Saving Face: A Solution to the Hidden Crisis for Life Insurance Policyholders

Alexander Braun*  Lauren H. Cohen†  Christopher J. Malloy‡  Jiahua Xu§

November 3, 2019

Abstract

The total face amount of terminated life insurance in the U.S. amounts to over two trillion dollars per annum. This implies a gigantic windfall profit for insurance carriers, who may keep the premiums paid until termination and avoid the death benefit payout. Little known is that many individuals could realize their policy’s value by selling it in the secondary market in exchange for a lump-sum markedly above the surrender value. We demonstrate how a contract provision requiring a mandatory offer in the secondary market could increase the welfare of policyholders without negatively affecting that of insurance companies.

*Institute of Insurance Economics, University of St. Gallen
†Harvard Business School
‡Harvard Business School
§Institute of Insurance Economics, University of St. Gallen
1 Introduction

Over 33 million life insurance policies have been terminated prematurely\(^1\) each year since 2008, corresponding to a combined face value of $2.5 trillion dollars (Figure 1). This is roughly equivalent to the real estate market loss of $2.3 trillion p.a. during the global financial crisis from 2007 to 2009 (Federal Reserve 2012). While the losses in the housing market have triggered hundreds of academic papers, media firestorms and industry responses, as well as numerous regulatory and policy changes, there has been no such reaction to the apparently commensurate policyholder losses in the life insurance market. What many consumers do not know is that the intrinsic value of a life insurance policy can be realized through a sale to investors: a so-called life settlement. Between 2011 and 2016, the average price for a life insurance policy traded in the secondary market amounted to $340,000,\(^2\) which is over three times the size of the average American’s 401(k) savings ($104,300) (Campbell 2018). The amount policyholders forego in personal wealth through life insurance termination, deliberately or not, is staggering — hundreds of billions of dollars over the past decade alone.

This crisis has been largely hidden from the general public as the scale of the problem is not widely understood. We suggest a simple and low-cost regulatory change for rectification and demonstrate its positive impact on utilitarian social welfare. The proposed reform consists of three components. Firstly, a mandatory offer-for-sale treatment of life insurance contracts: whenever a policy is to be terminated, it would be placed in the secondary market to obtain bids. Secondly, a surrender penalty for negative-value policies, which protects the profitability of the carriers. Thirdly, a ban of cost-of-insurance (COI) increases for in-force policies so that insurers cannot discourage investors through retrospective premium hikes, particularly on the policies of health-impaired insureds. The new equilibrium would yield a fair sharing of economic value and prevent the current misdirection of billions of dollars of policyholder value every year.

The remainder of the paper is structured as follows: Section 2 reviews current practices and issues in the life insurance market; Section 3 proposes a reform to address the issues; Section 4 demonstrates the welfare-enhancing effect of the proposed reform; Section 5 discusses the operationalization of the reform; Section 6 concludes.

2 Current State of the Life Insurance Market

2.1 The secondary market for life insurance

The cash flow profile of a life insurance contract comprises two payment streams: the regular premiums by the policyholder and the death benefit, which is disbursed by the carrier when the insured has passed away. Hence, the policy price at any given point in time crucially depends on an insured’s life expectancy (Braun and Xu, 2019). The lower the latter the more expensive the policy, since premiums will have to be paid for a shorter period and the payout from the insurer will be received earlier. To estimate the life expectancy of the insured, investors usually have his medical history analyzed before making a purchase (Braun et al., 2012). Unlike viatical settlements, which involve terminally ill patients, life settlement investors focus on contracts of elderly individuals with a health impairment: from 2011 to 2016, the insured’s average age at settlement date was 76.3. Due to high transaction costs, investors also show a preference for large-face policies. Over the last five years, the average notional amount traded was as high as $1.9 million with a median of $1 million. Consequently, settled policies mainly tend to originate from California, New York, Florida and Texas, where older, wealthier people are more likely to retire and reside (Blake et al. 2013) (Figure 2).

\(^1\)This includes policies lapsed and surrendered. Source: The U.S. Life industry Briefing book of the SNL financial (S&P Global Market Intelligence).

\(^2\)Data from the independent life settlement consulting firm AA-Partners.
Currently, the total face amount of settled policies that are still in force is $100 billion (Roland 2016).

Since life expectancies for lapsed and surrendered policies are unavailable, it is difficult to empirically assess how many of those would have qualified for a secondary sale. Instead, the size of the untapped life settlement market is commonly approximated by the total face amount of ordinary life insurance terminated by seniors above 65, which equaled $148 billion in 2017 (see Figure 3). If this potential had been fully realized, the life settlement market would have grown 50-fold by now.

How much of this value could the original policyholders have expected to obtain through a secondary sale? Considering that the average empirical transaction price per unit of face amount is 18% of the policy’s face value, we arrive at an estimate of $26.6 billion. That is to say, had those policyholders sold their policies to life settlement investors instead of terminating them, they would have received an aggregate $26.6 billion in cash. To put it into perspective, this amount exceeds the drug invoice spending for long-term care and home health care in the U.S. in 2017, which amounted to $20.6 billion (IQVIA 2018).

Two major reasons for the apparent underdevelopment of the life settlement market are the low awareness among policyholders and a lack of trust on the side of investors. To some extent the industry is still troubled by its past. Legacy issues such as prominent fraud cases or stranger-originated life insurance (STOLI) policies mass produced by life settlement intermediaries in the mid-2000s still resonate today (Braun et al. 2018). However, since the market matured and became increasingly regulated, misconducts are now less common. Consequently, the strongest headwind currently originates from a questionable pricing practice in the primary U.S. life insurance market, the shortcomings of which will be discussed below.

### 2.2 Lapse-supported pricing

Lapse-supported pricing means that insurance carriers set competitively low premiums to gain market share, expecting that they will not have to disburse some of the death benefits (Richmond 2012). The idea is that the firms may at least break even on the transaction, should their assumptions regarding termination and mortality be accurate. Therefore, realized policy termination rates are a key driver of insurer profitability. The specific effect, however, depends on the insureds’ characteristics. Termination on the part of young and healthy insureds with long life expectancies harms carriers, as the former would ordinarily pay premiums for an extended period and receive death benefits in the distant future. In other words, their expected discounted premium payments exceed the expected discounted death payouts. From the perspective of the policyholder, these contracts therefore exhibit a negative economic value at termination, which turns them into an asset for the insurer. Termination by more senior and unhealthy insureds with short life expectancies, in contrast, leads to windfall profits, as the firms are relieved of the obligation to pay out the death benefit in the near future — significant amounts compared to the foregone revenues (Leimberg et al. 2006). Because their expected discounted premium payments are lower than the expected discounted death benefit, these policies exhibit a positive economic value and thus represent a liability for the insurer.

Carriers who extensively employ lapse-supported pricing are vulnerable to the unanticipated persistency of policies from individuals with impaired health and below-average life expectancies (Bakos and Parankirinathan 2006). Exactly those contracts are attractive for life settlement investors. Consequently, The secondary market creates an adverse selection problem, implying that “bad risks”

---

3The values for other years are similar. It should be noted that not all policies from insureds above 65 are economically viable as a life settlement, and this results in overestimation. On the other hand, policies from younger insureds and group policies are occasionally also settled, which leads to some underestimation. We suppose those two effects offset each other to some extent.

4The average price has been derived from AA-Partners’ database.
remain on the insurers’ books. Gottlieb and Smetters (2016) estimate a negative profit margin of -12.8% for life insurance carriers in the absence of policy lapsation. The dollar change in insurers’ net incomes can be approximated as follows. In 2017, U.S. carriers received $138 billion in life insurance premiums and paid out $107 billion in benefits. With $24 billion of operational expenses deducted, the net income amounted to $7 billion. This net income has been at a stable level for the past 5 years and is expected to remain so, absent any drastic market events. However, if the annual $26.6 billion in policy value estimated above were to be paid out to insureds who terminated their contracts, carriers would not only have their profit reduced, but they would actually make a loss (Figure 4).

This part of the insurers net income represents a covert value transfer from policyholders to shareholders. Owing to the zero-sum nature of the insurance business, policy terminators have surrendered valuable assets to “assist” insurance companies in achieving their current level of profitability. From the policyholders’ perspective lapse-supported pricing therefore leads to unjust wealth redistribution. Peer subsidization is a fundamental concept of insurance: through pooling, the fortunate support the unfortunate. With lapse-supported pricing, however, ill and senior policyholders — the unfortunate, who terminate their valuable policies that would otherwise benefit their estate — become the subsidizers of the competitively low premiums of other policyholders (Atmeh 2011).

2.3 Retrospective premium adjustment

The last couple of years have witnessed multiple COI-hike incidents initiated by leading insurance companies such as AXA, Transamerica Lincoln National and Voya (Hanson 2017). Premium increases ranged between 5% and an incredible 600% and mostly affected large-face universal and whole life policies of senior insureds, precisely those which are sought after in the secondary market. Evidently, COI raises on these contracts are a means to deter investors, since they may cause tremendous losses in their portfolios. Apart from the persistent low interest rates, which decrease insurers’ asset management results and drain the policyholders’ cash accounts, lapse-supported pricing is viewed as a major reason for premium instability. For example, John Hancock recently announced to increase premiums “as a result of changes in [their] expectations of […] lapse experience” (Jeffrey Leonard v. John Hancock 2018). In other words, carriers that did not factor the adverse selection effect caused by the secondary market into their pricing are exposed to unexpected losses and may try to remedy this ex-post, despite the risk of lawsuits and reputational damage.

3 Recapturing value through regulatory change

Partial Equilibrium Theory calls for a perfect secondary market to ensure that a policy’s price is permanently pegged to its economic value. To converge towards this goal, we suggest to change the status quo by means of a minimally-invasive regulatory intervention, comprising three pillars:

A. It is mandatory to offer policies in the secondary market, if they are threatened by termination.
B. For policies with a negative economic value, carriers can charge a matching surrender penalty.
C. Premiums of in-force policies may not be raised due to higher-than expected persistency rates.

Data collected from SNL.
Note that termination-based planning can also be found outside the insurance industry. Airline companies, for example, oversell a flight assuming a certain number of no-shows.
Rule A: Mandatory offer for secondary sale

A contractual provision that automatically places a contract in the secondary market when it is threatened to be terminated would help policyholders monetize the associated value. Through this provision, life insurance products will become more attractive, since customers can trade out of their policies and avoid losses due to early termination more often. Mandating policies to be offered for secondary sale circumvents the monopsony problem faced by policyholders who are unaware of the life settlement market.

Rule B: Surrender penalty for policies with a negative economic value

If introduced on its own, the prior provision could hurt insurance companies through a downward jump in lapsation rates. More specifically, there would be no more windfall profits to counter losses associated with the termination by young and healthy insured, whose policies exhibit a negative economic value and are thus an asset for the insurer. Hence, from an overall welfare perspective, a sustainable solution calls for additional measures to avoid a negative impact on the carriers’ profitability. We therefore suggest to abandon lapse assumptions in premium pricing and instead link the termination of policies with a negative economic value to an equally-sized surrender penalty.

Rule C: Prohibition of persistency-induced premium increases

Individuals who purchase insurance policies are often unaware of, or underestimate, the risk of a premium increase. Our first suggested provision will only work if it is prohibited to raise the COI of in-force contracts to counter higher-than-expected persistence rates in a portfolio. Otherwise, carriers are able to set competitively low lapse-based premiums to attract customers and then provoke policy termination through retrospective rate hikes (“shock lapse”). There is an evident moral hazard issue involved, since insurers benefit from (i) higher revenues and (ii) a termination of death benefit entitlements of those policyholders who cannot or do not want to pay the higher premiums. A provision to prohibit persistency-induced COI increases for in-force policies is especially critical in the context of life settlements. Without it, carriers can target policy groups most commonly seen in the life settlement market, making those contracts costlier, or even uneconomical, to keep in force.

4 Impact of the Regulatory Change

Next, we analyze the impact of the proposed three-pillar regime, mandatory offer for secondary sale (“MSS”), in a theoretical framework based on the following set of assumptions:7

i) The primary insurance market is perfectly competitive,

ii) The secondary insurance market is perfectly competitive,

iii) Insurance carriers and investors can diversify and are therefore risk neutral,

iv) Information on the insureds’ mortality profiles is symmetrical,

v) Markets are frictionless, i.e. we have no transaction costs, sales provisions, taxes etc.

---

7Some of these assumptions do not hold in reality, but allow us to concentrate our analysis on the monopsony problem instead of other issues. Frictions, e.g., play a major role, as numerous intermediaries are involved in a life settlement transaction (Braun et al., 2018).
Let $PP$ denote policy price, which has two meanings: a) if the insured terminates his contract, then $PP$ represents the surrender value $SV$ paid by the carrier; b) if the insured sells the contract in the secondary market, then $PP$ represents the highest bid from the investors. The difference between $PP$ and the policy’s economic value $EV$ is the price-markup $\Delta$:

$$\Delta := PP - EV. \quad (1)$$

In case i), $\Delta$ can be positive, zero, or negative (“price-markdown”). In case ii), $\Delta$ depends on the intensity of competition in the secondary market as reflected by the investors’ target returns for policies with given characteristics (Braun and Xu, 2019). At time $t = 0$, when an individual enters into a contract with a carrier, assumption i) ensures fair pricing ($PP_0 = EV_0$). If, in addition, there is no upfront payment, i.e., the present value of expected future premiums equals the present value of expected death benefits, we have $PP_0 = EV_0 = 0$. Finally, owing to assumption ii), the introduction of MSS ensures $PP \equiv EV$ or $\Delta^{MSS} \equiv 0$ so that insureds can always monetize the economic value of a policy.

### 4.1 Effect on pricing

#### Status quo

Consider homogeneous cohorts of insureds with the same age and mortality profile as well as identical policies. Each period, they decide between selling/surrendering for the payoff $PP$ or keeping the policy in force and therefore retaining the economic value $EV$. The latter is composed of:

1. the deterministic premium payment $\{\pi_t\}_{t \in \mathbb{N}_0}$ at time $t$ (for coverage between $t$ and $t + 1$).
2. the time-discounted expected value of future payoffs and economic values.

Assume that cash flows only occur at the start of a period, i.e. at $t = 0, 1, 2, ...$. The following cash flow tree illustrates the average payoff of policies from a cohort:

---

**Diagram:**

- $EV_t = -\pi_t + \rho \left( l_t \left( PP_{t+1} + (1 - p_t)PP_{t+1} \right) + (1 - l_t) \right)$

where

- $\rho$: time discount factor
- $\{l_t\}_{t \in \mathbb{N}_0}$: survival rate between time $t$ and $t + 1$.
- $\{p_t\}_{t \in \mathbb{N}_0}$: persistency rate between time $t$ and $t + 1$.

We can thus express $EV_t$ in a recursive manner:
Inserting $PP_{t+1} = EV_{t+1} + \Delta_{t+1}$ (see Equation 1) yields:

$$EV_t = -\pi_t + \rho \left\{ l_t \left[ p_t EV_{t+1} + (1 - p_t)(EV_{t+1} + \Delta_{t+1}) \right] + (1 - l_t) \right\}$$

$$= \rho \left\{ l_t \left[ EV_{t+1} + (1 - p_t)\Delta_{t+1} \right] + (1 - l_t) \right\} - \pi_t. \quad (2)$$

We can also transform Equation 2 into a non-recursive format (details see Appendix):

$$EV_t = \sum_{j=t}^{\infty} \frac{\rho \left[ l_j (1 - p_j)\Delta_{j+1} + (1 - l_j) \right] - \pi_j}{l_j} \rho^{j-t} \prod_{i=t}^{j} l_i. \quad (3)$$

**Under MSS**

Plugging $\Delta_t^{MSS} = 0, \forall t$ into Equation 3, we obtain:

$$EV_t^{MSS} = \sum_{j=t}^{\infty} \frac{\rho (1 - l_j) - \pi_t^{MSS}}{l_j} \rho^{j-t} \prod_{i=t}^{j} l_i. \quad (4)$$

Carriers hold the exact inverse positions of the insureds. In other words, their liability is the negative value of the policy’s economic value ($-EV_t$). To ensure that MSS does not affect profitability, premiums of newly-issued policies may be set such that the economic value at inception remains the same with or without MSS:

$$EV_0^{MSS} = EV_0$$

$$\Rightarrow \sum_{j=0}^{\infty} \frac{\rho (1 - l_j) - \pi_t^{MSS}}{l_j} \rho^{j-t} \prod_{i=t}^{j} l_i =$$

$$\sum_{j=0}^{\infty} \frac{\rho \left[ l_j (1 - p_j)\Delta_{j+1} + (1 - l_j) \right] - \pi_j}{l_j} \rho^{j-t} \prod_{i=t}^{j} l_i$$

One possible solution would be:

$$\forall t, \pi_t^{MSS} = \pi_t - \rho l_t (1 - p_t)\Delta_{t+1}$$

$$\begin{cases} < \pi_t, & \text{if } \Delta_{t+1} > 0 \\ = \pi_t, & \text{if } \Delta_{t+1} = 0 \\ > \pi_t, & \text{if } \Delta_{t+1} < 0 \end{cases}$$

Absent MSS, carriers are the monopsonistic buyer of policies, whose price is the surrender value $SV_t$, i.e. $\Delta_{t+1} = SV_{t+1} - EV_{t+1}$. If $SV_{t+1} > EV_{t+1}$, then MSS-based premiums $\pi_t^{MSS}$ would be set lower than current lapse-based premiums $\pi_t$; if $SV_{t+1} < EV_{t+1}$, then $\pi_t^{MSS}$ would be set higher than $\pi_t$.

Despite these adjustments in dollar premiums, MSS does not alter the economic values of the contracts, since increased premiums coincide with better prospective payoffs $PP$ for the customer (Doherty, O’Dea, and Singer, 2004). Our proposal should therefore not affect insurance demand. Moreover, since Equation 4 does not include persistency rates, carriers are freed from lapse-based policy pricing. This, in turn, makes retrospective premium adjustments due to erroneous termination assumptions unnecessary.
4.2 Effect on welfare

Policyholder utility

Let \( u \) denote the utility (in monetary units) that an insured derives from life insurance coverage. Two conditions must hold at the outset: i) \( u \geq PP \) so that the individual buys the policy; ii) \( SV \leq PP \) so that insureds cannot lock in an arbitrage profit by acquiring a policy and surrendering it immediately. With the passage of time, both \( u \) and \( EV \) may change. For some insureds, keeping the policy in force remains worthwhile. For others, the opportunity cost becomes too high and \( u \) falls below \( PP \). Consequently, those insureds will choose to terminate the contract. In contrast to policyholders, carriers and investors are risk neutral (see assumption ii), basing their decisions solely on \( EV \). Tables 3 and 4 summarize the surpluses or deficits of the various parties, associated with the issuance, persistence, surrender or sale of a policy.

A policy’s utility \( u \) is positively related to its economic value \( EV \), but they are not necessarily equal. Define \( u^e \) as the non-monetary excess utility that an insured derives from a policy over and above its economic value:

\[
u^e := u - EV
\]  

\( u^e \) can be positive, zero, or negative (“deficient utility”) and depends on the insured’s personal situation. Factors that influence \( u^e \) include, but are not limited to, the insured’s family situation, liquidity situation and bequest motive. Ceteris paribus, \( u^e \) will be higher if the insured has adequate cash, and lower if the insured experiences a liquidity crisis. \( u^e \) will also be lower if the insured’s beneficiary becomes financially independent, and conversely higher if the beneficiary is financially dependent and the insured’s bequest motive is strong.

Based on Equation 1 and Equation 5 we obtain

\[
u - PP = u^e - \Delta.
\]  

This leads to the following lemma:

**Lemma 1.** \( \Delta \) and \( u^e \) defined as above.

At the time of issuance of a policy
1. if \( u^e > \Delta \), an insured acquires the policy;
2. if \( u^e < \Delta \), an insured does not acquire the policy;
3. if \( u^e = \Delta \), an insured is indifferent whether or not to acquire the policy.

If the policy is owned by an insured, then
1. if \( u^e > \Delta \), an insured keeps the policy;
2. if \( u^e < \Delta \), an insured abandons the policy;
3. if \( u^e = \Delta \), an insured is indifferent with regard to keeping or abandoning the policy.

The proof is self-explanatory.

Supply and demand of policies

Since a cohort is per our definition homogeneous in age and mortality profile, all of its policies exhibit the same economic value \( EV \), get the same markup \( \Delta \) and can thus be sold/surrendered for the same price \( PP \). The utility of each policy in a cohort, on the contrary, is idiosyncratic and varies based on the situation of the insured. Let \( m \) be the total number of policies in a cohort, each of which is associated with a utility \( \{u_i\}_{i=1,2,...,m} \) to its holder. The function \( F_u(x) \) denotes the number of policies surrendered or sold at a given price \( x \):

\[
F_u(x) = \sum_{i=1}^{m} 1_{u_i \leq x},
\]  

(7)
where 1 is the indicator function. Hence, the the cohort’s policy supply $Q$ at price level $PP$ can be expressed as follows:

$$Q = F_u(PP) = \sum_{i=1}^{m} 1_{u_i \leq PP},$$

(8)

Consistent with Equation 6 and Equation 8, the aggregate policyholder surplus for the cohort, denoted by $PhS$, can be expressed as:

$$PhS = \sum_{i=1}^{m} (PP - u_i) 1_{u_i \leq PP}$$

(9)

Furthermore, the marginal cost of acquiring an additional policy unit and hence the policy demand curve equals the prevailing price $PP \in \{SV; EV\}$. Unlike insureds, who attribute idiosyncratic utilities to their policies, investors are exclusively concerned about $EV$. Hence, the surplus derived from each acquired policy equals $EV - PP$ ($= -\Delta$). Given that the number of policies supplied by insureds is $Q$, the total buyer surplus, $BrS$, amounts to:

$$BrS = (EV - PP)Q.$$  

(10)

**Utilitarian social welfare without MSS**

In the absence of MSS, the surrender value $SV$ can deviate from the economic value $EV$ ($\Delta \neq 0$). Let $PP'$ denote the market clearing price in this case. Without access to the secondary market, insureds who wish to terminate their contracts have to do so by surrendering it to the carrier. The latter possesses monopsony power and can set any level of $SV$ for “buying back” a policy (Doherty and Singer 2003; Kohli 2006). In the following, we analyze the associated welfare effects.

**Figure 5** illustrates a cohort of insureds with health impairments and life expectancies that are lower than originally expected. Those individuals will pass away soon, leaving only few premium inflows before the death benefit becomes due. Their policies exhibit a positive economic value ($EV > 0$) and are therefore a liability for the insurer. The latter offers an inadequate surrender value ($SV < 0$) compared to price $EV$ at all ($EV - SV < 0$). The light-green triangular area bounded by lines $Q = 0$, $u = SV$ and curve $PP(Q) = F_u^{-1}$ represents the policyholder surplus $PhS$. Moreover, the buyer surplus $BrS = (EV - SV)Q$ ($> 0$) is shown as a grey rectangle bounded by lines $Q = 0$, $u = SV$, $u = EV$ and $Q = F_u(SV)$. Finally, the red triangle reflects the dead weight loss, caused by the fact that the buyers’ gain surplus at price $SV$ compared to price $EV$ is less than the sellers’ loss in surplus.

**Figure 6** shows a cohort of insureds who have developed little or no health impairment since policy issuance and therefore exhibit longer life expectancies than originally expected. Those individuals are likely to pay premiums for a long time before a death claim occurs. Their policies exhibit a negative economic value ($EV < 0$) and are thus an asset for the insurer. If the latter charges no surrender penalty ($SV > 0$) or an inadequate surrender penalty ($0 > SV > EV$), he suffers a deficit due to the loss of good risks from his book. The light-green triangular area bounded by lines $Q = 0$, $u = SV$ and curve $PP(Q) = F_u^{-1}$ represents the policyholder surplus $PhS$. The buyer deficit $BrS = (EV - SV)Q$ ($< 0$) is depicted as a grey rectangle bounded by lines $Q = 0$, $u = SV$, $u = EV$ and $Q = F_u(SV)$. Finally, the red triangle highlights the dead weight loss, which occurs since $-BrS > PhS$.

Finally, **Figure 7** highlights an extreme scenario that can be observed in cohorts of terminally ill insureds. Due to a very short life expectancy, those insureds’ policies have an extraordinarily high $EV$. The surrender value $SV$ offered by the carrier does not reflect the high $EV$ at all ($SV << EV$) such that no insured in the cohort is willing to terminate the policy. Instead, they prefer keeping their policies in force, knowing that the beneficiaries will receive the death benefit in short order. The red triangle shows the dead weight loss, which occurs because the market does not clear.
Utilitarian social welfare with MSS

The introduction of MSS ensures $SV = EV$ ($\Delta = 0$). Let $PP^*$ denote the associated market clearing price. Insureds now have the choice of surrendering or selling in the secondary market such that the carrier faces competition from investors. More specifically, if the surrender value is inadequate ($SV < EV$), then insureds will prefer the life settlement (Doherty and Singer 2003). After the purchase by an investor, a policy is kept in force until maturity. Investors would only terminate if $SV > EV$, which carriers will not offer. Therefore, the latter may as well grant the insured a surrender value of $EV$ right away. The loss of monopsony power eradicates the buyers’ surplus ($BrS = 0$). To see this, plug $PP^* = EV$ into Equation 10. As it turns out, however, MSS is still beneficial for the overall welfare situation across cohorts, since dead weight losses can be avoided altogether.

In all three cases discussed above, the new policyholder surplus $PhS$ equals the triangle bounded by $Q = 0$, $u = EV$ and curve $PP(Q) = F_u^{-1}$. For the unhealthy cohort in Figure 5, the convergence of $SV$ towards $EV$, increases the number of sold/surrendered policies and therefore grows the insured surplus. In addition, for the healthy cohort in Figure 6, the carrier is now able to penalize policy terminations with an amount up to the absolute value of $EV$, thus reducing the insured surplus. Finally, with a terminally ill cohort as shown in Figure 7, the policyholder surplus becomes positive.

To sum up, it should be emphasized that the real-world secondary market is neither a monopsony nor characterized by perfect competition. Each year, a number of insureds becomes aware of the possibility to sell instead of terminate and their policies are ultimately purchased by investors. Nevertheless, high information costs and other frictions remain. Hence, a regulatory change such as MSS is necessary to increase competition and eliminate economic inefficiencies. Without MSS, the carrier gains from the unhealthy terminators and loses from the healthy terminators. Even if it ultimately breaks even, those fortunate enough to maintain good health are additionally advantaged as they are subsidized by the unfortunate. MSS reverses this dynamic, redistributes wealth and increases overall welfare through the elimination of deadweight losses.

5 Further Considerations

5.1 Surrender penalty reserve

As discussed above, one cornerstone of MSS is the surrender penalty if the economic value of a policy is negative. For the desired effects to be realized, insureds must not be able to simply walk away from negative-value contracts without paying. There are several ways to tackle this issue. Carriers with a rigorous credit checking system can charge the insured through direct debit. Surrender penalty reserves are another possible solution to cover the loss incurred by an early termination. Those are saved in the insureds’ cash accounts in order to accumulate interest. The money belongs to insureds but is not at their disposal. It is held by the carrier and can only be released or used to cover the insured’s premium payment, if there is evidence for an appreciation of the policy’s economic value (e.g., through a deteriorated health situation). The carrier can contractually enforce deposits to gradually build up the reserve, take a sufficient amount up-front, or rely on single premium insurance (Gatzert, Hoermann, and Schmeiser 2009). This is the same as front loading of premiums, a practice that is widely employed by carriers to protect themselves against reclassification risk.

---

8 The insurer surplus stays at zero. Note that this explains the existence of free accelerated death benefit (ADB) riders for terminally ill. Carriers offering such an option lose nothing, but enhance the insureds’ welfare.

9 The is already a common practice. See Koutnik (2013) and Gatzert (2009).

10 Hence, the surrender penalty reserve is fundamentally no different from margin calls in derivative transaction.

11 A liquid and open secondary market eases policy trading, exposing carriers to higher reclassification risk.
Should the policy have a negative economic value upon termination, the reserve can be used to cover the actual surrender penalty and any excess amount will be refunded to the insured. In case of a positive economic value, all funds can be returned. The reserve equals the “cash value” of a policy, which is influenced by the money that the policyholder deposits in his account, either voluntarily (to benefit from tax deferral), or contractually demanded by the carrier (for potential surrender penalty deduction).¹²

As corollary of the up-front deposit, insurance demand may be affected. For consumers who lack liquidity, it adds to the cash drain and may thus deter them from purchasing a policy (see e.g. Cowley and Cummins 2005; Daily, Hendel, and Lizzeri 2008). However, it also has upsides. Firstly, some consumers are willing to use a cash account for deferral of income tax (Gelfond 2009; Doherty, O’Dea, and Singer 2004). Secondly, the idea can block “freeloaders” with no strong insurance desire, who sign insurance contracts when there is little or no premium charge at the beginning, but terminate as soon as the low premium period ends. Those consumers are likely to be attracted by a premium financing scheme with some years of free coverage after issuance. Due to their low persistency, they would not be profitable for carriers.

5.2 Carrier buy-back

Carriers can compete against life settlement investors by offering a fair surrender value. Their direct participation in secondary sales would increase market liquidity and efficiency. It is likely that own buy backs would be preferable to them, since in-force policies owned by investors cause ongoing operational costs. In addition, carriers should have a cost advantage over third-party purchasers because they can simply retire a policy without having to maintain it and thereby incurring servicing fees. Due this cost advantage, active secondary market participation of carriers might actually discourage investors from bidding. Yet, MSS still ensures that carriers do not regain monopsony power, as investors will return if the surrender value deviates too far from the economic value.

Evans, Russell, and Sager (2013) suggest a contractual provision that gives carriers the right of first refusal to match any viable life settlement offer, implying an additional second-mover advantage. Such a provision could jeopardize market competition. The life settlement business would become unattractive, if buyers went through all the effort of pricing policies, only to have carriers disrupt the deals in the last minute. An attempt to undermine a life settlement transaction was, e.g., undertaken by insurer John Hancock. Consequently, Coventry First LLC, the life settlement buyer, filed a complaint against the insurer (Hersch 2011). The case was ultimately settled with John Hancock paying $2 million to Coventry and sent a pro-competitive signal (Horowitz 2013).

5.3 Accelerated death benefit

Today, most life insurance companies offer some form of accelerated death benefit (ADB) option: as a stand-alone insurance product, as a rider attached to a life insurance, or directly embedded in the policy. There are several ways for a carrier to pay out the ADB. For so-called terminal illness ADBs, the NAIC model act requires carriers to provide a lump sum option. The lump sum payment can be in the form of (i) a pre-specified portion of the death benefit, (ii) the present value of the total death benefit, taking into account the insureds’ life expectancy, or (iii) a lien against the death benefit (Spurrier 1997; Schmidt 1997). Version (ii) works most similar to a life settlement from the perspective of an insured, whereas (iii) resembles a life settlement from the perspective of a carrier,

¹²Note that this conceptual “surrender penalty reserve” is not to be confused with the existing “terminal reserve” owned by the carrier. Both are designed as a precautionary measure for a contingent liability. However, the former is reserved for a policyholder’s surrender payment to the carrier whereas the latter for a carrier’s death benefit payment to the beneficiary.
who continues to receive premiums until the insured’s death and is still liable to pay death benefit (with policy lien deducted) at the death event.

Terminal illness ADBs apply to insureds with a life expectancy of no longer than 12 months (Perez 2002; Gatzert, Hoermann, and Schmeiser 2009). They are often free and included as a default provision in a life insurance contract. With this option, a carrier maximally loses a one-year premium and the one-year time value of the death benefit. Given the low prevalence of terminal illness, this financial burden is insignificant and occurs infrequently. In fact, carriers can even benefit from the inclusion of this option, since it gives insureds flexibility, thus creating demand and a positive reputation effect. Therefore, most carriers are willing to offer the ADB for terminal illness option free of charge. ADB also exists for chronic or critical illnesses, sometimes referred to as long-term care benefit riders. Unlike terminal illness ADBs, these options are predominantly subject to additional premium charges.

The idea of the ADB originated in the 1980s during the onset of the HIV/AIDS epidemic, which also drove the advent of the viatical settlement market (Spurrier 1997). As more and more carriers adopted a free ADB for the terminally ill, viatical settlements became rarer. This demonstrates that carriers have an edge when competing against secondary market investors. Insureds seemed to prefer the ADB over a secondary sale, possibly because of a higher trust in their carrier and the desire to avoid the introduction of a third party. Hence, just as the success of the viaticals market made ADBs for terminally ill a standard byproduct of life insurance with little or no surcharge, a more competitive life settlement market induced by MSS could help expanding the reach of ADBs to a broader range of insureds such as health impaired seniors.

5.4 Regulation

Officially legalized in 1911, the life settlement market has become healthier and more regulated with successive legislation. Meanwhile, many states have begun to adopt laws that support the market. For example, Georgia forbids carriers from penalizing insurance agents for assisting policyholders to lawfully secure policy benefits. Florida stipulates that carriers suggest policyholders to consult professionals before making a change to their policy. California stipulates that carriers should not “make any false or misleading statement for the purpose of dissuading an owner or insured from a lawful life settlement contract” or “unreasonably delay effecting change of ownership or beneficiary with any life settlement contract lawfully entered into in this state or with a resident of this state.” All three laws helps to increase awareness of the secondary market and signal that legislators may be prepared for MSS as a natural next step in consumer protection.

5.5 Frictional costs

Insureds are the ultimate bearer of the medical underwriting and intermediary expenses for life settlements. Whether explicitly stated in a contract or not, the proceeds an insured receives from a policy sale are net of these costs. MSS envisions that upon consent of insureds who wish to terminate, carriers immediately place their policies into the secondary market. This reduces search costs and promotes disintermediation. Specifically, policy originators could most likely be eliminated from the transaction chain. MSS also aligns with the latest technology trends in the insurance sector (Braun

---

13Examples comprise the Accelerated Death Benefit Rider of William Penn and the Terminal Illness Death Benefit Advance Rider of Principal Life.
14Examples are the LifeEnhance Accelerated Benefits Rider of Lincoln and the Long-Term Care Services Rider of AXA.
15Grigsby v. Russell, No. 53. 222 U.S. 149 (1911), see https://www.leagle.com/decision/1911371222us1491348
16101120.5.(6) http://www.leg.state.co.us/preclics/1997/hbills97/HB1345.htm
18SEC. 6. 10113.3. (d) (n) https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=200920100SB98
and Schreiber 2017). New software and models allow the surrender value of a life insurance policy to be more easily adjusted to the insured’s personal health condition and enable an efficient data management, hence lowering underwriting and information costs in general. (Tempesta 2018).

6 Conclusion and Outlook

We reveal the enormous wealth transfers associated with surrender and lapsation of life insurance contracts and attempt to raise awareness of the life settlement market and its economic benefits. To this end, we propose a simple but robust regulatory reform: a mandated offer for secondary sale treatment of all life policies that are to be terminated. The regime could drastically increase the amount of wealth retained by U.S. seniors and, in turn, help fund to retirement, long-term care, and other issues surrounding rising life expectancies. As an accompanying measure, we envision the abandonment of lapse assumptions in premium pricing in combination with a morbidity-contingent surrender penalty scheme for contracts with a negative economic value. Finally, we suggest to ban persistency-induced COI hikes for in-force policies to preclude insurers from putting pressure on life settlement investors and discourage them from a purchase in the first place.

The central effect of the proposed regime is to peg a policy’s surrender value to its economic value, thus increasing market efficiency; avoiding wealth transfers from policyholders to shareholders, and enhancing overall utilitarian welfare through the elimination of dead weight losses. We believe that the regulatory change will not only protect the consumers’ interests but also benefit the life insurance industry as it increases its overall financial stability and enhances its reputation. In this spirit, carriers should begin to actively participate in the secondary market instead of antagonizing it.

Our study has a few limitations that we leave for future research. Firstly, tax issues related to life insurance are beyond the scope of this paper but should be investigated before an implementation of the reform. Secondly, in order to simplify the demonstration and isolate the core effects, our theoretical framework assumes risk neutrality of carriers and investors, independence of excess utility from the economic value of a policy, perfect competition, zero transaction costs, and the absence of information asymmetries. Evidently, the market mechanism under more relaxed assumptions merits further examination. Thirdly, insurance research is often limited by data availability and this case is no exception. Due to the opaque nature of the life settlement market, we are unable to quantify the welfare effects of the proposed reform at this time. Subsequent work may be in a better position to approach the topic from an empirical perspective.
References


Bayston, Darwin. 2016. “Are life insurance companies betraying their senior policyholders?” https://www.thinkadvisor.com/2016/02/20/are-life-insurance-companies-betraying-their-senio/.


Appendix

Data

We use two sets of samples to depict the U.S. life insurance landscape: (i) The U.S. Life industry Briefing book of the SNL financial (S&P Global Market Intelligence); and (ii) SOA & LIMRA - U.S. Individual Life Insurance Persistency studies. The SNL sample is representative, covering more than 90% of total life insurance business in the U.S., and up-to-date, reported on an annual basis. Each year from 1996 to 2017, the total face amount of policies issued, surrendered, lapsed during the year and in force at year-end is reported. The sums are broken down by policy format: (i) ordinary / individual insurance, (ii) group insurance which is often offered through the workplace, (iii) credit insurance which is primarily offered by lending companies and (iv) industrial insurance which is mainly for burial purposes. Ordinary policies are further broken down into two categories: (i) Whole Life & Endowment; (ii) Term-life. However, the data lack granularity: no further break-down (by e.g. insureds' age, policy face value, insurance duration) is available.

The data used for the SOA & LIMRA study provide much more insight on U.S. life insurance termination pattern. The study breaks down the sample policies by attained age, issue age, policy type, duration etc. In addition, the study provides lapse rate of different cohorts (e.g. the lapse rate term policies from 65-year old insureds). Unfortunately, the sample only involves individual life insurance business from 27 U.S. life carriers, and is not up to date, with the last observation year being 2009.

We use the granularity of SOA & LIMRA study and scale the numbers up to reach the aggregate shown in the SNL database. While the LIMRA sample is incomplete, to make sure that it is not skewed, we compare the age distribution indicated from this sample and from another independent study by SOA, and find that the two distributions are similar.
Transformation of Equation 2 to Equation 3

Equation 2 can be rewritten as:

\[
EV_t - \rho l_t EV_{t+1} = \rho [l_t(1 - p_t) \Delta_{t+1} + (1 - l_t)] - \pi_t
\]

Let

\[
A_t := \rho l_t \quad (11)
\]
\[
B_t := \rho [l_t(1 - p_t) \Delta_{t+1} + (1 - l_t)] - \pi_t \quad (12)
\]

We can further write down the following equation system:

\[
EV_t - EV_{t+1} A_t = B_t
\]
\[
EV_{t+1} A_t - EV_{t+2} A_t A_{t+1} = B_{t+1} A_t
\]
\[
EV_{t+2} A_t A_{t+1} - EV_{t+3} A_t A_{t+1} A_{t+2} = B_{t+2} A_t A_{t+1}
\]
\[
\vdots
\]
\[
EV_{t+n} \prod_{i=t}^{t+n-1} A_i - EV_{t+n+1} \prod_{i=t}^{n} A_{t+i} = B_{t+n} \prod_{i=t}^{t+n-1} A_i
\]
\[
= \frac{B_{t+n} \prod_{i=t}^{t+n-1} A_i}{A_{t+n}}
\]

By summing the LHS and the RHS of the equation system above, we have:

\[
EV_t - EV_{t+n+1} \prod_{i=t}^{n} A_i = \sum_{j=t}^{t+n} B_j \prod_{i=t}^{j} A_i
\]
\[
\lim_{n \to \infty} \left( EV_t - EV_{t+n+1} \prod_{i=t}^{n} A_i \right) = \lim_{n \to \infty} \sum_{j=t}^{t+n} B_j \prod_{i=t}^{j} A_{t+i}
\]
\[
EV_t - \lim_{n \to \infty} \left( EV_{t+n+1} \prod_{i=t}^{n} A_i \right) = \sum_{j=t}^{\infty} B_j \prod_{i=t}^{j} A_i
\]
\[
\therefore \lim_{n \to \infty} \left( EV_{t+n+1} \prod_{i=t}^{n} A_i \right) = 0 \quad \therefore EV_t = \sum_{j=t}^{\infty} B_j \prod_{i=t}^{j} A_i \quad (13)
\]

Plugging Equation 11 and Equation 12 back to Equation 13:

\[
EV_t = \sum_{j=t}^{\infty} \left\{ \rho [l_j(1 - p_j) \Delta_{j+1} + (1 - l_j)] - \pi_j \right\} \prod_{i=t}^{j} \rho l_i
\]
\[
= \sum_{j=t}^{\infty} \left\{ \rho [l_j(1 - p_j) \Delta_{j+1} + (1 - l_j)] - \pi_j \right\} \rho^{-t} \prod_{i=t}^{j} l_i
\]
Figure 1: Annual face amount terminated 1996-2017

Figure 2: Life settlement by state

Note: Life settlements are mostly concentrated in California, New York and Florida, where wealthy people reside.
Figure 3: Face amount of ordinary life policies terminated in 2017

Total: USD1,857bn

Note: The face amount of ordinary life policies terminated in 2017 totals 1,857 billion USD, consisting of 1,709 billion USD from insureds below age 65, and 148 billion USD from insureds above age 65.
Figure 4: Life settlement effect on insurers’ net income

Note: *General expenses do not include insurance commissions, taxes, licenses and fees. If primary insurers leave life settlement being fully exploited, their net income would be negatively impacted.

Table 1: Summary of used variables and notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>Policy price</td>
</tr>
<tr>
<td>EV</td>
<td>Policy economic value</td>
</tr>
<tr>
<td>Δ</td>
<td>Policy price mark-up, $Δ = PP - EV$</td>
</tr>
<tr>
<td>$u$</td>
<td>Policy utility to insured</td>
</tr>
<tr>
<td>$u^e$</td>
<td>Policy excess utility to insured, $u^e = u - EV$</td>
</tr>
<tr>
<td>$F_u(\cdot)$</td>
<td>Cumulative frequency function</td>
</tr>
<tr>
<td>$U$</td>
<td>Policy utility to society</td>
</tr>
<tr>
<td>PhS</td>
<td>Policyholders’ surplus</td>
</tr>
<tr>
<td>BrS</td>
<td>Buyers’ (carrier and investor’s) surplus</td>
</tr>
<tr>
<td>$\rho$</td>
<td>discount factor</td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>policy premium to be paid at time $t$</td>
</tr>
<tr>
<td>$l_t$</td>
<td>insured survival rate between time $t$ and $t+1$</td>
</tr>
<tr>
<td>$p_t$</td>
<td>policy persistency rate between time $t$ and $t+1$</td>
</tr>
</tbody>
</table>
Table 3: Utility depending on insured’s choice between leaving and taking a policy

<table>
<thead>
<tr>
<th></th>
<th>surplus from taking</th>
</tr>
</thead>
<tbody>
<tr>
<td>insured utility</td>
<td>$u - PP$</td>
</tr>
<tr>
<td>carrier utility</td>
<td>$PP - EV$</td>
</tr>
<tr>
<td>total utility</td>
<td>$u - EV$</td>
</tr>
</tbody>
</table>

Table 4: Utility depending on insured’s choice between persistency and termination of a policy

Panel A: carrier+investor utility

<table>
<thead>
<tr>
<th></th>
<th>persistency</th>
<th>termination</th>
<th>surplus from termination</th>
<th>way of termination</th>
</tr>
</thead>
<tbody>
<tr>
<td>carrier utility</td>
<td>$-EV$</td>
<td>$-PP$</td>
<td>$EV - PP$</td>
<td>terminated via carrier</td>
</tr>
<tr>
<td>investor utility</td>
<td>$0$</td>
<td>$EV - PP$</td>
<td>$EV - PP$</td>
<td>sold to investor</td>
</tr>
<tr>
<td>carrier+investor utility</td>
<td>$-EV$</td>
<td>$-PP$</td>
<td>$EV - PP$</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: insured and carrier+investor utility

<table>
<thead>
<tr>
<th></th>
<th>persistency</th>
<th>termination</th>
<th>surplus from termination</th>
</tr>
</thead>
<tbody>
<tr>
<td>insured utility</td>
<td>$u$</td>
<td>$PP$</td>
<td>$PP - u$</td>
</tr>
<tr>
<td>carrier+investor utility</td>
<td>$-EV$</td>
<td>$-PP$</td>
<td>$EV - PP$</td>
</tr>
<tr>
<td>total utility</td>
<td>$u - EV$</td>
<td>$0$</td>
<td>$EV - u$</td>
</tr>
</tbody>
</table>
Figure 5: Utility of policyholders and policy buyers: unhealthy insured cohort

\[ u, PP, SV \]

\[ \Delta \]

\[ EV \]

\[ SV \]

\[ 0 \]

\[ u_{\text{min}} \]

\[ Q \]

\[ PP^* \]

\[ PP' \]

\[ EV \]

\[ SV \]

\[ 0 \]

\[ u_{\text{min}} \]

\[ Q \]

\[ PP^* \]

\[ PP' \]

\[ \text{policyholder surplus} \]

\[ \text{insurer surplus} \]

\[ \text{deadweight loss} \]

Note: With an unhealthy insured cohort, MSS decreases insurance persistency, decreases carrier surplus (from positive to zero), and increases insured surplus.

Figure 6: Utility of policyholders and policy buyers: healthy insured cohort

\[ u, PP, SV \]

\[ \Delta \]

\[ EV \]

\[ SV \]

\[ 0 \]

\[ u_{\text{min}} \]

\[ Q \]

\[ PP^* \]

\[ PP' \]

\[ \text{policyholder surplus} \]

\[ \text{insurer deficit} \]

\[ \text{deadweight loss} \]

Note: With a healthy insured cohort, MSS increases insurance persistency, increases carrier surplus (from negative to zero), and decreases insured surplus.
Note: With a terminally ill insured cohort, MSS reduces insurance persistency, increases insured surplus (from zero to positive), and has no impact on carrier surplus.
Note: With $\Delta \equiv 0$, the insured abandons a policy at $t = 6$ and re-acquires a policy at $t = 8$. The insurance’s utility to the whole society, $U$, is thus maximized, equal to area $A+B+D+E$. If the policy is kept in force from $t = 0$ to $t = 19$, $U$ equals area $A+B-C+D+E$, smaller than $A+B+D+E$; if the termination occurs at $t = 3$ and re-acquisition at $t = 11$, $U$ equals area $A+E$, also smaller than $A+B+D+E$. 

Figure 8: Time series of discounted excess utility $\rho(t)u^e(t)$ of insurance