yield a negative relationship. However, it is possible that there exists an indirect relationship in case ψ itself is related to the lifetime distribution. For instance, the policyholder's risk aversion or wealth reflected in ψ may be positively linked to her propensity to survive—although for wealth such a relationship would arguably work in the opposite direction since more financially constrained individuals are more likely to settle. In any case, a negative relation between settling and dying will—directly or indirectly—originate from an informational asymmetry with respect to the time of death, and hence our basic empirical approach analyzes this relationship.

3 Data and Empirical Approach

To test for the negative relationship, we analyze the impact of settling on realized future lifetime based on individual survival data. Our primary dataset consists of $N = 53,947$ distinct lives underwritten by Fasano Associates (Fasano), a leading US LE provider, between beginning-of-year 2001 and end-of-year 2013. More precisely, we are given survival information for each individual and, particularly, the realized death times for individuals that died before January 1st, 2015. In addition to their lifetimes, we are given individual characteristics including sex, age, smoking status, primary impairment, as well as one or more life expectancy estimates at certain points in time. Here, the LE provider calculates the LE estimate by applying an individual mortality multiplier (*frailty factor*)—which is the result of the underwriting process—to a given proprietary mortality table. Therefore, we can use the LE estimate in combination with the underlying table (also provided by Fasano) to derive the mortality multiplier, and then use it to obtain the estimated hazard, $\hat{\mu}_t^{(i)}$, for each individual. All-in-all, there are $140,257$ LE evaluations, so many of the lives occur multiple times in the dataset. Since we are interested in the influence of informational frictions on the settlement decision, we focus on the earliest underwriting date for each individual since it serves as a proxy for the decision time.[8] Table 1 provides summary statistics.

This dataset contains LEs for policyholders that decided to settle (*close*) their policy, LEs for policyholders that walked away from a settlement offer, and LEs for individuals that were underwritten for different reasons, such as LEs for newly issued life insurances or for existing workers' compensation portfolios. The LE provider typically does not receive feedback on whether or not

[8] We discuss the impact of deviations between the earliest underwriting date and the settlement decision time in the context of our robustness analyses in Section 5.3.

	Average (Std. Dev.)			Count (Perc.)	
	Full	Closed		Full	Closed
Life Expectancy Estimate			*Male*		
	11.83	11.54		34,182	8,767
	(4.28)	(4.00)		(63.36%)	(66.31%)
Underwriting Age			*Observed Deaths*		
	75.10	75.40		13,418	3,415
	(7.43)	(6.50)		(24.87%)	(25.83%)

Table 1: Summary statistics for the entire dataset ("Full"; 53,947 lives, earliest observation date) and the subsample of closed cases ("Closed"; 13,221 lives, earliest observation date).

a policy closed, so that this aspect is unknown for our full dataset—and it is clearly unknown (not yet known) when compiling the initial life expectancy estimate. However, we also have access to a secondary dataset of overall $13,221$ lives underwritten by Fasano (and several policies not underwritten by Fasano) from individual investors as well as from a large service provider in the life settlement market. For this subsample of policyholders, we have the additional information that they settled their policy. We will refer to this secondary dataset as the subsample of *closed* cases, whereas we will refer to the rest as the *remaining* sample. Corresponding summary statistics are also provided in Table 1. We note that when relying on average face values, our sample of closed policies exceeds half of market share based on the estimate by Roland (2016)—and on its own far exceeds earlier estimates of total market size (e.g. by the research firm Conning, see Cohen (2013)). Thus, we cover a significant portion of the total life settlement market. Furthermore, since more than 90% of the sample comes from a third-party service provider that handles policy origination and policy servicing (premium payments, annual reviews, valuation, etc.) for a broad set of firms, our sample is not affected by idiosyncrasies of a single or a small number of investors.

Our empirical strategy follows studies of asymmetric information in primary insurance markets: We regress ex-post realized risk on ex-ante coverage (Cohen and Siegelman, 2010). If, conditional on all observed covariates, coverage has a positive and significant influence on risk, one can infer the existence of asymmetric information. In the setting of a secondary life market, risk is given by the realized death time, whereas (elimination of) coverage is given by the settlement decision. Thus, we analyze the impact of settling on hazard.

We first consider a conventional proportional hazards model (we alternatively analyze additive specifications to establish robustness in Section 5.1). More precisely, we assume the hazard for individual i, $\mu_t^{(i)}$, satisfies:

$$\mu_t^{(i)} = \beta_0(t) \times \exp\Bigg\{\beta_1 \ln(\hat{\mu}_t^{(i)}) + \beta_2 \ln(1+\mathrm{DOU}_i) + \beta_3 \ln(1+\mathrm{AU}_i) + \beta_4\,\mathrm{SE}_i + \sum_{j=1}^{15} \beta_{5,j}\,\mathrm{PI}_{i,j} + \sum_{j=1}^{2} \beta_{6,j}\,\mathrm{SM}_{i,j} + \gamma\,\mathrm{SaO}_i\Bigg\},\ 1 \le i \le N. \tag{4}$$

Here $\beta_0(t)$ is a non-parametric term. $\hat{\mu}_t^{(i)}$ is the estimated hazard recovered from the provider's LE assessment. DOU_i is the underwriting date, measured in years and normalized so that zero corresponds to January 1st, 2001. AU_i is the individual's age at underwriting, measured in years. SE_i is a sex dummy, zero for female and one for male. $\mathrm{PI}_{i,j}, j = 1, \ldots, 15$, are primary impairment dummies for various diseases.[9] $\mathrm{SM}_{i,j}, j = 1, 2$, are smoker dummies, where $\mathrm{SM}_{i,1} = 1$ for a smoker and $\mathrm{SM}_{i,2} = 1$ for an "aggregate" (unknown/uncertain smoking status) entry.

We include all covariates that are available for the full dataset in our regression (4) (Chiappori and Salanié (2000) and Dionne et al. (2001) emphasize the importance of incorporating all pricing-relevant variables in asymmetric information tests). The estimated hazard $\hat{\mu}_t^{(i)}$ serves both to capture the basic shape of the mortality curve over time and to pick up the information from the underwriting process. Hence, the coefficients for age, sex, primary impairments, etc. reflect residual effects beyond the LE provider's estimate. We include log-linear effects for underwriting date and age for ease of presentation and interpretation; specifications with dummies are provided in the Appendix (see columns [C] and [F] in Table A.1 and Figure A.1). We omit information that is only available for a fraction of the individuals in our basic regressions. However, we run checks including these variables and address the possible impact of omitted variables and sample selection issues in our robustness analyses (Sec. 5).

Finally, we include a Settled-and-Observed dummy SaO_i that is set to one for the subsample of closed cases and zero otherwise. We test for asymmetric information by inferring whether the estimate $\hat{\gamma}$ for the corresponding coefficient is *negative* and significant. Since the life expectancy for individual i is (Bowers et al., 1997):

$$LE_i = \mathbb{E}[\tau_i] = \int_0^\infty \underbrace{\exp\left\{-\int_0^t \mu_s^{(i)}\,ds\right\}}_{=\mathbb{P}(\tau_i > t)} dt,$$

where τ_i is the individual's remaining lifetime, a negative coefficient γ *increases* life expectancy, yielding the positive settlement-survival correlation indicative of asymmetric information (see Eq. (3) in Sec. 2). Note that the remaining cases include policyholders that rejected the settlement

[9]We do not list the primary impairments to protect proprietary information of our data supplier since they are not material to our results.

offer as well as individuals that settled but are not contained in our closed cases and individuals that were underwritten for other reasons. Thus, we actually compare closed cases relative to a mix of closed and non-closed cases, and analyzing the difference presents a more conservative test than when directly comparing closed versus non-closed cases.[10]

We rely on the conventional partial maximum likelihood method to estimate the coefficient vector (Cox, 1975). Column [A] in Table 2 presents the results. The estimated hazard $\hat{\mu}_t^{(i)}$ is highly significant, with a coefficient $\hat{\beta}_1$ of around 0.9—and, thus, close to one as would be the case for (ex-post) "perfect" estimates by the LE provider. Indeed, most of the primary impairment dummies are insignificant, suggesting that the provider's estimates adequately debit for the corresponding conditions, with a couple of exceptions. However, the regression results also show that underwriting date, age, sex, and smoking status have a significant influence beyond their inclusion in $\hat{\mu}_t^{(i)}$. For age and smoking status (both positive), this may be a consequence of $\hat{\beta}_1$ being less than one, whereas the positive impact of the underwriting date may originate from the estimates becoming more conservative (lower) over time. This is broadly consistent with analyses of the provider's performance in Bauer et al. (2017); we refer to that paper for corresponding details.

As for the Settled-and-Observed variable that is in the focus of our analysis, the corresponding coefficient estimate is negative and highly significant. More precisely, we find that for two individuals with otherwise the same observables that are both included in our dataset, the one that is known to have settled her policy will exhibit a $1 - e^{\hat{\gamma}} \approx 11.3\%$ *lower* hazard—and thus will, on average, live longer. Thus, we find a strong negative relationship between settlement and mortality, which indicates the existence of asymmetric information in the life settlement market. In particular, individuals possess private information on their survival prospects and make use of it in their settlement decision.

Aside from its relevance to the life settlement market, this result complements analyses of asymmetric information in primary life insurance markets, where several papers fail to find evidence for the existence of asymmetric information based on correlation tests (Cawley and Philipson, 1999; McCarthy and Mitchell, 2010). As discussed in detail by Finkelstein and Poterba (2014), these results may originate from (unobserved) related confounding factors such as risk aversion or wealth also affecting insurance decisions, or also from risk factors not included in the pricing—so that researchers may fail to reject the null hypothesis of symmetric information within a correlation test even if there exists private information about risk type. For example, underwriting is limited in certain segments of the primary market (such as life annuities) and regulation in some instances restricts factors that can be considered in pricing (such as gender or genetic

[10]As a consequence, our point estimate $\hat{\gamma}$ will need to be inflated to account for the mixed nature of the sample in order to present a suitable point estimate for the latter (closed vs. non-closed) comparison. Online Appendix A.1 provides more details on this issue and derives inflation formulae.

	[A]	[B]	[C]	[D]	[E]	[F]
$^1/_{14} \times \int_0^{14} \beta_0(t)\, dt$	0.0187	0.0184	0.0082	0.0682	0.0140	0.0230
Estimated hazard, $\hat{\mu}_t^{(i)}$	0.8986^{***}	0.8968^{***}	0.8873^{***}	0.8496^{***}	0.8914^{***}	0.8787^{***}
	(0.0101)	(0.0101)	(0.0170)	(0.0245)	(0.0103)	(0.0097)
Underwriting date, $\ln(1+\mathrm{DOU}_i)$	0.3101^{***}	0.3043^{***}	0.4008^{***}	0.0037	0.3032^{***}	0.2480^{***}
	(0.0277)	(0.0277)	(0.0559)	(0.1742)	(0.0280)	(0.0259)
Age at underwriting, $\ln(1+\mathrm{AU}_i)$	0.5796^{***}	0.5852^{***}	0.5282^{***}	0.3333^{*}	0.6273^{***}	0.5364^{***}
	(0.0827)	(0.0828)	(0.1458)	(0.1860)	(0.0846)	(0.0819)
Sex, SE_i	-0.1022^{***}	-0.0986^{***}	-0.1002^{***}	-0.1173^{**}	-0.1012^{***}	-0.0735^{***}
	(0.0196)	(0.0196)	(0.0338)	(0.0562)	(0.0198)	(0.0196)
Primary impairment 1, PI_1	-0.0268	-0.0414	0.6292	0.2387	0.0371	0.0958
	(0.2863)	(0.2863)	(1.0090)	(0.1740)	(0.2975)	(0.3242)
Primary impairment 2, PI_2	-0.0186	-0.0310	0.4725	0.1681	0.0382	0.1767
	(0.2894)	(0.2894)	(1.0095)	(0.1747)	(0.3009)	(0.3254)
Primary impairment 3, PI_3	0.1722	0.1618	0.8950	0.4860^{***}	0.2254	0.2311
	(0.2835)	(0.2835)	(1.0062)	(0.1812)	(0.2947)	(0.3219)
Primary impairment 4, PI_4	0.2012	0.1921	0.9694	0.7312^{***}	0.2449	0.3068
	(0.2854)	(0.2854)	(1.0067)	(0.1495)	(0.2967)	(0.3226)
Primary impairment 5, PI_5	0.0102	-0.0031	0.6858	0.4032^{***}	0.0532	0.1868
	(0.2812)	(0.2812)	(1.0051)	(0.1353)	(0.2925)	(0.3195)
Primary impairment 6, PI_6	-0.0235	-0.0348	0.7388	0.2516^{**}	0.0342	0.0582
	(0.2793)	(0.2793)	(1.0030)	(0.1281)	(0.2906)	(0.3181)
Primary impairment 7, PI_7	0.2090	0.1971	0.9404	0.5932^{***}	0.2642	0.3144
	(0.2788)	(0.2788)	(1.0025)	(0.1067)	(0.2901)	(0.3176)
Primary impairment 8, PI_8	0.2190	0.2072	1.0213	0.6967^{***}	0.2779	0.2622
	(0.2811)	(0.2811)	(1.0044)	(0.1622)	(0.2923)	(0.3197)
Primary impairment 9, PI_9	0.8307^{***}	0.8187^{***}	1.4796	0.9988^{***}	0.8908^{***}	0.8779^{***}
	(0.2820)	(0.2820)	(1.0054)	(0.1435)	(0.2932)	(0.3199)
Primary impairment 10, PI_{10}	0.1125	0.0999	0.8266	0.3712^{***}	0.1663	0.2048
	(0.2812)	(0.2812)	(1.0044)	(0.1393)	(0.2925)	(0.3199)
Primary impairment 11, PI_{11}	0.9088^{***}	0.8981^{***}	1.7950^{*}	1.5868^{***}	0.9548^{***}	1.0873^{***}
	(0.2837)	(0.2837)	(1.0061)	(0.1642)	(0.2950)	(0.3218)
Primary impairment 12, PI_{12}	0.4133	0.4003	1.1213	0.8346^{***}	0.4571	0.5142
	(0.2818)	(0.2818)	(1.0051)	(0.1560)	(0.2931)	(0.3203)
Primary impairment 13, PI_{13}	0.1131	0.1008	0.8708	0.5406^{***}	0.1654	0.2211
	(0.2790)	(0.2790)	(1.0026)	(0.0973)	(0.2903)	(0.3179)
Primary impairment 14, PI_{14}	0.3822	0.3702	1.0951	0.8518^{***}	0.4375	0.5144
	(0.2789)	(0.2789)	(1.0026)	(0.1101)	(0.2902)	(0.3178)
Primary impairment 15, PI_{15}	-0.1845	-0.1968	0.5275	-	-0.1321	-0.1300
	(0.2801)	(0.2801)	(1.0034)	-	(0.2913)	(0.3191)
Smoker, $\mathrm{SM}_{i,1}$	0.3743^{***}	0.3736^{***}	0.3892^{***}	0.4391^{***}	0.3757^{***}	0.3112^{***}
	(0.0429)	(0.0429)	(0.0737)	(0.1109)	(0.0434)	(0.0433)
"Aggregate" smoking status, $\mathrm{SM}_{i,2}$	0.2114^{***}	0.2120^{***}	0.2416^{***}	0.2884^{*}	0.2124^{***}	0.2250^{***}
	(0.0551)	(0.0551)	(0.0878)	(0.1581)	(0.0555)	(0.0550)
Face Value, $\ln(1+\ln(1+\mathrm{FV}))$				-0.1328		
				(0.1369)		
Settled-and-Observed, SaO_i	-0.1203^{***}	-0.4929^{***}	-0.6439^{***}	-0.3065^{**}	-0.3835^{***}	-0.1111^{**}
	(0.0200)	(0.0642)	(0.1047)	(0.1556)	(0.0659)	(0.0527)
Settled-and-Observed $\times$ trend,		0.2225^{***}	0.3940^{***}	0.2201^{**}	0.1638^{***}	0.0711^{**}
$\mathrm{SaO}_i \times \ln(1+t)$		(0.0360)	(0.0681)	(0.1047)	(0.0370)	(0.0327)
Log-likelihood value	$-134,032.02$	$-134,012.44$	$-41,600.90$	$-13,707.24$	$-132,337.92$	$-132,918.00$

Table 2: Proportional hazards survival regression results. Column [A]: Basic regression (Eq. (4)), earliest observation date; column [B]: With time trend, earliest observation date; column [C]: Only considering cases with known face value (in the remaining sample) and with time trend, earliest observation date; column [D]: Only considering cases with known face value (in the entire dataset) and with time trend, face value as covariate, earliest observation date; column [E]: Excluding cases with times of death within six months of underwriting (in the remaining sample) and with time trend, earliest observation date; column [F]: With time trend, latest observation date. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

information). In contrast, the evaluation of mortality for the pricing of life settlements is highly individualized. Furthermore, while risk aversion is a key driver for purchasing life insurance, the decision of whether or not to sell a policy for an affluent senior frequently is driven by investment or estate planing considerations—so that risk aversion may be less relevant. Therefore, our analysis may not be subject to the same confounding influences as purchasing coverage in the primary market, or at least not to the same extent. Our result that individuals possess private information is in line with He (2009) and Wu and Gan (2013), who find evidence for asymmetric information in primary life insurance when accounting for certain biases.

4 Time Trend of the Informational Asymmetry

To shed light on the characteristics of the informational friction, we derive non-parametric estimates of the *excess hazard* associated with settling over time since settlement. Here, by excess hazard, we mean the difference in the hazard rate between an arbitrary individual in the subsample of closed policies relative to an otherwise identical individual in the full sample. That is, the difference in mortality when knowing a policyholder settled relative to not having this information.

In line with our regression approach, we consider *proportional excess hazard*—which is also referred to as *multiplicative excess hazard* in the statistical literature. We rely on a repeated application of the proportional excess hazard estimator from Andersen and Vaeth (1989), which in turn is based on the well-known Nelson-Aalen non-parametric estimator. More precisely, we first adjust all hazard estimates from the LE provider $\hat{\mu}_t^{(i)}$ based on the survival experience in the full sample, and then derive the excess hazard to the adjusted hazard estimate based on the survival experience in the closed subsample (see Appendix A.2 for more details). Thus, the result is an estimate of the factor to be multiplied on the hazard rate for an arbitrary individual from the full dataset to give the hazard rate for an individual with the same observables but from the closed subsample, as a function of time since settlement.

Figure 1 shows the estimate (solid curve). Clearly, if the estimate had the shape of a horizontal line at one (horizontal dashed line) or if the horizontal line at one fell within the (point-wise) 95% confidence intervals (dashed curves), we would conclude that there is no (significant) impact of settling on an individual's hazard. The observation that the estimate is overall less than one illustrates the negative association between settling and hazard, in line with the regression result from the previous section, and therefore the existence of asymmetric information.

With an approximately 60-70% reduction in hazard, the impact of settling is very pronounced immediately after the settlement decision. However, the effect is wearing off over the course of about eight years. While the (point) estimate continues to increase after year eight followed by a sharp decrease in the last years, the confidence intervals become wider due to the limited data in

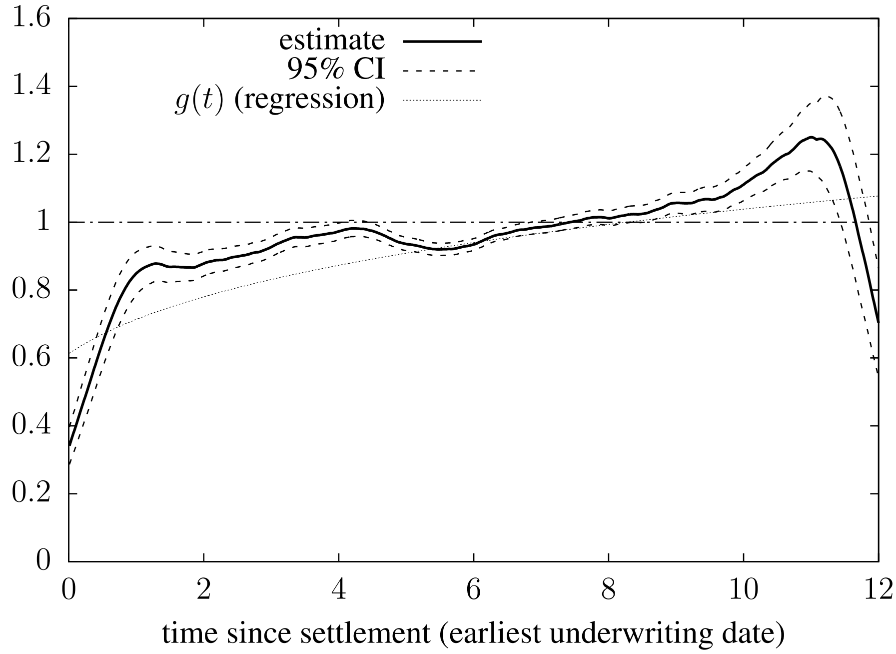


Figure 1: Non-parametric estimate of proportional excess hazard for an individual in the closed subsample relative to an individual in the full sample (solid curve), with point-wise 95% confidence intervals (dashed curves) and the corresponding trend line from the survival regression (dotted curve); earliest observation date.

this region, making it difficult to make an inference on the existence or the sign of the trend in the later years after settling. Hence, the key characteristic that emerges is a negative influence on hazard that is receding over time since settlement.

Survival regressions confirm this observation. We augment the basic specification from Equation (4) by a logarithmic time trend interacted with the Settled-and-Observed variable $\text{SaO}_i \times \ln(1+t)$ in the exponent. Column [B] in Table 2 presents the resulting estimates. The coefficients for the covariates that are not related to the settlement decision are similar to the basic specification in column [A]. The coefficient for the Settled-and-Observed dummy (intercept of the trend) again is negative and strongly significant, with its absolute value being more than four times that of the basic specification. Hence, in line with the non-parametric estimate, we find a very pronounced negative relationship between settling and hazard shortly after the settlement decision. The slope of the trend is highly significant and positive, implying that the influence of settling on hazard becomes weaker over time since settlement, which is again in line with with the pattern as observed in the non-parametric estimate.

Indeed, the regression model suggests a proportional excess hazard for individuals in the closed subsample relative to the remaining sample of the form:

$$g(t) = \exp\{-0.49\} \times (1+t)^{0.22} \approx 0.61 \times (1+t)^{0.22},$$

which we also plot in Figure 1 (dotted curve). In particular, the trend suggests a reduction of hazard of roughly 40% immediately after settlement but that the effect wears off zeroing after roughly 8 years, with a decreasing slope so that the effect after the 8-year time period is minor. The log-likelihood of the model also increases markedly when adding the time trend, compared with the basic specification. Alternative trend specifications (e.g., a linear trend in the exponent) yield similar conclusions, although corresponding model likelihoods are lower. We do not find significant results for a quadratic trend component. We refer to Online Appendix B for corresponding results.

5 Robustness

To demonstrate that our results do not originate from model misspecification and that they are not driven by biases, we conduct a series of robustness analyses. We commence by repeating our analyses under an additive model for the hazard, obtaining similar results. We then discuss omitted variables and sample selection, again concluding that our qualitative findings are robust.

5.1 Additive Model Specification

In addition to the proportional hazards assumption in Section 3, we alternatively consider an additive hazards regression model (Aalen et al., 2008, e.g.):

$$\mu_t^{(i)} = \beta_0(t) + \beta_1\,\hat{\mu}_t^{(i)} + \beta_2\,\mathrm{DOU}_i + \beta_3\,\mathrm{AU}_i + \beta_4\,\mathrm{SE}_i + \sum_{j=1}^{15}\beta_{5,j}\,\mathrm{PI}_{i,j} + \sum_{j=1}^{2}\beta_{6,j}\,\mathrm{SM}_{i,j} + \gamma\,\mathrm{SaO}_i, \quad (5)$$

where the variables are defined as in Equation (4). While less popular, the additive specification directly resembles the standard regression test for the coverage-risk correlation as described in Cohen and Siegelman (2010). We rely on the generalized least-squares (GLS) approach from Lin and Ying (1994) to estimate the coefficient vector and on their formula for the model likelihood. Column [A] in Table 3 presents the results for the basic model (5).

Similarly to the proportional model, the coefficients for underwriting age, sex, and the variables relating to smoking status are significant—although underwriting date is not significant here. Unlike the proportional model, however, five of the fifteen primary impairments are significant and all of the significant coefficients are positive. In turn, to balance these positive terms, the coefficient for the estimated hazard $\hat{\mu}_t^{(i)}$, while highly significant, with roughly 0.2 is now far away from one as would be the case for "perfect" estimates by the LE provider. This suggests that the proportional model may be better suited to capture residual effects.

Important for our focus, the coefficient for the Settled-and-Observed covariate again is negative and highly significant. Thus, we again find a strong negative relationship between settlement and hazard, indicating the existence of asymmetric information in the life settlement market.

When augmenting the basic specification (5) by a linear time trend interacted with the Settled-and-Observed covariate, $\text{SaO}_i \times t$, we obtain analogous effects as in the proportional model: The intercept more than doubles, and the coefficient for the time trend is positive and significant. The model likelihood also increases, and the coefficients for the non-settlement related variables are very similar to the basic model. Column [B] in Table 3 presents the corresponding estimates (see also Appendix B for alternative trend specifications with lower likelihood values). Hence, here our result that the influence of the informational friction is most pronounced right after settlement but wears off over time also appears robust.

To corroborate, we again derive non-parametric estimates of the *additive excess hazard* associated with settling. More precisely, similarly to the proportional excess hazard from Section 4, we estimate a function of time since settlement, which *added* to the hazard of an arbitrary individual in the full sample gives the hazard for an individual with the same observables from the closed subsample. Here, again our estimate is based on a repeated application of the corresponding (additive) excess hazard estimator from Andersen and Vaeth (1989), which in turn is based on the well-known Kaplan-Meier non-parametric estimator (see Appendix A.2 for more details).

Figure 2 shows the resulting estimate (solid curve). Clearly, here if the estimate had the shape of a horizontal line at *zero* (horizontal dashed line) or if the horizontal line at zero fell within the (point-wise) 95% confidence intervals (dashed curves), we would conclude that there is no significant impact of settling on an individual's hazard. The observation that the estimate is overall less than zero illustrates the negative association between settling and hazard, in line with the regression results. And, also in analogy with the proportional case and the regression with time trend, we find that the negative excess hazard is most pronounced in the earlier years after settlement and dissipates over approximately eight years. Similar to the proportional excess hazard case, we also plot in Figure 2 the additive excess hazard suggested from the survival regression, $g(t) = -0.01 + 0.0014 \times t$ (dotted line).

5.2 Omitted Variables

In preparing the offer price, the LS company will have access to additional information, beginning with the fact that the policy is for sale. Since our full dataset also includes individuals that were underwritten for different reasons than the intent to sell their policy, this could create a bias.

	[A]	[B]	[C]	[D]	[E]	[F]
$^1/_{14} \times \int_0^{14} \beta_0(t)\, dt$	-0.0090	-0.0090	-0.0102	-0.0038	-0.0094	-0.0108
Estimated hazard, $\hat{\mu}_t^{(i)}$	0.2032^{***}	0.2031^{***}	0.2551^{***}	0.2775^{***}	0.1959^{***}	0.2264^{***}
	(0.0067)	(0.0067)	(0.0120)	(0.0182)	(0.0066)	(0.0073)
Underwriting date, DOU$_i$	0.0001	0.0001	0.0017^{***}	0.0023^{***}	3.2×10^{-5}	-0.0008^{***}
	(0.0002)	(0.0002)	(0.0003)	(0.0007)	(0.0002)	(0.0002)
Age at underwriting, AU$_i$	0.0016^{***}	0.0016^{***}	0.0011^{***}	0.0007^{***}	0.0017^{***}	0.0019^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
Sex, SE$_i$	0.0043^{***}	0.0042^{***}	0.0014	-0.0002	0.0043^{***}	0.0054^{***}
	(0.0007)	(0.0007)	(0.0011)	(0.0016)	(0.0007)	(0.0008)
Primary impairment 1, PI$_1$	0.0131	0.0129	0.0419^{**}	0.0302^{***}	0.0157	0.0245
	(0.0150)	(0.0150)	(0.0206)	(0.0086)	(0.0145)	(0.0177)
Primary impairment 2, PI$_2$	0.0793^{***}	0.0791^{***}	0.0614^{**}	0.0515^{***}	0.0757^{***}	0.1074^{***}
	(0.0192)	(0.0192)	(0.0251)	(0.0197)	(0.0185)	(0.0216)
Primary impairment 3, PI$_3$	0.0029	0.0028	0.0325	0.0186^{***}	0.0058	0.0101
	(0.0149)	(0.0149)	(0.0202)	(0.0064)	(0.0143)	(0.0175)
Primary impairment 4, PI$_4$	0.0293^{*}	0.0292^{*}	0.0572^{***}	0.0499^{***}	0.0309^{**}	0.0414^{**}
	(0.0153)	(0.0153)	(0.0211)	(0.0100)	(0.0148)	(0.0180)
Primary impairment 5, PI$_5$	0.0284^{*}	0.0282^{*}	0.0505^{**}	0.0507^{***}	0.0287^{**}	0.0485^{***}
	(0.0150)	(0.0150)	(0.0204)	(0.0078)	(0.0144)	(0.0177)
Primary impairment 6, PI$_6$	-0.0042	-0.0043	0.0227	0.0049^{**}	-0.0008	0.0025
	(0.0147)	(0.0147)	(0.0200)	(0.0022)	(0.0142)	(0.0174)
Primary impairment 7, PI$_7$	0.0052	0.0051	0.0297	0.0130^{***}	0.0084	0.0140
	(0.0147)	(0.0147)	(0.0200)	(0.0024)	(0.0142)	(0.0174)
Primary impairment 8, PI$_8$	0.0044	0.0042	0.0328	0.0176^{***}	0.0077	0.0100
	(0.0148)	(0.0148)	(0.0201)	(0.0050)	(0.0143)	(0.0175)
Primary impairment 9, PI$_9$	0.0879^{***}	0.0878^{***}	0.0944^{***}	0.0741^{***}	0.0897^{***}	0.1045^{***}
	(0.0160)	(0.0160)	(0.0219)	(0.0129)	(0.0154)	(0.0185)
Primary impairment 10, PI$_{10}$	-0.0035	-0.0036	0.0228	0.0061^{**}	-0.0002	0.0038
	(0.0148)	(0.0148)	(0.0200)	(0.0024)	(0.0142)	(0.0174)
Primary impairment 11, PI$_{11}$	0.0486^{***}	0.0484^{***}	0.0775^{***}	0.0730^{***}	0.0500^{***}	0.0708^{***}
	(0.0154)	(0.0154)	(0.0212)	(0.0127)	(0.0148)	(0.0182)
Primary impairment 12, PI$_{12}$	0.0086	0.0084	0.0317	0.0157^{***}	0.0113	0.0174
	(0.0148)	(0.0148)	(0.0202)	(0.0044)	(0.0143)	(0.0175)
Primary impairment 13, PI$_{13}$	-0.0045	-0.0047	0.0232	0.0089^{***}	-0.0013	0.0035
	(0.0147)	(0.0147)	(0.0199)	(0.0016)	(0.0142)	(0.0174)
Primary impairment 14, PI$_{14}$	0.0152	0.0151	0.0382^{*}	0.0241^{***}	0.0183	0.0269
	(0.0148)	(0.0148)	(0.0200)	(0.0036)	(0.0142)	(0.0174)
Primary impairment 15, PI$_{15}$	-0.0141	-0.0143	0.0137	-	-0.0108	-0.0080
	(0.0147)	(0.0147)	(0.0199)	-	(0.0142)	(0.0174)
Smoker, SM$_{i,1}$	0.0297^{***}	0.0297^{***}	0.0311^{***}	0.0405^{***}	0.0292^{***}	0.0313^{***}
	(0.0026)	(0.0026)	(0.0048)	(0.0093)	(0.0026)	(0.0029)
"Aggregate" smoking status, SM$_{i,2}$	0.0115^{***}	0.0116^{***}	0.0138^{***}	0.0124	0.0115^{***}	0.0135^{***}
	(0.0025)	(0.0025)	(0.0042)	(0.0077)	(0.0025)	(0.0029)
Face Value, $\ln(1+\text{FV})$				-0.0031^{***}		
				(0.0005)		
Settled-and-Observed, SaO$_i$	-0.0049^{***}	-0.0107^{***}	-0.0132^{***}	-0.0084^{***}	-0.0087^{***}	-0.0046^{***}
	(0.0007)	(0.0012)	(0.0019)	(0.0028)	(0.0011)	(0.0015)
Settled-and-Observed $\times$ trend, SaO$_i \times t$		0.0014^{***}	0.0034^{***}	0.0023^{***}	0.0011^{***}	0.0011^{***}
		(0.0003)	(0.0005)	(0.0008)	(0.0003)	(0.0004)
Log-likelihood value	$-74,521.29$	$-74,304.23$	$-26,748.04$	-6872.19	$-75,599.54$	$-73,678.67$

Table 3: Additive hazards survival regression results. Column [A]: Basic regression (Eq. (5)), earliest observation date; column [B]: With time trend, earliest observation date; column [C]: Only considering cases with known face value (in the remaining sample) and with time trend, earliest observation date; column [D]: Only considering cases with known face value (in the entire dataset) and with time trend, face value as covariate, earliest observation date; column [E]: Excluding cases with times of death within six months of underwriting (in the remaining sample) and with time trend, earliest observation date; column [F]: With time trend, latest observation date. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

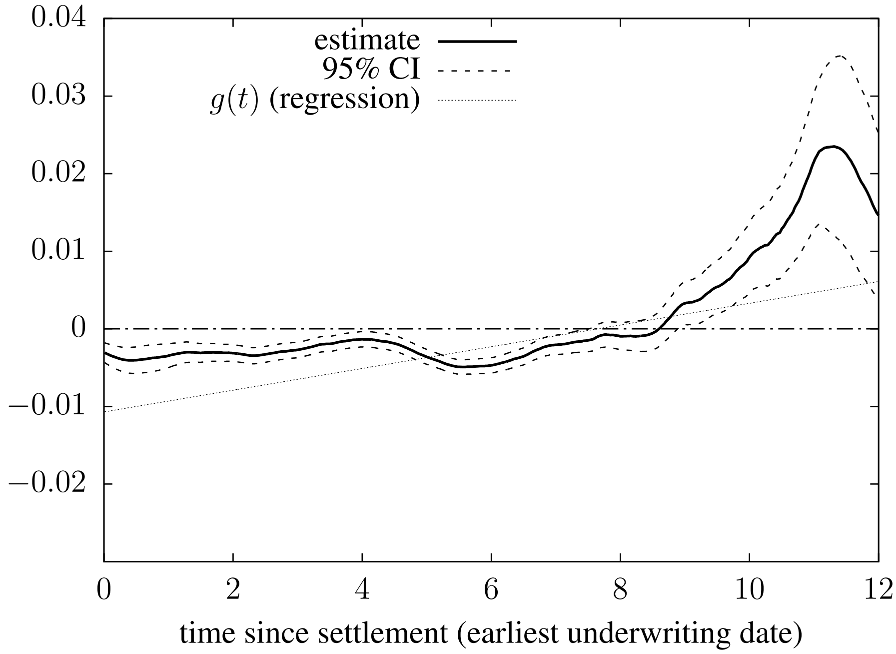


Figure 2: Non-parametric estimate of additive excess hazard for an individual in the closed subsample relative to an individual in the full sample (solid curve), with point-wise 95% confidence intervals (dashed curves) and the corresponding trend line from the survival regression (dotted curve); earliest observation date.

Furthermore, policy information such as the face value may proxy for unknown variables such as wealth, and the company may have available additional LE estimates from different LE providers or insights from their own experience.

To rule out endogeneity concerns, we first repeat the regression analyses when only considering the $7,832$ cases in the remaining sample for which we have information on the policy face value (but keeping all cases for the closed subsample). Since face value is only available for individuals participating in the life settlement market, this addresses the possibility of "participant in the life settlement market" to be a relevant omitted variable. Columns [C] of Tables 2 and 3 show the results for the proportional and the additive hazards model with time trend, respectively. Again, we find significant negative intercepts and significant positive slope coefficients for the time trend associated with the SaO variable. While the magnitude of the intercepts remains at a similar (yet slightly higher) level, the slopes are considerably higher than the corresponding estimates from columns [B] in both specifications, implying that the effect wears off over a shorter period.

To analyze the impact of the policy face value or potentially non-observed correlated variables on our findings, we further limit the dataset by also considering only the $2,672$ cases with known face values in the *closed* subsample—so that we can additionally include face value as a covariate in the survival regression ($\ln(1 + \ln(1 + \text{FV}))$ and $\ln(1 + \text{FV})$ in the proportional and the additive specification, respectively). We present the results in columns [D] of Tables 2 and 3 for the proportional and the additive hazards model with time trend, respectively.[11] Since we only have face value for a fraction of all (closed and remaining) policies, the standard errors are a lot larger. Nonetheless, the settlement-related variables are again significant with consistent signs for both specifications. This reinforces our main prediction of a negative and receding relationship between settling and hazard. We observe that for the proportional hazards regression, the coefficient for face value—while negative—is not significant whereas it is significant and negative in the additive specification. Hence, we have mild evidence that high face values are associated with longer realized lifetimes, indicating a wealth effect.

The LS company may have available additional pricing-relevant information that is unknown to *our* LE provider, particularly the underwriting results from different LE providers (typically there are at least two evaluations).[12] More precisely, we only have access to the LE provider's estimate $\hat{\mu}_t^{(i)}$ and not the LE used for pricing. To the extent that the difference is substantial, a second estimate may affect the pricing and thereby the decision to settle, giving rise to possible endogeneity and a potential bias.

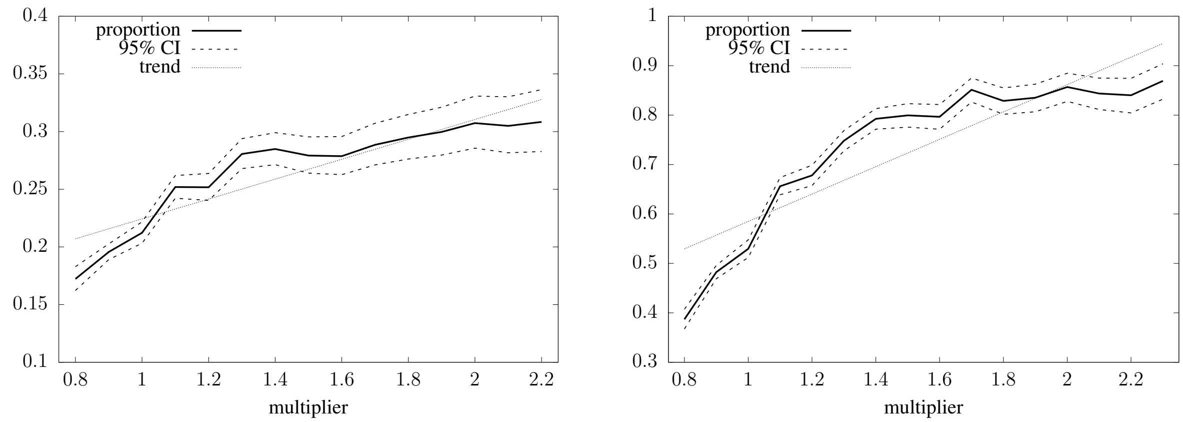
However, since we are primarily interested in the sign of the settlement coefficient, a positive (conditional) relationship between the omitted estimate and settlement yielding a positive bias will not be critical in view of our result whereas a negative relationship may pose problems.[13] It is important to note that there are two relevant influences: On the one hand, a relatively high second hazard estimate will typically lead to a higher offer price rendering settling more likely; on the other hand, a relatively high second estimate is indicative of a higher *true* hazard rate, which will make settling less likely for an unchanged offer price. Hence, in order to assess whether the relationship is positive or negative, the key question is whether or not the proclivity for settling increases in the estimate. Appendix A.3 corroborates this insight by working out a version of the simple model from Section 2 with uncertainty in the offer price originating from additional information on the mortality probability estimate. In line with the arguments here, the model shows that the average difference between the unconditional mortality probability and the mortality probability conditional on settling will be larger in the presence of additional information *if* the fraction of policyholders deciding to settle is increasing in the unknown mortality probability estimate.

We can assess this relationship in the context of the available estimate by analyzing the propor-

[11] Here, we encounter collinearity with respect to the primary impairments since every policyholder in the reduced dataset is affected by exactly one primary impairment. We address this by taking out PI_{15} in the regression.

[12] We emphasize that the relevant perspective is that of the LS company / the investor with the winning offer for the policy. Other parties such as the broker and, of course, the policyholder may have different information sets that could also be affected by the bidding process. We discuss the origin of the informational asymmetry in the next section.

[13] Consider e.g. the extreme and stylized case where the company has full information (such that the *true* coefficient γ will be zero) and the correlation between $\mu_t^{(i)}$ and SaO_i is 1 (-1). Then clearly the estimated $\hat{\gamma}$ from Eqs. (4) and (5) will be positive (negative).



(a) full sample

(b) sample with known policy face values

Figure 3: Proportion of policyholders that settled their policy as a function of the mortality multiplier (solid curve), with point-wise 95% confidence intervals (dashed curves); earliest observation date. Left panel: Proportion calculated based on the full sample. Right panel: proportion calculated based only on policies with known face values.

tion of policyholders that decided to settle their policy as a function of the corresponding mortality multiplier. As discussed in Section 3, this multiplier is used relative to a life table that accounts for basic characteristics such as gender, age, etc., so that it controls for observable characteristics and reflects the assessment of the LE provider. In Figure 3, we plot the proportion of policyholders within our subsample of closed policies, both relative to the full sample (left panel) and relative to the sample of policies with known face value (right panel). In constructing the figures, we consider bins of multipliers of lengths 0.1 and derive the proportion (solid curve) as well as 95% confidence intervals (dashed curves) based on a binomial assumption. Since we have many outliers with relatively (very) small or (very) large multipliers and the calculation of reliable proportions is difficult in this range, we disregard the 5% of the sample with the lowest multipliers and the 15% with the highest multipliers, so that the plots show 80% of the sample. As is evident from the figures and the trend lines (dotted lines), we find a generally positive relationship. This suggests our analysis is robust with regards to additional information on LEs.

5.3 Sample Selection

As indicated in Section 3, we use the earliest underwriting date as a proxy for the time the individual decides whether or not to settle her policy—which is the relevant point in time for exploiting the informational advantage. This is justified by the fact that the time delay between the time an LE estimate is ordered and the time when it is communicated is typically not very long; indeed, frequently investors have rules in place that they will not use an estimate for bidding if it is older than six months. However, a policyholder interested in settling might nonetheless decease during this waiting period and therefore not be included in the closed subsample—potentially contributing to the negative relationship between settlement and hazard that we observe.

To analyze whether the negative correlation is solely driven by this delay, we eliminate cases where the policyholder died within six months of the (earliest) underwriting date in our remaining sample. Thus, all policyholders that might have considered settling but died before having the opportunity will be excluded from the analysis. Since by doing so we also exclude policyholders that did settle but are not observed, policyholders that would not have settled, and individuals that did not even contemplate settling in the first place, and since being in the remaining subsample now implies a survival of at least six months, this prodedure even creates a bias *against* the hypothesis of a negative settlement-hazard correlation in our analysis. Nonetheless, as can be seen from Columns [E] in Tables 2 and 3, while the effect naturally decreases due to the aforementioned bias, the intercept and the slope of the settlement dummy are still highly significant and remain similar to the original results. Hence, we can exclude this mechanism as a primary driver for the observed correlation.

A related silent assumption for our analyses thus far is that a random sample of closed policies is a random sample of closed policies at settlement or, in other words, that being *observed* as one of the policies in the closed subsample is independent of survival. A potential problem with this assumption are so-called *tertiary trades*, that is policies that are resold at some point in time after the closing date. Clearly, a tertiary trade will only take place if the policyholder is alive, so that—assuming there exists an LE from our provider for the company that originally owned the policy—tertiary policies may be associated with longer observed lifetimes.[14]

To address this and other concerns relating to possible sample selection issues, we rerun the regression analyses using the latest observation date for each policy in the full dataset, i.e. we evaluate the impact of settling on survival experience relative to the last time the life was underwritten by our LE provider. Regression results are presented in columns [F] of Tables 2 and 3 for the proportional and additive specifications, respectively. The estimates for the non-settlement-related variables are similar to the earliest observation date (columns [B]). For the settlement-related variables, the qualitative observations are analogous, although—as is to be expected given the results on the time trend from Section 4—the effect is less pronounced since it weakens over time. In particular, the results from both specifications indicate that the effect wears off after approximately four years. Thus, while these estimates are less in line with our objective of studying the existence and pattern of private information when selling the policy, we are able to identify the residual effect—lending force to our results.

[14]To illustrate, assume Investor 1 buys policies from insureds A and B. Later, Investor 1 sells her portfolio to Investor 2. Assume that by then, B has died whereas A is still alive. If our closed subsample contained Investor 2's portfolio, we would know that A settled but would have no information on B.

6 Impact and Origin of the Informational Friction

To quantify the impact of settling on life expectancy, we provide example calculations based on our proportional hazards regression results. We face two difficulties. First, as discussed in Section 3, our regression estimates are based on analyses of the known closed policies relative to the remaining policies, with the latter including a mix of closed and non-closed cases. Since we are interested in the more direct closed versus non-closed comparison, we adjust our point estimate based on different parameter values of the (unknown) proportion of closed policies in the full sample p (cf. Footnote 10). More precisely, we inflate the coefficient γ based on the analysis in Appendix A.1 (Eq. (7)) and assume that the impact wears off over eight years according to the time trend in our regression model from Section 4.[15] As a reference, when relying on average face values, the market size estimate by Roland (2016) implies a proportion of roughly 47%. Second, our regressions give us estimates for the overall impact, but not for a specific individual. Hence, we rely on US population mortality data to evaluate the impact on average policyholders at different ages that are roughly in line with the aggregate statistics from our dataset (ages 70, 75, and 80).[16]

Table 4 presents results for representative US male policyholders. Appendix B presents additional results for female policyholders as well as for alternative specifications (time-constant effect, additive regression specification). The first column of the table presents LE estimates/increases based on the observed proportion of closed policies ($13,221/53,947 \approx 24.5\%$). Since the actual proportion can only be higher, these estimates provide lower bounds for the LE differences between settlers and non-settlers with identical observable characteristics. The remaining columns present LE results based on various assumptions of the proportion that range from 30% to 70%. Overall, our calculations suggest that the life expectancies for individuals that settled their policy exceed those of policyholders not settling their policy by between 0.35 to 0.81 years, or roughly 2.5% to 11% of the life expectancies. In particular, for a 75 year old policyholder and assuming that the proportion of closed cases in the full sample is 50%, we obtain roughly half a year of additional life expectancy relative to a non-settler's life expectancy of a little over 10 years. Of course, the results are based on specific assumptions and, as discussed in Section 5, there are aspects that may cause deviations in either direction. Nevertheless, these magnitudes suggest that asymmetric information has an economically significant impact on the life settlement market, and should be accounted for in market operations—e.g. in view of pricing and risk management.

Identifying the origin of the informational asymmetry is a difficult problem since different ex-

[15] As shown in Appendix B, assuming a time-constant effect as in the basic specification (4) yields a more pronounced impact on life expectancy, so our choice is conservative.

[16] The mortality data are taken from the Human Mortality Database; University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany); available at *www.mortality.org* or *www.humanmortality.de*. More precisely, we calculate life expectancies based on expected future survival probabilities, where we use the Lee and Carter (1992) method to produce forecasts.

	Proportion of closed policies (p)					
	24.5%	30%	40%	50%	60%	70%
Age 70 (non-adjusted LE 13.93*)*						
LE for settled cases	14.28	14.30	14.34	14.39	14.46	14.55
LE increase	0.35	0.37	0.41	0.46	0.53	0.62
Age 75 (non-adjusted LE 10.48*)*						
LE for settled cases	10.88	10.90	10.95	11.01	11.09	11.20
LE increase	0.40	0.42	0.47	0.53	0.61	0.72
Age 80 (non-adjusted LE 7.50*)*						
LE for settled cases	7.95	7.98	8.03	8.10	8.19	8.31
LE increase	0.45	0.48	0.53	0.60	0.69	0.81

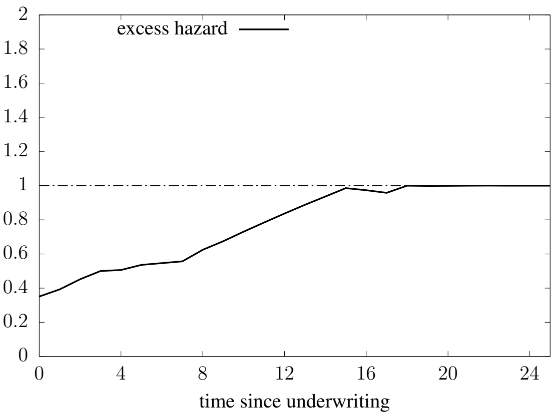
Table 4: Comparisons of average life expectancies between population-level and settled US male policyholders; proportional hazards model with time-weakening effect.

planations have similar empirical implications, particularly the positive risk-coverage relationship we observe (Chiappori and Salanié, 2013). However, different sources of asymmetric information may lead to different risk-coverage patterns over time. For instance, Abbring et al. (2003) rely on dynamic relationships to characterize asymmetric information in the context of experience ratings in automobile insurance.

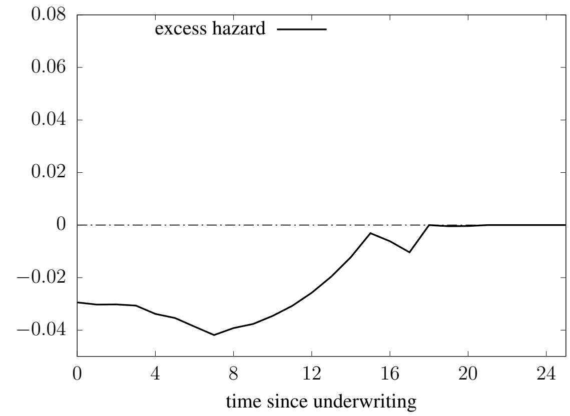
The temporary and subsiding influence of settlement over time resembles patterns in so-called *select-and-ultimate* life tables in actuarial studies that capture selection effects due to underwriting. To illustrate, in the left and right-hand panels of Figure 4 we plot the proportional and additive excess hazard, respectively, for a preferred male life underwritten at age 75 as a function of time since underwriting relative to ultimate hazard rates based on the Society of Actuaries 2001 Commissioner's Standard Ordinary (CSO) preferred life table.[17] Of course, here the "selection effect" comes from the underwriting process allowing insurers to use lower hazard rates in the *select* period, so the origin for the deviation is not an informational asymmetry. The relevant analogy is that insurers will only have information on the policyholder's health condition at the point of sale (time of underwriting), and the relevance of this information dissipates as time progresses, producing the converging pattern.

The shapes in Figure 4 are similar to our non-parametric estimates of proportional and additive

[17] For information on the 2001 CSO table, see `https://www.soa.org/Research/Experience-Study/ind-life/tables/intl-2001-cso-preferred-class-structure-mortality-tables.aspx`.



(a) proportional excess hazard



(b) additive excess hazard

Figure 4: Excess hazard for a male policyholder underwritten at age 75 as a function of time since underwriting for the Society of Actuaries 2001 Commissioner's Standard Ordinary (CSO) preferred life table. Left panel: Proportional excess hazard. Right panel: Additive excess hazard.

excess hazard due to asymmetric information in Figures 1 and 2, respectively. The key difference is that the selection effect in life insurance lasts fifteen to seventeen years, whereas our estimates show a vanishing impact of settlement after approximately eight years. Nonetheless, the analogy provides suggestive evidence for an informational advantage regarding the initial health condition that individuals select on in their settlement decision.

Alternative explanations for the positive settlement-survival correlation include hidden actions (*moral hazard*), indirect effects e.g. via risk aversion, or a possible information gain during the settlement process. In the present context, "moral hazard" may take the form of healthier lifestyle choices after relinquishing the life insurance coverage, seeking improved medical care using the proceeds from settling, or other positive changes in health-related behavior. If *permanent* changes in behavior were the sole driver for the informational asymmetry, two policyholders with exactly the same observable characteristics but only differing in their settling decision should display exactly the same hazard rate right up until settlement, and we would expect to see a diverging relationship thereafter. In particular, if there were differences in care or in lifestyle, we would arguably expect (at least) a *persistent* effect on hazard—in contrast to the *subsiding* pattern we identify.

Heterogeneity in underlying policyholder characteristics such as wealth or risk aversion can be the root cause for an informational asymmetry. While, as indicated in Section 2, heterogeneity in wealth is not likely to deliver the observed result, a (negative) correlation between settlement and hazard could arise due to differences in risk aversion. More precisely, risk aversion—which in the basic model from Section 2 is captured by the parameter ψ—may directly affect the decision to settle or it may lead individuals to hold more (relinquishable) life insurance in the first place, but may also affect survival prospects, e.g. by limiting engagement in risky activities or more engagement in preventative healthcare.[18] However, these explanations, and particularly the idea that risk aversion influences the purchase of insurance several years before the settlement decision, assume that risk aversion is—or policyholder characteristics more generally are—*persistent*. Thus, again, the converging pattern of the hazard rates after the settlement decision is not in line with these explanations, since persistent differences in characteristics should yield persistent differences in survival prospects.

The evolution of the difference in hazards also allows us to address the relevance of potential informational asymmetries that originate directly from the settlement process. More precisely, a wedge could arise from brokers forwarding shorter LEs to LS companies, or policyholders picking the highest among several bids for their policy (*winner's curse*)—which may be submitted by the LS company that received the lowest (combination of) LEs.[19] Consider the following thought experiment in opposition to this explanation: Suppose there are several identical distributed LEs with different associated multipliers but the broker only forwards the one with the highest multiplier and the "winning" LS company prepares a bid on this basis; now, assuming the multiplier is simply a relatively high random realization, the proportional excess hazard will be constant over time at a level below one and the additive excess hazard will necessarily need to diverge to sustain this constant proportional trend, in contrast to the converging pattern depicted in Section 4. We provide a more detailed discussion and a Monte Carlo implementation of this experiment in Appendix C. The results show that indeed this mechanism can deliver a negative correlation between settlement and hazard—we obtain negative and significant intercepts in all five simulations for the proportional hazards specification. However, the trend components in the proportional hazards regression specification are all insignificant, and they are all negative and significant for the additive hazards specification (see Table A.4 and Figure A.2 for the corresponding non-parametric estimates). These results are in stark contrast to the significantly positive results for the trend component from Tables 2 and 3.

Thus, while there are several conceivable aspects contributing to the informational asymmetry, the pattern over time is in line with adverse selection by policyholders based on private information regarding their near-term survival prospects.

[18] Note that here, since risk aversion is not known to the LS company, a negative correlation between settlement and hazard would still be driven by an informational asymmetry, although the mechanism is "indirect."

[19] While LS companies used to order LEs themselves, within current transactions typically the broker assembles several LEs and forwards them to the companies together with the policy information. While there is some discretion which estimates to forward, certain LS companies require LEs from specific providers.

7 Conclusion

In this paper, we show that in the secondary life insurance market, policyholders choosing to settle their policy, ceteris paribus, exhibit significantly longer (relative to their LE estimate) lifetimes—although the impact of settling on survival subsides over time. This documents the existence and relevance of private information regarding near-term survival prospects. Our findings are robust with respect to model specification and other sources for potential biases.

While the quantitative results are specific to our setting and particularly the population in view, we believe that our qualitative insights have broader repercussions. More precisely, in addition to complementing studies on informational asymmetries in primary insurance markets, our findings indicate that individuals in our sample are competent in assessing their *relative* survival prospects when prompted with relevant information, in a situation with significant monetary consequences. Here, by *relative* survival prospects we mean the appraisal of whether an individual expects to live longer or shorter than an average individual with a similar profile. This is in contrast to individuals' ability in predicting *absolute* life expectancies that seem to be subject to framing and other behavioral biases (Payne et al., 2013; and references therein). We believe that the former task may be more material for retirement planning given that individuals may be provided with background information or suitable default choices based on their profile.

The existence and the origin of informational frictions is also material for answering policy-relevant questions regarding the efficiency and welfare implications of the life settlement market. For addressing such questions, it will be necessary to consider equilibrium implications, accounting for barriers to participate in this market (Einav et al., 2010b) and repercussions on primary insurance (Daily et al., 2008; Fang and Kung, 2017). While addressing these issues is beyond the scope of this paper, our findings will inform the process of building and estimating corresponding equilibrium models (see e.g. Einav et al. (2010a)).

References

Aalen, Odd, Ørnulf Borgan, and Hakon K. Gjessing (2008). "Survival and event history analysis: a process point of view." Series Statistics for Biology and Health, Springer, New York.

Abbring, Jaap H., James J. Heckman, Pierre-André Chiappori, and Jean Pinquet (2003). "Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish?" *Journal of the European Economic Association*, 1: 512-521.

Akerlof, George A. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, 84: 488-500.

Andersen, Per Kragh, and Michael Vaeth (1989). "Simple parametric and nonparametric models for excess and relative mortality." *Biometrics*, 45: 523-535.

Arrow, Kenneth J. (1963). "Uncertainty and the welfare economics of medical care." *American Economic Review*, 53: 941-973.

Bauer, Daniel, Michael V. Fasano, Jochen Russ, and Nan Zhu (2017). "Evaluating Life Expectancy Evaluations." *North American Actuarial Journal*, forthcoming.

Bowers, Newton L., Hans U. Gerber, James Hickman, Donald A. Jones, and Cecil J. Nesbitt (1997). *Actuarial Mathematics*, 2nd Edition, The Society of Actuaries.

Cardon, James H., and Igal Hendel (2001). "Asymmetric information in health insurance: evidence from the National Medical Expenditure Survey." *RAND Journal of Economics*, 32: 408-427.

Cawley, John, and Tomas Philipson (1999). "An Empirical Examination of Information Barriers to Trade in Insurance." *American Economic Review*, 89: 827-846.

Chiappori, Pierre-André, and Bernard Salanié (2000). "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy*, 108: 56-78.

Chiappori, Pierre-André, and Bernard Salanié (2013). "Asymmetric Information in Insurance Markets: Predictions and Tests." *Handbook of Insurance,* 2nd edition (G. Dionne, ed.), 397-422. Springer, New York.

Cohen, Alma, and Liran Einav (2007). "Estimating risk preferences from deductible choice." *American Economic Review*, 97: 74588.

Cohen, Alma, and Peter Siegelman (2010). "Testing for Adverse Selection in Insurance Markets." *Journal of Risk and Insurance*, 77: 39-84.

Cohen, Lauren (2013). "Seeking Alpha in the Afterlife: CMG Life Services and the Life Settlement Industry." *Harvard Business School Case Study*, No. 9-214-011. Harvard Business School Publishing, Boston, MA.

Cox, D.R. (1975). "Partial Likelihood." *Biometrika*, 62: 269-276.

Cutler, David M., Amy Finkelstein, and Kathleen McGarry (2008). "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance." *American Economic Review*, 98: 157-162.

Daily, Glenn, Igal Hendel, and Alessandro Lizzeri (2008). "Does the Secondary Life Insurance Market Threaten Dynamic Insurance?" *American Economic Review Papers & Proceedings*, 98: 151-156.

Dionne, Georges, Christian Gouriéroux, and Charles Vanasse (2001). "Testing for evidence of adverse selection in the automobile insurance market: A comment." *Journal of Political Economy*, 109: 444-453.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2010a). "Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market." *Econometrica,* 78: 1031-1092.

Einav, Liran, Amy Finkelstein, and Jonathan Levin (2010b). "Beyond Testing: Empirical Models of Insurance Markets." *Annual Review of Economics,* 2: 311-336.

Elder, Todd E. (2013). "The Predictive Validity of Subjective Mortality Expectations: Evidence From the Health and Retirement Study." *Demography*, 50: 569-589.

Fang, Hanming, Michael P. Keane, and Dan Silverman (2008). "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy*, 116: 303-350.

Fang, Hanming, and Edward Kung (2017). "Life Insurance and Life Settlements: The Case for Health-Contingent Cash Surrender Values." Working Paper, University of Pennsylvania.

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

Finkelstein, Amy, and James Poterba (2004). "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market." *Journal of Political Economy*, 112: 183-208.

Finkelstein, Amy, and James Poterba (2014). "Testing for Asymmetric Information Using 'Unobserved Observables' in Insurance Markets: Evidence from the U.K. Annuity Market." *Journal of Risk and Insurance*, 81: 709-734.

Granieri, Vincent J., and Gregory P. Heck (2014). "Survival Characteristics of Three Senior Populations, with a Focus on Life Settlements." Paper presented at the SOA Living to 100 and Beyond Symposium.

He, Daifeng (2009). "The life insurance market: Asymmetric information revisited." *Journal of Public Economics*, 93: 1090-1097.

Lee, Ronald D., and Lawrence R. Carter (1992). "Modeling and Forecasting U. S. Mortality." *Journal of the American Statistical Association*, 87: 659-675.

Lin, D.Y., and Zhiliang Ying (1994). "Semiparametric analysis of the additive risk model." *Biometrika*, 81: 61-71.

McCarthy, David, and Olivia S. Mitchell (2010). "International adverse selection in life insurance and annuities." *Ageing in Advanced Industrial States,* Springer Netherlands, 119-135.

Payne, John W., Namika Sagara, Suzanne B. Shu, Kirstin C. Appelt, and Eric J. Johnson (2013). "Life expectancy as a constructed belief: Evidence of a live-to or die-by framing effect." *Journal of Risk and Uncertainty*, 46: 27-50.

Puelz, Robert, and Arthur Snow (1994). "Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market." *Journal of Political Economy*, 102: 236-257.

Roland, Rocky (2016). "Life Settlement Tertiary Market Dynamics." Paper presented at the Life Settlement Institutional Investor Conference.

Rothschild, Michael, and Joseph E. Stiglitz (1976). "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information." *Quarterly Journal of Economics*, 90: 630-649.

Stiglitz, Joseph E., and Andrew Weiss (1981). "Credit rationing in markets with imperfect information." *American Economic Review*, 71: 393-410.

Wu, Xi, and Li Gan (2013). "Multiple Dimensions of Private Information in Life Insurance Markets." NBER Working Paper No. 19629.

Zhu, Nan, and Daniel Bauer (2013). "Coherent Pricing of Life Settlements Under Asymmetric Information." *Journal of Risk and Insurance*, 80: 827-851.

Online Appendix to "Asymmetric Information in Secondary Insurance Markets: Evidence from the Life Settlement Market"

Daniel Bauer, Jochen Russ, Nan Zhu[1]

November 2017

This online appendix collects supplemental material for the paper "Asymmetric Information in Secondary Insurance Markets: Evidence from the Life Settlement Market." Part A provides relevant technical details, particularly a discussion of the impact of the mixed nature of the dataset discussed in Section 3 of the main text (Appendix A.1), the derivation of the non-parametric estimates of *excess hazard* used in Sections 4 and 5 of the main text (Appendix A.2), and an extension of the model from Section 2 with price uncertainty that is relevant for robustness analyses with regards to omitted variables in Section 5.2 of the main text (Appendix A.3). Part B presents additional results for the analyses in Sections 3, 4, and 5 in the main text. Part C presents a Monte Carlo experiment on excess hazard over time when brokers or policyholders "cherry-pick" among LEs or policy offers as discussed in Section 6 in the main text.

[1]Bauer: Department of Economics, Finance, and Legal Studies, University of Alabama, Tuscaloosa, AL 35487, dbauer@cba.ua.edu. Russ: Institut für Finanz- und Aktuarwissenschaften and Ulm University, Lise-Meitner-Straße 14, 89081 Ulm, Germany, j.russ@ifa-ulm.de. Zhu: Department of Risk Management, Pennsylvania State University, University Park, PA 16802, nanzhu@psu.edu.

A Technical Appendix

A.1 Impact of the Mixed Nature of the Remaining Subsample

As discussed in the main text, the estimate for the regression coefficient γ of the *Settled-and-Observed* variable generally does not provide a consistent estimate for the difference between closed and non-closed cases due to the mixed nature of the subsample of remaining policies. Put differently, since the remaining cases include both non-closed and closed cases, γ will not constitute a suitable adjustment for closed cases relative to individuals that did not settle their policy but will only amount to a fraction of the "true" difference, and therefore needs to be inflated. In what follows, we derive appropriate inflation formulas for the proportional hazards model and the additive hazards model.

Proportional Hazards Model

To illustrate, consider the following simplified version of our proportional hazards model (4):

$$\mu_t^{(i)} = \beta_0(t) \times \exp\left\{\gamma \, \text{SaO}_i\right\}. \tag{6}$$

Denote by Y_t all remaining observations at time t, by $Y_t^{(1)}$ all remaining settled/closed cases at time t (unobserved), $p_t = Y_t^{(1)}/Y_t$, and by $Y_t^{(2)}$ all remaining Settled-and-Observed cases at time t, $q_t = Y_t^{(2)}/Y_t$. Denote by $D_i(t)$ the "death counting process" for policyholder i (zero if alive, one if dead). Furthermore, denote by γ^{act} the *unknown* actual regression coefficient for the model in which the econometrician observes all settlement decisions—effectively replacing the Settled-and-Observed variable (SaO_i) by a corresponding Settled variable (Set_i) in (6), and by γ^{our} our coefficient based on the Settled-and-Observed cases only. For simplicity, assume further that at any time t, the probability that a settlement decision is observed is constant. Therefore, based on Lin and Ying (1994, Eq. (2.6)), γ^{act} and γ^{our} will be solutions to the following *(partial) score functions*, respectively:

$$\begin{aligned} 0 &= \sum_{i=1}^{N} \int_{t=0}^{\infty} \left[\text{Set}_i - \frac{p_t}{(1-p_t)\times\exp(-\hat{\gamma}^{act}) + p_t}\right] dD_i(t), \text{ and} \\ 0 &= \sum_{i=1}^{N} \int_{t=0}^{\infty} \left[\text{SaO}_i - \frac{q_t}{(1-q_t)\times\exp(-\hat{\gamma}^{our}) + q_t}\right] dD_i(t). \end{aligned}$$

Integrating and using the assumption that we obtain the number of Settled-and-Observed deaths by multiplying the number of Settled deaths by the corresponding proportion (${}^{q_t}/_{p_t}$), we obtain that:

$$(1-p) \times \exp(-\hat{\gamma}^{act}) + p \approx (1-q) \times \exp(-\hat{\gamma}^{our}) + q,$$

where p is the (unknown) overall proportion of *settled cases* and q is the (known) overall proportion of *Settled-and-Observed cases* in the portfolio, which for simplicity we assume are constant. Thus, under the assumptions above, a suitable estimator for the actual difference between closed and non-closed cases under the proportion hazards assumption is:

$$\hat{\gamma}^{act} \approx -\ln\left(\frac{q-p}{1-p} + \exp(-\hat{\gamma}^{our}) \times \frac{1-q}{1-p}\right), \tag{7}$$

where of course $\hat{\gamma}^{our}$ corresponds to the estimate from specification (4). In particular, for $\hat{\gamma}^{our} = 0$ we obtain $\hat{\gamma}^{act} = 0$, and in case $\hat{\gamma}^{our} < 0$ the actual coefficient $\hat{\gamma}^{act}$ needs to be further inflated ($\hat{\gamma}^{act} < \hat{\gamma}^{our} < 0$).

Additive Hazards Model

Similar to above, we consider the following simplified version of our additive hazards model (5):

$$\mu_t^{(i)} = \beta_0(t) + \gamma\,\mathrm{SaO}_i. \tag{8}$$

Using the same assumptions and notations as before, based on the estimates in Lin and Ying (1994, Eq. (2.8)) we have:

$$\frac{\hat{\gamma}^{act}}{\hat{\gamma}^{our}} = \frac{\int_0^\infty Y_t^{(2)}[1-q_t]\,dt}{\int_0^\infty Y_t^{(2)}[1-p_t]\,dt},$$

and again using the assumption of constant proportions we obtain:

$$\frac{\hat{\gamma}^{act}}{\hat{\gamma}^{our}} \approx \frac{(1-q)}{(1-p)}.$$

Thus, a suitable estimator for the actual difference between closed and non-closed cases under the additive hazards assumption is:

$$\hat{\gamma}^{act} \approx \hat{\gamma}^{our} \times \frac{(1-q)}{(1-p)}, \tag{9}$$

where $\hat{\gamma}^{our}$ corresponds to the estimate from specification (5). In particular, since the ratio $(1-q)/(1-p)$ is always greater than one, the inflated coefficient will again be greater (in its absolute value) than the one estimated from the mixed sample.

A.2 Development of the non-parametric estimators

Following the description in the main text, we derive non-parametric estimates for the *excess hazard* for policyholders that settled their policy as a function of time. To formalize our notions of *excess hazard*, assume we are given two individuals S and R with hazard rates $\{\mu_t^S\}_{t\geq 0}$ and $\{\mu_t^R\}_{t\geq 0}$, respectively, that differ only in the information regarding their settlement decision but are identical with respect to all observables. More precisely, assume that we know S settled her policy whereas the settlement decision for R is not known. Then we can define the *proportional excess hazard* $\{\alpha(t)\}_{t\geq 0}$ and the *additive excess hazard* $\{\beta(t)\}_{t\geq 0}$ via the following relationships:

$$\mu_t^S = \alpha(t) \times \mu_t^R \text{ and } \mu_t^S = \beta(t) + \mu_t^R.$$

Andersen and Vaeth (1989) provide non-parametric estimators for the proportional and additive excess hazard by relying on the Nelson-Aalen (N-A) estimator for $\int_0^t \alpha(s)\,ds$ and the Kaplan-Meier (K-M) estimator for $\int_0^t \beta(s)\,ds$, respectively. However, their approach relies on the assumption that the baseline mortality (μ_t^R in our specification) is known, whereas we only have available estimates $\{\hat{\mu}_t^{(i)}\}_{t\geq 0}$, $1 \leq i \leq N$, given by the LE provider. Therefore, for the estimation of the proportional excess hazard, we instead use the following three-step procedure that relies on a repeated application of the Andersen and Vaeth (1989) estimator:

1. We start with the specification:

$$\mu_t^{(i)} = A(t) \times \hat{\mu}_t^{(i)},\ 1 \leq i \leq N, \tag{10}$$

   and use the Andersen and Vaeth (1989) excess hazard estimator to obtain an estimate for A, say $\hat{A}$, based on the full dataset. Hence, $\hat{A}$ corrects systematic deviations of the given estimates based on the observed times of death (in sample). We set:

$$\bar{\mu}_t^{(i)} = \hat{A}(t) \times \hat{\mu}_t^{(i)},\ 1 \leq i \leq N,$$

   for the corrected individual baseline hazard rate.

2. We then use the specification:

$$\mu_t^{(i)} = \alpha(t) \times \bar{\mu}_t^{(i)} \tag{11}$$

   for individual i in the closed subsample. Note that if we used the full dataset to estimate α, we would obtain $\alpha(t) \equiv 1$ and $\int_0^t \alpha(s)\,ds$ would be a straight line with slope one. However, when applying (11) to the subsample of closed policies, the resulting estimate for α—or rather $\int_0^t \alpha(s)\,ds$—picks up the residual hazard information due to the settlement decision.

3. Finally, we derive an estimate for α from the cumulative estimate ($\int_0^t \alpha(s)\, ds$) using a suitable kernel function as in Wang (2005).

For the additive excess hazard, we proceed analogously replacing Equations (10) and (11) by:

$$\mu_t^{(i)} = B(t) + \hat{\mu}_t^{(i)} \text{ and } \mu_t^{(i)} = \beta(t) + \underbrace{\left[\hat{B}(t) + \hat{\mu}_t^{(i)}\right]}_{=\bar{\mu}_t^{(i)}},$$

respectively.

In the context of Figures 1, 2, and A.2, for the derivation of the derivatives in Step 3., we use the Epanechnikov kernel with a fixed bandwidth of one.

A.3 A Version of the Model from Section 2 with Price Uncertainty

Assume the LS company has access to an additional estimate for the insured's probability of death q that is not known to the econometrist, say θ. Here, we assume that the underlying probability measure $\mathbb{P}$ reflects all available information and, to simplify the presentation, we ignore uncertainty in ψ. Since we interpret θ as a signal for q, we assume (i) that a higher θ will result in a higher offer price, i.e. $\pi(\theta)$ is increasing, and (ii) that q is stochastically increasing in θ. Then it is easy to see that:

$$\mathbb{E}\left[q|q < \pi(\theta) + \psi, \theta\right] \text{ is increasing as a function of } \theta. \tag{12}$$

Indeed, it is sufficient to assume the weaker condition (12) holds, which solely indicates that the estimate for q conditional on a policyholder settling her policy is increasing in θ.

Now if the econometrist finds a negative correlation between settling and dying, in the context of this extended model this means:

$$\mathbb{E}\left[q|q < \pi(\theta) + \psi\right] < \mathbb{E}\left[q\right], \tag{13}$$

where the conditional expectation on the left-hand side incorporates all the information available to the econometrist (reflected in $\mathbb{P}$) and the observation that the policyholder settled. However, the question from the point of view of the LS company—which, as indicated in Footnote 13, is the relevant perspective—is whether there exists asymmetric information, indicated by a negative correlation, when incorporating all pricing-relevant information, particularly θ:

$$\mathbb{E}\left[q|q < \pi(\theta) + \psi, \theta\right] \overset{?}{<} \mathbb{E}\left[q|\theta\right]$$

for at least some choices of θ. When aggregating over all policyholders:

$$\mathbb{E}\left[\mathbb{E}\left[q|q<\pi(\theta)+\psi,\theta\right]\right] \overset{?}{<} \mathbb{E}\left[\mathbb{E}\left[q|\theta\right]\right] = \mathbb{E}\left[q\right]. \tag{14}$$

Therefore, the question of whether the observed relationship (13) provides definite evidence for the relevant relationship (14) depends on the relationship between the expectations on the left-hand sides of (13) and (14). In particular, the implication will hold if:

$$\mathbb{E}\left[\mathbb{E}\left[q|q<\pi(\theta)+\psi,\theta\right]\right] \leq \mathbb{E}\left[q|q<\pi(\theta)+\psi\right]. \tag{15}$$

We need the following lemma:

Lemma A.1. *Let X be a real random variable, g be an increasing function such that $\mathbb{E}\left[g(X)\right]=0$, and h be an increasing and positive function. Then $\mathbb{E}\left[g(X)\,h(X)\right]\geq 0$.*

Proof. Let $K = \text{argmax}_x\{g(x)\leq 0\}$. Then:

$$0=\mathbb{E}[g(X)]=\mathbb{E}\left[g(X)|X\leq K\right]\,\mathbb{P}(X\leq K)+\mathbb{E}\left[g(X)|X>K\right]\,\mathbb{P}(X>K).$$

Now clearly $g(X)\,h(K)\leq g(X)\,h(X)$ on $\{X\leq K\}$, so that

$$\mathbb{E}\left[g(X)\,h(K)|X\leq K\right]\leq \mathbb{E}\left[g(X)\,h(X)|X\leq K\right].$$

Similarly,

$$\mathbb{E}\left[g(X)\,h(K)|X>K\right]\leq \mathbb{E}\left[g(X)\,h(X)|X>K\right].$$

Thus,

$$\begin{aligned}
0 &= \mathbb{E}\left[g(X)\,h(K)|X\leq K\right]\,\mathbb{P}(X\leq K)+\mathbb{E}\left[g(X)\,h(K)|X>K\right]\,\mathbb{P}(X>K)\\
&\leq \mathbb{E}\left[g(X)\,h(X)|X\leq K\right]\,\mathbb{P}(X\leq K)+\mathbb{E}\left[g(X)\,h(X)|X>K\right]\,\mathbb{P}(X>K)\\
&= \mathbb{E}\left[g(X)\,h(X)\right].
\end{aligned}$$

$\square$

Now, by the tower property of conditional expectations, (15) is equivalent to:

$$\begin{aligned}
\frac{\mathbb{E}\left[\mathbb{E}\left[q\ \mathbf{1}_{\{q<\pi(\theta)+\psi\}}|\theta\right]\right]}{\mathbb{P}(q<\pi(\theta)+\psi)}-\mathbb{E}\left[\mathbb{E}\left[q|q<\pi(\theta)+\psi,\theta\right]\right] &\geq 0\\
\Leftrightarrow \mathbb{E}\left[\mathbb{E}\left[q|q<\pi(\theta)+\psi,\theta\right]\underbrace{\left(\frac{\mathbb{P}(q<\pi(\theta)+\psi|\theta)}{\mathbb{P}(q<\pi(\theta)+\psi)}-1\right)}_{=g(\theta)}\right] &\geq 0.
\end{aligned}$$

Since $\mathbb{E}\left[q|q<\pi(\theta)+\psi,\theta\right]$ is increasing as a function of θ by our assumption and since $\mathbb{E}\left[g(\theta)\right]=0$, with the lemma relationship (14) will hold if g is increasing. Note that g is an affine transformation of the proportion of policyholders deciding to settle given the estimate θ, so that the pivotal relationship is the increasingness of this proportion in θ. Conversely, the implication will go in the other direction, so that that the econometrist's analysis will potentially overstate the effect, if the proportion of policyholders settling their policy is decreasing in the estimate.

B Supplemental Results

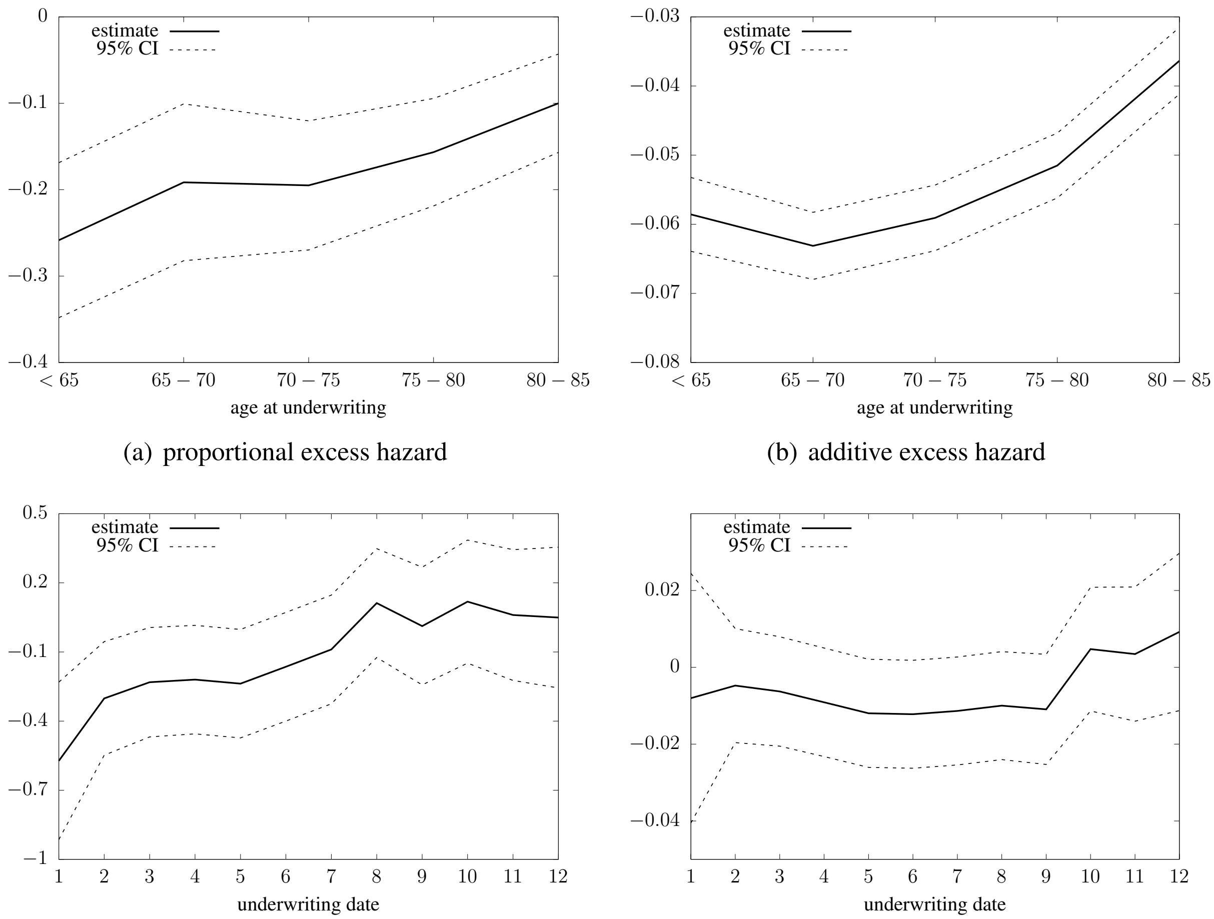
Table A.1 presents supplemental survival regression results based on proportional and additive hazards specifications. Columns [A] and [B] in the table show results for the earliest observation date with alternative time trend specification under the proportional hazards specification (adding a quadratic trend, $\text{SaO}_i \times \ln^2(1+t)$, in column [A] and adding a linear trend, $\text{SaO}_i \times t$, in column [B]). Columns [D] and [E] show comparable results under the additive hazards specification (adding a quadratic trend, $\text{SaO}_i \times t^2$, in column [D] and adding a logarithmic trend, $\text{SaO}_i \times \ln(1+t)$, in column [E]). As is evident from the table, the quadratic trend components fail to be significant in both specifications, whereas the alternative time trend specification in either case provides qualitatively similar conclusions on informational frictions, yet with lower likelihood values when compared with the default trend choices in the main text.

Columns [C] and [F] of Table A.1 present results when using dummy variables for underwriting date and age at underwriting in the proportional and additive hazards specification, respectively. Comparing the estimates to the baseline results from Tables 2 and 3, it is clear that using the (log-)linear trend did not have a significant impact on the results. In particular, there is little change in the settlement-related variables that are in the focus of our analysis. This is also illustrated by Figure A.1 that plots coefficients of the corresponding dummy variables, from which we note that a basic increasing trend assumption can capture the relevant shape.

Tables A.2 and A.3 present the results of the impact of settling on life expectancy for representative US male and female policyholders under various model specification, using the approximate inflation formulas as derived in Equations (7) and (9) in Appendix A.1. Among all cases, we estimate LE increases from 0.21 to 2.21 years for policyholders who choose to settle their policy. Percentage-wise, the relative change of LE for settled policyholders ranges from 2.8% to 16.0%. We note that the estimated LE increases are more pronounced under the time-constant trend assumption.

	[A]	[B]	[C]	[D]	[E]	[F]
$^1/_{14} \times \int_0^{14} \beta_0(t)\,dt$	0.0182	0.0189	0.4645	-0.0090	-0.0090	0.0061
Estimated hazard, $\hat{\mu}_t^{(i)}$	0.8966^{***}	0.8974^{***}	0.8804^{***}	0.2031^{***}	0.2031^{***}	0.1822^{***}
	(0.0101)	(0.0101)	(0.0125)	(0.0067)	(0.0067)	(0.0069)
Underwriting date, $\ln(1+\mathrm{DOU}_i)$	0.3048^{***}	0.3039^{***}	-	0.0001	0.0001	-
	(0.0277)	(0.0277)	-	(0.0002)	(0.0002)	-
Age at underwriting, $\ln(1+\mathrm{AU}_i)$	0.5880^{***}	0.5798^{***}	-	0.0016^{***}	0.0016^{***}	-
	(0.0828)	(0.0828)	-	(0.0001)	(0.0001)	-
Sex, SE_i	-0.0984^{***}	-0.0990^{***}	-0.0866^{***}	0.0042^{***}	0.0043^{***}	0.0049^{***}
	(0.0196)	(0.0196)	(0.0202)	(0.0007)	(0.0007)	(0.0007)
Primary impairment 1, PI_1	-0.0385	-0.0451	-0.0185	0.0129	0.0129	0.0122
	(0.2863)	(0.2863)	(0.2867)	(0.0150)	(0.0150)	(0.0150)
Primary impairment 2, PI_2	-0.0286	-0.0341	0.0416	0.0792^{***}	0.0792^{***}	0.0820^{***}
	(0.2894)	(0.2894)	(0.2902)	(0.0192)	(0.0192)	(0.0192)
Primary impairment 3, PI_3	0.1639	0.1590	0.2388	0.0028	0.0028	0.0055
	(0.2835)	(0.2835)	(0.2840)	(0.0149)	(0.0149)	(0.0149)
Primary impairment 4, PI_4	0.1945	0.1887	0.2404	0.0292^{*}	0.0293^{*}	0.0271^{*}
	(0.2854)	(0.2854)	(0.2858)	(0.0153)	(0.0153)	(0.0154)
Primary impairment 5, PI_5	-0.0005	-0.0063	0.0607	0.0282^{*}	0.0283^{*}	0.0279^{*}
	(0.2812)	(0.2812)	(0.2818)	(0.0150)	(0.0150)	(0.0150)
Primary impairment 6, PI_6	-0.0429	-0.0377	0.0385	-0.0043	-0.0043	-0.0022
	(0.2793)	(0.2793)	(0.2799)	(0.0147)	(0.0147)	(0.0148)
Primary impairment 7, PI_7	0.1994	0.1942	0.2793	0.0051	0.0051	0.0071
	(0.2788)	(0.2788)	(0.2794)	(0.0147)	(0.0147)	(0.0148)
Primary impairment 8, PI_8	0.2095	0.2042	0.2840	0.0043	0.0043	0.0063
	(0.2811)	(0.2811)	(0.2817)	(0.0148)	(0.0148)	(0.0148)
Primary impairment 9, PI_9	0.8207^{***}	0.8164^{***}	0.9004^{***}	0.0878^{***}	0.0878^{***}	0.0860^{***}
	(0.2820)	(0.2820)	(0.2825)	(0.0160)	(0.0160)	(0.0160)
Primary impairment 10, PI_{10}	0.1024	0.0967	0.1648	-0.0036	-0.0036	-0.0007
	(0.2813)	(0.2812)	(0.2818)	(0.0148)	(0.0148)	(0.0148)
Primary impairment 11, PI_{11}	0.9003^{***}	0.8952^{***}	0.9591^{***}	0.0484^{***}	0.0485^{***}	0.0476^{***}
	(0.2837)	(0.2837)	(0.2843)	(0.0154)	(0.0154)	(0.0154)
Primary impairment 12, PI_{12}	0.4029	0.3970	0.4702^{*}	0.0084	0.0085	0.0118
	(0.2818)	(0.2818)	(0.2824)	(0.0148)	(0.0148)	(0.0149)
Primary impairment 13, PI_{13}	0.1031	0.0978	0.1712	-0.0047	-0.0046	-0.0026
	(0.2790)	(0.2790)	(0.2796)	(0.0147)	(0.0147)	(0.0148)
Primary impairment 14, PI_{14}	0.3726	0.3671	0.4535	0.0151	0.0151	0.0172
	(0.2790)	(0.2789)	(0.2796)	(0.0148)	(0.0148)	(0.0148)
Primary impairment 15, PI_{15}	-0.1945	-0.1997	-0.1326	-0.0142	-0.0142	-0.0107
	(0.2801)	(0.2801)	(0.2808)	(0.0147)	(0.0147)	(0.0148)
Smoker, $\mathrm{SM}_{i,1}$	0.3739^{***}	0.3733^{***}	0.3604^{***}	0.0298^{***}	0.0297^{***}	0.0268^{***}
	(0.0429)	(0.0429)	(0.0432)	(0.0026)	(0.0026)	(0.0026)
“Aggregate” smoking status, $\mathrm{SM}_{i,2}$	0.2117^{***}	0.2123^{***}	0.2213^{***}	0.0116^{***}	0.0116^{***}	0.0105^{***}
	(0.0551)	(0.0551)	(0.0552)	(0.0025)	(0.0025)	(0.0025)
Settled-and-Observed, SaO_i	-0.5994^{***}	-0.3300^{***}	-0.4650^{***}	-0.0113^{***}	-0.0143^{***}	-0.0101^{***}
	(0.1189)	(0.0419)	(0.0643)	(0.0016)	(0.0016)	(0.0012)
Settled-and-Observed $\times$ trend,	0.3993^{**}		0.2127^{***}		0.0063^{***}	
$\mathrm{SaO}_i \times \ln(1+t)$	(0.1691)		(0.0361)		(0.0011)	
Settled-and-Observed $\times$ quadratic trend,	-0.0608					
$\mathrm{SaO}_i \times \ln^2(1+t)$	(0.0567)					
Settled-and-Observed $\times$ trend,		0.0411^{***}		0.0018^{**}		0.0014^{***}
$\mathrm{SaO}_i \times t$		(0.0071)		(0.0009)		(0.0003)
Settled-and-Observed $\times$ quadratic trend,				-4.1×10^{-5}		
$\mathrm{SaO}_i \times t^2$				(0.0001)		
Log-likelihood value	$-134,011.87$	$-134,015.18$	$-133,987.62$	$-74,322.44$	$-74,357.20$	-

Table A.1: Survival regression supplemental results. Column [A]: Proportional hazards assumption with additional quadratic trend, earliest observation date; column [B]: Proportional hazards assumption with linear time trend, earliest observation date; column [C]: Proportional hazards assumption with logarithmic time trend and dummy variables for underwriting date and age at underwriting, earliest observation date; column [D]: Additive hazards assumption with additional quadratic trend, earliest observation date; column [E]: Additive hazards assumption with logarithmic time trend, earliest observation date; column [F]: Additive hazards assumption with linear time trend and dummy variables for underwriting date and age at underwriting, earliest observation date. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.



(a) proportional excess hazard

(b) additive excess hazard

(c) proportional excess hazard

(d) additive excess hazard

Figure A.1: Survival regression parameter estimations (left panels: Proportional; right panels: Additive) of dummy variables for age at underwriting (top panels) and underwriting date (bottom panels), with point-wise 95% confidence intervals (dashed curves); earliest observation date.

	Proportion of closed policies (p)					
	24.5%	30%	40%	50%	60%	70%
Age 70 (non-adjusted LE 13.93*)*						
Proportional hazards; time-constant effect						
LE for settled cases	14.69	14.75	14.88	15.05	15.30	15.69
LE increase	0.76	0.82	0.95	1.12	1.37	1.76
Additive hazards; time-weakening effect						
LE for settled cases	14.37	14.41	14.49	14.60	14.77	15.07
LE increase	0.44	0.48	0.56	0.67	0.84	1.14
Additive hazards; time-constant effect						
LE for settled cases	14.60	14.65	14.78	14.96	15.23	15.70
LE increase	0.67	0.72	0.85	1.03	1.30	1.77
Age 75 (non-adjusted LE 10.48*)*						
Proportional hazards; time-constant effect						
LE for settled cases	11.13	11.17	11.28	11.43	11.64	11.98
LE increase	0.65	0.69	0.80	0.95	1.16	1.50
Additive hazards; time-weakening effect						
LE for settled cases	10.79	10.82	10.88	10.96	11.08	11.30
LE increase	0.31	0.34	0.40	0.48	0.60	0.82
Additive hazards; time-constant effect						
LE for settled cases	10.88	10.91	10.98	11.09	11.25	11.52
LE increase	0.40	0.43	0.50	0.61	0.77	1.04
Age 80 (non-adjusted LE 7.50*)*						
Proportional hazards; time-constant effect						
LE for settled cases	8.02	8.06	8.15	8.26	8.43	8.70
LE increase	0.52	0.56	0.65	0.76	0.93	1.20
Additive hazards; time-weakening effect						
LE for settled cases	7.71	7.73	7.77	7.83	7.91	8.05
LE increase	0.21	0.23	0.27	0.33	0.41	0.55
Additive hazards; time-constant effect						
LE for settled cases	7.72	7.74	7.78	7.83	7.92	8.07
LE increase	0.22	0.24	0.28	0.33	0.42	0.57

Table A.2: Comparisons of average life expectancies between population-level and settled US male policyholders; various model specifications.

	Proportion of closed policies (p)					
	24.5%	30%	40%	50%	60%	70%
Age 70 (non-adjusted LE 15.97*)*						
Proportional hazards; time-weakening effect						
LE for settled cases	16.25	16.26	16.29	16.34	16.39	16.46
LE increase	0.28	0.29	0.32	0.37	0.42	0.49
Proportional hazards; time-constant effect						
LE for settled cases	16.69	16.74	16.86	17.02	17.25	17.60
LE increase	0.72	0.77	0.89	1.05	1.28	1.63
Additive hazards; time-weakening effect						
LE for settled cases	16.48	16.52	16.61	16.75	16.95	17.29
LE increase	0.51	0.55	0.64	0.78	0.98	1.32
Additive hazards; time-constant effect						
LE for settled cases	16.80	16.87	17.02	17.25	17.59	18.18
LE increase	0.83	0.90	1.05	1.28	1.63	2.21
Age 75 (non-adjusted LE 12.25*)*						
Proportional hazards; time-weakening effect						
LE for settled cases	12.58	12.60	12.64	12.69	12.76	12.84
LE increase	0.33	0.35	0.39	0.44	0.51	0.59
Proportional hazards; time-constant effect						
LE for settled cases	12.87	12.92	13.02	13.16	13.35	13.66
LE increase	0.62	0.67	0.77	0.91	1.10	1.41
Additive hazards; time-weakening effect						
LE for settled cases	12.63	12.66	12.73	12.83	12.98	13.23
LE increase	0.38	0.41	0.48	0.58	0.73	0.98
Additive hazards; time-constant effect						
LE for settled cases	12.76	12.80	12.90	13.03	13.24	13.59
LE increase	0.51	0.55	0.65	0.78	0.99	1.34
Age 80 (non-adjusted LE 8.90*)*						
Proportional hazards; time-weakening effect						
LE for settled cases	9.29	9.31	9.35	9.41	9.49	9.59
LE increase	0.39	0.41	0.45	0.51	0.59	0.69
Proportional hazards; time-constant effect						
LE for settled cases	9.40	9.44	9.52	9.63	9.79	10.15
LE increase	0.50	0.54	0.62	0.73	0.89	1.15
Additive hazards; time-weakening effect						
LE for settled cases	9.16	9.18	9.23	9.30	9.40	9.58
LE increase	0.26	0.28	0.33	0.40	0.50	0.68
Additive hazards; time-constant effect						
LE for settled cases	9.19	9.21	9.26	9.34	9.45	9.64
LE increase	0.29	0.31	0.36	0.44	0.55	0.74

Table A.3: Comparisons of average life expectancies between population-level and settled US female policyholders; various model specifications.

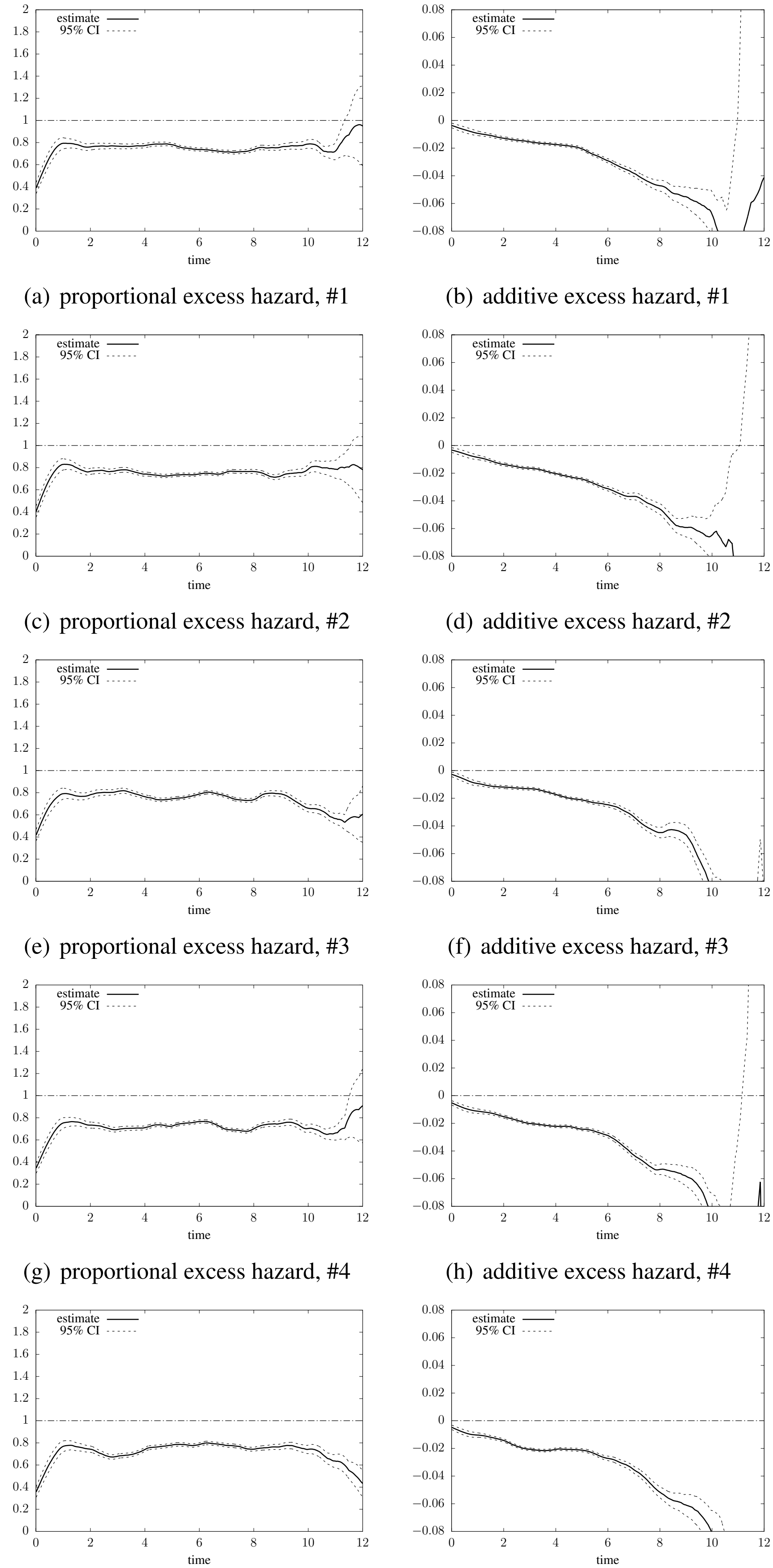
C Monte Carlo Experiment on Excess Hazard

To implement a Monte Carlo version of the thought experiment in Section 6, we first run a least-square regression of the logarithm of the mortality multipliers on the observable characteristics (excluding the multipliers themselves and the settlement-related variables). We also include significant interactions of the terms so that we have 41 covariates in total. Based on the regression model, we then derive projected life expectancies as well as the standard deviation of the error term. The projected life expectancies will then be used as the benchmark of the assessment in the Monte Carlo experiment, i.e. we assume these present the *true* life expectancies.

Now, following the logic from Section 6, assume that a fraction of all cases enter into a life settlement transaction and the brokers commissioned with the sale "cherry-pick" among the available LEs (multiplier estimates). More precisely, assume that for these transactions, several LEs from various LE providers will be obtained but only the shortest LE (highest multiplier) is submitted. Alternatively, we may assume that there are several offers from various LS companies that base their pricing on different mortality multipliers, and the one with the highest bidding price (corresponding to the highest multiplier estimate) will make the trade. Importantly, while in the context of this experiment we assume the policyholder does not have private information on her survival prospects, note that there still exists an informational asymmetry—the broker and/or policyholder will have more information than the *winning* LS company—but this asymmetry emerges in the transaction process.

Assume that each LE provider's estimate is based on the same projected mortality multiplier plus a varying error term (with mean of zero), according to our regression estimates. For simplicity, we assume that the submitted (highest) multiplier corresponds to the 75th percentile. Based on this logic, closed cases are systematically assessed with shorter LEs, whereas the remaining cases have no systematic deviation. We use the resulting multipliers to generate a hypothetical set of forecasts $\hat{\mu}_t^{(i)}$, $1 \leq i \leq N$, where we use the skewed (75th percentile) multiplier for the (randomly sampled) closed cases and the projected (median) multiplier for the remaining cases. Based on the simulated sample, we derive non-parametric estimators similarly as in Section 4.

Figure A.2 presents the results for five different simulated datasets, where as for our actual dataset we assume 13,221 out of the 53,947 policyholders are Settled-and-Observed. While the shapes and magnitudes differ between the simulated datasets, we observe that the proportional excess hazard roughly evolves according to a straight line below one, whereas the additive excess hazard diverges. This is consistent with the assertions in Section 6. While it is possible that there are systematic differences in the underwriting process between the LE providers, it is difficult to construct a situation that yields the observed patterns from Section 4 based on this selection process.



(a) proportional excess hazard, #1 (b) additive excess hazard, #1

(c) proportional excess hazard, #2 (d) additive excess hazard, #2

(e) proportional excess hazard, #3 (f) additive excess hazard, #3

(g) proportional excess hazard, #4 (h) additive excess hazard, #4

(i) proportional excess hazard, #5 (j) additive excess hazard, #5

Figure A.2: Monte-Carlo experiment of non-parametric estimates of excess hazard.

	Simulation				
	#1	#2	#3	#4	#5
Proportional hazards specification					
Settled-and-Observed, SaO_i	-0.3218^{***}	-0.2568^{***}	-0.2890^{***}	-0.4372^{***}	-0.3896^{***}
	(0.0596)	(0.0587)	(0.0589)	(0.0599)	(0.0598)
Settled-and-Observed $\times$ trend, $\mathrm{SaO}_i \times \ln(1+t)$	-0.0133	-0.0499	-0.0358	0.0204	0.0181
	(0.0347)	(0.0342)	(0.0343)	(0.0347)	(0.0347)
Additive hazards specification					
Settled-and-Observed, SaO_i	-0.0009	0.0001	0.0001	-0.0029^{**}	-0.0023^{**}
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Settled-and-Observed $\times$ trend, $\mathrm{SaO}_i \times t$	-0.0009^{***}	-0.0010^{***}	-0.0012^{***}	-0.0008^{***}	-0.0007^{**}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)

Table A.4: Survival regression results for the five Monte-Carlo simulated datasets. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.4 confirms the non-parametric findings by presenting corresponding survival regression results based on the same five simulated datasets, using both proportional and additive specifications with time trend. For simplicity we only show regressed coefficients for settlement-related covariates. We observe from the table consistent results across all simulation trials. In particular, for the proportional specification, the intercept of the trend starts at significantly negative values, however, the slope of the trend is insignificant and very close to zero, suggesting a persistent impact of settling on survival prospects. For the additive specification, the slope of the trend is no longer positive but significantly negative, which is again necessary to sustain the constant proportional trend as assumed in the skewing process of our Monte Carlo simulation.

References

Andersen, Per Kragh, and Michael Vaeth (1989). "Simple parametric and nonparametric models for excess and relative mortality." *Biometrics*, 45: 523-535.

Lin, D.Y., and Zhiliang Ying (1994). "Semiparametric analysis of the additive risk model." *Biometrika*, 81: 61-71.

Wang, Jane-Ling (2005). "Smoothing Hazard Rates." *Encyclopedia of Biostatistics*, 2005.