Life Insurance and Life Settlements: The Case for Health-Contingent Cash Surrender Values*

Hanming Fang† Edward Kung‡

January 30, 2018

Abstract

We investigate why life insurance policies in practice either do not have a cash surrender value (CSV), or have cash surrender values that are small and are not adjusted for health status. We show that including health-contingent CSVs to a life insurance contract causes a dynamic commitment problem, which makes it more costly up front for policyholders to purchase long term contracts (because some poor risks who would otherwise have lapsed can and will now capture the cash surrender value instead). To the extent that life insurance policyholders' incomes tend to increase over the course of the policy, policyholders are not willing to accept higher *ex ante* premium costs in return for the extra liquidity provided by the cash surrender value. Because health-contingent CSVs act in a similar way to a life settlement market, we also study the life insurers' equilibrium choice of cash surrender values in the presence of a life settlement market. We find that optimally chosen cash surrender values can partially mitigate the consumer welfare loss caused by the settlement market (as in Daily et al., 2008), but only if the cash surrender values are allowed to be contingent on health status.

Keywords: Life insurance, cash surrender values, life settlement market, reclassification risk insurance.

JEL Codes: G22, L11

---

*This paper supersedes our earlier NBER Working Paper (Fang and Kung, 2010a). We would like to thank the Editor, three anonymous referees, Jacques Crémer, Alessandro Lizzeri, George Mailath, Andrew Postlewaite, Wing Suen and seminar participants at Duke University, Hong Kong University, Peking University, University of Pennsylvania and the 2009 Econometric Society Summer Meeting in Boston for useful comments. Fang would also like to gratefully acknowledge the generous financial support from the National Science Foundation through Grant SES-0844845. All remaining errors are our own.

†Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104, and the NBER. Email: hanming.fang@econ.upenn.edu

‡Department of Economics, UCLA, 8343 Bunche Hall, Los Angeles, CA 90095. Email: ekung@econ.ucla.edu

1 Introduction

Life insurance policies are typically long-term contracts in which policyholders pay an annual premium in return for a guarantee that his/her beneficiaries will receive a sum of cash, called the death benefit, if the policyholder dies within the coverage period. Premiums are almost universally front-loaded, meaning that policyholders pay more than the actuarially fair cost of insurance in the early stage of the contract (when they are young), and less than the actuarially fair cost of insurance in later stages of the contract (when they are old and when their mortality risk has likely increased). Hendel and Lizzeri (2003) showed that long-term, front-loaded contracts serve a useful purpose by protecting policyholders against reclassification risk—that is, the risk of negative health shocks driving up the contemporaneous price of insurance, thus making insurance unaffordable in bad health states.

An interesting feature of most life insurance contracts is that if the policyholder lapses before the coverage period is over, the cash they receive for surrendering their policy (called the cash surrender value, henceforth CSV) is usually zero or else very small.[1] Moreover, cash surrender values are typically not adjusted for the health status of the policyholder. To see why these features are puzzling, consider that most contracts are front-loaded. This implies that later in the life of a policy, the expected payout of benefits exceeds the net present value of premiums. Lapsation with no cash surrender value therefore represents a pure profit to the insurer, and even more so if the policyholder has impaired health status. In a competitive marketplace, one might expect the policyholder to be able to extract some value from the contract, rather than lapsing it for zero private returns.

It has been suggested by Doherty and Singer (2003) and the Deloitte (2005) that health-contingent cash surrender values may face regulatory difficulties. While this may indeed be the case, we have not been able to find any specific regulations that explicitly prohibit the writing of health-contingent cash surrender values in life insurance contracts. In response to the inquiries we sent to the North Carolina Department of Insurance, regulators said that they were not aware of any such regulations, either in North Carolina or in other states. Another possibility, also suggested by Doherty and Singer (2003), is that there may be large administrative costs to implementing health-contingent surrender values. This is certainly a possibility, but it begs the question of why a life settlement market has emerged precisely to take advantage of the gap between the actuarial

[1] The Life Insurance Settlement Association (LISA) estimates that the average surrender value is only 3-5% of the policy's face value. See http://www.lisa.org/content/13/What-is-a-Life-Settlement.aspx/.

value of a life insurance policy and its cash surrender value.[2] A third possibility is that if surrender values are health-contingent, then policyholders can game the system by pretending to be of poor health. Since it is likely easier for a healthy person to pretend to be sick, than for a sick person to pretend to be healthy, verification costs may be high.[3]

In this paper, we study the role of surrender values in life insurance contract design. We find that when there is no life settlement market, the life insurance contracts that would emerge in equilibrium will not contain cash surrender values. This result holds regardless of the presence of any regulatory, administrative, or verification costs in implementing health-contingent cash surrender values. Thus, the observed lack of health-contingent surrender values is an equilibrium outcome of consumers' optimizing behavior.

Our analysis follows closely the Hendel and Lizzeri (2003) (henceforth HL) model of life insurance contracts. HL studied a model in which consumers' mortality risks change over time, and these changes are symmetrically observed by both the consumer and the life insurance companies. Because the consumers' mortality risks change over time, they face reclassification risk, which is the risk that a deterioration in health makes it more costly for them to obtain life insurance on the spot market. HL showed that, in equilibrium, a competitive life insurance market will offer dynamic contracts that insure against reclassification risk by charging a higher premium up front, in exchange for fixed premiums that do not depend on mortality risk later. Thus, the front-loaded contracts that are so widely prevalent in the life insurance industry emerge in the equilibrium of the HL model.

However, the HL model did not allow for the possibility that life insurance may no longer be needed in the future, nor for the possibility of a cash surrender value to be specified as part of the contract. In this paper, we expand on the HL model by allowing for possible bequest-motive loss and surrender values that are endogenously chosen. We find that having a positive surrender value introduces a dynamic commitment problem in which some ex post poor risks who would have otherwise dropped out of the pool for exogenous reasons, instead capture the surrender value. This makes it more costly for life insurers, who operate competitively, to provide reclassification risk insurance via front loading. Specifically, for any level of premiums guaranteed in the second stage

[2] A life settlement is a financial transaction in which a policyholder sells his/her policy to a third party for more cash than the surrender value offered by the policy itself. The third party subsequently assumes responsibility for all future premium payments, and is entitled to the death benefits if the original policyholder dies within the coverage period. The industry is young but growing rapidly, with purchases of about $2.57 billion in face value in 2013 (LISA).

[3] To be sure, the life settlement market faces the same issue, and verification costs could explain why the settlement market tends to target people who are already aged, and why the initial development of the settlement market started with viatical transactions that focused entirely on policyholders who were terminally ill.

of the contract, a higher up front premium must be charged in the first stage of the contract if the contract contains a positive surrender value. If the policyholder's income is rising over the term of the contract—which is likely to be the case for most life insurance purchasers—this represents a transfer of wealth from a low income state to a high income state. Because of this, the policyholder is not willing to accept a higher up-front premium cost in return for the extra liquidity provided by the cash surrender value. Our results thus help explain why cash surrender values are not typically observed in the life insurance industry.

The logic underpinning our analysis relies on the assumption that life insurance companies can perfectly commit to *not* buy back policies from individuals whose health has deteriorated. It also assumes the absence of a life settlement market that would purchase policies for which there is a gap between the actuarial value and the cash surrender value (and thus keep these policies in the pool). In reality, a life settlements market has emerged precisely to take advantage of that gap. We therefore also analyze the effects that a life settlement market would have on the equilibrium of our model. We find that if surrender values are zero, then the presence of a life settlement market distorts the form of the optimal contract. In particular, the life settlement market reduces the amount of reclassification risk through front-loading, reducing *ex ante* consumer welfare. Intuitively, this happens because even though life insurers can commit to zero cash surrender values, policyholders who no longer need their policies *ex post* cannot commit not to sell their policies on the settlement market. The original insurer is thus required to honor some policies that otherwise would have been lapsed or surrendered for less than the actuarial value. In a competitive setting, this increased cost will have to be passed on to consumers through higher first period premiums. Since the only reason to front-load in this model is to reduce dynamic reclassification, any loading above that amount must be inefficient. Our analysis of the welfare effect of the settlement market mirrors the analysis of Daily et al. (2008) (henceforth DHL) and echoes the arguments expressed in Deloitte (2005).[4,5]

[4]Our results and the results of DHL should not be taken to mean that the life settlement market is unequivocally bad for consumers. Both of our analyses assume deterministic growing income, and that policyholders lapse and surrender for purely exogenous reasons (such as loss of bequest motive). If income is risky, then a life settlement market can provide increased liquidity in a state of high marginal utility. This possibility is recognized in DHL and also discussed in Doherty and Singer (2003), Singer and Stallard (2005). If life insurance purchasers have incorrect beliefs about future income, health, or bequest shocks, then the life settlement market can also enhance consumer welfare by providing liquidity in unanticipated states (see Gottlieb and Smetters (2016); Fang and Wu (2017)).

[5]Empirically, a key determinant of the welfare effects associated with life settlements is the degree to which life insurance premiums are supported by lapses, and the motivation for lapsing. Gatzert (2010) documents that lapsation rates are high (on the order of 7% per year) and Gatzert et al. (2009) estimates that lapsation contributes significantly to insurer profits. Fang and Kung (2010b) argue that lapsation among the elderly (the target demographic of life settlements) is mostly driven by loss of bequest motive.

Finally, we study how the equilibrium choice of life insurance contracts, and in particular the choice of the cash surrender values, might respond to the presence of a settlement market. We find an interesting and policy relevant result, which is that endogenously chosen, health-contingent cash surrender values can partially, but not completely, mitigate the welfare loss associated with the settlement market, especially if there are less frictions in surrendering one's policy than in a life settlement transaction.[6] However, if cash surrender values are not allowed to depend on the policyholder's health status, then the equilibrium cash surrender value will be zero, and the equilibrium contracts and allocations will be identical to the case in which there is a settlement market but no cash surrender values. If indeed there are regulatory barriers to offering life insurance policies with health-contingent cash surrender values, then our results suggest that, given the emergence of the life settlement market, it would be wise to take a second look at whether such regulations are justified. If such regulations do not exist, then it will be interesting to observe whether life insurance companies will begin to offer health-contingent cash surrender values in response to the life settlement market.

Our paper contributes to a growing academic literature studying the impact of lapsation and the settlement market on life insurance. Gatzert et al. (2009) uses an actuarial model to study the impact of life settlements on insurers' lapsation profits, finding substantial declines. Daily et al. (2008) studies the impact of the life settlement market on the ability for life insurers to front load. Zhu and Bauer (2013) and Januário and Naik (2014) study the role of adverse selection in the settlement market. We contribute to this literature by showing how cash surrender values would be chosen in equilibrium, both with and without a secondary market, and by highlighting the important role of mortality heterogeneity in the contract design.

Our model itself follows closely the model of Daily et al. (2008) (DHL), which itself is an extension of the HL model to consider the effect of introducing a life settlement market on life insurance contracts when there is a possibility of bequest motive loss. We contribute to DHL in two specific ways. First, we formally consider the endogenous response of the primary insurer to the settlement market through the offer of health-contingent surrender values in the contract design. The DHL model did not allow for pre-specified cash surrender values in the contract. We show that the welfare loss incurred by the settlement market can be mitigated by the introduction of health-contingent surrender values. Our second contribution to DHL is that, because we are

[6]Frictions in life settlement transactions can arise, for example, from marketing costs. This friction is unlikely to arise in a surrender because the life insurance companies already know who their policyholders are.

not constrained by the space requirements of a note (as DHL are), we are able to provide a more complete analysis of their model. In particular, we formalize the welfare argument of DHL which, while intuitively correct, assumed that "when insurance companies respond to the presence of a settlement market, the outcome is an increase in consumption in the event in which it was already the highest, and a reduction in consumption in period 1 when it was already fairly low." We show that the settlement market can actually unravel dynamic reclassification insurance, resulting in an equilibrium which is a sequence of spot contracts. In this equilibrium, consumption is actually higher in the first period when income is low, but welfare is still reduced because of the lack of insurance against reclassification.

One limitation of our approach is that we assume fully rational and forward-looking agents. This is a useful benchmark, but certain empirical facts do not seem to support a fully rational model. For example, lapse rates seem too high and lapsers usually have poorer health, which runs counter to rational behavior. This has motivated a literature in which behavioral biases are also analyzed. Gottlieb and Smetters (2016) investigate the equilibrium life insurance contracts where lapsation is motivated by a negative income shock and consumers are overconfident in the sense that they place zero probability on the event of experiencing the liquidity shock. They show that the settlement market can help smooth consumption and increase consumer welfare. Fang and Wu (2017) studies how the life settlement market may affect consumer welfare when consumers are overconfident about their bequest motives or about their mortality rates. They show that life settlement market may improve consumer welfare by imposing discipline on how much the life insurance companies are able to exploit consumers' biased beliefs.

More broadly speaking, our research contributes to the literature on dynamic contracts and market incompleteness. Our results suggest that secondary markets for dynamic contracts can result in dynamic inefficiencies by eroding the commitment power of agents. Moreover, the results suggest that adding options to contracts does not necessarily move the market towards greater efficiency, unless those options can be conditioned on all the states relevant to the exercise of the option. In our context, the cash surrender value is much like an option on the life insurance policy, but it is not useful unless it can be conditioned on health status.

The remainder of the paper is structured as follows. In Section 2 we present our baseline model with endogenous cash surrender values but no settlement market. In Section 3, we extend the baseline model to include a life settlement market, but do not allow for cash surrender values. In Section 4, we examine how cash surrender value will be chosen in the presence of a settlement market, and show that the usefulness of cash surrender values depends crucially on whether they can be made contingent on health. In Section 5 we consider the case where policyholders' bequest motives may diminish instead of totally disappear. In Section 6 we summarize our findings and discuss directions for future research. All proofs are collected in the appendix.

2 A Model of Life Insurance with Endogenously Chosen cash surrender values

2.1 The Model

Health, Income and Bequest Motives

Consider a perfectly competitive market for life insurance that includes risk averse individuals (policyholders) and risk neutral life insurance companies. There are two periods. In the first period, the policyholder has a probability of death $p_1 \in (0, 1)$ known to both himself and the insurance companies. In the second period, the policyholder has a new probability of death $p_2 \in (0, 1)$, which is randomly drawn from a continuous and differentiable c.d.f. $\Phi(\cdot)$ with a corresponding density $\phi(\cdot)$. A consumer's period 2 health state p_2 is not known in period 1, but p_2 is learned by both the insurance company and the consumer (and is thus common knowledge) at the start of period 2.

The policyholder's income stream is $y-g$ in period 1 and $y+g$ in period 2, where y is interpreted as the mean life cycle income and $g > 0$ captures the income growth over the periods. Both y and g are assumed to be common knowledge.

The policyholder has two sources of utility: his own consumption should he live, and his dependents' consumption should he die. When the policyholder is alive, he derives utility $u(c)$ from consuming $c \geq 0$. If he dies, then he derives utility $v(c)$ if his dependents consume $c \geq 0$. $u(\cdot)$ and $v(\cdot)$ are both strictly concave and twice differentiable.

In period 2, there is a chance the policyholder loses his bequest motive.[7] We denote the probability of bequest motive loss by q. The bequest motive is realized at the same time as the period 2 health state, but unlike the realization of health status, the bequest motive is private information to the policyholder and cannot be contracted upon. If the policyholder loses his bequest motive, then he does not receive utility $v(\cdot)$ from his dependents' consumption when he

[7] A loss of bequest motive could result from death of a spouse, or from divorce, or from changes in the circumstances of the intended beneficiaries of the life insurance policy. For example, the policyholders children may have graduated from college and found well-paying jobs.

dies. We assume that the policyholder does not save, and that the preferred method of ensuring a bequest is through life insurance.[8]

Timing, Commitment, and Contracts

Now we provide more details about the timing of events. At the beginning of period 1, after learning the period 1 health state p_1, the consumer may purchase a long term contract from an insurance company. A long term contract specifies a premium and face value (the amount of death benefits) for period 1, (Q_1, F_1), a menu of health-contingent premiums and face values $(Q_2(p_2), F_2(p_2))$ for each period 2 health state, and a menu of cash surrender values $S_2(p_2)$ for each period 2 health state.[9]

After purchasing a contract, the consumer pays Q_1 in premiums. He then consumes his remaining income, given by $y - g - Q_1$. With probability p_1 he dies. If he dies, his dependents consume the face amount of the policy, F_1. If he lives, then at the start of period 2, both the insurance company and the policyholder learn the period 2 health state p_2, and the policyholder learns whether or not he has a bequest motive. The policyholder then has three options: 1) he can continue with his contract by paying the premium $Q_2(p_2)$. In this case he consumes $y + g - Q_2(p_2)$ and if he dies, his dependents receive $F_2(p_2)$. 2) He can surrender his policy for $S_2(p_2)$ and purchase a new policy (Q, F) on the spot market. In this case, he consumes $y + g + S_2(p_2) - Q$, and if he dies his dependents consume F. 3) He can surrender his policy and remain uninsured. In this case he consumes $y + g + S_2(p_2)$ and his dependents consume nothing if he dies.[10] Figure 1 illustrates the timing of information arrival and decisions in our model.

2.2 Equilibrium Contracts

To characterize the equilibrium contract, we first consider the actions of a policyholder in the second period. If the policyholder loses his bequest motive, then his best course of action is to surrender his policy and remain uninsured. If the policyholder retains his bequest motive, then he can either keep his policy or surrender it and purchase a new policy. If he keeps his policy, his

[8] This would indeed be the case for policyholders whose income is increasing over the coverage period.

[9] It may seem strange to allow for health-contingent premiums and face amounts, but such contracts do exist in the life insurance market. For example, some types of annual renewable term policies will award you a premium discount if you prove your good health. Moreover, as we will see, absent a secondary market, the equilibrium outcome can be replicated with contracts that are not contingent on second period health.

[10] This choice is also equivalent to surrendering the policy and purchasing a new contract with premium and face amount $(0, 0)$.

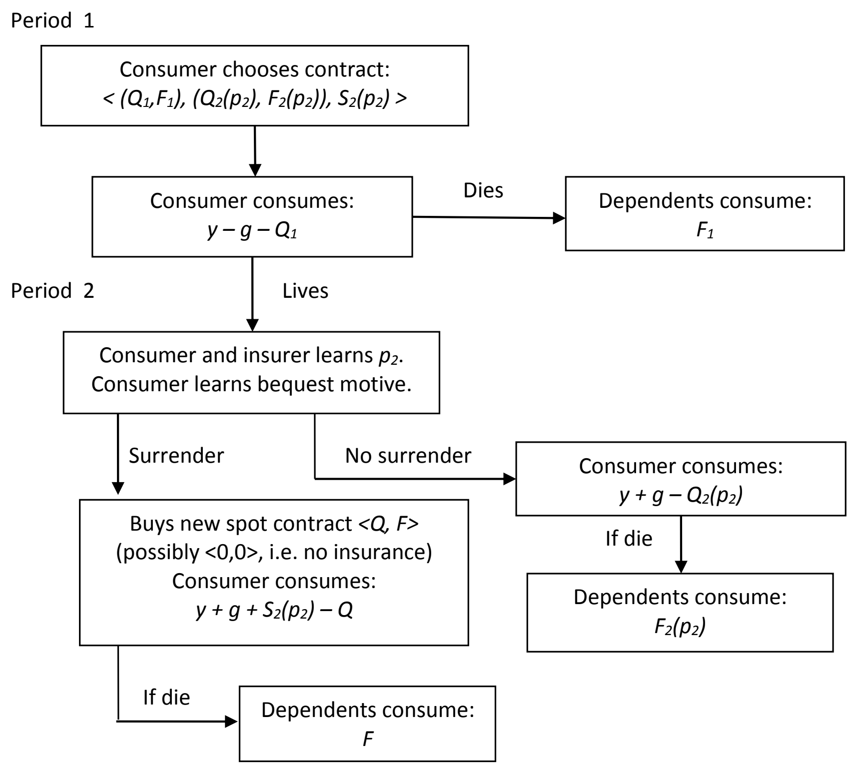

Figure 1: Model Timing

expected utility is:

$$u(y + g - Q_2(p_2)) + p_2 v(F_2(p_2)) \tag{1}$$

If he surrenders and repurchases on the spot market, his expected utility is:

$$\max_{Q,F} \; u(y + g + S_2(p_2) - Q) + p_2 v(F)$$
$$\text{s.t.} \;\; p_2 F - Q = 0 \tag{2}$$

The constraint in (2) simply requires that competition drives the actuarial value of any spot contract to zero.

We will assume, without loss of generality, that *in equilibrium* the policyholder who retains his bequest motive will always keep the policy rather than surrender and repurchase. To see why this does not result in any loss of generality, let $(Q^*(p_2), F^*(p_2))$ be the solution to (2). If the policyholder surrenders and repurchases, he consumes $y + g + S_2(p_2) - Q^*(p_2)$ and his dependents consume $F^*(p_2)$ if he dies. This outcome could have been replicated by choosing

$Q_2(p_2) = Q^*(p_2) - S_2(p_2)$ and $F_2(p_2) = F^*(p_2)$.[11] Moreover, by choosing $Q_2(p_2)$ and $F_2(p_2)$ in this way, the firm is indifferent between whether the policyholder surrenders or continues the policy. Since any outcome that can be achieved via surrender and repurchase can be replicated by an appropriate choice of second period contract terms that will not induce surrender and repurchase, we constrain the search for the optimal contract to such contracts.[12]

Under perfect competition, the equilibrium contract must maximize *ex ante* consumer welfare subject to the above constraint, together with a zero profit constraint and a constraint that the cash surrender value $S(p_2)$ is non-negative. We require the cash surrender value to be non-negative because policyholders cannot commit to contracts requiring them to pay the insurance company in the case of voluntary surrender. Specifically, the equilibrium contract must solve:

$$\max \; u(y - g - Q_1) + p_1 v(F_1) + (1 - p_1) \int \left\{ (1 - q) \begin{bmatrix} u(y + g - Q_2(p_2)) \\ +p_2 v(F_2(p_2)) \end{bmatrix} + qu(y + g + S(p_2)) \right\} d\Phi(p_2) \tag{3a}$$

$$\text{s.t. } Q_1 - p_1 F_1 + (1 - p_1) \int \{(1 - q) [Q_2(p_2) - p_2 F_2(p_2)] - qS(p_2)\} \, d\Phi(p_2) = 0, \tag{3b}$$

$$p_2 F_2(p_2) - Q_2(p_2) - S_2(p_2) \geq 0 \text{ for all } p_2, \tag{3c}$$

$$S(p_2) \geq 0 \text{ for all } p_2, \tag{3d}$$

The first order conditions for an optimum with respect to Q_1, F_1, $Q_2(p_2)$, $F_2(p_2)$, and $S_2(p_2)$ are:

$$u'(y - g - Q_1) = \mu \tag{4a}$$

$$v'(F_1) = \mu \tag{4b}$$

$$u'(y + g - Q_2(p_2)) = \mu - \frac{\lambda(p_2)}{(1 - p_1)(1 - q)\,\phi(p_2)} \tag{4c}$$

$$v'(F_2(p_2)) = \mu - \frac{\lambda(p_2)}{(1 - p_1)(1 - q)\,\phi(p_2)} \tag{4d}$$

$$u'(y + g + S(p_2)) = \mu + \frac{\lambda(p_2)}{(1 - p_1)q\phi(p_2)} - \frac{\gamma(p_2)}{(1 - p_1)q\phi(p_2)} \tag{4e}$$

[11] This raises the possibility that $Q_2(p_2)$ might be negative. We place no restriction on the sign of $Q_2(p_2)$, but in equilibrium we will find that it is always chosen to be positive for any p_2. In general, relaxing the constraints in an optimization problem do not result in any loss of generality unless the constraint would have been binding.

[12] This assumption manifests itself in constraint (3c) below. Relaxing this assumption will introduce a multiplicity of optimal contracts, some with equilibrium surrender and repurchase, but outcomes will be the same in each equilibrium (i.e. consumption of the policyholder and the dependents, and profits to the insurance company, will be the same in every state.)

where $\mu > 0$, $\lambda(p_2) \geq 0$, and $\gamma(p_2) \geq 0$ are the Lagrange multipliers.

It is easy to see that constraint (3d) must bind for all p_2. If it were slack for some p_2, then $\gamma(p_2) = 0$. This implies that $u'(y + g + S(p_2)) \geq u'(y - g - Q_1)$, which is impossible. The intuition is clearly illustrated by the first order conditions. The consumer wants to equalize marginal utility between states as much as possible, but the cash surrender value is only received in the state that already has the lowest marginal utility. The optimal decision then is to push cash surrender value down to its lower bound.

Proposition 1. *In the absence of a life settlement market, equilibrium life insurance contracts will not include a positive cash surrender value.*

Proposition 1 is consistent with the empirical observation that most Term Life policies (which are pure life insurance products) do not include cash surrender values. It may seem at odds, however, with the observation that many Permanent Life policies do specify a cash surrender value. The industry has sometimes advertised the cash surrender value as a redemption of front loaded premiums, but in practice surrender payouts are extremely small.[13] Moreover, in the industry's marketing literature, the term "surrender value" is applied even to the investment component of Permanent Life products that bundle life insurance with investment. In such cases the surrender value would more appropriately be called the account value.[14] In our current model, policyholders have no motivation to save because their income growth is positive over the life cycle (perhaps due to credit constraints, etc), and so the equilibrium contract contains no investment component and more closely aligns with Term Life products. If income growth is negative, the equilibrium contract may contain a positive, non-health-contingent cash surrender value as policyholders use the life insurance contract as a savings device.[15]

Since cash surrender values are zero in equilibrium, we can simplify the optimization problem in (3) by imposing $S_2(p_2) = 0$ for all p_2. The problem then reduces to a model that is similar to the model of Hendel and Lizzeri (2003). In fact, the equilibrium contracts can be characterized in much the same way as in Proposition 1 of HL.

Proposition 2. ***(Proposition 1 of Hendel and Lizzeri 2003)*** *The equilibrium contract satisfies the following:*

[13]Table 4 in Gatzert (2010) shows that the average surrender value paid per voluntarily terminated policy in 2006 was less than \$4; this despite Permanent Life having significant market share.

[14]This is the view of Gilbert and Schultz (1994).

[15]A contract with: $S_2(p_2) = -2g - Q_1 = S > 0$, $Q_2(p_2) = -S$, $F_2(p_2) = v'^{-1}(u'(y+g+S))$, $F_1 = v'^{-1}(u'(y+g+S))$ will satisfy all the first order conditions (4a)-(4e) when $g<0$. Q_1 would then be chosen to satisfy the zero profit constraint.

1. *In each period and in each health state for which there is a bequest motive, premium and face value are chosen to equalize the marginal utility of consumption and the marginal utility of dependents' consumption:*

$$u'\left(y-g-Q_1\right)=v'\left(F_1\right) \tag{5}$$

$$u'\left(y+g-Q_2(p_2)\right)=v'\left(F_2(p_2)\right) \text{ for all } p_2. \tag{6}$$

2. *There is a period 2 threshold health state, p_2^* such that for all $p_2 \leq p_2^*$, the period 2 premiums are actuarially fair, and for all $p_2 > p_2^*$, the period 2 premiums are constant with respect to health status, and therefore actuarially favorable to the policyholder.*

3. *For any q, there is a threshold $\hat{g}$ such that when the income growth parameter g is smaller than $\hat{g}$, then p_2^* is strictly less than 1. Thus, reclassification risk insurance will always be provided for individuals with low income growth.*

Figure 2 illustrates the profile of period 2 premiums with respect to health state when there is no settlement market. Note that $Q_2^{FI}(p_2)$ is defined as the *actuarially fair level of premium payment* that solves (6) (i.e. the level of premium that satisfies both $pF-Q=0$ and $u'(y+g-Q)=v'(F)$). We see that for a set of *ex post* healthy consumers, i.e. those with $p_2 \leq p_2^*$, the premium is set to the actuarially fair level. For a set of *ex post* unhealthy consumers, i.e. those with $p_2 \geq p_2^*$, the premium is actuarially favorable. It also follows then, from the zero profit constraint 3b, that it must be the case that in equilibrium $Q_1 > Q^{FI}(p_1)$; that is, in equilibrium, policyholders are paying a higher than actuarially fair premium in the first period. This is exactly the phenomenon of *front loading.*

It is instructive to consider the intuition for why the contract is structured in this way. In a world of full commitment, the policyholder would ideally prefer to insure against all reclassification risk by writing a contract that does not depend on period 2 health states at all. However, policyholders with better than expected health realizations will not be able to commit to the second period terms. If premiums were completely constant across health states, then the healthiest of policyholders will prefer to surrender their policy and purchase a policy with better terms on the spot market.[16]

[16] Indeed, a long term contract with second period premiums that are constant in health realizations will replicate the outcome of the equilibrium described in Proposition 2. *Ex post* healthy policyholders will simply let the policy lapse and repurchase on the spot market. As we will see, the ability for the health-contingent equilibrium contract to be replicated by a non-health-contingent contract is a feature does not carry over when a settlement market is introduced.

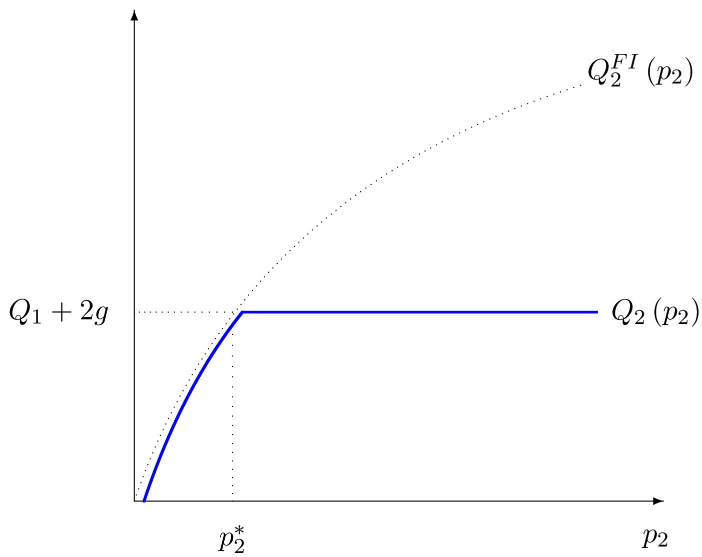

Figure 2: Equilibrium Period 2-Premium Profile $Q_2(p_2)$

Figure 2 also illustrates one of the mechanisms by which cash surrender values increase the upfront cost of insurance. Without a cash surrender value, some poor health risks, with mortality risk greater than p_2^*, will lose their bequest motive and let their policies lapse. Because these poor health risks have lower than actuarially fair premiums, the life insurance company stands to save a lot of money when they lapse. If on the other hand there is a positive cash surrender value, these individuals who would have otherwise lapsed will instead capture the cash surrender value. The cost of the cash surrender value will then have to be passed on in the form of higher upfront premiums in the first period. The reasoning outlined here is the same reasoning for why a life settlement market can reduce *ex ante* consumer welfare. Even if cash surrender values are set to zero, the settlement market can act in the place of cash surrender values, and because consumers cannot commit to not participating in the settlement market, this raises the upfront cost of insurance. We explore this line of reasoning more formally in the next section.

3 Introducing the Life Settlement Market

We now study the equilibrium life insurance contract in the presence of a settlement market. Policyholders may now sell their policies on the life settlement market in period 2. We assume that, like the life insurance companies, life settlement firms can observe and contract on the policyholder's second period mortality risk p_2. If the policyholder's second period premium and face value are $Q_2(p_2)$ and $F_2(p_2)$, then the policy can be sold for a fraction $\beta \leq 1$ of its actuarial value on the settlement market (i.e. the policyholder receives $\beta\left(p_2 F_2(p_2) - Q_2(p_2)\right)$). β can represent either the degree of competition in the secondary market, or the amount of fees or commissions required by the settlement firms, or any other frictions associated with the settlement market.[17] We assume that the settlement market operates at the beginning of period 2, just after the mortality risk and bequest motive are learned, but before any life insurance premiums are paid and before consumption occurs. This is an innocuous assumption because perfect competition among life insurance companies ensures that the expected net present value of period 1 contracts are zero, and thus there is no surplus to be recovered on the settlement market for period 1 contracts. To better contrast the role of cash surrender values, we first consider the case in which cash surrender values are restricted to be zero.

3.1 Equilibrium Contracts with a Settlement Market

To characterize the equilibrium contract in the presence of a settlement market, we assume that policyholders participate in the life settlement market if and only if they lose their bequest motive. This assumption does not change the equilibrium outcome for the same reasoning as discussed in Section 2.2. If the policyholder retains his bequest motive, any allocation that can be achieved by selling his policy to the settlement market and repurchasing insurance on the spot market, could have been achieved with the appropriate contract choice in period 1. The equilibrium contract is therefore of a form $\langle (Q_1^s, F_1^s), (Q_2^s(p_2), F_2^s(p_2)) : p_2 \in (0,1) \rangle$, and is chosen to solve:[18]

$$\max \; u(y - g - Q_1^s) + p_1 v(F_1^s) \tag{7a}$$

$$+ (1-p_1) \int \left\{ (1-q) \begin{bmatrix} u(y+g-Q_2^s(p_2)) \\ +p_2 v(F_2^s(p_2)) \end{bmatrix} + q u(y+g+\beta V_2^s(p_2)) \right\} d\Phi(p_2)$$

$$\text{s.t. } Q_1^s - p_1 F_1^s + (1-p_1) \int \left[Q_2^s(p_2) - p_2 F_2^s(p_2) \right] d\Phi(p_2) = 0, \tag{7b}$$

$$p_2 F_2^s(p_2) - Q_2^s(p_2) \geq 0 \text{ for all } p_2, \tag{7c}$$

[17] Currently, the life settlement industry typically offers about 20% of the death benefits to sellers after commissions and fees. Since β is relative to the actuarial value of the policy, and not the death benefit, the plausible range of β is around 0.4 to 0.6 (see Life Insurance Settlement Association (2006)).

[18] The s superscript is used to denote the equilibrium contract in the presence of a settlement market.

where $V_2^s(p_2) \equiv p_2^s F_2^s(p_2) - Q_2^s(p_2)$ is defined as the actuarial value of the second period contract terms. Constraint (7c) is required because consumers cannot commit to contracts with negative actuarial value in period 2. If the policy has negative actuarial for some health state in period 2, then the policyholder would simply let the policy lapse and purchase a new policy on the spot market.[19]

It is important to emphasize that the solution to (7) is *not* the same as the solution to (3) with the restriction that $S_2(p_2) = \beta V_2^s(p_2)$. Although setting $S_2(p_2) = \beta V_2^s(p_2)$ would restore the objective function in (7a) to be identical to that in (3a), the zero profit constraints would still be different: in (3b), $\beta V_2^s(p_2)$ wold enter the zero profit constraint in place of $S_2(p_2)$; but in (7b), the full $V_2^s(p_2)$ enters. To put it differently, even if policyholders are selling their policies at a discount relative to the actuarial value, the life insurance company is still liable for the full value of every policy sold to the settlement market.

Proposition 3. *The equilibrium contract satisfies the following:*

1. *In each period and in each health state for which there is a bequest motive, premium and face value are chosen to equalize the marginal utility of consumption and the marginal utility of dependents' consumption:*

$$u'\left(y - g - Q_1^s\right) = v'\left(F_1^s\right) \tag{8}$$

$$u'\left(y + g - Q_2^s(p_2)\right) = v'\left(F_2^s(p_2)\right) \text{ for all } p_2. \tag{9}$$

2. *There is a period 2 threshold health state p_2^{s*} such that for all $p_2 \leq p_2^{s*}$, the period 2 premiums are actuarially fair, and for all $p_2 > p_2^{s*}$, the period 2 premiums are actuarially favorable to the policyholder. Unlike in the case without settlement markets, the second period premiums are increasing with respect to mortality risk.*

3. *There is a threshold $\hat{q}$ such that, if $q > \hat{q}$, then for any $g > 0$, $p_2^{s*} = 1$. That is, if the probability of bequest motive loss is high enough, the equilibrium contract is simply a sequence of spot contracts.*

The equilibrium contract in the presence of a settlement market exhibits many similarities to the equilibrium contract without a settlement market, but also has two key differences. The first

[19] Analogously to constraint (3c) above, we can constrain our search for the optimal contract to contracts where such lapsation does not occur, because any outcome where lapsation occurs when the policyholder has bequest motive can be replicated by a contract in which the lapsation does not occur.

similarity is that marginal utilities are still equalized between the policyholder and his dependents' consumption in states with a bequest motive. This similarity is very natural to occur because the firms care only about the actuarial value of the contract, and not the allocation between premium and face amount. Thus, within any health state, there is always one degree of freedom for the policyholder to adjust premiums and face values appropriately to equalize the marginal utilities between his own and his dependents' consumption. The second similarity is that there is a mortality risk threshold above which reclassification risk insurance is provided.

The first key difference between the equilibrium contracts with and without settlement markets is the form in which reclassification risk insurance is provided. Figure 3 shows the profile of second period premiums with respect to mortality risk. We can see that in the presence of a settlement market, reclassification risk insurance no longer takes the form of guaranteed flat premiums in the second period. Instead, reclassification risk insurance is now provided in the form of *premium discounts* relative to the spot market premium. An interesting corollary of this result is that the equilibrium allocation in the presence of a settlement market can no longer be replicated by a non-health-contingent contract. Premiums and face values *must* be health-contingent. As the settlement market continues to grow and become a more important player in the life insurance market, it will interesting to see if life insurance policies with health-contingent premiums become more popular.

The second key difference between equilibrium contracts with and without settlement markets is the extent to which the market is even capable of providing reclassification risk insurance. Notice that the threshold health states, p_2^* and p_2^{s*} are not necessarily the same. In fact, as Proposition 3 shows, if the probability of bequest motive loss is too high, no reclassification risk insurance can be provided at all. The equilibrium contract will be equivalent to a sequence of spot contracts. This illustrates a potentially severe consequence of the life settlement market: it can lead to the unraveling of dynamic contracts. Clearly, consumer welfare is reduced if such an unraveling occurs. However, as has been shown in Daily et al. (2008), a more general welfare result can be shown. In the context of our model, the presence of a settlement market is generically welfare reducing, even if there are no inefficiencies in the settlement market itself (i.e. if $\beta = 1$).

Proposition 4. *Ex ante consumer welfare is reduced by the presence of a settlement market.*

The argument we use in proving Proposition 4 is as follows: for any contract that is feasible in the presence of a settlement market, we construct a different contract that is feasible in a world without

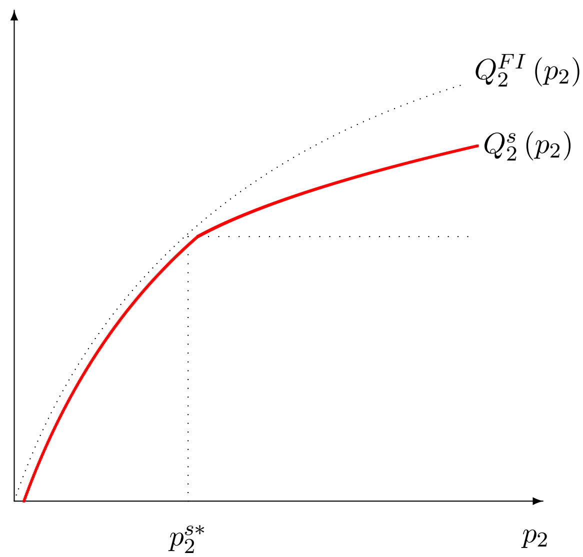

Figure 3: Equilibrium Period-2 Premium Profile $Q_2^s(p_2)$: The Life Settlement Market Case.

a settlement market. The constructed contract offers identical coverage in the second period, but at a lower first period premium. The gain in welfare from the lower first period premium must be compared to the loss in welfare from the reduced consumption in second period non-bequest states. Since marginal utility is higher in the first period, this results in an overall welfare gain. Thus, for any feasible contract in the world with a settlement market, we can construct a contract in the world without a settlement market that results in higher *ex ante* expected utility.

Proposition 4 formalizes an intuitive argument provided in Proposition 2 of Daily et al. (2008). They argued that the settlement market effectively transfers resources from period 1 when income is low to period 2 when income is high. Such transfers, due to the concavity of the utility function, are welfare reducing. The informal argument provided in their paper hinges on the hypothesis that the equilibrium first period premium is higher with a settlement market than without. This hypothesis does not hold in general. An extreme example of when the hypothesis fails is provided in Proposition 3. When q is sufficiently large and there is a settlement market, the insurance market can only offer spot contracts, which implies that the first period premium is $Q_1^s = Q^{FI}$, the actuarially fair premium. In contrast, if q is large but g is very small, reclassification risk insurance is still offered when there is no settlement market. In this case, $Q_1 > Q^{FI}$ due to front loading. Thus, for sufficiently high q and small g, the equilibrium first period premium is actually *lower* in the presence of a settlement market. Nevertheless, consumer welfare is still reduced because of the unraveling of dynamic contracts.

Many would consider the emergence of a settlement market as a form of market completion (e.g. Doherty and Singer (2003)). After all, consumers who lose their bequest motives in period 2 can share the surplus in the actuarial value of their policy with the settlement firm, something they could not do when there is no settlement market. So at a first glance, the welfare result in Proposition 4 is somewhat surprising. However, from Lipsey and Lancaster (1956), we know that once we depart from complete markets, the second best solution may not be the one with the least degree of market incompleteness.[20] In our context, market incompleteness due to lack of commitment power and an inability to contract on bequest motives exists regardless of the settlement market. Therefore, moving towards "more completeness" by introducing the settlement market is not necessarily second best.

Another way to think of the welfare result is that the settlement market weakens the consumer's ability to commit to not asking for a return of their front loaded premiums in the event that they lose their bequest motive. Without a life settlement market, the life insurance company is the monopsonist buyer of surrendered or lapsed policies, and can commit fully to *not* buy them, even if it is in their best *ex post* interest to do so. With the introduction of a settlement market, the commitment power of the monopsonist, which was earlier enforcing commitment among consumers, becomes eroded.[21]

4 Life Insurance and Life Settlements with Endogenous cash surrender values

4.1 Health-contingent cash surrender values

So far, we have analyzed equilibrium life insurance contracts in an environment without a settlement market and with endogenously chosen cash surrender values, and in an environment with a settlement market but without cash surrender values. One can loosely think of the former as a model of life insurance markets before the innovation of the life settlement markets, and of

[20]For another example, Levin (2001) showed that in an Akerlof lemons model, greater information asymmetries between buyers and sellers do not necessarily reduce the equilibrium gains from trade.

[21]These are statements concerning the new equilibrium after the primary insurers respond to the life settlement firms. In a more dynamic setting, initial cohorts whose policies were purchased before the existence of a secondary market (and thus have full reclassification risk insurance), can benefit from the newly added option to settle, while their insurers will likely suffer.

the latter as a model of life insurance markets in their current state: with a settlement market but without endogenous cash surrender values. We now turn our attention to the case in which there is a settlement market, but cash surrender values are also endogenously chosen and allowed to be health-contingent.

Contracts are now of the form $\langle (Q_1^{ss}, F_1^{ss}), (Q_2^{ss}(p_2), F_2^{ss}(p_2), S_2^{ss}(p_2)) : p_2 \in (0,1) \rangle$, and chosen to solve the following:[22]

$$\max \ u(y - g - Q_1^{ss}) + p_1 v(F_1^{ss}) \tag{10a}$$

$$+ (1 - p_1) \int \left\{ (1-q) \begin{bmatrix} u(y + g - Q_2^{ss}(p_2)) \\ +p_2 v(F_2^{ss}(p_2)) \end{bmatrix} + qu(y + g + S_2^{ss}(p_2)) \right\} d\Phi(p_2)$$

$$\text{s.t. } Q_1^{ss} - p_1 F_1^{ss} + (1 - p_1) \int \{(1-q)\left[Q_2^{ss}(p_2) - p_2 F_2^{ss}(p_2)\right] - qS_2^{ss}(p_2)\} \, d\Phi(p_2) = 0, \tag{10b}$$

$$p_2 F_2^{ss}(p_2) - Q_2^{ss}(p_2) \geq 0 \text{ for all } p_2, \tag{10c}$$

$$S_2^{ss}(p_2) - \beta V_2^{ss}(p_2) \geq 0 \text{ for all } p_2 \tag{10d}$$

We have assumed that in equilibrium, the cash surrender value is always chosen to be at least as great as $\beta V_2^{ss}(p_2)$, the amount that could be obtained on the settlement market. This is an innocuous assumption. The insurance company will never set the cash surrender value to be lower than what could be obtained on the settlement market because by offering just an ϵ more, the insurance company can repurchase the policy for $\beta V_2^{ss}(p_2) + \epsilon$. This is preferable to letting the policy be sold to the settlement market, in which case the insurance company is liable for $V_2^{ss}(p_2)$.

Proposition 5. *In the presence of a settlement market, equilibrium health-contingent cash surrender values will equal the amount that can be obtained from the settlement market:* $S_2^{ss}(p_2) = \beta V_2^{ss}(p_2)$.

Proposition 5 shows that if insurance companies are allowed to compete with settlement firms by offering health-contingent cash surrender values, they will do so in a way so as to just barely undercut the settlement firms. This is not surprising given our results from Sections 2 and 3. When $S_2^{ss}(p_2) = \beta V_2^{ss}(p_2)$, policyholders are indifferent between surrendering the contract and selling it on the settlement market. If $\beta < 1$, however, the life insurance companies clearly benefit from a surrender rather than a sale. Therefore, by surrendering their contracts to the insurance companies

[22]The subscript ss is chosen to denote the equilibrium contract with endogenous cash surrender values in the presence of a life settlement market.

instead of selling them, policyholders can obtain any second period allocation for a lower first period premium.[23] When a settlement market exists, policyholders are clearly better off when they are allowed to endogenously choose health-contingent cash surrender values.

Proposition 6. *When a settlement market exists, consumer welfare is higher (strictly higher if* $\beta < 1$*) when life insurance companies can offer health-contingent cash surrender values.*

Although Proposition 6 shows that endogenously chosen, health-contingent cash surrender values can reduce the welfare loss caused by the settlement market, the welfare loss cannot be eliminated completely. The reason for this is quite clear. As shown in Proposition 1, consumers would ideally set the cash surrender values to zero, but the presence of a settlement market makes it so that such a commitment is not possible. The presence of a settlement market forces the credible lower bound of the cash surrender value up to β times the actuarial value of the policy. The welfare gain attributed to cash surrender values comes only through the increased efficiency of surrender as opposed to sale on the settlement market. We summarize this as:

Proposition 7. *Ex ante consumer welfare is lower when there is a settlement market than when there is no settlement market.*

4.2 Non-health contingent cash surrender values

We now consider equilibrium life insurance contracts when the cash surrender values can be chosen endogenously, but are restricted *not* to depend on health status. Whether due to regulations or other reasons, this seems to be the most relevant case to real world life insurance markets. We already know that health contingency of cash surrender values is irrelevant when there is no settlement market. But if there is a settlement market, will a non-health-contingent cash surrender value help mitigate the welfare loss caused by the settlement market?

Contracts are now of the form $\langle (Q_1, F_1, S), (Q_2(p_2), F_2(p_2)) : p_2 \in (0,1) \rangle$, and are chosen to

[23] This also implies that, in equilibrium, no policyholder will actually participate on the settlement market. Nevertheless, the threat of the settlement market is itself enough to place a lower bound on the chosen cash surrender values.

solve:[24]

$$
\begin{aligned}
\max\ & u(y-g-Q_1)+p_1 v(F_1)+(1-p_1)(1-q)\int_0^1 \left[u(y+g-Q_2(p_2))+p_2 v(F_2(p_2))\right]d\Phi(p_2) \\
& +(1-p_1)q\int_{S\geq \beta V_2(p_2)} u(y+g+S)d\Phi(p_2) \qquad (11a)\\
& +(1-p_1)q\int_{S<\beta V_2(p_2)} u(y+g+\beta V_2(p_2))d\Phi(p_2) \\
\text{s.t. } & V_2(p_2)\geq S,\ \text{for all } p_2, \qquad (11b)\\
& S\geq 0, \qquad (11c)\\
& Q_1-p_1F_1=(1-p_1)(1-q)\int_0^1 V_2(p_2)d\Phi(p_2)+(1-p_1)q\int_{S\geq\beta V_2(p_2)} S d\Phi(p_2) \\
& +(1-p_1)q\int_{S<\beta V_2(p_2)} V_2(p_2)d\Phi(p_2) \qquad (11d)
\end{aligned}
$$

To understand the above problem, let us first explain the constraints. Constraint (11b) is the analog of constraints (10d) and (3c). These constraints require the actuarial value of the contract terms for any period 2 health state to be at least equal to the cash surrender value. As before, this requirement reflects the consumer's inability to commit. If the actuarial value of the contract was less than the cash surrender value, the consumer would surrender the contract and repurchase better insurance on the spot market. But the outcome of such an action can be replicated by an appropriate choice of contract terms in period 1, so we simply assume that contract terms are chosen such that these actions do not occur for individuals with a bequest motive. Constraint (11c) simply requires that the cash value be non negative, because the consumer cannot commit to paying the insurance company when surrendering the policy. Constraint (11d) is the zero profit condition reflecting perfect competition in the market. The first integral in the right hand side of (11d) is the expected loss the insurance company suffers from consumers who retain their bequest motive. The second integral in the RHS of (11d) is the expected loss the insurance company suffers from consumers who lose their bequest motive and find it optimal to surrender the policy back to the original insurer. The third integral in the RHS of (11d) is the expected loss the insurer suffers from consumers who lose their bequest motive but find it optimal to sell the policy on the settlement market.

Now let us explain the objective function. The first integral in (11a) is the expected second period utility to consumers with a bequest motive, for whom constraint (11b) ensures that they

[24] We omit superscripts here both for notational clarity and because the equilibrium outcome will be equivalent to the case with a settlement market and no cash surrender values.

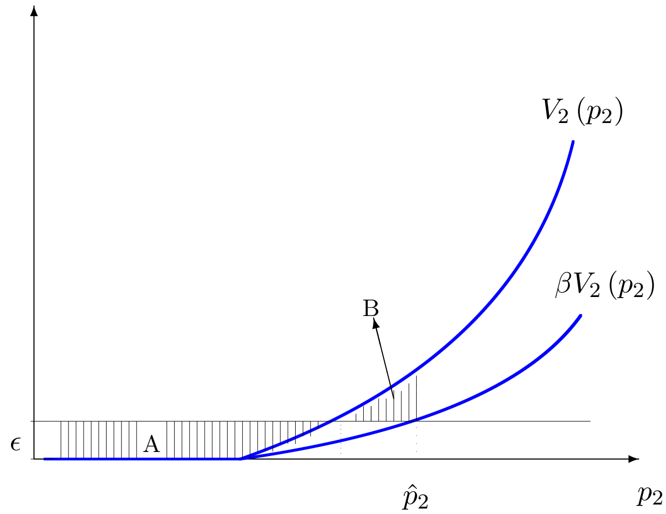

Figure 4: The Effect of Increasing S by $\epsilon > 0$ on Primary Insurer's Period-2 Profit

remain with the original contract terms. The second integral is the expected second period utility for consumers who lose their bequest motive and find it optimal to surrender their contract back to the insurance company. The third integral is the expected second period utility for consumers who lose their bequest motive and find it optimal to sell their contract on the settlement market.

Note that problem (11) is substantially more complicated than problem (10) because now the policyholders who lose their bequest motive must choose whether to sell the policy on the settlement market or surrender the policy back to the insurer. However, using a rather intuitive perturbation argument, we can prove the following result:

Proposition 8. *In the presence of a settlement market, if cash surrender values are not allowed to be health contingent, then the equilibrium contract will not contain a positive cash surrender value.*

To understand the intuition for Proposition 8, it is useful to consider the effect of raising S from 0 to ϵ on the firm's second period profits. In Figure 4, the curve labeled $V_2(p_2)$ depicts the period 2 actuarial value with respect to health state p_2, under the equilibrium of Section 3.1. The curve labeled $\beta V_2(p_2)$ depicts the settlement firm's payment with respect to p_2. If the primary insurer raises the non-health-contingent cash surrender value from 0 to ϵ, policyholders with period 2 health in regions A and B, and who no longer have a bequest motive, will surrender their policies to the primary insurer for a payment of $\epsilon > \beta V_2(p_2)$. The area labeled A captures the loss in profits from the change, since the firm will be paying these consumers ϵ after the change, whereas before the change they were paying $V_2(p_2) < \epsilon$. The area labeled B captures the gain in profits from

		Cash cash surrender value				
		None		Non-Health Contingent		Health Contingent
Settlement Market	No	A	$=$	A	$=$	A
						$\vee$
	Yes	C	$=$	C	$\leq$	B

Note: Inequality is strict when $\beta < 1$

Table 1: Comparison of Consumer Welfare Across Market Regimes: A Summary

the change, since the firm was paying $V_2(p_2)$ before the change, but they are paying $\epsilon < V_2(p_2)$ after the change. As is clear from the graph, area A is first order proportional to ϵ, while area B is second order proportional to ϵ. When ϵ is small, the firm's second period losses increase as a result of increasing S from 0 to ϵ. In order to maintain zero profit, the insurance company has to increase the first period premium Q_1. It is easy to see that the utility cost for the consumer of increasing the first period premium is $u'(y - g - Q_1)$. It turns out that the utility gain for the consumer when S increases from 0 to ϵ is captured by $(1 - p_1)qu'(y + g)\Phi(\hat{p}_2)$ where $\hat{p}_2$ is defined by $V_2(\hat{p}_2) = \epsilon$. The marginal utility gain in the second period is thus smaller than the marginal loss caused by the increase in the first period premium, so the tradeoff is welfare reducing. Similar perturbation arguments can be used to show that marginally decreasing S from any positive level is always welfare improving. Thus, the optimal $S^* = 0$. Proposition 8 tells us that when insurance companies are offering cash surrender values that do not depend on health, then such an option is essentially useless as a response to the settlement market. Thus, the consumer welfare is the same as if cash surrender values were not specified at all.

Table 1 summarizes our welfare results for the various regimes we have so far analyzed. We see that the introduction of a settlement market is unambiguously welfare reducing within the context of our setup. We also find that without a settlement market, cash surrender values will not be used at all, regardless of whether they can be made contingent on health. When a settlement market is present, however, the health contingency of cash surrender values becomes very important. In particular, if cash surrender values are allowed to be health contingent, then they can be chosen in such a way as to reduce the welfare loss caused by the settlement market. If they are not allowed to be health-contingent, then they will not be used at all.

5 Diminishing Bequest Motives

So far, we have considered the case in which policyholders, with some probability, may *completely* lose bequest motive in the second period. In this section, we consider how the contracting problem changes when bequest motive *diminishes* instead of disappears. Specifically, we assume that, with probability $q \in (0,1)$, a policyholder's bequest motive goes from $v(F)$ to $\delta v(F)$, where $0 < \delta < 1$. We assume that whether a policyholder's bequest motive experiences a reduction is his/her private information. By the Revelation Principle, the insurance company designs a contract contingent on bequest motive, but the contract is subject to incentive compatibility constraints. The contract becomes $\langle Q_1, F_1, Q_2(p_2), F_2(p_2), Q_2'(p_2), F_2'(p_2)\rangle$ where $Q_2'(p_2)$ and $F_2'(p_2)$ are the premium and death benefit when bequest motive is diminished. The optimal contracting problem is:

$$\max \; u(y-g-Q_1) + p_1 v(F_1) + (1-p_1)\int \left\{ (1-q)\begin{bmatrix} u(y+g-Q_2(p_2)) \\ +p_2 v(F_2(p_2)) \end{bmatrix} + q \begin{bmatrix} u(y+g-Q_2'(p_2)) \\ +p_2 \delta v(F_2'(p_2)) \end{bmatrix} \right\} d\Phi(p_2) \tag{12}$$

$$\text{s.t. } Q_1 - p_1 F_1 + (1-p_1)\int \left\{ (1-q)\begin{bmatrix} Q_2(p_2) \\ -p_2 F_2(p_2) \end{bmatrix} + q \begin{bmatrix} Q_2'(p_2) \\ -p_2 F_2'(p_2) \end{bmatrix} \right\} d\Phi(p_2) = 0 \tag{13a}$$

$$V_2(p_2) \equiv p_2 F_2(p_2) - Q_2(p_2) \geq 0 \text{ for all } p_2 \tag{13b}$$

$$V_2'(p_2) \equiv p_2 F_2'(p_2) - Q_2'(p_2) \geq 0 \text{ for all } p_2 \tag{13c}$$

$$u(y+g-Q_2(p_2)) + p_2 v(F_2(p_2)) \geq u(y+g-Q_2'(p_2)) + p_2 v(F_2'(p_2)) \text{ for all } p_2 \tag{13d}$$

$$u(y+g-Q_2'(p_2)) + p_2 \delta v(F_2'(p_2)) \geq u(y+g-Q_2(p_2)) + p_2 \delta v(F_2(p_2)) \text{ for all } p_2, \tag{13e}$$

where constraint (13a) is the zero profit constraint; constraints (13b) and (13c) are the no-lapsation constraints as discussed in Section 2; and constraints (13d) and (13e) are the incentive compatibility constraints for the types of policyholders whose bequest motive maintains and diminishes, respectively.

Implementation via health-contingent surrender values

It is difficult to characterize the equilibrium contract for the problem described in (12). However, we are able to show that the equilibrium contract can be implemented by a standard contract with a health-contingent surrender value. Intuitively, what happens is that when bequest motive changes, the policyholder will surrender the contract for its surrender value and repurchase on the spot market, thus allowing the policyholder to downsize the policy while still retaining positive actuarial value and thus reclassification risk insurance. Formally, we have the following result:

Proposition 9. *Suppose that the equilibrium contract that solves Problem (12) is given by* $\langle Q_1, F_1, Q_2(p_2), F_2(p_2), Q_2'(p_2), F_2'(p_2)\rangle$. *It can be implemented by a standard non-bequest contingent contract with a health-contingent cash surrender value,* $\langle \hat{Q}_1, \hat{F}_1, \hat{Q}_2(p_2), \hat{F}_2(p_2), \hat{S}_2(p_2)\rangle$, *specified by:*

$$\langle \hat{Q}_1, \hat{F}_1\rangle = \langle Q_1, F_1\rangle$$

$$\langle \hat{Q}_2(p_2), \hat{F}_2(p_2)\rangle = \begin{cases} \langle Q_2(p_2), F_2(p_2)\rangle & \textit{if } V_2(p_2) \geq V_2'(p_2) \\ \langle Q_2'(p_2), F_2'(p_2)\rangle & \textit{otherwise} \end{cases}$$

$$\hat{S}_2(p_2) = \min\{V_2(p_2), V_2'(p_2)\}$$

where $V_2(p_2), V_2'(p_2$ *are defined in (13b) and (13c) respectively.*

Welfare effects of a settlement market

The contracting problem above assumed that there was no life settlement market. Now consider the effect of a settlement market, and for simplicity assume the settlement market is efficient ($\beta = 1$). The settlement market will reduce *ex ante* consumer welfare because it makes the incentive compatibility constraints more binding. In particular, the objective function remains the same but we now have the additional constraints:

$$V_2(p_2) = V_2'(p_2) \tag{14}$$

$$Q_2(p_2), F_2(p_2) = \arg\max_{Q,F} u(y + g - Q) + p_2 v(F) \text{ s.t. } p_2 F - Q = V_2(p_2) \tag{15}$$

$$Q_2'(p_2), F_2'(p_2) = \arg\max_{Q,F} u(y + g - Q) + p_2 \delta v(F) \text{ s.t. } p_2 F - Q = V_2'(p_2) \tag{16}$$

If constraint (14) were violated, the policyholder would simply pretend to be whatever bequest state had higher actuarial value and settle the contract in the secondary market. If constraint (15) or (16) were violated, then the policyholder would similarly settle the contract and then re-optimize the choice of premium and face amount. Since the introduction of a settlement market simply adds constraints to the optimization problem, it reduces *ex ante* consumer welfare.

It is useful to note that, from Proposition 9, implementation of the optimal equilibrium contract requires health-contingent cash surrender values. If health-contingency for surrender values is not possible, for one of the various reasons we discuss in the paper, then the optimal contract would have to be implemented directly through a menu of premiums and death benefits. If this is not feasible, i.e. due to menu costs, then the optimal contract cannot be implemented, and there may be some scope for the life settlement market to be welfare improving. We see that as a potentially interesting extension of the model, but we leave it to future research.

6 Conclusion

In this paper we began by investigating why life insurance policies in practice do not contain cash surrender values that are reflective of the policies' actuarial value. We show that competitive life insurance companies would choose in equilibrium *not* to include positive cash surrender values because doing so would exacerbate a dynamic commitment problem. Policyholders would want to commit not to exercise the cash surrender value, since it occurs in a state of relatively low marginal utility. But when the second period arrives, they cannot credibly commit to not exercise it.

We then extend our model to study the effects of a life settlement market on the structure of life insurance contracts and on consumer welfare. We replicate and expand on the results in Daily et al. (2008), finding that the introduction of a life settlement market changes the nature through which reclassification risk insurance is provided. In particular, we find that the equilibrium contract in the presence of a settlement market is health-contingent, and cannot be replicated by any contracts that are not health-contingent (unlike the case in which there is no settlement market). We also find that in the context of our model, the settlement market generically leads to lower *ex ante* consumer welfare. In the most extreme case, the presence of a settlement market can unravel the market for dynamic contracts to a sequence of spot contracts with no insurance against reclassification risk at all.

We also examine how endogenously chosen cash surrender values can play a role in the life insurers' response to the settlement market. We show that allowing for health-contingent cash surrender values improves consumer welfare, but consumers are still worse off than if there were no settlement market. We also show that if cash surrender values are not allowed to depend on health, then they will again be chosen to be zero, even in the presence of a settlement market. This surprising result has policy relevance, because to our knowledge, current life insurance markets do not offer any policies with health-contingent cash surrender values, whether due to regulations or other reasons.

Finally, we show that in environments where policyholders' bequest motives may diminish instead of disappear, health-contingent cash surrender value may also play an important role in implementing the optimal equilibrium contract using standard non-bequest contingent contracts.

Taking the above results into account, our research suggests that life insurance markets may respond to the life settlement market through a greater provision of health-contingent contracts, whether through health-contingent premiums or cash surrender values, and such contracts should become more popular as the life settlement market grows.

References

Daily, Glenn, Igal Hendel, and Alessandro Lizzeri, "Does the Secondary Life Insurance Market Threaten Dynamic Insurance?," *American Economic Review Papers and Proceedings*, 2008, *98 (2)*, 151–156.

Deloitte, "The Life Settlement Market: An Actuarial Perspective on Consumer Economic Value," 2005.

Doherty, Neil A. and Hal J. Singer, "The Benefits of a Secondary Market for Life Insurance Policies," *Real Property, Probate and Trust Journal*, 2003, *38* (3), 449–478.

Fang, Hanming and Edward Kung, "How Does Life Settlement Affect the Primary Life Insurance Market?," NBER Working Paper 25761, 2010.

_ and _ , "Why Do Life Insurance Policyholders Lapse? Liquidity Shocks vs. Loss of Bequest Motives," Working Paper, University of Pennsylvania and UCLA, 2010.

_ and Zenan Wu, "Life Insurance and Life Settlement Markets with Overconfident Policyholders," Working Paper, University of Pennsylvania and Peking University, 2017.

Gatzert, Nadine, "The Secondary Market for Life Insurance in the United Kingdom, Germany, and the United States: Comparison and Overview," *Risk Management and Insurance Review*, 2010, *13* (2), 279–301.

_ , Gudrun Hoermann, and Hato Schmeiser, "The Impact of the Secondary Market on Life Insurers' Surrender Profits," *Journal of Risk and Insurance*, 2009, *76* (4), 887–908.

Gilbert, Jersey and Ellen Schultz, *Consumer Reports Life Insurance Handbook*, Consumer Reports Books: Yonkers, NY, 1994.

Gottlieb, Daniel and Kent Smetters, "Lapse-Based Insurance," Working Paper, Olin Business School, Washington University in St. Louis and Wharton School, University of Pennsylvania, 2016.

Hendel, Igal and Alessandro Lizzeri, "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics*, 2003, *118 (1)*, 299–327.

Januário, Alfonso V. and Narayan Y. Naik, "Testing for adverse selection in life settlements: The secondary market for life insurance policies," Working Paper, London Business School, 2014.

Levin, Jonathan, "Information and the Market for Lemons," *The RAND Journal of Economics*, 2001, *32* (4), 657–666.

Life Insurance Settlement Association, "Data Collection Report, 2004-2005," 2006.

Lipsey, Richard G and Kevin Lancaster, "The General Theory of Second Best," *The Review of Economic Studies*, 1956, *24* (1), 11–32.

Singer, Hal J. and Eric Stallard, "Reply to 'The Life Settlement Market: An Actuarial Perspective on Consumer Economic Value'," Criterion Economics L.L.C., 2005.

Zhu, Nan and Daniel Bauer, "Coherent Pricing of Life Settlements Under Asymmetric Information," *Journal of Risk and Insurance*, 2013, *80* (3), 827–851.

Appendix: Proofs

Proof of Proposition 1.

Proof. See main text. □

Proof of Proposition 2.

Proof. The statement of the optimization problem is given by the objective function (3a), and constraints (3b) and (3c), with $S_2(p_2) = 0$ for all p_2. Part 1 of the Proposition follows directly from the first order conditions (4a)-(4d).

To prove part 2 of the Proposition, let $\mathcal{B}$ be the health states for which constraint (3c) binds and let $\mathcal{NB}$ be the health states for which constraint (3c) does not bind. We first show that if $p_2 \in \mathcal{B}$ and $p_2' \in \mathcal{NB}$, then $p_2 < p_2'$ and $Q_2(p_2) \leq Q_2(p_2')$. Complementary slackness conditions require that $\lambda(p_2) \geq 0$ and $\lambda(p_2') = 0$. The first order conditions then imply that:

$$u'\left(y + g - Q_2(p_2)\right) = \mu - \frac{\lambda(p_2)}{(1-p_1)q\phi(p_2)} \leq u'\left(y + g - Q_2(p_2')\right) = \mu$$

Since u' is decreasing, it must be that $Q_2(p_2) \leq Q_2(p_2')$. The conditions in (5)-(6) also imply that $F_2(p_2) \geq F_2(p_2')$.

To show that $p_2 < p_2'$, suppose the contrary. Since $p_2' \in \mathcal{NB}$ implies that $Q_2(p_2') < p_2' F_2(p_2')$. Therefore:

$$Q_2(p_2) \leq Q_2(p_2') < p_2' F_2(p_2') \leq p_2 F_2(p_2)$$

where the last inequality follows from the postulated $p_2 \geq p_2'$ and the fact that $F_2(p_2) \geq F_2(p_2')$. Thus, $Q_2(p_2) < p_2 F_2(p_2)$, which contradicts $p_2 \in \mathcal{B}$.

We have shown that if $p_2 \in \mathcal{B}$ and $p_2' \in \mathcal{NB}$, then $p_2 < p_2'$. This immediately implies the existence of a p_2^* such that premiums are actuarially fair if $p_2 \leq p_2^*$, and actuarially favorable if $p_2 > p_2^*$. To see that $Q_2(p_2)$ is constant for $p_2 > p_2^*$, simply note that $\lambda(p_2) = 0$ for $p_2 \in \mathcal{NB}$. Therefore, by the first order condition (4c), $u'(y + g - Q_2(p_2)) = \mu$, and so $Q_2(p_2)$ is constant in p_2.

Finally, to prove part 3 of the Proposition, suppose that $p_2^* = 1$, so that premiums are actuarially fair for all second period health states. The first order conditions imply that $u'\left(y + g - Q_2(p_2)\right) \leq u'\left(y - g - Q_2\right)$ for all p_2, and thus $Q_2(p_2) \leq Q_1 + 2g$ for all p_2. However, for any $p_2 > p_1$, it must be the case that $Q_2(p_2) > Q_1$, because premiums are actuarially fair. Thus, when g is sufficiently small, it is impossible to have $Q_2(p_2) \leq Q_1 + 2g$ for all p_2. □

Proof of Proposition 3.

Proof. The first order conditions of the optimization problem are the following:

$$u'(y - g - Q_1^s) = \mu \tag{17}$$

$$v'(F_1^s) = \mu \tag{18}$$

$$(1-q)\,u'(y + g - Q_2^s(p_2)) + \beta q u'(y + g + \beta V_2^s(p_2)) = \mu - \frac{\lambda(p_2)}{(1-p_1)\phi(p_2)} \tag{19}$$

$$(1-q)\,v'(F_2^s(p_2)) + \beta q u'(y + g + \beta V_2^s(p_2)) = \mu - \frac{\lambda(p_2)}{(1-p_1)\phi(p_2)} \tag{20}$$

where $\mu > 0$ is the Lagrange multiplier of constraint (7b) and $\lambda(p_2) \geq 0$ is the Lagrange multiplier of constraint (7c). Part 1 of the Proposition follows directly from these first order conditions.

The proof for the existence of a threshold p_2^{s*} follows the same steps as in the proof of Proposition 2. To see that $Q_2^s(p_2)$ is increasing in p_2 for $p_2 > p_2^{s*}$, we first rewrite the first order conditions in the following manner:

$$\begin{aligned}
&(1-q)\,u'(y + g - Q_2^s(p_2)) + \beta q u'(y + g + \beta V_2^s(p_2)) = u'(y - g - Q_1^s),\\
&v'(F_2^s(p_2)) = u'(y + g - Q_2^s(p_2)),\\
&V_2^s(p_2) = p_2 F_2^s(p_2) - Q_2^s(p_2).
\end{aligned}$$

Taking derivatives with respect to p_2 for each equation, we obtain:

$$\begin{aligned}
&(1-q)\,u''(y + g - Q_2^s(p_2))\frac{dQ_2^s}{dp_2} = \beta^2 q u''(y + g + \beta V_2^s(p_2))\frac{dV_2^s}{dp_2}\\
&v''(F_2^s(p_2))\frac{dF_2^s}{dp_2} = -u''(y + g - Q_2^s(p_2))\frac{dQ_2^s}{dp_2}\\
&\frac{dV_2^s}{dp_2} = F_2^s(p_2) + p_2\frac{dF_2^s}{dp_2} - \frac{dQ_2^s}{dp_2}
\end{aligned}$$

Solving for dQ_2^s/dp_2, we obtain:

$$\frac{dQ_2^s}{dp_2} = \frac{F_2^s(p_2)}{\frac{(1-q)u''(y+g-Q_2^s(p_2))}{\beta^2 q u''(y+g+\beta V_2^s(p_2))} + \left[1 + p_2\frac{u''(y+g-Q_2^s(p_2))}{v''(F_2^s(p_2))}\right]},$$

which is strictly positive if $q > 0$.

Finally, we prove part 3 of the Proposition, the potential for unraveling. Once again, let $\mathcal{NB}^s$ denote the set of health states for which constraint (7c) is non-binding. If $\mathcal{NB}^s$ is not empty, then for any $p_2 \in \mathcal{NB}^s$ the contract terms must satisfy:

$$(1-q)\,u'(y+g-Q_2^s(p_2)) + \beta q u'(y+g+\beta V_2^s(p_2)) = u'(y-g-Q_1^s), \tag{21}$$

which can be rewritten as:

$$(1-q)\left[u'(y+g-Q_2^s(p_2)) - \beta u'(y+g+\beta V_2^s(p_2))\right] = u'(y-g-Q_1^s) - \beta u'(y+g+\beta V_2^s(p_2)). \tag{22}$$

First note that the zero profit condition (7b) implies that if $\mathcal{NB}^s$ is non empty, then we must have $Q_1^s \geq Q_1^{FI}$, where Q_1^{FI} was defined in the text as the level of premium that solves:

$$u'\left(y-g-Q_1^{FI}\right) = v'\left(F_1^{FI}\right)$$
$$p_1 F_1^{FI} - Q_1^{FI} = 0$$

Notice that Q_1^{FI} does not depend on q, but is decreasing in g. Let $\bar{g}$ be the upper bound of the values that g can take, and let $\underline{Q}_1^{FI}$ denote the actuarially fair premium at $g = \bar{g}$. Therefore the right hand side (RHS) of (22) is bounded below, for any $g > 0$, by:

$$RHS > u'(y-\underline{Q}_1^{FI}) - \beta u'(y).$$

Now examine the left hand side (LHS) of (22). We will consider two cases. For the first case, suppose that $\lim_{x\to 0} u'(x) \equiv u'(0) < \infty$. Because $Q_2^s(p_2)$ is always smaller than $y+g$ in equilibrium, we have that

$$LHS = (1-q)\left[u'(y+g-Q_2^s(p)) - \beta u'(y+g+\beta V_2^s(p))\right] < (1-q)u'(0).$$

Thus if

$$q > \hat{q} \equiv 1 - \frac{u'(y-\underline{Q}_1^{FI}) - \beta u'(y)}{u'(0)}$$

then the LHS of (22) will always be smaller than the RHS. Thus, equation (21) can never be satisfied for any p_2, and $\mathcal{NB}^s$ must be empty.

For the second case, suppose that $\lim_{x\to 0} u'(x) = \infty$. Since $p_2 \in \mathcal{NB}^s$, we have that $p_2 F_2^s(p_2) -$

$Q_2(p_2) > 0$. This implies that:

$$u'\left(y + g - Q_2^s(p_2)\right) < v'\left(\frac{Q_2^s(p_2)}{p_2}\right) \tag{23}$$

Notice that the LHS of (23) is increasing as $Q_2^s(p_2)$ varies from 0 to $y+g$, and the RHS is decreasing in $Q_2^s(p_2)$ over the same interval. If $u'(y+g) \geq v'(0)$ then (23) cannot be satisfied for any value of $Q_2^s(p_2)$, so $\mathcal{NB}^s$ must be empty. So let us consider the case in which $u'(y+g) < v'(0)$, so that at $Q_2^s(p_2) = 0$, the LHS of (23) is less than the RHS. Since $u'(0) = \infty$, we know that at $Q_2^s(p_2) = y + g$, the LHS of (23) is greater than the RHS. Because the LHS is continuous and monotonically increasing in $Q_2^s(p_2)$ and the RHS is continuous and monotonically decreasing in $Q_2^s(p_2)$, there must exist some x such that the LHS and the RHS are equal to each other at $Q_2^s(p_2) = x$. For each p_2 and g let $x(p_2; g)$ denote this quantity. $Q_2^s(p_2)$ must be bounded above by $x(p_2; g)$. Now let us write $x(g) \equiv \sup_{p_2 \in \mathcal{NB}^s} x(p_2; g)$ and $\bar{u}' \equiv \max_g u'\left(y + g - x(g)\right)$. We hence have:

$$\begin{aligned}
LHS &= (1-q)\left[u'(y+g-Q_2^s(p)) - \beta u'(y+g+\beta V_2^s(p))\right] \\
&< (1-q)u'\left(y+g-Q_2^s(p_2)\right) \\
&\leq (1-q)u'(y+g-x(p_2;g)) \\
&\leq (1-q)u'(y+g-x(g)) \\
&\leq (1-q)\bar{u}'
\end{aligned}$$

Thus, if

$$q > \hat{q} \equiv 1 - \frac{u'\left(y - \underline{Q}_1^{FI}\right) - \beta u'(y)}{\bar{u}'}$$

then the LHS of (22) will always be smaller than the RHS and (21) can never be satisfied for any p_2. So $\mathcal{NB}^s$ must be empty. □

Proof of Proposition 4.

Proof. We will show that for any feasible contract of problem (7), we can construct a feasible contract for problem (3) that makes the consumers better off *ex ante*.

Let $C^s = \langle (Q_1^s, F_1^s), \{(Q_2^s(p_2), F_2^s(p_2)) : p_2 \in [0,1]\}\rangle$ be a feasible contract for problem (7) when there is a settlement market. Thus, $Q_1^s - p_1 F_1^s = (1-p_1)\int V_2^s(p_2)d\Phi(p_2)$, where $V_2^s(p_2) \equiv p_2 F_2^s(p_2) - Q_2^s(p_2)$.

Now consider a contract $\hat{C} \equiv \left\langle (\hat{Q}_1, F_1^s), (Q_2^s(p_2), F_2^s(p_2)) : p_2 \in [0,1] \right\rangle$ where $\hat{Q}_1$ is given by:

$$\hat{Q}_1 - p_1 F_1^s = (1-p_1)(1-q)\int V_2^s(p_2)d\Phi(p_2).$$

Since $q \in (0,1)$, we know that $\hat{Q}_1 < Q_1^s$. That is, $\hat{C}$ is exactly the same contract as C^s except that the first period premium is decreased from Q_1^s until the zero profit condition for the no-settlement-market case (3b) holds. It is easy to show that $\hat{C}$ is a feasible contract for problem (3).

We will now show that $\hat{C}$ in a world without settlement market is better than C^s in a world with settlement market. To see this, let

$$\begin{aligned} W^s(C^s) = {} & p_1 v(F_1^s) + u(y-g-Q_1^s) + \\ & (1-p_1)\int \left\{ \begin{array}{c} (1-q)\left[p_2 v(F_2^s(p_2)) + u(y+g-Q_2^s(p_2))\right] \\ +qu(y+g+\beta V_2^s(p_2)) \end{array} \right\} d\Phi(p_2) \end{aligned}$$

denote the expected consumer welfare associated with contract C^s in a world with the settlement market. Let

$$\begin{aligned} W\left(\hat{C}\right) = {} & p_1 v(F_1^s) + u(y-g-\hat{Q}_1) + \\ & (1-p_1)\int \left\{ \begin{array}{c} (1-q)\left[p_2 v(F_2^s(p_2)) + u(y+g-Q_2^s(p_2))\right] \\ +qu(y+g) \end{array} \right\} d\Phi(p_2) \end{aligned}$$

denote the expected consumer welfare associated with contract $\hat{C}$ in a world without the settlement market. Note that

$$\begin{aligned} W\left(\hat{C}\right) - W^s(C^s) &= u(y-g-\hat{Q}_1) - u(y-g-Q_1^s) \\ &\quad - (1-p_1)q\int \left[u(y+g+\beta V_2^s(p_2)) - u(y+g)\right] d\Phi(p_2) \\ &\geq u(y-g-\hat{Q}_1) - u(y-g-Q_1^s) \\ &\quad - (1-p_1)q\left[u\left(y+g+\beta\int V_2^s(p_2)d\Phi(p_2)\right) - u(y+g)\right] \end{aligned}$$

where the inequality follows from Jensen's inequality due to the concavity of $u(\cdot)$. Further note that:

$$
\begin{aligned}
& q\left[u\left(y+g+\beta\int V_2^s(p_2)d\Phi\left(p_2\right)\right)-u\left(y+g\right)\right] \\
& = qu\left(y+g+\beta\int V_2^s(p_2)d\Phi\left(p_2\right)\right)+\left(1-q\right)u(y+g)-u(y+g) \\
& \leq u\left(y+g+\beta q\int V_2^s(p)d\Phi\left(p\right)\right)-u(y+g),
\end{aligned}
$$

where again the inequality follows from Jensen's inequality. Thus,

$$
\begin{aligned}
W\left(\hat{C}\right)-W^s\left(C^s\right) \geq\ & u(y-g-\hat{Q}_1)-u(y-g-Q_1^s) \\
& -\left(1-p_1\right)\left[u\left(y+g+\beta q\int V_2^s(p_2)d\Phi\left(p_2\right)\right)-u(y+g)\right]
\end{aligned}
$$

First note that $Q_1^s-\hat{Q}_1=(1-p_1)q\int V_2^s(p_2)d\Phi\left(p_2\right)$. By the continuous function theorem, we know that there exists $\delta_1\in(0,1)$ such that

$$
\begin{aligned}
& u(y-g-\hat{Q}_1)-u(y-g-Q_1^s) \\
= \ & u'\left(y-g-Q_1^s+\delta'\left(Q_1^s-\hat{Q}_1\right)\right)\left(Q_1^s-\hat{Q}_1\right).
\end{aligned}
$$

Similarly, there exists $\delta_2\in(0,1)$ such that

$$
\begin{aligned}
& \left(1-p_1\right)\left[u\left(y+g+\beta q\int V_2^s(p_2)d\Phi\left(p_2\right)\right)-u(y+g)\right] \\
= \ & \left(1-p_1\right)\left[u'\left(y+g+\delta_2\beta q\int V_2^s(p_2)d\Phi\left(p_2\right)\right)\beta q\int V_2^s(p_2)d\Phi\left(p_2\right)\right] \\
= \ & u'\left(y+g+\delta_2\beta q\int V_2^s(p_2)d\Phi\left(p_2\right)\right)\beta\left(Q_1^s-\hat{Q}_1\right).
\end{aligned}
$$

Hence

$$
\begin{aligned}
& W\left(\hat{C}\right)-W^s\left(C^s\right) \\
\geq \ & \left[u'\left(y-g-Q_1^s+\delta'\left(Q_1^s-\hat{Q}_1\right)\right)-\beta u'\left(y+g+\delta_2\beta q\int V_2^s(p_2)d\Phi\left(p_2\right)\right)\right]\left(Q_1^s-\hat{Q}_1\right)\geq 0
\end{aligned}
$$

where the last inequality will be strict if $Q_1^s-\hat{Q}_1$ is strictly positive, i.e., if there is dynamic reclassification risk insurance under contract C^s.

Now let C^s be the equilibrium contract in the presence of the settlement market. The above argument shows that the contract $\hat{C}$ constructed through a simple reduction of first period premium is feasible for the problem without the settlement market; and $\hat{C}$ provides weakly (or strictly, if C^s offers some dynamic insurance) higher expected utility to the consumers for the case without settlement market than C^s would provide for consumers with settlement market. Because $\hat{C}$ is only a candidate contract for the case without settlement market, the equilibrium contract in that case must provide no lower expected consumer welfare than $\hat{C}$. □

Proof of Proposition 5.

Proof. The first order conditions for the solution to problem (10) are:

$$u'\left(y-g-Q_1^{ss}\right)=\mu \tag{24a}$$

$$v'\left(F_1^{ss}\right)=\mu \tag{24b}$$

$$u'\left(y+g-Q_2^{ss}(p_2)\right)=\mu-\frac{\lambda(p_2)}{(1-p_1)(1-q)\phi(p_2)}-\frac{\beta\gamma(p_2)}{(1-p_1)(1-q)\phi(p_2)} \tag{24c}$$

$$v'\left(F_2^{ss}(p_2)\right)=\mu-\frac{\lambda(p_2)}{(1-p_1)(1-q)\phi(p_2)}-\frac{\beta\gamma(p_2)}{(1-p_1)(1-q)\phi(p_2)} \tag{24d}$$

$$u'\left(y+g+S^{ss}(p_2)\right)=\mu+\frac{\gamma(p_2)}{(1-p_1)q\phi(p_2)} \tag{24e}$$

where $\mu>0$, $\lambda(p_2)\geq 0$, and $\gamma(p_2)\geq 0$ are respectively the Lagrange multipliers for constraints (10b), (10c) and (10d).

From these conditions, we see that constraint (10d) must bind for all p_2 because otherwise $\gamma(p_2)=0$, which implies that $u'(y+g+S_2^{ss}(p_2))=u'(y-g-Q_1^{ss})$, which is impossible. □

Proof of Propositions 6 and 7.

Proof. The proofs are similar to the proof of Proposition 4 so we omit them for brevity. □

Proof of Proposition 8.

Proof. The Lagrangian for problem (11) is:

$$\begin{aligned}\mathcal{L} &= u(y-g-Q_1)+p_1v(F_1)+(1-p_1)(1-q)\int_0^1\left[u(y+g-Q_2(p_2))+pv(F_2(p_2))\right]d\Phi(p_2)\\ &+(1-p_1)q\int_{S\geq\beta V_2(p_2)}u(y+g+S)d\Phi(p_2)+(1-p_1)q\int_{S<\beta V_2(p_2)}u(y+g+\beta V_2(p_2))d\Phi(p_2)\\ &+\int_0^1\lambda(p)\left[Q_2(p_2)-p_2F_2(p_2)+S\right]d\Phi\left(p_2\right)+\gamma S\\ &+\mu\left[\begin{array}{c}(Q_1-p_1F_1)+(1-p_1)(1-q)\int_0^1\left[Q_2(p_2)-p_2F_2(p_2)\right]d\Phi(p_2)\\ -(1-p_1)q\int_{S\geq\beta V_2(p_2)}Sd\Phi(p_2)+(1-p_1)(1-q)\int_{S<\beta V_2(p)}\left[Q_2(p)-p_2F_2(p_2)\right]d\Phi(p_2)\end{array}\right]\end{aligned} \tag{25}$$

where $\{\lambda(p_2)\leq 0: p_2\in[0,1]\},\gamma\geq 0,\mu\geq 0$ are respectively the Lagrange multiplier for constraints (11b), (11c), and (11d).

Using standard arguments, we can show that under the optimum, $V_2(\cdot)$ must be continuous and monotonically increasing in p_2, with $V_2(p_2)>0$ for some p_2 if there is some dynamic reclassification risk insurance in equilibrium. Thus we know that for every $S\geq 0$ with S sufficiently small, there exists a $\hat{p}_2$ such that $\beta V_2(p_2)\geq S$ if and only if $p_2\geq\hat{p}_2$ where $\beta V_2(\hat{p}_2)=S$. Thus from the Implicit Function Theorem, we have:

$$\frac{d\hat{p}_2}{dS}=\frac{1}{\beta V_2'(\hat{p}_2)}. \tag{26}$$

Therefore, the Lagrangian (25) can be rewritten as:

$$\begin{aligned}\mathcal{L} &= u(y-g-Q_1)+p_1v(F_1)+(1-p_1)(1-q)\int_0^1\left[u(y+g-Q_2(p_2))+pv(F_2(p_2))\right]d\Phi(p_2)\\ &+(1-p_1)q\int_0^{\hat{p}_2}u(y+g+S)d\Phi(p_2)+(1-p_1)q\int_{\hat{p}_2}^1u(y+g+\beta V_2(p_2))d\Phi(p_2)\\ &+\int_0^1\lambda(p)\left[Q_2(p_2)-p_2F_2(p_2)+S\right]d\Phi\left(p_2\right)+\gamma S\\ &+\mu\left[\begin{array}{c}(Q_1-p_1F_1)+(1-p_1)(1-q)\int_0^1\left[Q_2(p_2)-p_2F_2(p_2)\right]d\Phi(p_2)\\ -(1-p_1)q\int_0^{\hat{p}_2}Sd\Phi(p_2)+(1-p_1)(1-q)\int_{\hat{p}_2}^1\left[Q_2(p)-p_2F_2(p_2)\right]d\Phi(p_2)\end{array}\right].\end{aligned} \tag{27}$$

Applying the Leibniz rule and (26), we have that the derivative of the Lagrangian (27) with respect to S, evaluated at the optimum (superscripted by $*$), is

$$\frac{\partial \mathcal{L}}{\partial S} = (1-p_1)q\int_0^{\hat{p}_2^*} u'(y+g+S^*)d\Phi(p_2) + (1-p_1)qu(y+g+S^*)\frac{\phi(\hat{p}_2^*)}{\beta V_2^{*\prime}(\hat{p}_2^*)}$$
$$- (1-p_1)qu(y+g+\beta V_2^*(\hat{p}_2^*))\frac{\phi(\hat{p}_2^*)}{\beta V_2^{*\prime}(\hat{p}_2^*)} + \int_0^1 \lambda(p_2)d\Phi\left(p_2\right) + \gamma$$
$$-\mu(1-p_1)q\int_0^{\hat{p}_2^*} d\Phi(p_2) - \mu(1-p_1)qS^*\frac{\phi(\hat{p}_2^*)}{\beta V_2^{*\prime}(\hat{p}_2^*)}$$
$$-\mu(1-p_1)q\left[Q_2^*(\hat{p}_2^*) - \hat{p}_2^* F_2^*(\hat{p}_2^*)\right]\frac{\phi(\hat{p}_2^*)}{\beta V_2^{*\prime}(\hat{p}_2^*)}. \tag{28}$$

Since by definition, $\beta V_2^*(\hat{p}_2^*) = S^*$, (28) simplifies to:

$$\frac{\partial \mathcal{L}}{\partial S} = (1-p_1)qu'(y+g+S^*)\Phi(\hat{p}_2^*) + \int_0^1 \lambda(p_2)d\Phi\left(p_2\right) + \gamma$$
$$- \mu(1-p_1)q\Phi(\hat{p}_2^*) + \mu(1-p_1)q(1-\beta)V_2^*(\hat{p}_2^*)\frac{\phi(\hat{p}_2^*)}{\beta V_2^{*\prime}(\hat{p}_2^*)}. \tag{29}$$

We now argue that $\frac{\partial \mathcal{L}}{\partial S}$ is strictly negative when S deviates from 0 to a small $\varepsilon > 0$. To see this, note that in the ε-neighborhood of $S = 0$, we have $\gamma = 0$, $\lim_{s\to\varepsilon=0^+} V_2^*\left(\hat{p}_2^*\right) = \varepsilon$, thus

$$\lim_{s\to\varepsilon=0^+}\frac{\partial \mathcal{L}}{\partial S} = (1-p_1)q\left[u'(y+g) - \mu\right]\Phi(\hat{p}_2^*\left(0\right)) + \int_0^1 \lambda(p_2)d\Phi\left(p_2\right),$$

where $\hat{p}_2^*\left(0\right) = \lim_{\omega\to 0^+} \hat{p}_2\left(\varepsilon\right)$ and $\hat{p}_2\left(\varepsilon\right)$ solves $\beta V_2\left(\hat{p}_2\left(\varepsilon\right)\right) = \varepsilon$. Note that the first order condition with respect to Q_1 implies that $u'\left(y-g-Q_1^*\right) = \mu > u'(y+g)$ and that $\lambda\left(p_2\right) \leq 0$ for all p_2, we have:

$$\lim_{s\to\varepsilon=0^+}\frac{\partial \mathcal{L}}{\partial S} < 0.$$

The same argument can be used to show that if the optimal S^* was strictly positive, a deviation of S from S^* to $S^* - \varepsilon$ will be strictly preferred. Thus the optimal S^* must be equal to 0. □

Proof of Proposition 9.

Proof. We will proceed by a series of four lemmas:

Lemma A1. *If $V_2(p_2) \geq V_2'(p_2)$, then $Q_2'(p_2), F_2'(p_2)$ maximizes $u(y+g-Q) + p_2\delta v(F)$ subject to $p_2F - Q = V_2'(p_2)$.*

Proof. If it were not the case, then one can choose an alternative contract $\hat{Q}_2'(p_2), \hat{F}_2'(p_2)$ with the same actuarial value $V_2'(p_2)$ that improves utility in the diminished bequest state. It is possible that under the new contract IC constraint (13d) becomes violated. If that is the case, also set $\hat{Q}_2(p_2), \hat{F}_2(p_2) = \hat{Q}_2'(p_2), \hat{F}_2'(p_2)$, which now improves utility in the high bequest state as well, while simultaneously reducing saving costs for the insurer, which can be passed to the new contract as a lower first period premium. $\square$

Lemma A2. *If $V_2(p_2) \leq V_2'(p_2)$, then $Q_2(p_2), F_2(p_2)$ maximizes $u(y+g-Q) + p_2 v(F)$ subject to $p_2 F - Q = V_2(p_2)$.*

Proof. Logic is the same as above. $\square$

Lemma A3. *If $V_2(p_2) \geq V_2'(p_2)$, then $u(y+g-Q_2(p_2)) + p_2 v(F_2(p_2))$ is greater than or equal to $\max_{Q,F} u(y+g-Q) + p_2 v(F)$ s.t. $p_2 Q - F = V_2'(p_2)$.*

Proof. Let Q^*, F^* maximize $u(y+g-Q) + p_2 v(F)$ s.t. $p_2 Q - F = V_2'(p_2)$. If $u(y+g-Q_2(p_2)) + p_2 v(F_2(p_2)) < u(y+g-Q^*) + p_2 v(F^*)$, then simply choose an alternative contract with $\hat{Q}_2(p_2), \hat{F}_2(p_2) = Q^*, F^*$. This improves utility in the high bequest state and even saves costs to the insurer. Moreover, because of Lemma Lemma A1, IC constraint (13e) is not violated because $Q_2'(p_2), F_2'(p_2)$ already maximizes utility in the low-bequest state subject to actuarial value $V_2'(p_2)$. $\square$

Lemma A4. *If $V_2(p_2) \leq V_2'(p_2)$, then $u(y+g-Q_2'(p_2)) + p_2 \delta v(F_2'(p_2))$ is greater than or equal to $\max_{Q,F} u(y+g-Q) + p_2 \delta v(F)$ s.t. $p_2 Q - F = V_2(p_2)$.*

Proof. The logic is the same as above. $\square$

Lemmas Lemma A1-Lemma A4 imply that the optimal contract can be implemented with a non-bequest contingent contract, $\langle \hat{Q}_1, \hat{F}_1, \hat{Q}_2(p_2), \hat{F}_2(p_2), \hat{S}_2(p_2) \rangle$, with a health-contingent surrender value. Intuitively, the surrender value will equal the lowest actuarial value across bequest states, i.e. $\hat{S}_2(p_2) = \min\{V_2(p_2), V_2'(p_2)\}$, and policyholders in that bequest state will surrender and repurchase on the spot market. Lemmas Lemma A1 and Lemma A2 imply that such surrender and repurchase does replicate the optimal contract, because surrender and repurchase solves:

$$\max_{Q,F} u(y+g-Q+V_2'(p_2)) + p_2 \delta v(F) \text{ s.t. } p_2 F - Q = 0$$

which is equivalent to:

$$\max_{Q,F} u(y+g-Q) + p_2\delta v(F) \text{ s.t. } p_2F - Q = V_2'(p_2)$$

(or replace δ with 1 and $V_2'(p_2)$ with $V_2(p_2)$ if the high-bequest state is surrendering). Moreover, Lemmas Lemma A3 and Lemma A4 imply that in the bequest state with the higher actuarial value, the policyholder would not want to surrender and repurchase. Thus, the optimal contract can be implemented by the non-bequest contingent contract with a health-contingent cash surrender value as specified in Proposition 9. $\square$