Adverse Selection in Secondary Insurance Markets: Evidence from the Life Settlement Market *

Daniel Bauer†
Department of Risk Management and Insurance, Georgia State University
35 Broad Street, 11th Floor; Atlanta, GA 30303; USA

Jochen Russ
Institut für Finanz- und Aktuarwissenschaften and Ulm University
Lise-Meitner-Straße 14; 89081 Ulm; Germany

Nan Zhu
Department of Mathematics, Illinois State University
100 N University Street; Normal, IL 61790; USA

February 2014

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 the average difference between realized lifetimes and estimated life expectancies for a sub-sample of settled policies relative to the entire sample. We find a significant positive difference indicating private information on mortality prospects. Using non-parametric estimates for the excess mortality and survival regressions, we show that the informational advantage is temporary and wears off over five to six years. We argue this is in line with adverse selection on an individual's condition, which has important economic consequences for the life settlement market and beyond.

JEL classification: D12; G22; J10

Keywords: Life Settlements; Life Expectancy; Asymmetric Information; Adverse Selection

---

*We are grateful for support from Fasano Associates, especially to Mike Fasano. Moreover, we thank seminar participants at Georgia State University and, particularly, Stephen Shore for valuable comments. Bauer and Zhu also gratefully acknowledge financial support from the Society of Actuaries under a CAE Research Grant.

†Corresponding author. Phone: +1-(404)-413-7490. Fax: +1-(404)-413-7499. Email addresses: dbauer@gsu.edu (Bauer); j.russ@ifa-ulm.de (Russ); nzhu@ilstu.edu (Zhu).

1 Introduction

Although seminal theoretical contributions have emphasized the significance of asymmetric information for a market's functioning since the 1960s (Arrow, 1963; Akerlof, 1970; Rothschild and Stiglitz, 1976), the corresponding empirical literature has flourished only relatively recently.[1] 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). One market segment that offers the benefits of insurance data but is not subject to confounding factors—or at least, not to the same confounding factors—is the secondary insurance market, yet this aspect to our knowledge has not been explored thus far.

In this paper, we provide evidence for asymmetric information in the secondary life insurance market, the market for so-called *life settlements*. Within such a transaction, a policyholder sells—or *settles*—her life-contingent insurance benefits for a lump sum payment, where the offered price depends on an individualized evaluation of her survival probabilities by a third party *life expectancy provider* (LE provider). In particular, using the full dataset of a large US LE provider, we show that individuals have significant private information regarding their mortality prospects that is particularly pronounced in the early period after settling but is "decreasing" in the sense that the informational advantage wears off over about five to six years. We argue that such a pattern is in line with "hidden information" with respect to the current health state but does not cohere with "hidden actions" that affect future survival probabilities. As such, we can characterize the informational friction as *adverse selection* rather than *moral hazard*.[2]

Of course, 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). However, while naturally our quantitative results are specific to our setting, we believe our qualitative insights lend themselves to draw broader conclusions. On the one hand, with regards to primary life insurance markets, our results reinforce the insight that individuals have private information on their lifetime distribution and make use of it, despite the mixed evidence from corresponding studies. In particular, our characterization supports the proposition from various authors that informational asymmetries in these markets originate from adverse selection rather than moral hazard, which previously were based on "intuitive insight rather than on rigorous evidence" (He, 2009). On the

[1] See, e.g., Puelz and Snow (1994), Cawley and Philipson (1999), Chiappori and Salanié (2000), Cardon and Hendel (2001), Finkelstein and Poterba (2004), Cohen (2005), Finkelstein and McGarry (2006), or Cohen and Einav (2007).

[2] Our findings are in line with a recent industry study by Granieri and Heck (2013) that postdates the first draft 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."

other hand, our results suggest that individuals more broadly are competent in predicting their relative survival prospects, possibly in contrast to their ability in predicting absolute life expectancies that may be subject to framing and other behavioral biases (Payne et al., 2013). We believe that the former task may be more material for retirement planning given that individuals may be provided with background information on "typical" life expectancies or suitable "default" choices for average individuals.

Our empirical strategy relies on comparisons of mortality experience for the entire sample, consisting of policyholders who settled their policies and those who did not, with a subsample in which everyone is known to have settled their policy. Figure 1, the construction of which we describe in much more detail in Section 3, illustrates our basic result. Here, we plot multiplicative "excess mortality" for individuals that we know have settled their policy over time since the person's life expectancy was evaluated (blue solid curve). Naturally, if this curve had the shape of a horizontal line at one (black dotted line) or if the horizontal line at one fell within the 95% confidence intervals (red dashed curves), we would conclude that there is no (significant) effect of settling on an individual's mortality pattern. The observation that this curve is overall less than one, particularly over the early years after underwriting, implies that there is a negative association between settling and the instantaneous mortality probability. Thus, policyholders who settled their policies, on average, live longer. This result corresponds to the so-called *positive correlation test* for asymmetric information that tests the basic prediction of a positive relationship between (ex-post) risk and purchasing insurance coverage (Chiappori and Salanié, 2000, 2013), although in a secondary market setting, the mechanism is "reversed." Here, a policyholder will be more inclined to *sell* her policy if she is a "*low risk*"—i.e., if she has a low probability of dying. Hence, the observed negative relationship documents the existence of asymmetric information in the life settlement market.

An alternative (and simpler) approach to establishing this negative relationship is to compare the average difference in realized partial life spans (to date) and estimated temporary life expectancies (DTLE) for the entire sample and the subsample of settled policies. We find that for the entire sample, the average DTLE is around one month. This indicates that overall the life expectancy evaluations are fairly accurate and, at least for the LE provider in view, do not exhibit a considerable positive bias as opposed to findings by other authors that may have been based on data from different LE providers (Gatzert, 2010). In contrast, the DTLE for the settled cases is significantly greater and amounts to around four months, also demonstrating the negative relationship between mortality and settling under asymmetric information.

In addition to the existence of asymmetric information, Figure 1 illustrates the pattern over time, which in turn provides insights on the characteristics of the informational friction.[3] More

[3]The idea to consider dynamic relationships to characterize asymmetric information already occurs in Abbring et

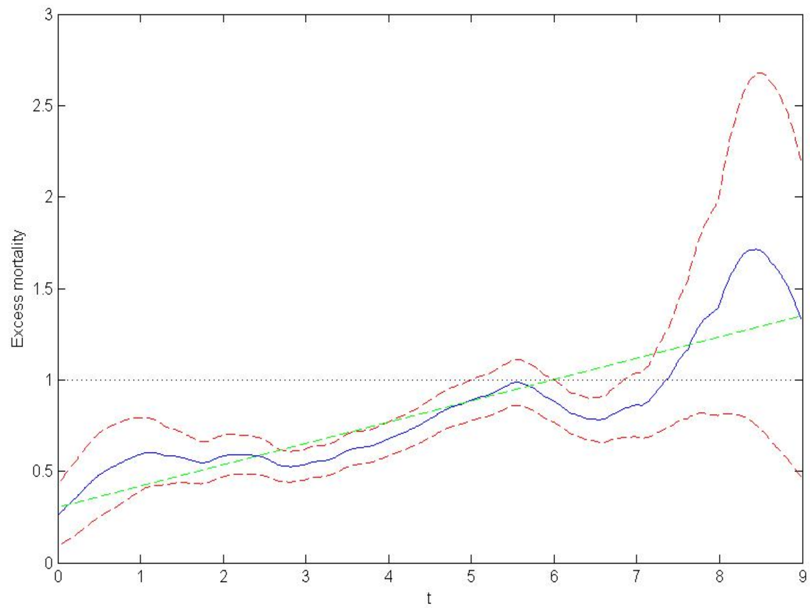


Figure 1: Excess mortality for settled subsample

precisely, we observe that the informational advantage is particularly pronounced in the early years after underwriting but that it is decreasing in the sense that it wears off over five to six years. Such a structure is akin to so-called *select-and-ultimate* life tables in actuarial or demographic studies that capture temporary selection effects, and thus is in line with (adverse) selection on the initial health condition. In contrast, if the negative relationship were driven by hidden actions (moral hazard) such as a healthier lifestyle choice or other changes in behavior after settling, we would expect to see a steady or increasingly pronounced relationship over time.

To ascertain that our results are not driven by idiosyncrasies of our sample of settled policies, we run detailed semi-parametric survival regressions controlling for observables. In particular, we include the estimated force of mortality from the LE provider as well as all observable characteristics such as the time of underwriting, the primary impairment, etc. as explanatory variables.[4] Several—though not all—of the (known) primary impairment dummies fail to be significant, indicating that the LE provider's estimates correctly account for these. In contrast, having settled the policy (unknown to the LE provider at the time of underwriting) has a significant negative effect on the force of mortality, particularly when also considering a linear trend. The slope of the trend is significant and positive, implying that the effect becomes weaker over time. Hence, the results of the survival regressions reinforce the insights from the aggregate non-parametric estimate shown

al. (2003) in the context of experience ratings in automobile insurance.

[4]Chiappori and Salanié (2000) and Dionne et al. (2005) emphasize the importance of accounting for all pricing-relevant observed covariates.

in Figure 1: We find evidence for asymmetric information, though the informational advantage wears off over time—in line with adverse selection on the initial condition.

To appraise the quantitative impact, we then rely on the regression results to calculate the implications of the settlement decision on an "average" individual's life expectancy. We find that the effect is considerable and, for a 75-year old male, amounts to around two years when considering the basic (constant) effect and around eight months when accounting for the temporary pattern, depending on the total proportion of settled policies in our dataset.

However, assessing the economic impact of adverse selection on the life settlement market will not be possible without putting in place more structure on the policyholder's decision process (Einav et al., 2007, 2010a). In particular, for answering policy-relevant questions regarding efficiency and welfare implications of this market, it is necessary to consider and estimate equilibrium models that also account for barriers to participate in this market (Einav et al., 2010b) and repercussions on primary insurance (Daily et al., 2008; Fang and Kung, 2010; Zhu and Bauer, 2011). While such questions are beyond the scope of this paper, we believe they are intriguing problems for future research.

Relationship to the Literature

Of course, there is an array of papers analyzing the existence of information asymmetries in insurance markets (see Footnote 1 for a selection). For life contingencies, Finkelstein and Poterba (2002, 2004) establish that there exists asymmetric information in the market for life annuities, whereas the evidence for life insurance is mixed (Cawley and Philipson, 1999; He, 2009; McCarthy and Mitchell, 2010; Wu and Gan, 2013). Cohen and Siegelman (2010) posit that these seemingly contrasting findings may be explained by a positive relationship between income and insurance coverage: A higher wealth or income may be negatively associated with mortality risk but positively associated with insurance coverage. Thus, it is not clear whether the evidence is due to confounding factors that are not priced (wealth, risk aversion) or a "true" informational advantage. In contrast, the pricing of life settlements is highly individualized, and wealth effects should, if they are present at all, imply that sicker people are more likely to settle (since sick people may be more wealth-constraint). Hence, our findings present crisper evidence that individuals possess—and make use of—superior information regarding their mortality prospects. Moreover, we complement the above studies in that we are able to provide insights on the characteristics of the informational friction that are in line with adverse selection.

More broadly, our results provide new evidence on individuals' ability to forecast their own mortality prospects. Recent contributions to the behavioral literature provide negative results in this direction, with forecasts differing according to the framing of corresponding questions and being subject to various biases (Elder, 2013; Payne et al., 2013; and references therein). However, these studies consider forecasts of absolute life expectancies. Our results indicate that individuals appear to fare better when evaluating their relative mortality prospects, which may be the more material task for retirement planning given that individuals can be provided with background information on population mortality averages or "default" choices that are suitable for average individuals.

2 A Simple Model

This section presents a simple one-period model to provide the intuition for our empirical analysis of the existence of asymmetric information in the life settlement market. 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 before time one is $\mathbb{P}(\tau < 1) = q$, where τ is the time of death.

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 present value of her contract:

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

where r is the risk-free interest rate and ψ characterizes the policyholder's proclivity for settling. The latter may originate from risk-averse policyholder preferences with a bequest motive as in Zhu and Bauer (2013) or from liquidity constraints. Here, we simply use ψ to capture deviations from a value-maximizing behavior, under which the market may collapse due to a "lemons problem" as in Akerlof (1970).

Thus, from the policyholder's perspective, the question of whether or not to settle the policy based on Equation (1) is deterministic. However, this may not be the case from the perspective of the life settlement company offering to purchase the policy since it may have imperfect information with respect to q and/or ψ.[5] More precisely, assume that the policyholder has private information on the mortality probability q but the life settlement company solely observes the expected value, $\mathbb{E}[q]$, across the entire population (potentially conditional on various observable characteristics such as medical impairments). Then, we obtain for the mortality probability conditional on the observation that the policyholder settled her policy:

$$\begin{aligned} \mathbb{P}(\tau < 1|\Delta = 1) = \mathbb{E}\left[q|\Delta = 1\right] &= \mathbb{E}\left[q\,\middle|\pi > q\, e^{-r} - \psi\right] \\ &= \mathbb{E}\left[q\,|q < (\pi + \psi)\, e^{r}\right] \leq \mathbb{E}[q] = \mathbb{P}(\tau < 1). \end{aligned} \tag{2}$$

[5]Of course, such an information 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.

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:

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

Therefore, this is simply a version of the well-known *correlation test* for the presence of asymmetric information that tests if (ex-post) risk and insurance coverage are positively related (Chiappori and Salanié, 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—i.e., if she has a low probability of dying.

The intuition for this result is quite straightforward: As indicated in (2), if the policyholder has private insights on her lifetime distribution, she will gladly agree to beneficial offers from her perspective while she will walk away from bad offers. Hence, a pool of settled policies, on average, will display "longer" life expectancies than the entire population of policyholders, controlling for observables.

Asymmetric information with respect to ψ alone, e.g. arising from liquidity constraints, does not yield this 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 wealth reflected in ψ may be positively linked to her propensity to survive, although such a relationship would arguably work in the opposite direction. In any case, a negative relationship will—directly or indirectly—originate from an information asymmetry with respect to the time of death.

To test for its existence, we therefore simply analyze the impact of settling on realized lifetimes. If mortality probabilities are significantly lower—or, alternatively, if aveage lifetimes are considerably higher—for policyholders who settled their policy, we will be able to conclude that there exists private information on mortality prospects.

3 Data and Basic Econometrical Approach

3.1 Data

Our primary dataset consists of 78,571 life expectancy evaluations underwritten by Fasano Associates (Fasano), a leading US LE provider, between beginning-of-year 2001 and end-of-year 2011. More precisely, for each record, aside from individual characteristics including sex, age, smoking status, and primary impairment, we are given a life expectancy estimate (LE) at a certain point in time. Here, the LE is calculated by applying a given individual mortality multiplier (frailty

	Average (Std Dev)		Count
Life Expectancy Estimate		*Male*	
	11.86		33,299
	(4.27)		(63.30%)
Underwriting Age		*Observed Deaths*	
	75.20		7,552
	(7.25)		(14.26%)

Table 1: Summary statistics for the entire 52,603 cases; earliest observation date.

	Average (Std Dev)		Count
Life Expectancy Estimate		*Male*	
	10.01		597
	(3.21)		(60.12%)
Underwriting Age		*Observed Deaths*	
	73.53		114
	(19.51)		(12.22%)

Table 2: Summary statistics for the closed 933 cases; earliest observation date.

factor)—which is the result of the underwriting process—on a specified standard mortality table. Hence, we can use this information to derive the entire estimated lifetime distribution for each record. We eliminate duplicates by either considering the earliest or the latest underwriting date for each individual. This leaves 52,603 distinct individuals. Beyond LEs, we are also given realized times of death for the individuals that had died before January 1st 2012. Thus, by comparing estimated and realized lifetimes, we can assess the quality of the LEs. Table 1 provides summary statistics on the LEs for the earliest observation date.

The full dataset contains LEs for policyholders that decided to settle their policies (so-called *closed* cases), LEs for policyholders that walked away from a settlement offer, and LEs for individuals that were underwritten for different reasons. Typically, the LE provider does not receive feedback of whether or not a policy closed, so that this aspect is unknown for our full dataset. However, we also have access to portfolio information for three life settlement investors consisting of overall 933 policies underwritten by Fasano. Hence, for this small subsample of individuals, we have the additional information that they settled their policies. We will refer to this set as the *subsample of closed policies*, whereas we will refer to the rest of the sample as the *remaining cases*. Table 2 provides corresponding summary statistics for the closed subsample, also based on the earliest observation date.

We thus can compare the quality of the LEs for the closed subsample and the entire sample to analyze whether there exists a significant (positive) difference, as suggested by the asymmetric information test described in the previous section. For this purpose, we need an aggregate statistic to assess the quality of a sample of (heterogeneous) LEs. The next section introduces the *average difference in realized lifetimes and projected life expectancies* (DLE) and the *average difference in realized partial lifetimes and projected temporary life expectancies* (DTLE) as suitable candidates.

3.2 Assessing Life Expectancy Estimates

Assume we knew the time of death for each individual in a given sample of LEs. Then, we could calculate the difference between the realized lifetime and the given LE in each case. For a given individual, of course this would be a random variable due to the intrinsic randomness of death. However, under the hypothesis that the LEs are accurate, the average of the (mean-zero) random variables with bounded variance should converge to zero by the law of large numbers. We refer to this average as the *average difference in realized lifetimes and projected life expectancies* (DLE).

The key issue with this approach within our setting is that our death data are right-censored. That is, we only observe times of death that occurred before January 1st 2012, whereas for other individuals we solely know that they are still alive at this cut-off date. However, since we are given the entire estimated lifetime distribution, we are able to rely on an alternative concept from actuarial science and demographic research, namely the *temporary life expectancy*. More precisely, for each individual i in a sample of LEs of size N, denote by $T^{(i)}$ the time of death distributed with force of mortality $\{\mu_t^{(i)}\}_{t\geq 0}$, $1 \leq i \leq N$. Then, at the cut-off date, we observe the right-censored version of the time of death $\bar{T}^{(i)} = \min\{T^{(i)}, t_i\}$, where t_i is the (given) difference between the cut-off date and the time of underwriting, all measured in years. Its expected value, the so-called *temporary life expectancy*, is given by the integral of the survival probabilities until time t_i (Bowers et al., 1997):

$$e_{\overline{t_i}|}^{(i)} = \mathbb{E}\left[\bar{T}^{(i)}\right] = \int_0^{t_i} \exp\left\{-\int_0^t \mu_s^{(i)}\, ds\right\} dt,\ 1 \leq i \leq N.$$

Denote the estimated force of mortality for individual i supplied by the LE provider by $\{\hat{\mu}_t^{(i)}\}_{t\geq 0}$, and denote the corresponding estimated temporary life expectancy by:

$$\hat{e}_{\overline{t_i}|}^{(i)} = \int_0^{t_i} \exp\left\{-\int_0^t \hat{\mu}_s^{(i)}\, ds\right\} dt,\ 1 \leq i \leq N.$$

Then, under the hypothesis that the estimated lifetime distributions are accurate, the difference in realized partial lifetime and estimated temporary life expectancy for individual $i = 1, \ldots, N$:

$$\bar{T}^{(i)} - \hat{e}_{\overline{t_i}|}^{(i)},$$

is a zero-mean random variable with bounded variance.[6] Thus, the average:

$$\text{DTLE}_N = \frac{1}{N}\sum_{i=1}^{N}\left[\bar{T}^{(i)} - \hat{e}_{\overline{t_i}|}^{(i)}\right],$$

which we will refer to as the *average difference in realized partial lifetimes and projected temporary life expectancies* (DTLE), will converge to zero by Kolmogorov's Law of Large Numbers. Moreover, with Lyapunov's Central Limit Theorem (Billingsley, 1995), we obtain:

$$\frac{\text{DTLE}_N}{\frac{1}{N}S_N^{TR}} = \frac{1}{S_N^{TR}}\sum_{i=1}^{N}\left[\bar{T}^{(i)} - \hat{e}_{\overline{t_i}|}^{(i)}\right] \to N(0,1), \tag{4}$$

where $(S_N^{TR})^2 = \sum_{i=1}^N \text{Var}[\bar{T}^{(i)}] \approx \sum_{i=1}^N (\bar{T}^{(i)} - \hat{e}_{\overline{t_i}|}^{(i)})^2$. In particular, we can rely on (4) to draw inference on the quality of the LEs.

Table 3 provides DTLE calculations for our entire dataset, the subsample of closed cases, and the remaining cases. The significance stars indicate the deviation from zero under the hypothesis

[6]It is important to note that this hypothesis is considerably stronger than the previous hypothesis that only the LEs are accurate. In particular, our tests rely on the entire estimated distributions whereas Fasano typically only supplies the expected values to their clients. Thus, our tests might be too stringent to appraise the quality of their estimates.

	Estimates		
	(1) All Cases	(2) Closed Cases	(3) Remaining Cases
Earliest observation			
N	52603	933	51670
DTLE	0.1646***	0.4580***	0.1593***
$\frac{1}{N}S_N^{TR}$	(0.0050)	(0.0368)	(0.0051)
Latest observation			
N	52603	933	51670
DTLE	0.0634***	0.1466***	0.0619***
$\frac{1}{N}S_N^{TR}$	(0.0047)	(0.0292)	(0.0048)

Table 3: Difference in projected and realized temporary life expectancies (DTLEs) in years for various subsamples.

that the estimated lifetime distributions are accurate. We find that the DTLEs for the entire sample amount to less than one month or slightly less than two months, depending on the observation time. Given that LEs are only provided up to full months and that there is some arbitrariness in the rounding procedure, a difference of less than one month as within the latest underwriting date is almost negligible for practical purposes—although it is still statistically significant due to the very large number of observations.[7]

In contrast, the DTLEs for the closed cases are almost two and five months for the latest and the earliest observation date, respectively. In particular, the (asymptotically Normal) difference between the closed and the remaining sample is positive and highly significant, implying that individuals that settled their policy in the secondary market, on average, live relatively longer. In view of the discussion from Section 2, these findings provide evidence for the existence of asymmetric information regarding mortality prospects in the life settlement market—which is the central result in this paper.

Another potential explanation for this finding is that the subsample of closed policies differs in some significant way from the full sample. Before addressing this issue in Section 4 by running detailed survival regressions that control for all available observables, in the remainder of this section, we analyze in more detail the time pattern of the difference in mortalities.

[7]The reason that the earliest observation date produces a slightly larger difference in part is explained by scale effects since obviously lives are observed over a longer time period. However, we also found that the estimates appear to have become more accurate over time.

3.3 Non-parametric Estimation of Excess Mortality

The results in the previous subsection indicate that, ceteris paribus, mortality is significantly lower for individuals that settled their policy. In what follows, we derive the "excess mortality" for policyholders that settled as a function of time to gain insight on the characteristics of the informational friction. To illustrate what we mean by "excess mortality," assume we are given two individuals S and R with forces of mortality $\{\mu_t^S\}_{t\geq 0}$ and $\{\mu_t^R\}_{t\geq 0}$, respectively, that differ only in the information regarding their settlement decision but are otherwise identical. More precisely, assume that we know S settled her policy whereas the settlement decision for R is not known. Then we can define the *multiplicative excess mortality* $\{\alpha(t)\}_{t\geq 0}$ and the *additive excess mortality* $\{\beta(t)\}_{t\geq 0}$ 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 multiplicative and additive excess mortality by relying on the most popular non-parametric survival estimators, namely the Nelson-Aalen (N-A) estimator for $\int_0^t \alpha(s)\,ds$ and 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 $\{\hat{\mu}_t^{(i)}\}_{t\geq 0}$ given by the LE provider, $1 \leq i \leq N$. Therefore, for the estimation of the multiplicative excess mortality, 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:

$$\mu_t^{(i)} = A(t) \times \hat{\mu}_t^{(i)},\ 1 \leq i \leq N, \tag{5}$$

   and use the Andersen and Vaeth (1989) excess mortality 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" force of mortality.

2. We then use the specification:

$$\mu_t^{(i)} = \alpha(t) \times \bar{\mu}_t^{(i)} \tag{6}$$

   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 (6) to the subsample of closed policies, the resulting estimate for α—or rather $\int_0^t \alpha(s)\,ds$—picks up the residual mortality information due to the settlement decision.

3. Finally, we derive an estimate for α itself from the cumulative estimate using a suitable kernel function as in Wang (2005).

For the additive excess mortality, we proceed analogously replacing Equations (5) and (6) by:

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

respectively.

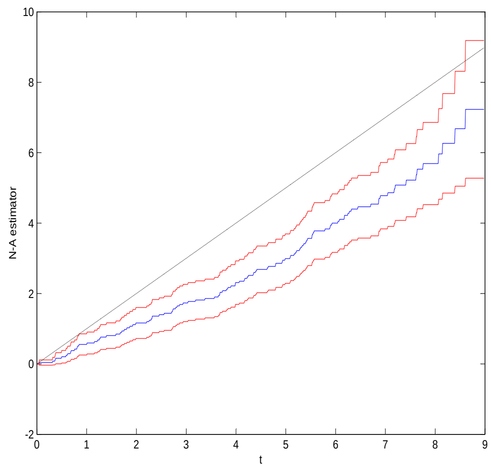
Figure 2 presents the results for the earliest observation date.[8] More specifically, panels (a) and (b) illustrate the N-A and the K-M estimate for the cumulative excess mortalities, including 95% confidence intervals. Panels (c) and (d) plot the corresponding excess mortality coefficients α and β also including 95% confidence intervals, where we use the Epanechnikov kernel with a fixed bandwidth of one in their derivation.

In accordance with the results from the previous subsection, we find that the multiplicative excess mortality is mostly significantly less than one (panel (c)), i.e. individuals that settled their policy on average live longer. This can also be inferred from the cumulative version (panel (a)) that generally has a slope of less than one. Moreover, we observe a clear time pattern: The excess mortality in panel (c) is smallest right after underwriting but increases over time and is no longer significantly different from zero after roughly five years. In other words, the impact of settling on the policyholder's force of mortality is temporary and wears off over time, until there is no longer a significant effect after five to six years.

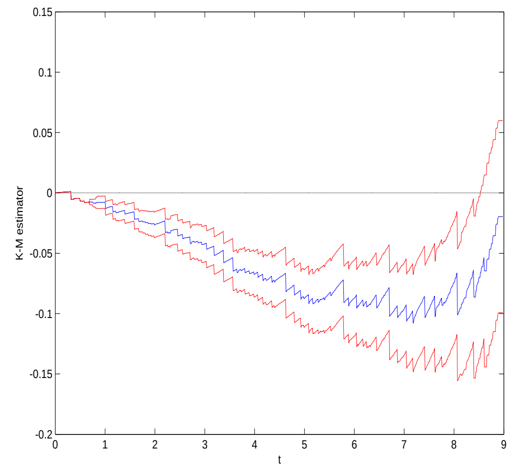
These findings are confirmed by the additive excess mortality, which also illustrates the negative association between settling and mortality since it is overall significantly less than zero (panel (d)). A possible exception is the last observation year, where we see a slight positive effect, though the overall effect on long-term survival probabilities is still negative since the cumulative multiplicative excess mortality is less than zero (panel (b)).[9] Moreover, we again generally find an increasing trend over time with the impact wearing off and losing significance after five to six years. The observation that the trend over the first year appears to be negative while the multiplicative trend is increasing can be reconciled by the observation that the force of mortality is a rapidly increasing function.

[8] Corresponding results for the latest observation date are provided in Figure 4 in Appendix A. The quantitative observations are analogous.

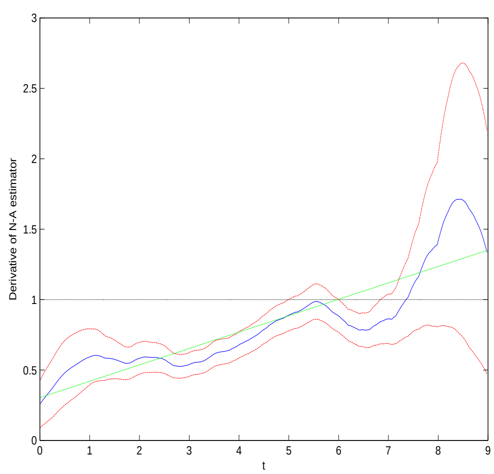
[9] Obviously, we obtain t-year survival probabilities for the closed cases by multiplying the corresponding t-year "baseline" survival probability by $\exp\{-\int_0^t \beta(s)\,ds\}$.



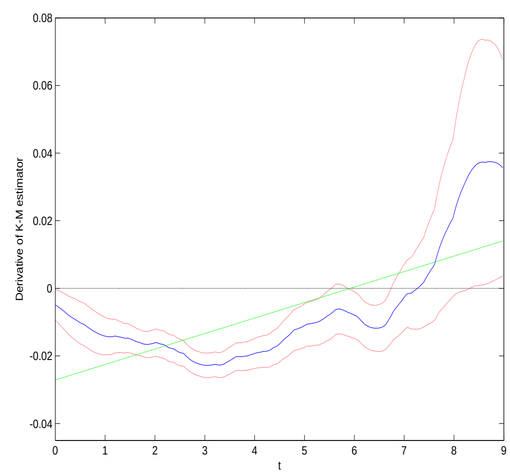
(a) N-A estimator ($\int_0^t \alpha(s)ds$)



(b) K-M estimator ($\int_0^t \beta(s)ds$)

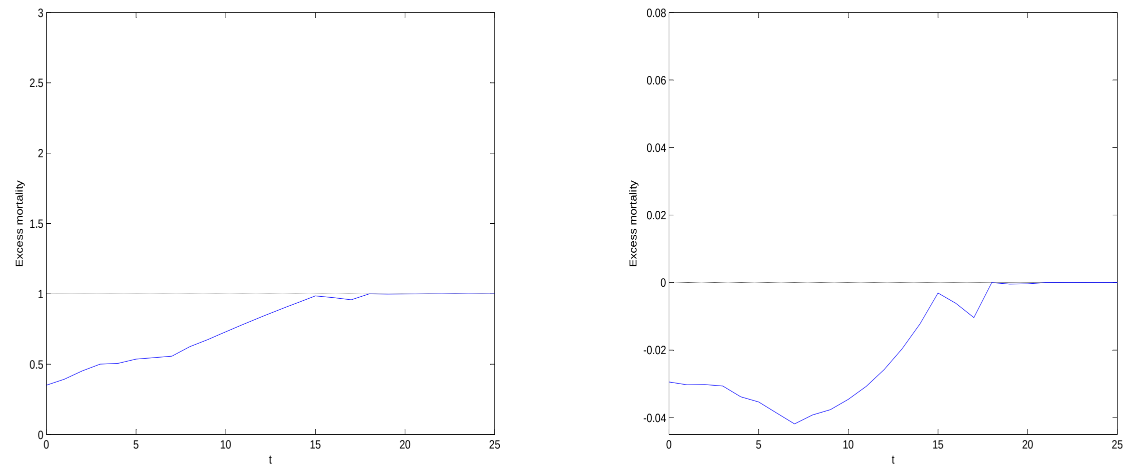


(c) Derivative of N-A estimator ($\alpha(t)$)



(d) Derivative of K-M estimator ($\beta(t)$)

Figure 2: Non-parametric estimators; earliest observation date.



(a) Multiplicative excess mortality (b) Additive excess mortality

Figure 3: Excess mortality for the CSO 2001 preferred life table.

This temporary and subsiding pattern is akin to so-called select-and-ultimate life tables used in life insurance to account for temporary selection effects due to mandatory health examinations. More precisely, the “select” part of the mortality table shows relatively lower mortality probabilities in the early contract years according to the health classification when applying for the insurance, although the difference gets smaller over time and finally the mortality probabilities approach an “ultimate” level. For comparison, panels (a) and (b) of Figure 3 illustrate the multiplicative and additive excess mortality after the life insurance underwriting process relative to ultimate mortalities for a “preferred” 75-year old male according to the *Commissioners Standard Ordinary* (CSO) 2001 life table that is used for life insurance reserving according to US regulations. When comparing Figures 2 and 3, it is apparent that the overall structure is very similar, although the selection period due to mandatory health examinations in life insurance is much longer (around twenty years). In particular, we observe a steadily increasing multiplicative excess mortality whereas there is an initial reversal for the additive excess mortality due to the rapidly increasing property of the force of mortality.

This analogy suggests an informational advantage regarding the initial health condition that individuals select on in their settlement decision. If, in contrast, the difference in mortalities for policyholders that settled their policy were driven by hidden actions such as a change in behavior or lifestyle after settling, we would expect to see a persistent or even increasing impact on the force of mortality. As such, the characteristics of the excess mortality are in line with adverse selection rather than moral hazard—which is a second central result in this paper.

As previously indicated, the next section corroborates our results by running survival regressions that control for observables.

4 Survival Regression Analysis

When testing for asymmetric information in primary insurance markets, a common approach is to regress ex-post realized risk on ex-ante coverage as in (Cohen and Siegelman, 2010):

$$\text{Risk}_i = \alpha + \beta \times X_i + \gamma \times \text{Coverage}_i + \epsilon_i, \tag{7}$$

where X_i is a vector of covariates observed by the insurer, and—ideally—by the econometrician. One can then infer the existence of asymmetric information if γ is significantly greater than zero, i.e. if there is a positive relationship between risk and purchase of coverage. As discussed in Section 2, a corresponding prediction in secondary insurance markets is a negative relationship between risk and settlement. Thus, in what follows, we test for asymmetric information by regressing ex-post mortality risk on observables and the settlement decision.

4.1 Basic Specification

Since we observe right-censored survival data, conventional techniques based on linear (OLS) regression as for Equation (7) are not applicable. Instead, we rely on survival regression. However, in order to include the mortality estimate $\{\hat{\mu}_t^{(i)}\}$ by the LE provider as a covariate since it may have (additional) predictive power, conventional multiplicative specifications as within the Cox proportional hazard model are not feasible either. Thus, we rely on the following additive semi-parametric specification:

$$\mu_t^{(i)} = \beta_0(t) + \beta_1\, \hat{\mu}_t^{(i)} + \beta_2\, \text{DOU}_i + \beta_3\, \text{AU}_i + \beta_4\, \text{SE}_i + \sum_{j=1}^{15} \beta_{5,j}\, \text{PI}_{i,j} + \sum_{j=1}^{2} \beta_{6,j}\, \text{SM}_{i,j} + \gamma_1\, \text{SaO}_i. \tag{8}$$

Here, DOU_i is the underwriting date, measured in years and normalized so that zero corresponds to January 1st 2001—the starting date of our sample. 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. $\text{PI}_{i,j}, j = 1, \ldots, 15$, are primary impairment dummies for various diseases.[10] No dummy is activated for blank entries. $\text{SM}_{i,j}, j = 1, 2$, are smoker dummies, where $\text{SM}_{i,1} = 1$ for smoker and $\text{SM}_{i.2} = 1$ for an "aggregate" entry. No dummy is activated for non-smokers. We omit information that is only available for a fraction of the individuals (states of residence, face amount). Finally, we include a settled-and-observed dummy SaO_i that is set to one for the subsample of closed cases. Clearly, we test for asymmetric information by inferring whether γ_1 is negative.

Semi-parametric additive regression models such as our specification (8) have been proposed

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

and discussed by Lin and Ying (1994), and they are special cases of the more general semi-parametric models considered by McKeague and Sasieni (1994) and the non-parametric additive regression models by Aalen (1989). For its estimation, we rely on the generalized least-squares approach from Lin and Ying (1994), who also provide a formula for the log-likelihood value.

Column A in Table 4 presents the regression results for our basic model (8) and the earliest observation date. We find that the estimated force of mortality is highly significant but the coefficient is considerably different from one, as one would expect in case of a perfect fit of the estimates by the LE provider.[11] In addition, various of the other characteristics are statistically significant, including the underwriting date, age, sex, smoking status, and some of the primary impairments. Note that this does not necessarily mean that the supplied estimates do not accurately account for these covariates because the leading coefficient β_1 is less than one.

As for the settled-and-observed covariate, the corresponding coefficient is negative and significant at the 99% level, i.e. individuals that settled have a lower mortality probability. Hence, again we find strong evidence for the existence of asymmetric information, also after controlling for observables. This reinforces our conclusions from Section 3.

Column A of Table 6 in Appendix A provides results for an alternative specification in which we only keep the significant covariates. The result for the settlement decision remains unchanged. Moreover, we present corresponding results for the latest observation date in Table 7, also in Appendix A. Again, the quantitative conclusions are identical.

4.2 Time Trend

To analyze the pattern of the excess mortality due to settling over time, we augment the basic specification (8) by three different time trends: A basic linear trend $[+\gamma_2 \, \mathrm{SaO}_i \, t]$; a quadratic trend $[+\gamma_2 \, \mathrm{SaO}_i \, t + \gamma_3 \, \mathrm{SaO}_i \, t^2]$; and a logarithmic trend $[+\gamma_4 \, \mathrm{SaO}_i \, \log\{t+1\}]$. Columns B, C, and D in Table 4 display the results for the earliest observation date.[12]

The coefficients for the non-settlement related covariates remain essentially unchanged relative to the basic specification. The basic settled-and-observed dummy again is negative and strongly significant in all cases. Moreover, we find that a basic linear trend is statistically significant at the 90% level and notably increases the log-likelihood.[13] In particular, the basic level γ_1 decreases,

[11] As indicated in Footnote 6, this criterion may be too stringent since Fasano typically only supplies estimated life expectancies. As we show in the previous section, (temporary) life expectancy are relatively accurate. Also, this assessment is based on the entire observation period and does not speak to the accuracy of current LEs. Additional analyses suggest that the quality has improved over time.

[12] Corresponding results for the latest observation date are provided in Appendix A. The quantitative conclusions are again identical.

[13] Within our approach, parameters are obtained from a generalized least squares estimation procedure, so that results not necessarily coincide with maximum likelihood estimates as formally required for a likelihood ratio test. However, the test statistic would be significant at the 95% level.

	[A]	[B]	[C]	[D]
Estimated force of mortality	0.2636^{***}	0.2636^{***}	0.2636^{***}	0.2636^{***}
	(0.0113)	(0.0113)	(0.0113)	(0.0113)
Underwriting date	$-5.8 \times 10^{-4***}$	$-5.8 \times 10^{-4***}$	$-5.9 \times 10^{-4***}$	$-5.8 \times 10^{-4***}$
	(2.4×10^{-4})	(2.4×10^{-4})	(2.4×10^{-4})	(2.4×10^{-4})
Age at underwriting	$9.2 \times 10^{-4***}$	$9.2 \times 10^{-4***}$	$9.2 \times 10^{-4***}$	$9.2 \times 10^{-4***}$
	(7.8×10^{-5})	(7.8×10^{-5})	(7.8×10^{-5})	(7.8×10^{-5})
Sex	0.0028^{***}	0.0028^{***}	0.0028^{***}	0.0028^{***}
	(7.8×10^{-4})	(7.8×10^{-4})	(7.8×10^{-4})	(7.8×10^{-4})
PI_1	-0.0017	-0.0017	-0.0017	-0.0017
	(0.0099)	(0.0099)	(0.0099)	(0.0099)
PI_2	0.0740^{***}	0.0741^{***}	0.0741^{***}	0.0741^{***}
	(0.0182)	(0.0182)	(0.0182)	(0.0182)
PI_3	-0.0055	-0.0055	-0.0055	-0.0055
	(0.0096)	(0.0096)	(0.0096)	(0.0096)
PI_4	0.0143^{*}	0.0143^{*}	0.0144^{*}	0.0143^{*}
	(0.0106)	(0.0106)	(0.0106)	(0.0106)
PI_5	0.0168^{**}	0.0168^{**}	0.0168^{**}	0.0168^{**}
	(0.0099)	(0.0099)	(0.0099)	(0.0099)
PI_6	-0.0152^{*}	-0.0152^{*}	-0.0152^{*}	-0.0152^{*}
	(0.0093)	(0.0093)	(0.0093)	(0.0093)
PI_7	-0.0079	-0.0079	-0.0079	-0.0079
	(0.0093)	(0.0093)	(0.0093)	(0.0093)
PI_8	-0.0092	-0.0092	-0.0092	-0.0092
	(0.0094)	(0.0094)	(0.0094)	(0.0094)
PI_9	0.0731^{***}	0.0731^{***}	0.0731^{***}	0.0731^{***}
	(0.0116)	(0.0116)	(0.0116)	(0.0116)
PI_{10}	-0.0148^{*}	-0.0148^{*}	-0.0148^{*}	-0.0148^{*}
	(0.0094)	(0.0094)	(0.0094)	(0.0094)
PI_{11}	0.0367^{***}	0.0367^{***}	0.0367^{***}	0.0367^{***}
	(0.0108)	(0.0108)	(0.0108)	(0.0108)
PI_{12}	-0.0016	-0.0015	-0.0015	-0.0015
	(0.0095)	(0.0095)	(0.0095)	(0.0095)
PI_{13}	-0.0146^{*}	-0.0146^{*}	-0.0146^{*}	-0.0146^{*}
	(0.0093)	(0.0093)	(0.0093)	(0.0093)
PI_{14}	0.0026	0.0026	0.0026	0.0026
	(0.0094)	(0.0094)	(0.0094)	(0.0094)
PI_{15}	-0.0217^{***}	-0.0217^{***}	-0.0217^{***}	-0.0217^{***}
	(0.0094)	(0.0094)	(0.0093)	(0.0094)
Smoker	0.0272^{***}	0.0272^{***}	0.0272^{***}	0.0272^{***}
	(0.0030)	(0.0030)	(0.0030)	(0.0030)
“Aggregate” smoking status	0.0067^{***}	0.0067^{***}	0.0067^{***}	0.0067^{***}
	(0.0028)	(0.0028)	(0.0028)	(0.0028)
Settle-and-observed	-0.0132^{***}	-0.0185^{***}	-0.0149^{***}	-0.0202^{***}
	(0.0024)	(0.0038)	(0.0051)	(0.0050)
Settle-and-observed$\times t$		0.0017^{*}	-0.0013	
		(0.0013)	(0.0039)	
Settle-and-observed$\times t^2$			4.0×10^{-4}	
			(5.6×10^{-4})	
Settle-and-observed$\times \log(t+1)$				0.0054^{*}
				(0.0043)
Log-likelihood value	-35205.56	-35203.18	-35203.08	-35204.74

Table 4: Linear survivor regression analysis – all covariates; earliest observation date.

indicating a—relative to the baseline specification—stronger effect for mortalities immediately after underwriting that wears off over time. Hence, our results are in line with the findings from Section 3.3, where we discuss that such a pattern conforms with adverse selection on the initial health state.

Similarly, a logarithmic time trend is significant but yields a lower likelihood value. In contrast, the quadratic specification does not yield significant coefficients and only very slightly increases the likelihood value relative to the linear trend specification. The results for the settlement decision remain unchanged.

4.3 Qualitative Impact

Of course we can rely on the regression coefficients γ_1 and γ_2 to derive an adjustment for survival probabilities—and, thus, life expectancies—for individuals in the closed subsample. However, this will not correspond to a suitable adjustment for closed cases relative to individuals that did not settle their policies since, as detailed in Section 3.1, our remaining sample contains LEs for both. Thus, it is necessary to inflate our estimates to account for the "commingled" nature of our remaining sample of LEs, where of course the inflation rate depends on the total proportion of settled policies.

To illustrate and to derive the appropriate inflation rate, consider the following simplified version of our additive hazard model (8):

$$\mu_t^{(i)} = \beta_0(t) + \gamma_1 \,\mathrm{SaO}_i.$$

Denote by N_t all (remaining) observations at time t, by $N_t^{(1)}$ all (remaining) settled cases at time t, $p_t = N_t^{(1)}/N_t$, and by $N_t^{(2)}$ all (remaining) observed settled cases at time t, $q_t = N_t^{(2)}/N_t$. Furthermore, denote by γ_1^{OBS} the *unknown* estimate for the model in which the econometrician observes all settlement decisions, and by γ_1^{ACT} the *actual* estimate based on observed cases only. Assume further that at any time t, the probability that a settlement decision is observed is a constant $\pi \in (0,1]$. Therefore, $\pi\, N_t^{(1)} = \mathbb{E}(N_t^{(2)})$, and we have, based on the estimates in Lin and Ying (1994):

$$\frac{\gamma_1^{OBS}}{\gamma_1^{ACT}} = \frac{\int_0^\tau N_t^{(2)}[1-q_t]\,dt}{\int_0^\tau N_t^{(2)}[1-p_t]\,dt},$$

which suggests that

$$\frac{\gamma_1^{OBS}}{\gamma_1^{ACT}} \approx \frac{(1-q)}{(1-p)}.$$

Here p is the overall proportion of *settled cases* and q is the overall proportion of *observed settled cases* in the portfolio, which for simplicity we assume are constant. Thus, we are able to derive an inflated version of the coefficient via:

$$\gamma_1^{OBS} = \gamma_1^{ACT} \times \frac{(1-q)}{(1-p)}, \tag{9}$$

where of course γ_1^{ACT} corresponds to the estimate from specification (8). In particular, since the ratio $(1-q)/(1-p)$ is always greater than one, the inflated coefficient will clearly be greater than the one estimated from the "commingled" sample.

To illustrate the effect, we use Equation (9) to adjust our estimates from the previous subsection, where we set q to $933/52,603$ according to the size of our closed subsample, and we use different assumptions on p from 20% to 50%, according to rough guesses by our data supplier. Based on the adjusted estimates, we then derive life expectancies for a 75-year old US male policyholder (cf. Table 1).[14] In addition, we also calculate life expectancies for an adjusted version of the specification with a linear trend, where we adjust γ_1 according to (9) and γ_2 such that the intersection with the time-axis remains the same, i.e. we assume the effect wears off over the same time period.

Table 5 presents the results, where for comparison the non-adjusted life expectancy is 10.48 years. The upper block provides calculations for the constant adjustment. We find that the impact of the adjustment is considerable, amounting to between 1.40 and 2.38 years, which corresponds to between 13% and 23% of the total life expectancy. In contrast, based on the adjusted linear trend, we obtain differences between 0.54 and 0.88 years, which still correspond to between 5% and 8% of the life expectancy. These magnitudes suggest that adverse selection may have a considerable impact on the life settlement market, and that asymmetric information should be accounted for in market operations and assessments.

5 Conclusion

In this paper, we provide evidence for asymmetric information in the life settlement market. More precisely, we find that individuals that decided to settle their policy display significantly longer (temporary) lifetimes—as predicted by a basic model in which individuals have private information on their mortality prospects. This finding is confirmed by survival regressions that control for observable characteristics.

In addition, we derive non-parametric estimates of the excess mortality for individuals that settled as a function of time. We find that the difference is particularly pronounced in the years

[14]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.

	p			
	20%	30%	40%	50%
Adjusted Constant				
average LE for settled subset	11.88	12.10	12.41	12.85
LE difference	1.40	1.63	1.93	2.38
Adjusted Linear Trend				
average LE for settled subset	11.02	11.09	11.20	11.35
LE difference	0.54	0.62	0.72	0.88

Table 5: Difference in average life expectancies between settled and non-settled policyholders.

immediately after underwriting but wears off over time and vanishes after five to six years. We argue such a pattern coheres with adverse selection on the initial health state, but not with moral hazard. Again, survival regressions controlling for observables confirm the results.

Our results have various implications for the life settlement market and beyond. From a broad perspective, our findings provide new positive evidence for individuals' ability to assess their own (relative) mortality prospects, in a situation where their actions have considerable monetary consequences. This is in contrast to recent studies from the behavioral literature. For the life settlement market, the existence of adverse selection has ramifications for basic operations (e.g., pricing the transactions, Zhu and Bauer, 2013) as well as for policy-relevant questions regarding efficiency and welfare implications. While addressing the latter issues is beyond the scope of this paper, building and estimating corresponding equilibrium models that account for adverse selection present intriguing problems for future research.

References

Aalen, Odd (1989). "A linear regression model for the analysis of lifetimes." *Statistics in Medicine*, 8: 907-925.

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." *The 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*, 1989: 523-535.

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

Billingsley, Patrick (1995). *Probability and Measure*, 3rd Edition, New York: John Wiley and Sons.

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.). Springer, New York.

Cohen, Alma (2005). "Asymmetric Information and Learning in the Automobile Insurance Market." *Review of Economics and Statistics*, 87: 197-207.

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.

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*, 98: 151-156.

Dionne, Georges, Mathieu Maurice, Jean Pinquet, and Charles Vanasse (2005). "The Role of Memory in Long-Term Contracting with Moral Hazard: Empirical Evidence in Automobile Insurance." Working Paper.

Einav, Lian, Amy Finkelstein, and Paul Schrimpf (2007). "The welfare cost of asymmetric information: evidence from the U.K. annuity market." NBER Working Paper No. 13228.

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

Einav, Lian, Amy Finkelstein, and Jonathan Levin (2010)b. "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 (2010). "How Does Life Settlement Affect the Primary Life Insurance Market?" NBER Working Paper No. 15761.

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 (2002). "Selection effects in the United Kingdom individual annuities market." *The Economic Journal*, 112: 2850.

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.

Gatzert, Nadine (2010). "The Secondary Market for Life Insurance in the U.K., Germany, and the U.S.: Comparison and Overview." *Risk Management and Insurance Review*, 13: 279-301.

Granieri, Vincent J., and Gregory P. Heck (2013). "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.

McKeague, Ian W., and Peter D. Sasieni (1994). "A partly parametric additive risk model." *Biometrika*, 81: 501-514.

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.

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

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

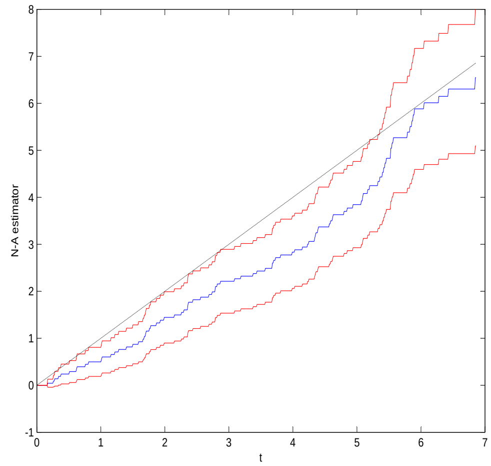
Wu, Xi, and Li Gan (2013). "Multiple Dimensions of Private Information in Life Insurance Markets." NBER Working Paper No. 19629.

Zhu, Nan, and Daniel Bauer (2011). "On the Economics of Life Settlements." Proceedings of the Risk Theory Society – 2011.

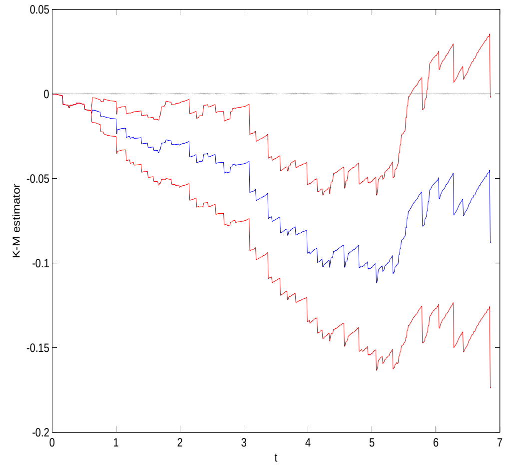
Zhu, Nan, and Daniel Bauer (2013). "Coherent Pricing of Life Settlements Under Asymmetric Information." *Journal of Risk and Insurance*, 80: 827-851.

Appendix

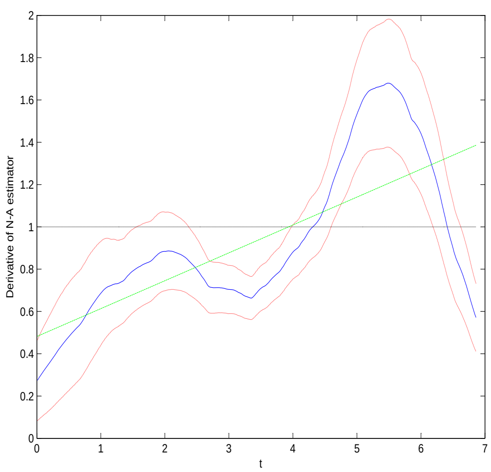
A Additional Tables and Figures



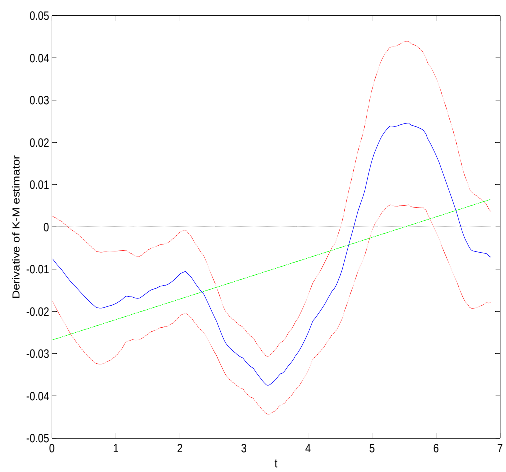
(a) N-A estimator ($\int_0^t \alpha(s)ds$)



(b) K-M estimator ($\int_0^t \beta(s)ds$)



(c) Derivative of N-A estimator ($\alpha(t)$)



(d) Derivative of K-M estimator ($\beta(t)$)

Figure 4: Non-parametric estimators; latest observation date.

	Case A	Case B	Case C	Case D
Estimated force of mortality	0.2640^{***}	0.2640^{***}	0.2640^{***}	0.2640^{***}
	(0.0113)	(0.0113)	(0.0113)	(0.0113)
Underwriting date	$-5.5 \times 10^{-4**}$	$-5.5 \times 10^{-4**}$	$-5.5 \times 10^{-4**}$	$-5.5 \times 10^{-4**}$
	(2.4×10^{-4})	(2.4×10^{-4})	(2.4×10^{-4})	(2.4×10^{-4})
Age at underwriting	$9.1 \times 10^{-4***}$	$9.1 \times 10^{-4***}$	$9.1 \times 10^{-4***}$	$9.1 \times 10^{-4***}$
	(7.9×10^{-5})	(7.9×10^{-5})	(7.9×10^{-5})	(7.9×10^{-5})
Sex	0.0027^{***}	0.0027^{***}	0.0027^{***}	0.0027^{***}
	(7.7×10^{-4})	(7.7×10^{-4})	(7.7×10^{-4})	(7.7×10^{-4})
PI_2	0.0780^{***}	0.0780^{***}	0.0780^{***}	0.0780^{***}
	(0.0157)	(0.0157)	(0.0157)	(0.0157)
PI_4	0.0184^{***}	0.0184^{***}	0.0184^{***}	0.0184^{***}
	(0.0053)	(0.0053)	(0.0053)	(0.0053)
PI_5	0.0209^{***}	0.0209^{***}	0.0209^{***}	0.0209^{***}
	(0.0037)	(0.0037)	(0.0037)	(0.0037)
PI_6	-0.0110^{***}	-0.0110^{***}	-0.0110^{***}	-0.0110^{***}
	(0.0011)	(0.0011)	(0.0011)	(0.0011)
PI_9	0.0773^{***}	0.0773^{***}	0.0773^{***}	0.0773^{***}
	(0.0071)	(0.0071)	(0.0071)	(0.0071)
PI_{10}	-0.0106^{***}	-0.0106^{***}	-0.0106^{***}	-0.0106^{***}
	(0.0014)	(0.0014)	(0.0014)	(0.0014)
PI_{11}	0.0408^{***}	0.0408^{***}	0.0408^{***}	0.0408^{***}
	(0.0056)	(0.0056)	(0.0056)	(0.0056)
PI_{13}	-0.0104^{***}	-0.0104^{***}	-0.0104^{***}	-0.0104^{***}
	(9.2×10^{-4})	(9.2×10^{-4})	(9.2×10^{-4})	(9.2×10^{-4})
PI_{15}	-0.0175^{***}	-0.0175^{***}	-0.0175^{***}	-0.0175^{***}
	(8.4×10^{-4})	(8.4×10^{-4})	(8.4×10^{-4})	(8.4×10^{-4})
Smoker	0.0279^{***}	0.0279^{***}	0.0279^{***}	0.0279^{***}
	(0.0030)	(0.0030)	(0.0030)	(0.0030)
"Aggregate" smoking status	0.0071^{***}	0.0071^{***}	0.0071^{***}	0.0071^{***}
	(0.0028)	(0.0028)	(0.0028)	(0.0028)
Settle-and-observed	-0.0131^{***}	-0.0185^{***}	-0.0149^{***}	-0.0202^{***}
	(0.0024)	(0.0038)	(0.0051)	(0.0050)
Settle-and-observed$\times t$		0.0017^{*}	-0.0013	
		(0.0013)	(0.0039)	
Settle-and-observed$\times t^2$			4.0×10^{-4}	
			(5.6×10^{-4})	
Settle-and-observed$\times \log(t+1)$				0.0055^{*}
				(0.0043)
Log-likelihood value	-35787.52	-35783.13	-35783.30	-35785.63

Table 6: Linear survivor regression analysis – reduced covariates, earliest observation date.

	Case A	Case B	Case C	Case D
Estimated force of mortality	0.3105^{***}	0.3104^{***}	0.3104^{***}	0.3104^{***}
	(0.0119)	(0.0119)	(0.0119)	(0.0119)
Underwriting data	-0.0020^{***}	-0.0020^{***}	-0.0020^{***}	-0.0020^{***}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Age at underwriting	0.0012^{***}	0.0012^{***}	0.0012^{***}	0.0012^{***}
	(8.7×10^{-5})	(8.7×10^{-5})	(8.7×10^{-5})	(8.7×10^{-5})
Sex	0.0040^{***}	0.0040^{***}	0.0040^{***}	0.0040^{***}
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
PI_1	-0.0080	-0.0080	-0.0079	-0.0079
	(0.0166)	(0.0166)	(0.0166)	(0.0166)
PI_2	0.0757^{***}	0.0757^{***}	0.0757^{***}	0.0757^{***}
	(0.0230)	(0.0230)	(0.0230)	(0.0230)
PI_3	-0.0096	-0.0096	-0.0096	-0.0096
	(0.0163)	(0.0163)	(0.0163)	(0.0163)
PI_4	0.0145	0.0145	0.0145	0.0145
	(0.0171)	(0.0171)	(0.0171)	(0.0171)
PI_5	0.0150	0.0150	0.0150	0.0150
	(0.0166)	(0.0166)	(0.0166)	(0.0166)
PI_6	-0.0213^{*}	-0.0213^{*}	-0.0213^{*}	-0.0213^{*}
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
PI_7	-0.0121	-0.0121	-0.0121	-0.0121
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
PI_8	-0.0151	-0.0152	-0.0152	-0.0152
	(0.0162)	(0.0162)	(0.0162)	(0.0162)
PI_9	0.0713^{***}	0.0713^{***}	0.0713^{***}	0.0713^{***}
	(0.0176)	(0.0176)	(0.0176)	(0.0176)
PI_{10}	-0.0195	-0.0196	-0.0196	-0.0196
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
PI_{11}	0.0422^{***}	0.0422^{***}	0.0422^{***}	0.0422^{***}
	(0.0173)	(0.0173)	(0.0173)	(0.0173)
PI_{12}	-0.0046	-0.0046	-0.0046	-0.0046
	(0.0163)	(0.0163)	(0.0163)	(0.0163)
PI_{13}	-0.0186	-0.0187	-0.0187	-0.0187
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
PI_{14}	0.0015	0.0015	0.0015	0.0015
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
PI_{15}	-0.0269^{**}	-0.0269^{**}	-0.0269^{**}	-0.0269^{**}
	(0.0161)	(0.0161)	(0.0161)	(0.0161)
Smoker	0.0283^{***}	0.0283^{***}	0.0283^{***}	0.0283^{***}
	(0.0034)	(0.0034)	(0.0034)	(0.0034)
"Aggregate" smoking status	0.0093^{***}	0.0092^{***}	0.0093^{***}	0.0093^{***}
	(0.0032)	(0.0032)	(0.0032)	(0.0032)
Settle-and-observed	-0.0084^{**}	-0.0165^{***}	-0.0193^{***}	-0.0207^{***}
	(0.0038)	(0.0056)	(0.0071)	(0.0074)
Settle-and-observed$\times t$		0.0034^{*}	0.0063	
		(0.0022)	(0.0056)	
Settle-and-observed$\times t^2$			-4.7×10^{-4}	
			(8.5×10^{-4})	
Settle-and-observed$\times \log(t+1)$				0.0113^{*}
				(0.0070)
Log-likelihood value	-33564.05	-33560.25	-33563.22	-33562.65

Table 7: Linear survivor regression analysis – all covariates; latest observation date.