

Guaranteed renewable life insurance under demand uncertainty

Hoy, Michael; Mirza, Afrasiab; Sadanand, Asha

DOI:
[10.1111/jori.12320](https://doi.org/10.1111/jori.12320)

License:
Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Hoy, M, Mirza, A & Sadanand, A 2020, 'Guaranteed renewable life insurance under demand uncertainty', *Journal of Risk and Insurance*. <https://doi.org/10.1111/jori.12320>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Guaranteed renewable life insurance under demand uncertainty

Michael Hoy¹ | Afrasiab Mirza²  | Asha Sadanand¹

¹Department of Economics and Finance,
University of Guelph, Guelph, Ontario,
Canada

²Department of Economics, University of
Birmingham, Birmingham, UK

Correspondence

Afrasiab Mirza, Department of Economics,
University of Birmingham, JG Smith
Building, Edgbaston, Birmingham, UK B15
2TT.

Email: a.mirza@bham.ac.uk

Abstract

Guaranteed renewability (GR) is a prominent feature in many health and life insurance markets. We develop a model that includes unpredictable (and unobservable) fluctuations in demand for life insurance as well as changes in risk type (observable) over individuals' lifetimes. The presence of demand type heterogeneity leads to the possibility that optimal GR contracts may have a renewal price that is either above or below the actuarially fair price of the lowest risk type in the population. Individuals whose type turns out to be high risk but low demand renew more of their GR insurance than is efficient due to the attractive renewal price. This results in imperfect insurance against reclassification risk. Although a first-best efficient contract is not possible in the presence of demand type heterogeneity, the presence of GR contracts nonetheless improves welfare relative to an environment with only spot markets.

KEYWORDS

demand uncertainty, guaranteed renewability, insurance, reclassification risk

JEL CLASSIFICATION

D8; D86; G22

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Journal of Risk and Insurance* published by Wiley Periodicals LLC on behalf of American Risk and Insurance Association

1 | INTRODUCTION

Guaranteed renewability (GR), which is a prominent feature in health and life insurance markets, provides an opportunity for individuals to insure against reclassification risk. This works as follows. Consider a set of *ex ante* identical individuals each of whom purchases an initial 10-year term life insurance contract with a view of possibly purchasing a subsequent policy at the end of the term. By the end of the contract period, some insureds may have discovered that their mortality status has changed. If this change is observable to insurers, then the price for a new insurance contract will reflect that change in risk. Individuals recognize *ex ante* that their risk type may change over time and so prefer to avoid the prospect of premium risk associated with stochastic mortality prospects. GR contracts contain a promise to offer a subsequent insurance policy at the expiry date of the first contract at a price that is independent of any changes in mortality risk. The premium for the implicit insurance against reclassification risk is embedded in the first contract (earlier period) through an extra premium assessment—a phenomenon known as front loading. This allows insurers to offer insurance to those individuals who turn out to be higher risk types in the subsequent 10-year period at a price below their actuarially fair rate. As long as the amount of front loading is sufficient, the added profit from the first (period) contract compensates for insurers' losses from the second (period) contract.¹

In our paper we focus on the implications for GR (and long-term) insurance to ameliorate premium risk when individuals face uncertainty over future changes in both mortality risk and insurance needs. We consider an environment where individuals face no capital market imperfections (they can borrow or save at the risk-free rate) nor other impediments such as the existence of resettlement or viatical market opportunities that can thwart GR insurance to fully protect against reclassification risk. We develop a two-period model of insurance in which individuals are homogeneous in the first period and hold the same beliefs about the likelihood of becoming a low- or high-demand type in the second period. An important feature of insurance demand is how it changes over the life cycle. As noted in Hong and Rios-Rull (2012), average demand follows a life cycle pattern that rises from young adulthood to “around age 45 for males and 35–40 for females” (based on 1990 data). They also show, in their figure 1 (p. 3705), that there is substantial variation in demand across individuals at all ages and especially around the peak level of demand. This means that to have an ideal amount of coverage for premium risk in the future, one may have to hold more insurance than is optimal early in life (i.e., for the younger part of the life cycle where demand tends to be increasing). This turns out to be a critical factor in determining the extent to which GR can provide insurance against reclassification risk. We allow second-period demand to be higher or lower than first-period demand for either or both demand types. Moreover, we also allow for the possibility that demand does not vanish over time. Each individual's risk type also evolves over time in a similar manner; that is, individuals have the same mortality risk in the first period but their mortality risk diverges in the second period. Moreover, in period 1 individuals hold the same beliefs about the evolution of their risk type for period 2.²

¹A similar phenomenon may be reflected in short-term versus long-term insurance contracts (e.g., 10 vs. 20 years) with longer contracts providing insurance against reclassification risk through front loading to keep premiums later in the contract sufficiently low to avoid lapsation by better risks.

²In a similar environment but without differential demand types, Peter, Richter, and Steinorth (2015) consider the implications of individuals learning imperfectly about their risk type over time with this information being private. Fei, Fluet, and Schlesinger (2013) also use a model that features demand uncertainty but that does not include risk type uncertainty nor any dynamic features of insurance demand present in our model of GR insurance.

There are many potential sources of (evolving) demand type heterogeneity relevant to life insurance. The amount of coverage an individual desires at any point in time is affected by a number of factors, including marital status, income of the insured individual, number of children, earning options for other family members, expenditure requirements for the survivor family should death of the individual occur, and the insured's pure (altruistic) preferences among other factors. All of these can change over time.³ Some of these characteristics are unobservable to the insurer and others, while observable, typically have idiosyncratic and unobservable implications on individuals' preferences for insurance. We treat demand type as unobservable to the insurer or as noncontractible. Insureds, in contrast, learn about their demand preferences and change their valuation of insurance accordingly. This uncertainty in future demand represents a challenge to individuals when deciding how much GR insurance is appropriate to purchase at a given point in their lifetime, which in turn compromises the ability of GR insurance contracts to protect consumers against reclassification risk. Moreover, given the noncontractible nature of demand risk, the combination of variations in both morality risk and demand risk creates a type of adverse selection problem, as described below.

One important feature of interest regarding the performance of GR insurance is contract lapsation. If the renewal terms are not sufficiently attractive to people who discover they have become relatively low risk, then they will have an incentive to opt out of the first contract at or before the expiry date and not purchase a subsequent contract at the agreed upon price. Moreover, if the renewal price is below the actuarially fair rate for high-risk types, which it must be to provide protection against reclassification risk, then those with low insurance demand who turn out to be high risk will wish to renew more insurance than is efficient. This means the second-period contract will have a disproportionate share of demand from high-risk types which creates a stress on the degree of front loading required to make GR insurance financially sustainable.⁴

Our paper contributes to the literature on GR insurance by providing an explicit welfare analysis of a two-period model of decision making based on expected utility preferences which evolve over time.⁵ Individuals may find their preference for insurance either rises or falls for the later (second) period under consideration. The various possible demands for second-period insurance may not align with first-period insurance needs and so the only way for an individual to hold the optimal amount of GR insurance from the second-period perspective may be to over-insure in the first period. This scenario would be expected during (typically earlier) periods of life when future insurance demand tends to increase (on average). We illustrate how these two factors compromise the effectiveness of GR insurance to protect against premium risk. We also see how these patterns influence the structure of premiums of GR insurance; that is, both the degree of front loading and the renewal price.

It has been shown (e.g., Hendel & Lizzeri, 2003; Pauly, Kunreuther, & Hirth, 1995) that, in ideal settings, GR insurance or long-term insurance with sufficient front loading of premiums

³See the survey by Zietz (2003), and particularly tables 2 and 3, for empirical evidence on the effect of personal characteristics on the demand for life insurance. Some of these characteristics would typically change stochastically over a person's lifetime. Using data from the Health and Retirement Study, Fang and Kung (2012, pp. 4, 5) demonstrate that income and health shocks are relatively more important than bequest motive shocks in explaining lapsation when policyholders are young, but as they age, the bequest motive shocks play a more important role.

⁴Evidence of this phenomenon in a health insurance market is provided by Hofmann and Browne (2013).

⁵In this paper we focus on life insurance, although the basic principles of GR insurance apply also to health insurance.

can fully eliminate reclassification risk.⁶ Following Harris and Holmstrom (1982), who develop a model of one-sided commitment between firms and their workers, Hendel and Lizzeri (2003) explore the implications of one-sided commitment in an insurer—consumer model with insurers being the party that can commit. Hendel and Lizzeri (2003) show that more front loading is consistent with increased efficiency as it generates more “commitment” on the part of households to renew their GR. However, in the presence of demand fluctuations, this is no longer true and we show that less commitment and less front loading can improve welfare. Alternatively, Frick (1998) demonstrates how capital market imperfections can destroy the potential for GR insurance to provide complete protection against reclassification risk. An imperfect capital market is but one market characteristic that can limit the ability of GR insurance to offer protection against reclassification risk.⁷ In different contexts, Polborn, Hoy, and Sadanand (2006) and Daily, Hendel, and Lizzeri (2008) show that if there is uncertainty about future insurance needs as well as risk type and individuals have access to settlement markets where they can sell their previously purchased (long-term) insurance coverage, then GR insurance contracts cannot completely eliminate premium risk.

Fang and Kung (2020) use a model which is very similar to that of Daily et al. (2008) to investigate why most life insurance policies have little to no cash surrender value as well as to analyze the welfare implications of a life settlement market. Given our goal of investigating the pricing and welfare implications of GR term insurance contracts, we use a very different contract structure. In particular, we assume that insurers do not base renewal prices on risk type. This is a common feature of such products.⁸ We also make a number of other different assumptions, the most important of which involve a more general underlying preference structure for consumers as well as allowing (costless) saving or borrowing. The implications of these differences are discussed at appropriate places in our paper.

In our paper, as in Fang and Kung (2020) and Daily et al. (2008), policyholders who let their insurance policies lapse (partially or fully), do so as a result of low realized (shocks) to their bequest motive. It is important to recognize other reasons for lapsing and also to investigate implications of individuals not behaving as fully rational and forward looking agents. Fang and Wu (2017) consider the effect of consumers being overconfident about their bequest motives or mortality rates in the presence of a life settlement market. Gottlieb and Smetters (2019) provide evidence that a majority of observed lapses are due to individuals either forgetting to pay premiums or underestimating the need for money in the future. Moreover, without imposing restrictions on the contract space, they develop two theoretical models with behavioral consumers that imply observed policy characteristics not well explained by models of fully rational consumers.

The rest of the paper is organized as follows. The next section presents the basic model while Section 3 characterizes the first-best (social) optimum as well as the characteristics of the

⁶In an early contribution to this literature, Cochrane (1995) developed an interesting alternative approach to protecting individuals from reclassification risk by introducing lump-sum severance payments for individuals whose risk deteriorates at the end of a period of insurance cover. The ideal setting in Hendel and Lizzeri (2003) corresponds to sufficiently low growth in income relative to importance of front loading that individuals are willing to accept sufficient front loading to cover even protection against reclassification risks for relatively good health states. This follows from their results (ii) and (iii) of Proposition 1 and conditions explained in footnote 13, p. 310.

⁷For a more complete discussion, see Peter et al. (2015).

⁸For example, from Canadian Life and Health Insurance Association—“Guaranteed renewable policies: A feature of an individual insurance policy where the insurance company guarantees to renew the insurance at the end of a certain period, regardless of any changes in your health.”

allocation when (only) spot insurance markets are available and when GR insurance is also available. The main welfare analysis is provided in Section 4.⁹ Section 5 provides conclusions.

2 | MODEL

We consider an economy populated by a measure one of ex ante (identical) individuals who buy life insurance and live at most three periods. Each such individual has a family associated with him. In case of the individual's death, we refer to his associated family as the *survivor family* while in any period that he lives we refer to his associated family as the *whole family*. No other members of the family may die. Preferences relate to those of the insurance buyer, albeit he takes his family members' well-being into account. For simplicity, we assume he is the only income earner in the family and receives income y_1 at the beginning of period 1. If he survives to period 2, he receives a further y_2 at the beginning of period 2. His risk and demand type evolve over time. Each individual has a probability of death of p , $0 < p < 1$, in the first period of life. If an individual survives the first period, then his probability of death in the second period depends on whether he is a high- or low-risk type. We describe risk type by index $i \in \{L, H\}$ for low- and high-risk type, respectively, with associated probabilities p_L, p_H where $0 < p_L < p_H < 1$. Moreover, we assume all risk types have a higher mortality in period 2 than in period 1 ($0 < p < p_L < p_H < 1$; see Hendel & Lizzeri, 2003).

The individuals (and associated families) are homogeneous in all respects in the first period and discover their risk type associated with second-period mortality at the beginning of period 2. Insurers also observe individuals' risk type and so there is no asymmetric information in this regard. However, individuals also discover their demand type at the beginning of period 2 which insurers do not observe.¹⁰ In period 1 individuals perceive their prospects about both risk type and demand type development according to the actual population portions of q_i , $i \in \{L, H\}$ for risk type and r_j , $j \in \{l, h\}$ for demand type where $i \in \{L, H\}$ represents low- and high-risk type while $j \in \{l, h\}$ represents low and high-demand type. Unlike the papers mentioned earlier, we allow low-demand types to have some positive bequest motive, albeit less than for high-demand types. We do not explore the implications of them having zero bequest motive since such individuals would not be in the market for period 2 insurance be it from the spot market or through a GR policy. Adding individuals who would have zero bequest motive, should they turn out to be low-demand type, would not change the qualitative nature of our results. Risk and demand type are not correlated (i.e., the probability of an individual being risk type i and demand type j is $q_i \cdot r_j$). These differing preferences (demand type) for life insurance in period 2 are reflected in the felicities for death state consumption in period 2 as described below.¹¹

So, period 2 decisions depend on both the individual's risk and demand type, characterized by the pair ij , with $i \in \{L, H\}$ and $j \in \{l, h\}$. In cases where confusion may occur, we index the time period and the state (life or death) using superscripts. We refer to the death state by D and the life state by N (i.e., not death). Thus, consumption in period 2 for a person of type ij is

⁹See the Supporting Information Appendix for a set of simulation results that help to further an understanding of how the combination of demand and risk type uncertainty affects the structure of contracts and the welfare effects of GR contracts.

¹⁰It is equivalent to alternatively assume demand characteristics are observable but not contractible.

¹¹Note that one could instead introduce demand heterogeneity through different felicities in the life state. This would have similar effects as in our model. Note that such an example may be a liquidity shock associated with the life state of the world. See also Gottlieb and Smetters (2019) for a model of life insurance where agents face liquidity shocks.

represented by C_{ij}^{2D} in the death state and C_{ij}^{2N} in the life (i.e., non-death) state. We write their felicity for consumption in the life state for period t as $u_t(\cdot)$, $t = 1, 2$. Their felicities in the period 2 death state, which depend on demand type j are modeled by the function $v_2(\cdot; \theta_j)$ where $v'_2(x; \theta_h) \equiv \frac{\partial v(x; \theta_h)}{\partial x} > v'_2(x; \theta_l) \equiv \frac{\partial v(x; \theta_l)}{\partial x}$ for all $x > 0$ with an abuse of notation in our use of primes to convey the partial derivative with respect to the first argument. The latter captures that high-demand types want more insurance than low-demand types. The functions u_2 and v_2 satisfy the usual assumptions for risk averters (i.e., $u'_2, v'_2 > 0$ and $u''_2, v''_2 < 0$), as well as the standard Inada conditions.¹²

Individuals have homogeneous preferences in period 1 with their felicity in the life state being $u_1(\cdot)$ and that in the death state being $v_1(\cdot)$, the latter of which is meant to reflect the insurance purchaser's perspective on the survivor family's future utility (including prospects for period 2 and beyond).¹³ This can naturally be different from the felicity in the death state of period 2. Similar to the above notation for period 2, consumption in the life and death states of period 1 are C^{1N} and C^{1D} , respectively.¹⁴ Again, we assume that period 1 felicities satisfy the usual conditions for risk averters and the corresponding Inada conditions.

Unlike Daily et al. (2008), and Fang and Kung (2020), we allow households to save or borrow at the risk-free rate (which we normalize to 1). This is important as it limits the role of GR contracts in our model to the provision of reclassification risk rather than confounding it with intertemporal consumption smoothing.

Timing of information revelation and taking of decisions is as follows. At the beginning of period 1 individuals receive income y_1 and decide on: (a) the amount of spot insurance to hold for period 1 (L^1), (b) amount of GR insurance (L^{1G}), and (c) the amount of savings, s , which we allow to be negative to capture borrowing.¹⁵ $L^1 + L^{1G}$ is the insurance coverage in period 1 and savings is also available to the survivor family should the insured die in period 1. If death occurs, it happens at the end of period 1. Note that s is not deducted from consumption in the death state as the survivor family gets to use savings from period 1 into the future. The felicity v_1 reflects continuation utility for this survivor family. L^{1G} is the amount of that coverage that could be renewed at a guaranteed (predetermined) rate in the second period should the insured survive to period 2. We let π^1 be the price of first-period spot insurance. We assume risk neutral insurers in a competitive environment with no administrative costs. Insurers can fully commit to long-term contracts. Thus, since coverage from first-period spot insurance expires at the end of period 1, competition leads to $\pi^1 = p$ (i.e., actuarially fair insurance).

The front loading of GR insurance allows an individual the option to renew at a price which earns the insurer expected losses. This implies that the unit price of this coverage, π^{1G} , must

¹²Intertemporal discounting can also be incorporated into the model through appropriate choices of felicities in period 1 and 2. The Inada conditions preclude the complete loss of bequest motives which maybe unrealistic but such individuals would not renew or purchase any insurance so this assumption is without loss of generality.

¹³This is an indirect utility based on how the family's circumstances will evolve should the income earning insurance buyer die in the first period. The family may be expected to evolve in the sense that a surviving spouse has uncertain prospects of generating income in period 2 (as well as the remainder of period 1) and so on. That is, death felicities should be interpreted as continuation utilities. This simplistic "main breadwinner" sort of model could be transformed to one with two earners and two potential insurance buyers. However, that would lead to a much more complicated model and, we believe, not significantly improved insights.

¹⁴Note that since individuals are homogenous in period 1, there is no subscript pair ij attached to these consumptions.

¹⁵In an intertemporal model, insurance purchases shift consumption in a current period into any loss state of a future period and therefore creates in some degree consumption smoothing, albeit state contingent consumption smoothing. As shown by Hofmann and Peter (2015), if one omits the savings decision in such a model, the role of insurance (reimbursement for financial losses) becomes contaminated with the motive for income smoothing.

exceed p , the expected unit cost of providing first-period insurance cover (i.e., front loading). This is explained in greater detail later. As is the case for period 1, income y_2 is received at the beginning of the period while if death occurs, this happens at the end of the period. At the beginning of period 2, the spot insurance from period 1 expires and individuals learn about their risk type (i) and their demand type (j). Insurers know the risk type of insureds but not their demand type or rather, if they do know their demand type, they are not able to contract upon that information. Each insured then chooses how much GR insurance that was purchased in period 1 (L^{1G}) to renew (L_{ij}^{2G}) at the predetermined (guaranteed) price of π^{2G} . This amount will depend on both risk and demand type with (obviously) $L_{ij}^{2G} \leq L^{1G}$. Importantly, following Hendel and Lizzeri (2003) we assume that individuals cannot commit to long-term contracts. Therefore, they may opt out of their GR contracts and purchase spot insurance (L_{ij}^2) at the risk type-specific price ($\pi_i^2 = p_i$). Note that if $\pi^{2G} > p_L$, low-risk types would not renew any of their GR insurance from period 1.¹⁶ Expected utility from the perspective of the beginning of period 1 is

$$EU = pv_1(C^{1D}) + (1 - p)u_1(C^{1N}) + (1 - p) \sum_i \sum_j q_i r_j \left[p_i v_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u_2(C_{ij}^{2N}) \right], \quad (1)$$

where

$$\begin{aligned} C^{1N} &= y_1 - s - \pi^1 L^1 - \pi^{1G} L^{1G}, \\ C^{1D} &= y_1 + (1 - \pi^1) L^1 + (1 - \pi^{1G}) L^{1G}, \\ C_{ij}^{2N} &= y_2 + s - \pi_i^2 L_{ij}^2 - \pi^{2G} L_{ij}^{2G}, \\ C_{ij}^{2D} &= y_2 + s + (1 - \pi_i^2) L_{ij}^2 + (1 - \pi^{2G}) L_{ij}^{2G}, \end{aligned}$$

with constraints

$$0 \leq L^1, 0 \leq L^{1G}, 0 \leq L_{ij}^{2G} \leq L^{1G}, 0 \leq L_{ij}^2.$$

In both periods, the death state consumption and felicity reflect the continuation values for the survivor family. The life state consumption and felicity in period 1 reflect the values for the whole family including the main breadwinner who in this case survives (at least) to the end of period 2. The consumption and felicity in the second-period life state reflect continuation values for the whole family, including the main breadwinner, for period 2 and beyond. Note that, for $s > 0$, saving is deducted from first-period consumption in the life state but is added back into consumption in the first-period death state (since it is available to the survivor family). If the main breadwinner survives to period 2, then saving from period 1 is available for consumption in either the life or death state. If $s < 0$, the amount borrowed is available for consumption in the life state. We assume that if the main breadwinner dies, the amount borrowed must be repaid (by his estate) and so borrowing does not augment consumption in the first-period death

¹⁶We also assume that low-risk types renew all of their L^{1G} if $\pi^{2G} = p_L$, the spot price for low-risk types in period 2. This is of no consequence since competition means π_L^2 is equal to the low-risk type loss probability which means the lapsation-renewal decision has no consequence on insurer profits and hence on π^{1G} or π^{2G} .

state. Should the main breadwinner survive to period 2, any amount borrowed must of course be repaid in that period in either life or death state.

3 | ALLOCATIONS UNDER FIRST-BEST, SPOT, AND GR CONTRACTS

Before examining the allocations that may be supported with GR contracts, it is instructive to examine two important alternative allocations: the first-best and spot markets only. These are useful to clarify both the value of GR and also its shortcomings.

3.1 | Benchmark: First-best

The first-best allocation is obtained by maximizing ex ante utility (i.e., from the perspective of individuals in period 1) subject to a set of aggregate resource constraints, one for each period.¹⁷ These resource constraints simply require that the total consumption across types and states in each period be equal to the total available resources. The first-best allocation is the solution to:

$$\max_{C^{1D}, C^{1N}, s, \{C_{ij}^{2D}, C_{ij}^{2N}\}} EU \equiv p v_1(C^{1D}) + (1 - p) u_1(C^{1N}) + (1 - p) \left[\sum_i \sum_j q_i r_j (p_i v_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u_2(C_{ij}^{2N})) \right] \text{ s.t.} \quad (2)$$

$$y_1 \geq p C^{1D} + (1 - p)(C^{1N} + s), \quad (3)$$

$$y_2 + s \geq \sum_i \sum_j q_i r_j \left[(p_i C_{ij}^{2D} + (1 - p_i) C_{ij}^{2N}) \right]. \quad (4)$$

As the objective function is strictly concave and the constraints are linear, the solution to this problem is unique, and can be characterized by the first-order conditions. This statement applies to all subsequent optimization problems.

Proposition 1. *The social optimum is characterized by:*

1. *Marginal utilities in all time/state contingent scenarios are equal across all risk and demand type combinations:*

$$v'_1(C^{1D}) = u'_1(C^{1N}) = v'_2(C_{ij}^{2D}; \theta_j) = u'_2(C_{ij}^{2N}), \quad \text{for all } ij \in \{H, L\} \times \{h, l\}. \quad (5)$$

2. *Consumption in the life state or death state for each period is independent of risk type.*

¹⁷As there is a measure one of individuals, aggregates are equal to per capita values.

3. *Second-period consumption level in the death state for high-demand types exceeds that for low-demand types (but is independent of risk type, as noted above).*

Proof. See Appendix B.1. □

It follows directly from this proposition that the first-best allocation requires marginal utilities of consumption in each time and state to be equalized. This is an application of the fundamental theorem of risk-bearing. This implies that, for a given demand type, consumption in the period 2 death state is the same for both risk types and likewise for the period 2 life state consumption. However, since at any given consumption level the marginal utility of consumption of high-demand types exceeds that for low-demand types, the first-best allocation requires that consumption in the death state be higher for the high-demand type. This is easily established as

$$v'_2(C_{il}^{2D}; \theta_l) = v'_2(C_{ih}^{2D}; \theta_h) < v'_2(C_{il}^{2D}; \theta_h) \Rightarrow C_{ih}^{2D} > C_{il}^{2D}. \quad (6)$$

If it were feasible, one way to decentralize the first-best allocation would be to allow individuals to write contracts at time 1 that offer transfers contingent on their realized demand and risk type at time 2. Such contracts replicate the social planner's ability to effectuate transfers across agents at time 2.¹⁸ Given unobservability of demand type, it is not possible to implement the first-best, as those with low demand have incentives to claim to be high demand to obtain more consumption in the death state (and equal consumption levels in all other situations). This is akin to Mirrlees (1971) where the unobservability of individual productivity precludes the ability to implement the first-best.

3.2 | Spot markets only

Now consider the equilibrium choices of individuals when only spot insurance is available in period 2. Determining each individual's optimal consumption requires first solving the second-period optimization problem for each individual conditional on risk and demand type, which is known at that point in time, conditional on a given set of first-period choices (i.e., for s and L^1). We then use the value functions from the second-period optimization problem to determine optimal values for decision variables relating to the first period.

Second-period choice problem is, given type ij :

$$Z_{ij}^{spot} = \max_{L_{ij}^2} p_i v_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u_2(C_{ij}^{2N}), \quad (7)$$

where

$$C_{ij}^{2N} = y_2 + s - \pi_i^2 L_{ij}^2, \quad (8)$$

¹⁸The first-best may also be decentralized by a tax and transfer scheme that is type contingent.

$$C_{ij}^{2D} = y_2 + s + (1 - \pi_i^2)L_{ij}^2, \quad (9)$$

which leads to the first-order condition:

$$p_i v'_2(C_{ij}^{2D}; \theta_j)(1 - \pi_i) - (1 - p_i)u'_2(C_{ij}^{2N})\pi_i = 0.$$

When spot market prices are actuarially fair (i.e., $\pi_i^2 = p_i$), we have

$$v'_2(C_{ij}^{2D}; \theta_j) = u'_2(C_{ij}^{2N}), \quad (10)$$

in other words, ex post efficiency prevails.

Let Z_{ij}^{spot} be the value function relating to the second-period optimization problem. Since no GR insurance is available to purchase in period 1 for potential renewal in period 2, it follows that the only decision variable from period 1 that carries over to period 2 is s . Note that $Z_{ij}^{spot}(s)$ is strictly concave given our assumptions regarding the period 2 felicities, and via the envelope theorem we obtain:

$$\frac{\partial Z_{ij}^{spot}(s)}{\partial s} = p_i u'_2(C_{ij}^{2N}) + (1 - p_i)v'_2(C_{ij}^{2D}; \theta_j) = u'_2(C_{ij}^{2N}).$$

We go back to the first-period choice problem to complete the description of the optimal plan. In the first period, households choose savings and spot purchases to maximize expected utility:

$$\max_{s, L^1} EU = p v_1(C^{1D}) + (1 - p)u_1(C^{1N}) + (1 - p) \sum_i \sum_j q_i r_j Z_{ij}^{spot}(s), \quad (11)$$

where

$$C^{1N} = y_1 - s - \pi^1 L^1, \quad C^{1D} = y_1 + (1 - \pi^1)L^1. \quad (12)$$

First-order conditions are:

$$\frac{\partial EU}{\partial L^1} = p v'_1(C^{1D})(1 - \pi^1) + (1 - p)u'_1(C^{1N})(-\pi^1) = 0, \quad (13)$$

$$\frac{\partial EU}{\partial s} = (1 - p)u'_1(C^{1N})(-1) + (1 - p) \sum_i \sum_j q_i r_j u'_2(C_{ij}^{2N}) = 0. \quad (14)$$

Competition ensures first-period insurance is actuarially fair, $\pi^1 = p$, and so we get

$$v'_1(C^{1D}) = u'_1(C^{1N}), \quad (15)$$

$$u'_1(C^{1N}) = \sum_i \sum_j q_i r_j u'_2(C_{ij}^{2N}). \quad (16)$$

The last equation shows that the optimal savings amount equalizes the marginal utility of consumption in the first-period life state to the expected marginal utility of consumption in the second-period life state. Thus, marginal utilities will generally not be equal over time, confirming that spot insurance does not insure individuals against reclassification risk.

Proposition 2 (Characterization of Allocation Under Spot Insurance Only). *If the only markets for insurance in both periods is spot insurance, then it follows that:*

1. *Ex post efficiency (in period 2) prevails; that is, for a given risk type, demand type combination, marginal utility of consumption in the death state is equal to marginal utility in the life state.*
2. *Consumption in the life and death states in period 2 are not independent of risk type. Conditional on a given demand type, high-risk types have lower consumption in both life and death states of the world than do low-risk types. (This follows from the fact that high-risk types face a higher price of insurance.)*
3. *The period two consumption level in the death state for high-demand types of a given risk type is higher than that for low-demand types.*

Proof. See Appendix B.2. □

3.3 | GR insurance contracts

We now examine the model of primary interest; that is, the one where GR insurance is available. Information assumptions are the same as in the preceding model. In this case, however, in the second period the individuals hold an amount of GR insurance (L^{1G}) that they purchased in the first period. They may renew any amount of this ($L_{ij}^{2G} \leq L^{1G}$) in period 2 at the predetermined price π^{2G} . Individuals may also purchase spot insurance in period 2 (L_{ij}^2) which, since insurers also observe risk type, is priced at the risk type-specific actuarially fair price (p_i). Formally, individuals solve the following problem:

$$\max_{L^{1G}, s, \{L_{ij}^2, L_{ij}^{2G}\}} EU \equiv p v_1(C^{1D}) + (1-p) u_1(C^{1N}) + (1-p) \left[\sum_i \sum_j q_i r_j (p_i v_2(C_{ij}^{2D}; \theta_j) + (1-p_i) u_2(C_{ij}^{2N})) \right] \text{ s.t.} \quad (17)$$

$$C^{1N} = y_1 - s - \pi^1 L^1 - \pi^{1G} L^{1G}, \quad (18)$$

$$C^{1D} = y_1 + (1 - \pi^1) L^1 + (1 - \pi^{1G}) L^{1G}, \quad (19)$$

$$C_{ij}^{2N} = y_2 + s - \pi_i^2 L_{ij}^2 - \pi^{2G} L_{ij}^{2G}, \quad (20)$$

$$C_{ij}^{2D} = y_2 + s + (1 - \pi_i^2) L_{ij}^2 + (1 - \pi^{2G}) L_{ij}^{2G}, \quad (21)$$

$$L_{ij}^{2G} \leq L^{1G}, \quad (22)$$

$$\pi^{1G} L^{1G} = p L^{1G} + (1-p) \sum_i \sum_j q_i r_j (p_i - \pi^{2G}) L_{ij}^{2G}. \quad (23)$$

where the final constraint is the zero-profit condition on GR contracts for insurers.

We solve the dynamic optimization problem in two steps, solving backwards. In Step 1, we specify the second-period expected utility function for an individual of type ij , defined as Z_{ij} , conditional on arbitrary levels of first-period choice variable (s, L^1, L^{1G}). We find the optimal conditions from optimization in period 2. In Step 2, we use the value functions from the

second-period optimization problem, Z_{ij} , to determine optimal values for decision variables relating to the first period.

Second-period choice problem is, given type ij :

$$Z_{ij} = \max_{L_{ij}^2, L_{ij}^{2G}} p_i v_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u_2(C_{ij}^{2N}), \quad (24)$$

where

$$C_{ij}^{2N} = y_2 + s - \pi_i^2 L_{ij}^2 - \pi^{2G} L_{ij}^{2G}, \quad (25)$$

$$C_{ij}^{2D} = y_2 + s + (1 - \pi_i^2) L_{ij}^2 + (1 - \pi^{2G}) L_{ij}^{2G}, \quad (26)$$

$$L_{ij}^{2G} \leq L^{1G}. \quad (27)$$

We denote the multipliers for each type pair's constraint by λ_{ij} . The first-order conditions with respect to the choice variables are:

$$L_{ij}^2: p_i v'_2(C_{ij}^{2D}; \theta_j)(1 - \pi_i^2) - (1 - p_i) u'_2(C_{ij}^{2N}) \pi_i^2 = 0, \quad (28)$$

$$L_{ij}^{2G}: p_i v'_2(C_{ij}^{2D}; \theta_j)(1 - \pi^{2G}) - (1 - p_i) u'_2(C_{ij}^{2N}) \pi^{2G} - \lambda_{ij} = 0, \quad (29)$$

$$\lambda_{ij} (L^{1G} - L_{ij}^{2G}) = 0. \quad (30)$$

For the scenario with only spot insurance available, the resource constraints are trivial. That is, spot insurance is actuarially fair in each period which means $\pi^1 = p$, since individuals have the same first-period mortality risk. In period 2 the spot market price is $\pi_i^2 = p_i$, $i \in \{L, H\}$, since insurers observe risk type. There is an additional resource constraint for GR insurance since front loading must be sufficient to cover any second-period costs associated with any risk types renewing at a rate that is more favorable than their actuarially fair rate (e.g., for $\pi^{2G} < p_i$). The extent to which the first-period contract must be front loaded (i.e., the difference $\pi^{1G} - p$) depends on the extent to which the renewal price falls below the actuarial cost of providing risk types with insurance as well as the amount of L^{1G} that is purchased and amounts that will be renewed in equilibrium by risk types of both low and demand type.

Although insureds who turn out to be low-demand types but are of high-risk type may not renew all of L^{1G} , they have an incentive to renew more than would a low-demand type who is also of low-risk type since the price is more favorable to them. This means that low-demand types who are high-risk types typically end up with more second-period insurance coverage than their low-demand-low-risk counterparts.¹⁹ From the characterization of the social optimum, we know this cannot be efficient and so insureds would prefer contracts that are designed so this does not happen. However, once a person knows he is of high-risk type, he cannot “resist” renewing more insurance than is necessarily efficient even though, from the ex ante perspective, everyone would like to prevent such an outcome. This “over renewal” by HL – types creates undesirable adverse selection costs which must be covered by a combination

¹⁹For this to happen depends on both how different is the desired demand of these two types of individual as well as on how much GR insurance L^{1G} they hold entering the second period.

of increasing the front loading and/or the second-period renewal price compared to what would be required if such inefficient behavior could be controlled. The following equation explains this additional resource constraint which ensures zero expected profits for insurers offering GR insurance:

$$\pi^{1G} L^{1G} = p L^{1G} + (1 - p) \sum_i \sum_j q_i r_j (p_i - \pi^{2G}) L_{ij}^{2G}.$$

Note that the LHS of this equation represents the total revenue from sales of GR insurance in period 1. The first term of the RHS is the expected cost of insurance claims of GR insurance in period 1 while the second term is the sum of net expected costs of claims from all possible risk and demand types who pay π^{2G} to renew amount L_{ij}^{2G} of their holdings of GR insurance.

The zero-profit condition can also be written as follows:

$$\pi^{1G} = p + (1 - p) \sum_i \sum_j q_i r_j (p_i - \pi^{2G}) \frac{L_{ij}^{2G}}{L^{1G}}. \quad (31)$$

There are several important points regarding this constraint with some admittedly obvious. First, the amount of front loading per contract, as measured by the difference $\pi^{1G} - p$, is increasing with the (average) fraction of GR insurance holdings from the first period that is in fact renewed in the second period. It is also increasing in the amount of effective subsidy ($p_i - \pi^{2G}$) to each risk type i .²⁰ An increase in π^{1G} will affect the demand for GR insurance (L^{1G}) and so affect the amount of front loading that is required through the ratio $\frac{L_{ij}^{2G}}{L^{1G}}$. L^{1G} is also naturally a function of π^{2G} since GR insurance is more attractive the lower is its renewal price. This means that the way to control adverse selection problems arising from those who become low demand but high risk is not simply through increasing the renewal price as changes in both prices π^{1G} and π^{2G} affects the desirability of GR insurance.

To gain further insights into drivers of GR renewals, note that we write the first-order condition for L_{ij}^{2G} as follows:

$$v'_2(C_{ij}^{2D}; \theta_j) - u'_2(C_{ij}^{2N}) = \left(\frac{\pi^{2G} - p_i}{\pi^{2G}(1 - p_i)} \right) v'_2(C_{ij}^{2D}; \theta_j) + \frac{\lambda_{ij}}{\pi^{2G}(1 - p_i)}. \quad (32)$$

Equation (32) is helpful in understanding a number of possible scenarios to be discussed more fully below. Consider, for example, an individual who is both high risk and low demand and so renews some but not all of first-period GR ($0 < L_{HI}^{2G} < L^{1G}$). For such a person the shadow value of L^{1G} is zero ($\lambda_{HI} = 0$) and so the second term on the RHS of Equation (32) is zero. With the renewal price for high-risk types being below their actuarially fair rate ($\pi^{2G} < p_H$), the RHS of Equation (32) is negative; that is, $v'_2(C_{ij}^{2D}; \theta_j) < u'_2(C_{ij}^{2N})$ and so this person ends up in a position of overinsurance. In this sense, the renewal price is mispriced from the ex post (period 2) perspective and there are adverse selection costs created in the renewal market for GR. If an individual renews all of his first-period GR, then the shadow value of GR is positive ($\lambda_{HI} > 0$) and so the second term of the RHS of Equation (32) is positive, mitigating the influence of “mispricing” that leads to overinsurance. Because of the existence of spot markets, an

²⁰Clearly, there will be no market if the renewal price equals or exceeds the actuarially fair cost of insurance of high-risk types (i.e., if $\pi^{2G} \geq p_H$).

individual will never end up with too little insurance from the perspective of second-period consumption choice. However, whenever second-period spot markets are active, it follows that individuals are not fully protected against reclassification risk since high-risk types face a higher spot price.

We write value functions (indirect utilities) from this exercise as $Z_{ij}(L^{1G}, s; \pi^1, \pi_i^2, \pi^{2G})$. Using the envelope theorem, it follows that

$$\frac{\partial Z_{ij}}{\partial L^{1G}} = \lambda_{ij} \quad \text{for all } i, j, \quad (33)$$

$$\frac{\partial Z_{ij}}{\partial s} = p_i v'_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u'_2(C_{ij}^{2N}) \quad \text{for all } i, j. \quad (34)$$

This implies that for types that fail to fully renew their GR in period 2, increasing the quantity of GR ex ante has no impact on their welfare.

We now turn our attention to the period 1 optimization problem to complete the description of the optimal plan.

$$\max_{s, L^1} EU = p v_1(C^{1D}) + (1 - p) u_1(C^{1N}) + (1 - p) \left[\sum_i \sum_j q_i r_j Z_{ij}(\cdot) \right], \quad (35)$$

where

$$C^{1N} = y_1 - s - \pi^1 L^1 - \pi^{1G} L^{1G}, \quad (36)$$

$$C^{1D} = y_1 + (1 - \pi^1) L^1 + (1 - \pi^{1G}) L^{1G}. \quad (37)$$

First-order conditions are:

$$\frac{\partial EU}{\partial L^1} = p v'_1(C^{1D})(1 - \pi^1) - (1 - p) u'_1(C^{1N}) \pi^1 = 0, \quad (38)$$

$$\frac{\partial EU}{\partial L^{1G}} = p v'_1(C^{1D})(1 - \pi^{1G}) - (1 - p) u'_1(C^{1N}) \pi^{1G} + (1 - p) \sum_i \sum_j q_i r_j \lambda_{ij} = 0, \quad (39)$$

$$\frac{\partial EU}{\partial s} = -(1 - p) u'_1(C^{1N}) + (1 - p) \sum_i \sum_j q_i r_j \left[p_i v'_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u'_2(C_{ij}^{2N}) \right] = 0. \quad (40)$$

We can re-write the first-order condition on savings as follows:

$$v'_1(C^{1D}) = u'_1(C^{1N}) = \sum_i \sum_j q_i r_j \left[p_i v'_2(C_{ij}^{2D}; \theta_j) + (1 - p_i) u'_2(C_{ij}^{2N}) \right]. \quad (41)$$

This demonstrates that the optimal savings (or borrowing if negative) amount allows households to smooth consumption over time by equalizing marginal utility of consumption in the first-period life state to the expected marginal utility of consumption in the second period.

We can also re-write the first-order condition with respect to L^{1G} as follows:

$$v'_1(C^{1D}) - u'_1(C^{1N}) = \underbrace{-\frac{(1-p)\sum_i \sum_j q_i r_j \lambda_{ij}}{(1-p)\pi^{1G}}}_{\text{re-classification risk effect}} + \underbrace{\left(\frac{\pi^{1G} - p}{(1-p)\pi^{1G}} \right)}_{\text{mispricing effect}} v'_1(C^{1D}). \quad (42)$$

As for Equation (32), there is a mispricing effect illustrated in Equation (42) but from the perspective of period 1 insurance purchasing. Due to front loading ($\pi^{1G} > p$), the second term on the RHS being positive implies GR is too expensive for their first-period insurance needs and they would purchase too little. However, as long as the first-period spot insurance is active it follows that $v'_1(C^{1D}) = u'_1(C^{1N})$. The first term, which is negative if $\lambda_{ij} > 0$ for at least some ij pair, reflects the value of GR in providing protection against reclassification risk (i.e., at least some of $\lambda_{ij} > 0$). However, if expected demand for insurance in the second period is large relative to first-period demand, the value for GR insurance in providing protection against reclassification risk can lead to excessive first-period insurance coverage; that is, if all first-period insurance needs are more than met by L^{1G} (overinsurance in period 1) we would have $L^1 = 0$ and $v'_1(C^{1D}) < u'_1(C^{1N})$. The important conclusions are summarized in the proposition below:

Proposition 3. *Characterization of Allocation with GR Insurance Available*

If there are markets for both spot and GR insurance, then it follows that

1. Ex post efficiency in period 2 will generally not prevail. In particular, marginal utility in the death state may be less than marginal utility in the life state for high-risk types who are also low-demand types (overinsurance).
2. Consumption in the life and death states in period 2 are not necessarily independent of risk type. Conditional on a given demand type, high-risk types may have lower consumption in both life and death states of the world than do low-risk types. This follows if second-period spot purchases are non-zero due to the fact that high-risk types face a higher spot price of insurance.
3. The period two consumption level for high-demand types of a given risk type is at least as high as that for low-demand types.

There are various patterns of spot and GR insurance purchases that can arise. One important consideration is the amount of first-period insurance cover desired relative to second period for both demand types. We return to this point later. For now, suppose first-period demand for insurance is not so low as to deter sufficient purchases of GR to provide protection against reclassification risk (should the individual become a high-risk type). As is well known in the literature, if there is a single demand type and no financial friction or other burden, each individual will purchase just enough GR in period 1 to completely insure against reclassification risk. In such a case, no spot market purchases are made in the second period. Provided the condition on insurance demand in period 1 being sufficiently high relative to insurance demand in period 2 is met, this is also the case in our model with no demand differences. But more generally, individuals ex ante anticipate both possibilities of being low- and high-demand type as well as being either high or low-risk type. Suppose in anticipation of being a high-risk-high-demand type, an individual purchases sufficient GR insurance in period 1 that he fulfills his insurance desires through renewing at a favorable price and so doesn't access the spot market in period 2. This would provide him with full protection against reclassification risk as in the case

of a single demand type. Ex ante, however, the individual realizes that even if he becomes a high-risk type, he may have low demand and so, in this case, the required amount of insurance to avoid needing to access the spot market to fulfill his low-demand type insurance needs imposes an undesirable cost (through front loading). Therefore, the individual will not wish to hold as much GR insurance as he would if he knew he would be a high-demand type with probability one. If the amount of GR is strictly between the amount that would provide the “desired” or efficient amount of coverage for low- and high-demand types, then in period 2 the high-demand types would renew all of his first-period GR insurance and also purchase some spot insurance. The low-demand type will face a renewal price that is more favorable than his actuarially fair price and so renew more than what would provide him with the efficient level. This is a source of adverse selection in the renewal market.

Continuing with the set of circumstances described above, suppose that in the optimal contract the renewal price exceeds the spot market price for low-risk types. This may happen in equilibrium due to the adverse selection pressure on the renewal market as described above. Then those individuals who become low-risk types, regardless of demand type, will not renew any of their GR insurance purchased in period 1 (i.e., they choose to let their GR insurance lapse) and fulfill their second-period insurance needs through spot market transactions. Therefore, conditional on demand type, low-risk types in this scenario end up with equal marginal utility of consumption in life and death states. However, overall ex post efficiency is not obtained across all types. Moreover, high-demand-high-risk types end up satisfying some part of their insurance needs through purchases on the spot market at a higher price than do high-demand-low-risk types. Therefore, consumption in life and death states depends on risk type (second result of Proposition 3).

There are more factors that can arise in an equilibrium with GR contracts available and further impact conditions required for ex post efficiency. Although these are not without interest, the more important welfare measure is ex ante utility which of course is always at least weakly higher than with spot markets only. Welfare considerations are investigated more fully in the following section. However, some factors that affect both pricing and welfare implications are important to characterize an allocation when GR insurance is available. Conditional on risk type, high-demand types have the same marginal utility of consumption in the life state as do low-demand types, but higher marginal utility of consumption in the death state. From an ex ante perspective, an individual would therefore benefit from transfers from low to high-demand types for either risk type. As mentioned earlier, this can be achieved between high and low-demand types who are both low risk by having the renewal price lower even than the actuarially fair rate for low-risk types. Such pricing helps high-demand types who purchase (renew) a greater amount and so are advantaged by the low price. This may occur in equilibrium if demand type differences are sufficiently strong. However, any price lower than the actuarially fair price for high-risk types encourages high-risk-low-demand types to “renew too much” and this imposes (undesirable) upward pressure on front loading of first-period GR contracts. Given these two conflicting forces, the equilibrium renewal price may be below or above the actuarially fair rate for low-risk types.

Relatively low demand for first-period insurance is another factor that can compromise the value of GR markets in mitigating reclassification risk and also affect the prices (purchase price as well as renewal price) and quantity of GR insurance. The price of first-period GR contains one component to cover claims in period 1 and another factor representing the front loading required to cover losses due to subsidizing insurance for high-risk types in period 2. Suppose the amount of second-period coverage desired to be renewed from first-period GR insurance

purchases is large (balancing low- and high-demand possibilities) to protect against reclassification risk. Given the unattractiveness of a large amount being paid for a large quantity of first-period insurance, be it spot or GR, the ability of GR to satisfy the needs for protection against reclassification risk will be compromised. Individuals will not purchase very much GR in this scenario and high-risk types may end up with lower consumption than low-risk types of either demand type, due to higher spot market prices for high-risk types and this will have greater impact for high-demand types. It is even possible that all types will purchase some spot insurance in period 2 which would mean equal marginal utility of consumption in life and death states for any given type. However, avoiding wedges in these marginal utilities does not of course mean that markets are working efficiently from an ex ante welfare perspective. In fact, the high level of activity on spot markets suggests ex ante welfare is not being enhanced very effectively by the presence of GR insurance in such a scenario.

4 | WELFARE ANALYSIS OF GR CONTRACTS

Upon comparing Propositions 1–3, it appears that there are at least as many tendencies toward inefficiency when GR insurance is available compared to the situation in which only spot markets are available. However, the presence of GR insurance allows for individuals who turn out to be high-risk types to obtain some insurance coverage at a price below the actuarially fair rate. This ameliorates the inefficiency of reclassification risk (i.e., the pushing apart of consumption levels of any given demand type in both states of period 2 due to risk-based pricing in period 2 spot markets). However, GR insurance may also lead to the phenomenon that low-demand types who are also high-risk types will renew so much of their GR insurance that they end up with greater consumption in the death state of period 2 than that of low-demand but low-risk types. This reflects a type of ex post inefficiency (see the second statement of Proposition 1).

Proposition 4. *In the presence of both demand and risk heterogeneity, the equilibrium with GR and spot markets is always inefficient relative to first-best. GR contracts can achieve a first-best efficient allocation if and only if:*

1. *There is no heterogeneity of demand types. (i.e., demand for insurance is identical across individuals in period 2.)*
2. *The renewal price is sufficiently attractive and thus front loading sufficiently high that renewing GR is (weakly) preferable to purchasing spot insurance in period 2 for all individuals and there is no lapsation.*
3. *Demand for insurance is nonincreasing overtime.*

Proof. See Appendix B.3. □

The reason that one cannot have any heterogeneity of demand type is straightforward. A first-best allocation requires equality of the marginal utility of consumption in the second-period death for all (four) types. Providing full protection against reclassification risk for high-demand-high-risk types conflicts with ensuring low-demand-high-risk types do not end up with too much insurance. This is the source of the adverse selection in the renewal market and the reason that the second-period renewal price may be higher than the rate for low-risk types.

If there is only one demand type, then it is possible to achieve a first-best efficient allocation through GR insurance. This is achieved if it is optimal (in all senses) for individuals to hold the unique amount of first-period GR insurance such that when it is fully renewed by high-risk types (at the price equal to the actuarially fair rate for low-risk types) we end up with equal marginal utility of consumption in both life and death states for both risk types. The “in all senses” proviso is that such an amount, call it L^* , would not lead to overinsurance in period 1. If we refer to L^* as the required demand for second-period insurance to avoid any positive amount of reclassification risk, then this level of demand must not be greater than the optimal amount of first-period insurance. If L^* is too large, then “solving” the second-period optimal allocation problem through GR insurance creates an inefficient allocation for period 1 (i.e., overinsurance). As noted earlier, the average level of life insurance demand over the life cycle displays increasing demand when individuals are young (i.e., under 45 for males and under 40 for females). The two factors described above, therefore, suggest important limitations for the role of simple GR contracts for mitigating reclassification risk. As Proposition 5 demonstrates, however, introducing GR into a model with only spot markets will improve welfare and possibly substantially so.

Proposition 5. *Making GR insurance available alongside spot markets is always strictly welfare enhancing in the presence of reclassification risk.*

Proof. See Appendix B.4. □

It is obvious that, absent any pecuniary externalities that matter, adding a new market or type of contract will lead to a weak welfare improvement. We describe below the intuition behind the condition that ensures a strict welfare improvement is obtained. Although welfare may improve due to the addition of GR contracts when only demand type heterogeneity is present, this is not always the case. We explain this assertion as well.

The intuition for the proof is straightforward. Consider starting from a position of only spot market insurance being available with demand for first-period insurance being positive (although possibly “small”). Consider substituting a small amount of first-period spot insurance with GR which is renewable at a price at least slightly below the second-period spot price for high-risk types but above the spot price for low-risk types. This GR insurance will be renewed only by high-risk types. The envelope theorem guarantees that there is no first-order effect on welfare. However, there is a transfer of consumption from both risk types in period 1 (due to a small amount of front loading) to only high-risk types who have a higher marginal utility of consumption in period 2 due to their higher loss probability. High-risk types also face a higher price in the period 2 market for spot insurance. This transfer represents a first-order improvement in welfare.

Note that the above set of steps does not work if there is only demand type heterogeneity. The reason is that, to transfer consumption from a lower to a higher marginal utility state (i.e., from low to high-demand types in period 2) requires that the renewal price for GR insurance be less than the actuarially fair price in period 2 for all consumers (i.e., since there is only one risk type). Following the above steps, making a small amount of GR insurance available which is renewable at a price below the actuarially fair price of insurance in period 2 will lead to all consumers (i.e., both low and high-demand types) renewing this insurance. This implies a transfer of consumption from all types in period 1 to all types in period 2. Moreover, the effect of such a scheme will be undone by all consumers who would reduce savings in

period 1 to “re-establish” their privately optimal choices. Therefore, GR will not always provide a welfare improvement when there is only demand type heterogeneity.

A sufficient condition that will guarantee a welfare improvement in these circumstances (i.e., when only demand type heterogeneity persists) is that the demand for insurance in period 2 by low-demand types be less than first-period demand. In this circumstance, replacing an amount of first-period insurance equal to first-period spot demand plus a small amount extra with GR insurance that can be renewed at a price marginally below the actuarially fair price for second-period insurance (which is the same for all consumers) will transfer consumption from low-demand types to high-demand types who have higher marginal utility of income in the death state. This transfer will induce a first-order increase in welfare. This sufficient condition, however, is not to be taken lightly. Recall that demand for insurance on average is rising for “young” individuals and so even low-demand types may have higher demand in period 2 of our model than in period 1. The proof of this result is available from the authors upon request.

It is useful at this point to compare our assumptions and results to those of Daily et al. (2008) and Fang and Kung (2020). In their models, “low-demand types” are essentially “zero demand types.” That is, low-demand types are individuals who lose entirely their bequest motive and so have zero willingness to pay for any quantity of second-period life insurance. They adopt a contract structure which requires the insured to renew all or none of his first-period holdings of insurance. Moreover, they allow renewal terms to depend on risk (health) type. These assumptions allow for straightforward implementation of a no lapsation constraint since their low-demand types place zero value on the amount of second-period life insurance which therefore can be aimed solely at high-demand types. In our model, conditional on risk type, the first-best contracts for high and low-demand types must cost the same but offer higher coverage to high-demand types. Unlike “zero demand types,” our low-demand types place at least some positive value on second period life insurance coverage and so prefer the first-best contract of high-demand types to their own first-best contract. This generates an incentive compatibility constraint that creates efficiency problems.²¹ Furthermore, we do not allow contract renewals to be priced differentially according to realized risk type and individuals may hold GR contracts which allow for renewing part of their overall coverage rather than forcing “all or nothing” renewal decisions. The result is that high-risk types who are of low-demand type will wish to renew a greater amount of their first-period GR purchases than is efficient.

Although our goal is not to find the optimal (second best) design of GR or long-term contracts, absent any restrictions on contract space, but rather to uncover some of the properties of existing real world GR contracts, it is interesting to consider what sort of variations in contract terms could facilitate improved welfare. We do not engage in a comprehensive analysis of this question here, but some hints at improvements in contract structure follow from the above discussion. With our more general preference structure, “all or nothing” renewal terms could not generally generate first-best contracts. But suppose low-demand types place sufficiently small value on second-period life insurance that requiring insureds to either renew all or nothing of their coverage reduces adverse selection costs by discouraging low-demand types from in fact renewing. The welfare loss to high-risk but low-demand types forgoing renewals may be smaller than this welfare gain. However, individuals could purchase a pair of GR contracts in period 1 with one contract having a small level of coverage which those who become low-demand types would want to fully renew while those discovering that they are

²¹Note that this would still be an issue in our model if we were to adopt the “all or nothing” contract structure used in the Daily et al. (2008) and Fang and Kung (2020).

high-demand types would wish to fully renew both contracts. This strategy by insureds would undo the benefits of an all or nothing renewal requirement. Given that life insurance contracts are typically not sold under conditions of exclusivity (i.e., insureds are allowed to purchase as many contracts from as many insurers as they like)²² such strategies by insureds cannot be overlooked.

Since allowing for any additional instrument in contract design is likely to allow for some improvement in welfare, it is clear that introducing risk (health) type-specific renewal terms of insurance purchased in the first period may improve welfare. Similarly, further possible welfare improvements can be realized if, in addition, we allow in the second-period menus of risk type-specific contracts to enable sorting of demand types. However, it can be shown that for our model, allowing such flexibility in renewal terms could not generally lead to a first-best outcome. The reason is that the incentive compatibility constraint that must be satisfied to ensure that, for a given risk type, the low-demand type prefers his first-best contract to that of high-demand types cannot be satisfied since the cost of each of these first-best contracts must be the same and low-demand types will always value (at least a little) the higher coverage involved in the first-best contract of high-demand types. Some distortion is required.

5 | CONCLUSIONS

We have developed a two-period model of life insurance in which individuals face uncertainty over future changes in both mortality risk and insurance needs (bequest motives). In the first-period individuals are identical in all respects, including their current insurance needs and beliefs about how their risk and demand type will evolve in the second period. We allow for spot markets in each period as well as GR (or long-term insurance) that can be purchased in the first period. In the second period, individuals may become either a high or low-demand type as well as a high or low-risk type. GR insurance offers the potential to ameliorate reclassification risk. As in Pauly et al. (1995) and other previous work, we find that when individuals face only future risk type uncertainty, GR may allow individuals to fully insure against reclassification risk and achieve a first-best efficient allocation. However, if insurance demand is increasing over time, then one can only fully insure against reclassification risk by holding more (GR) insurance in period 1 than is desirable. The result is that a first-best efficient allocation is not possible.

Moreover, if there is also uncertainty about future demand, individuals do not know how much GR insurance would be ideal to hold for possible use in period 2. This creates problems for the efficiency of the renewals market. To offer protection against reclassification risk, the renewal price must be less than the actuarially fair price for (period 2) high-risk types. Therefore, individuals who turn out to be low-demand but high-risk type will face a renewal price that is below their actuarially fair price and will renew too much and so end up over-insured. This creates a type of adverse selection problem which leads to inefficiency. As a result, unlike in the model of Pauly et al. (1995), we have shown through simulations that the optimal GR contract may involve a renewal price which exceeds the actuarially fair price for (period 2) low-risk types and so leads to lapsation by some insureds. Through a series of propositions, we have shown that, although a first-best efficient outcome is not possible when there is both risk

²²Empirical evidence suggests linear pricing is a reasonable assumption in life insurance markets. See Cawley and Philipson (1999) and Pauly et al. (2003). This feature allows for such a strategy involving “partial renewal.”

and demand type uncertainty, adding GR contracts to spot contracts does improve social welfare (i.e., *ex ante* utilities).

We demonstrate through the use of simulations that, as in Fei et al. (2013), GR insurance can be effective in smoothing consumption across demand types and so can improve welfare even if there is no reclassification risk (i.e., individuals are of homogeneous risk type). The source of the welfare improvement is that first-period purchases of GR with favorable (“subsidized”) renewal terms allows for shifting second-period death state consumption toward those with higher marginal utility of consumption (i.e., toward high-demand types). However, as noted above, in the presence of both demand and risk type it may be that the optimal GR contract involves a renewal price above the actuarially fair price for low-risk types who allow their policies to lapse and purchase their second-period insurance needs from the spot market. In this scenario, which is not considered in Fei et al. (2013), high-demand types who are also low-risk types are excluded from the benefits of the “subsidy.” Although Polborn et al. (2006) consider a two-period model with both demand and risk type uncertainty, they do not model insurance needs in the first period. Therefore, as with Fei et al. (2013), they do not capture the importance of the interaction of these characteristics with the possibility of increasing demand over time which creates an additional obstacle for GR or long-term insurance to improve welfare.

Our model has shown the importance of identifying consumers who have higher marginal utility of consumption due to taste differences in regard to bequest motive and due to risk type (i.e., high-demand types of a given risk type have higher marginal utility in the death state as do high-risk types of a given demand type). We also demonstrate the importance of life cycle effects in demand for insurance. There are other reasons for changing preferences over time for insurance, including health shocks or income shocks which may increase or decrease marginal utility in the life state versus death states. Our analysis shows the importance of explicitly modeling such changes when analyzing welfare implications of GR or long-term versus short-term insurance contracts. Future work should use such explicit modeling strategies as reflective of circumstances both for understanding contracts and for any regulations that may be of interest (e.g., (partially) enforced GR of health insurance contracts as in the Affordable Care Act).

ACKNOWLEDGMENTS

Earlier versions of this paper were presented at meetings of the Western Risk and Insurance Association (2016), American Risk and Insurance Association (2016) and European Group of Risk and Insurance Economists (2017) as well as at the Munich Risk and Insurance Center (2016), International Conference on Game Theory (2016), ETH Zurich (2017), and CERNA Paris Mines (2017). We thank participants at those presentations for many useful comments. In particular, we thank Renaud Bourlès, James Hilliard, Annette Hofmann, Johannes Jasperen, Wanda Mimra, Richard Peter, Casey Rothschild, and George Zanjani. Any errors or shortcomings are the responsibility of the authors only.

ORCID

Afrasiab Mirza  <http://orcid.org/0000-0002-9035-5109>

REFERENCES

- Cawley, J., & Philipson, T. (1999). An empirical examination of information barriers to trade in insurance. *American Economic Review*, 89, 827–46.
- Cochrane, J. H. (1995). Time-consistent health insurance. *Journal of Political Economy*, 103, 445–473.
- Daily, G., Hendel, I., & Lizzeri, A. (2008). Does the secondary life insurance market threaten dynamic insurance? *American Economic Review: Papers and Proceedings*, 98, 151–156.
- Fang, H., & Kung, E. (2012). *Why do life insurance policyholders lapse? The roles of income, health and bequest motive shocks* (NBER Working Papers 17899). Cambridge, MA: National Bureau of Economic Research, Inc.
- Fang, H., & Kung, E. (2020). Life insurance and life settlements: The case for health-contingent cash surrender values. *Journal of Risk and Insurance*, 87, 7–39.
- Fang, H., & Wu, Z. (2017). *Life insurance and life settlement markets with overconfident policyholders* (NBER Working Papers 23286). Cambridge, MA: National Bureau of Economic Research, Inc.
- Fei, W., Fluet, C., & Schlesinger, H. (2013). Uncertain bequest needs and long-term insurance contracts. *The Journal of Risk and Insurance*, 82, 125–148.
- Frick, K. D. (1998). Consumer capital market constraints and guaranteed renewable insurance. *Journal of Risk and Uncertainty*, 16, 271–278.
- Gottlieb, D., & Smetters, K. (2019). *Lapse-based insurance*. Philadelphia: Wharton School, University of Pennsylvania. (Working Paper).
- Harris, M., & Holmstrom, B. (1982). A theory of wage dynamics. *Review of Economic Studies*, 49, 315–333.
- Hendel, I., & Lizzeri, A. (2003). The role of commitment in dynamic contracts: Evidence from life insurance. *Quarterly Journal of Economics*, 118, 299–328.
- Hofmann, A., & Browne, M. (2013). One-sided commitment in dynamic insurance contracts: Evidence from private health insurance in Germany. *Journal of Risk and Uncertainty*, 46, 81–112.
- Hofmann, A., & Peter, R. (2015). Self-insurance, self-protection, and saving: On consumption smoothing and risk management. *Journal of Risk and Insurance*, 83, 719–734.
- Hong, J. H., & Rios-Rull, J-V (2012). Life insurance and household consumption. *American Economic Review*, 102, 3710–3730.
- Mirrlees, J. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38, 175–208.
- Pauly, M. V., Kunreuther, H., & Hirth, R. (1995). Guaranteed renewability in insurance. *Journal of Risk and Uncertainty*, 10, 143–156.
- Pauly, M. V., Witters, K. H., Subramanian-Viswanathan, K., Lemaire, J., Hershey, J. C., Armstrong, K., & Asch, D. A. (2003). *Price elasticity of demand for term life insurance and adverse selection* (NBER Working Papers 9925). Cambridge, MA: National Bureau of Economic Research, Inc.
- Peter, R., Richter, A., & Steinorth, P. (2015). Yes, no, perhaps? Premium risk and guaranteed renewable insurance contracts with heterogeneous incomplete private information. *The Journal of Risk and Insurance*, 83, 363–385.
- Polborn, M. K., Hoy, M., & Sadanand, A. (2006). Advantageous effects of regulatory adverse selection in the life insurance market. *The Economic Journal*, 83, 327–354.
- Zietz, E. N. (2003). An examination of the demand for life insurance. *Risk Management and Insurance Review*, 6, 159–191.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Hoy M, Mirza A, Sadanand A. Guaranteed renewable life insurance under demand uncertainty. *J Risk Insur*. 2020;1–29.

<https://doi.org/10.1111/jori.12320>

APPENDIX A

FIGURE A1 Period 1.

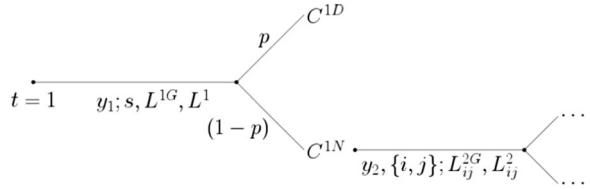
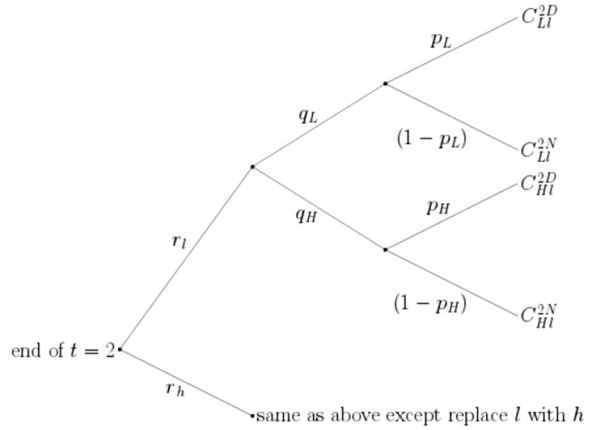


FIGURE A2 Period 2.



APPENDIX B: PROOFS

B. 1 Proof of Proposition 1

Denoting the Lagrange multipliers on the resource constraints in each period by λ_1 , and λ_2 , the necessary optimality conditions for an interior socially optimal allocation are:

$$C^{1D}: pv'_1(C^{1D}) = p\lambda_1, \quad (\text{B1})$$

$$C^{1N}: (1-p)u'_1(C^{1N}) = (1-p)\lambda_1, \quad (\text{B2})$$

$$s: (1-p)\lambda_1 = \lambda_2, \quad (\text{B3})$$

$$C_{ij}^{2D}: (1-p)q_i r_j p_i v'_2(C_{ij}^{2D}; \theta_j) = q_i r_j p_i \lambda_2, \quad (\text{B4})$$

$$C_{ij}^{2N}: (1-p)q_i r_j (1-p_i) u'_2(C_{ij}^{2N}) = q_i r_j (1-p_i) \lambda_2. \quad (\text{B5})$$

Combining (B1) and (B2) we obtain $v'_1(C^{1D}) = u'_1(C^{1N}) = \lambda_1$, and similarly combining (B4) and (B5) we obtain $v'_2(C_{ij}^{2D}; \theta_j) = u'_2(C_{ij}^{2N}) = \frac{\lambda_2}{1-p}$. Then, using (B3), we have

$$v'_1(C^{1D}) = u'_1(C^{1N}) = v'_2(C_{ij}^{2D}; \theta_j) = u'_2(C_{ij}^{2N}) \quad \text{for all pairs } (i, j) \in \{H, L\} \times \{h, l\}.$$

This implies that, for a given demand type, consumption in the period 2 death state is the same for both risk types and likewise for the period 2 life state consumption. However, consumption in the death state is higher for the high-demand type than for the low-demand type. This is easily established as

$$v'_2(C_{il}^{2D}; \theta_l) = v'_2(C_{ih}^{2D}; \theta_h) < v'_2(C_{il}^{2D}; \theta_h) \Rightarrow C_{ih}^{2D} > C_{il}^{2D}. \quad (B6)$$

Note also that the relationship between the period 2 death state consumption levels according to demand type is independent of risk type.

B. 2 Proof of Proposition 2

When only spot markets for insurance are available, the first-order conditions for the households are:

$$\begin{aligned} L^1: & p v'_1(C^{1D})(1 - \pi^1) + (1 - p) u'_1(C^{1N})(-\pi^1) = 0, \\ s: & -u'_1(C^{1N}) + \sum_i \sum_j q_i r_j \left(v'_2(C_{ij}^{2D}; \theta_j) \left(1 + (1 - \pi_i^2) \right) \frac{\partial L_{ij}^2}{\partial s} + u'_2(C_{ij}^{2N}) \left(1 - \pi_i^2 \frac{\partial L_{ij}^2}{\partial s} \right) \right) = 0, \\ L_i^2: & p_i v'_2(C_{ij}^{2D}; \theta_j) (1 - \pi_i^2) + (1 - p_i) u'_2(C_{ij}^{2N}) (-\pi_i^2) = 0. \end{aligned}$$

Actuarially fair spot insurance contracts require $\pi^1 = p$ and $\pi_i^2 = p_i$. Then, using the first and last conditions above we obtain:

$$\begin{aligned} v'_1(C^{1D}) &= u'_1(C^{1N}), \\ v'_2(C_{ij}^{2D}; \theta_j) &= u'_2(C_{ij}^{2N}) \quad \text{for all pairs } (i, j) \in \{H, L\} \times \{h, l\}. \end{aligned}$$

That is, marginal utilities are equated across life and death for all types ex post and also ex ante.

Now, differentiating the first condition with respect to p_i for a given level of savings, and solving for the change in insurance purchases we obtain:

$$\frac{\partial L_{ij}^2}{\partial p_i} = \frac{(v''_2(C_{ij}^{2D}; \theta_j) - u''_2(C_{ij}^{2L})) L_{ij}^2}{(1 - p_i) v''_2(C_{ij}^{2D}; \theta_j) + p_i u''_2(C_{ij}^{2L})}. \quad (B7)$$

For a given demand level θ_j , this is positive whenever returns to consumption diminish at a faster rate in the death state. Given that we assume this, higher risk types buy more insurance. Therefore, high-risk types consume less in both the life and death states than low-risk types as they also face higher prices.

Finally, differentiating the first-order condition with respect to θ_j for a given level of savings, and solving for the change in insurance purchases we obtain:

$$\frac{\partial L_{ij}^2}{\partial \theta_j} = \frac{-v'_2}{p_i (1 - \pi_i^2)^2 v''_2 + (1 - p_i) (\pi_i^2)^2 u''_2} > 0. \quad (B8)$$

This says that for a given risk level, higher demand types buy more insurance and therefore have more consumption than low-demand types.

B. 3 Proof of Proposition 4

Proof. We will first show that in the presence of fluctuations in demand, the use of GR contracts alone to insure against mortality risk is inefficient so that individuals always have incentives to use spot markets. Then, as long as spot markets are active, we will show that the equilibrium is inefficient as individuals do not receive full insurance against reclassification risk. However, we begin by demonstrating two important results that we will use throughout.

First, it is clear that if demand differences exist then in period 2 at the common price π^{2G} , higher demand types will want to purchase more coverage than low-demand types. Formally, suppose that low-demand types renew $L_{il}^{2G} \leq L^{1G}$ units of their GR. Then, the difference between their marginal utilities across life and death is always smaller at this amount of coverage than for the high-demand types:

$$\begin{aligned} & v'_2(y_2 + s + (1 - \pi^{2G})L_{il}^{2G}; \theta_l) - u'_2(y_2 + s - \pi^{2G}L_{il}^{2G}) \\ & < v'_2(y_2 + s + (1 - \pi^{2G})L_{il}^{2G}; \theta_h) - u'_2(y_2 + s - \pi^{2G}L_{il}^{2G}), \end{aligned}$$

as $\theta_h > \theta_l$ so high-demand types want more coverage.

Second, in period 2 all risk types of a given demand type want to purchase the exact same coverage at a common price if feasible. To see this note that full insurance is obtained for a type ij individual in period 2 by renewing $L_{ij}^{2G} \leq L^{1G}$ units of GR when

$$v'_2(y_2 + s + (1 - \pi^{2G})L_{ij}^{2G}; \theta_j) = u'_2(y_2 + s - \pi^{2G}L_{ij}^{2G}).$$

Then clearly L_{ij}^{2G} will differ across demand types but not across risk types as the above equation is independent of p_i . Now, when renewal of GR contracts are the sole means of obtaining insurance against mortality risk ex post, there are three possible outcomes in period 2:

1. Low-demand types fully renew: $\lambda_{il} > 0$. This implies that high-demand types also fully renew. However, as they want more insurance than low-demand types, they are under-insured and purchase additional spot insurance. Formally, when all types fully renew we have:

$$\begin{aligned} 0 & \leq v'_2(y_2 + s + (1 - \pi^{2G})L^{1G}; \theta_l) - u'_2(y_2 + s - \pi^{2G}L^{1G}) \\ & < v'_2(y_2 + s + (1 - \pi^{2G})L^{1G}; \theta_h) - u'_2(y_2 + s - \pi^{2G}L^{1G}), \end{aligned}$$

so high-demand types have an incentive to purchase spot insurance as they are under-insured when they only renew GR. This case arises when demand for insurance increases overtime for all types.

2. High-demand types do not fully renew: $\lambda_{ih} = 0$. This implies that low-demand types also do not fully renew their GR contracts as they demand less insurance. However, they purchase too much GR as is cheap. To see this, note that via (32) the foc on L_{il}^{2G} implies:

$$\begin{aligned} & v'_2(y_2 + s + (1 - \pi^{2G})L_{il}^{2G}; \theta_l) - u'_2(y_2 + s - \pi^{2G}L_{il}^{2G}) \\ &= \frac{\pi^{2G} - p_i}{\pi^{2G}(1 - p_i)} v'_2(C_{ij}^{2D}; \theta_j), \end{aligned} \quad (B9)$$

whenever $\lambda_{il} = 0$. However, note that if there is sufficient front loading $\pi^{2G} > p_i$ and therefore low-demand types are over-insured. This case arises when demand for insurance decreases overtime for all types.

3. High-demand types fully renew but low-demand types do not. In this case, we again have overinsurance by low-demand types for the exact same reason as in the previous case. This case arises when demand is increasing for high-demand types but decreasing for low-demand types over time.

To see the inefficiencies in the first-period note that in Case 1 above, (42) can be combined with the zero-profit condition to yield:

$$(1 - p) \sum_i \sum_j q_i r_j (p_i - \pi^{2G}) [v'_1(C^{1D}) - v'_2(C_{ij}^{2D}; \theta_j)] < 0,$$

whenever demand is decreasing over time for all types. This implies $v'_1(C^{1D}) - u'_1(C^{1N}) < 0$ or that there is overinsurance ex ante as excessive amounts of GR is purchased in period 1. Suppose, instead that we are in Case 2, then (42) implies

$$v'_1(C^{1D}) - u'_1(C^{1N}) = \frac{\pi^{1G} - p}{p(1 - \pi^{1G})} u'_1(C^{1N}) > 0,$$

as all the multipliers are zero and $\pi^{1G} > p$ due to front loading. This implies that individuals are under-insured and have incentives to purchase additional spot coverage ex ante.

Finally, note that if we are in Case 1 above (e.g., high-demand types desire more coverage in period 2), they always fully renew their GR and purchase additional spot insurance. The latter implies that consumptions depend on risk as the additional coverage is purchased at different prices for different risk types with the same demand. Hence, in equilibrium, individuals are not fully insured against reclassification risk. Now, suppose we are in Case 2 above (e.g., low-demand types desire less coverage in period 2), then the amount of GR low-demand types renew depends on their risk type. This is obvious by noting that the RHS of Equation (B9) depends on p_i . The same is true for Case 3. Hence, individuals do not receive full coverage against reclassification risk with demand fluctuations.

Now, to see the second-part of the result, note that from Proposition 1 first-best efficiency requires $v'_1(C^{1D}) = u'_1(C^{1N}) = v'_2(C_{ij}^{2D}; \theta_j) = u'_2(C_{ij}^{2N})$. Since spot prices differ by risk type, this means that all insurance in period 2 must be acquired through renewal

of L^{1G} (i.e., $L_{ij}^2 = 0$). This requires that $\pi^{2G} \leq p_L$ in order that all types (including the lowest risk type) at least weakly prefer to renew their holding of GR rather than access the spot market for period 2 insurance needs. Full insurance in period 2 requires the RHS of (32) be zero. Thus, we must have

$$v'_2(C_{ij}^{2D}; \theta_j) - u'_2(C_{ij}^{2N}) = \left(\frac{\pi^{2G} - p_i}{\pi^{2G}(1 - p_i)} \right) v'_2(C_{ij}^{2D}; \theta_j) + \frac{\lambda_{ij}}{\pi^{2G}(1 - p_i)} = 0, \quad (\text{B10})$$

and so

$$\left(\frac{\pi^{2G} - p_i}{\pi^{2G}(1 - p_i)} \right) v'_2(C_{ij}^{2D}; \theta_j) + \frac{\lambda_{ij}}{\pi^{2G}(1 - p_i)} = 0 \quad (\text{B11})$$

which implies that

$$\lambda_{ij} = (p_i - \pi^{2G})v'_2(C_{ij}^{2D}; \theta_j) \geq 0, \quad \forall i, j. \quad (\text{B12})$$

Since $\pi^{2G} \leq p_i$ for all i , it follows that $\lambda_{ij} \geq 0$ with strict inequality applying to all but the lowest risk type. Without loss of generality, we can assume that if $\pi^{2G} = p_L$ then $L - \text{types}$ (and hence all types) will renew all of L^{1G} . (For $L - \text{types}$, this follows by considering $\pi^{2G} = p_L - \varepsilon$ for $\varepsilon \rightarrow 0^+$ and relying on insurance demand being continuous in price. For all other risk types, $\lambda_{ij} > 0$ which implies $L_{ij}^{2G} = L^{1G}$). No lapsation and no second-period spot market activity means

$$C_{ij}^{2D} = y_2 + s + (1 - \pi^{2G})L^{1G}, \quad C_{ij}^{2N} = y_2 + s - \pi^{2G}L^{1G}. \quad (\text{B13})$$

Therefore, we can write $C_{ij}^{2D} = C^{2D}$ and $C_{ij}^{2N} = C^{2N}$, $\forall i, j$. It follows that $v'_2(C^{2D}, \theta_j) = u'_2(C^{2N})$, $\forall j$ which is possible only if θ_j does not vary with j ; that is $\theta_j = \theta$ for some $\theta > 0$.

We have now shown all conditions for period 2 that are required for first-best efficiency are met. We now need to consider conditions required for the first-period allocation to satisfy efficiency, for intertemporal efficiency to hold, and for the resource constraint to be satisfied. First-best efficiency also requires $v'_1(C^{1D}) = u'_1(C^{1N})$. Except for the possibility of a corner solution, $v'_1(C^{1D}) = u'_1(C^{1N})$ means $L^1 > 0$. Due to the requirements of no lapsation and no second-period spot market purchases, this means that demand for insurance in period 1 ($L^1 + L^{1G}$) must exceed (or in the case of $L^1 = 0$ be equal to) demand for insurance in period 2. This confirms requirement 3 of the proposition. We now need to check that the above conditions ensure intertemporal efficiency and satisfaction of the resource constraint. $v'_1(C^{1D}) = u'_1(C^{1N})$ implies that the RHS of (42) is zero; that is

$$-(1 - p) \sum_i \sum_j q_i r_j \lambda_{ij} + (\pi^{1G} - p) v'_1(C^{1D}) = 0 \quad (\text{B14})$$

which implies

$$(\pi^{1G} - p)v'_1(C^{1D}) = (1 - p) \sum_i \sum_j q_i r_j \lambda_{ij}. \quad (B15)$$

Using $\lambda_{ij} = (p_i - \pi^{2G})v'_2(C^{2D}; \theta)$ gives us

$$(\pi^{1G} - p)v'_1(C^{1D}) = (1 - p)v'_2(C^{2D}; \theta) \sum_i \sum_j q_i r_j (p_i - \pi^{2G}). \quad (B16)$$

Intertemporal efficiency implies $v'_1(C^{1D}) = v'_2(C^{2D}; \theta)$ and so we have

$$\pi^{1G} = p + (1 - p) \sum_i \sum_j q_i r_j (p_i - \pi^{2G}). \quad (B17)$$

The zero-profit condition is

$$\pi^{1G}L^{1G} = pL^{1G} + (1 - p) \sum_i \sum_j q_i r_j (p_i - \pi^{2G})L_{ij}. \quad (B18)$$

Therefore, no lapsation ($L_{ij}^{2G} = L^{1G}$) implies the above two equations are consistent; that is, the resource constraint is satisfied. This completes the proof. \square

B. 4 Proof of Proposition 5

Proof. Consider spot markets only equilibrium and let $\hat{L}^1, \hat{L}_{ij}^2, \hat{s}$ denote the corresponding equilibrium values. Recall that as spot markets are active, marginal utilities between life and death are always equated. Suppose now that an ε unit of GR is offered with some front loading ($\pi^{1G} > p$), and a second-period renewal price such that $p_L < \pi^{2G} < p_H$. Such a contract is always feasible by making π^{2G} arbitrarily close to (but less than) p_H and therefore π^{1G} arbitrarily close to (but above) p for any set of model parameters and ε . Moreover, such a contract is fully renewed by high risks as GR is cheaper relative to spot—they will substitute some spot for GR. However, such a contract does not affect the behavior of low risks—they continue to purchase the same amount of spot as before.

Then, $\lambda_{Hj} = (p_H - \pi^{2G})v'_2(y_2 + \hat{s} + (1 - p_H)\hat{L}_{Hj}^2; \theta_j)$ and $\lambda_{Lj} = 0$, and the change in welfare from the marginal unit of GR is non-negative if:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{\partial EU}{\partial L^{1G}} \Big|_{L^{1G} = L_{Hj}^{2G} = \varepsilon} &= p(1 - \pi^{1G})v'_1(y_1 + (1 - p)\hat{L}_1) - (1 - p)\pi^{1G}u'_1(y_1 - \hat{s} - p\hat{L}_1) \\ &+ (1 - p)(q_H r_h \lambda_{Hh} + q_L r_l \lambda_{Hl}) > 0. \end{aligned} \quad (B19)$$

Note that the zero-profit condition on the GR contract implies:

$$\pi^{1G} - p = (1 - p)q_H(p_H - \pi^{2G}), \quad (B20)$$

as it is fully renewed by high risks only. Using this expression and the definitions of λ_{ij} we obtain

$$\begin{aligned}
 & p(1 - \pi^{1G})v'_1(y_1 + (1 - p)\hat{L}_1) - (1 - p)\pi^{1G}u'_1(y_1 - \hat{s} - p\hat{L}_1) + (1 - p)(q_H r_h \lambda_{Hh} + q_L r_l \lambda_{Hl}) \\
 &= \sum_j q_H r_j (p_H - \pi^{2G}) \left[v'_2(y_2 + \hat{s} + (1 - p_H)\hat{L}_{Hj}^2; \theta_j) - v'_1(y_1 + (1 - p)\hat{L}_1) \right] \\
 &= q_H (p_H - \pi^{2G}) \sum_j r_j \left[v'_2(y_2 + \hat{s} + (1 - p_H)\hat{L}_{Hj}^2; \theta_j) - v'_1(y_1 + (1 - p)\hat{L}_1) \right].
 \end{aligned}$$

Finally, using the first-order condition on savings, we have

$$\lim_{\varepsilon \rightarrow 0} \frac{\partial EU}{\partial L^{1G}} \bigg|_{L^{1G} = L_{Hj}^{2G} = \varepsilon} = \frac{q_H (p_H - \pi^{2G}) q_L \sum_j r_j [v'_2(y_2 + \hat{s} + (1 - p_H)\hat{L}_{Hj}^2; \theta_j) - v'_2(y_2 + \hat{s} + (1 - p_L)\hat{L}_{Lj}^2; \theta_j)]}{1} > 0, \quad (\text{B21})$$

as whenever there is reclassification risk, high risks have higher marginal utility in the death state than low risks. \square