Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

What's Wrong with Annuity Markets? [2021044]

Stéphane Verani and Pei Cheng Yu

2021-044

Please cite this paper as:
Verani, Stéphane, and Pei Cheng Yu (2021). "What's Wrong with Annuity Markets? [2021044]," Finance and Economics Discussion Series 2021-044. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2021.044.

What's Wrong with Annuity Markets?*

Stéphane Verani[†] Pei Cheng Yu[‡]

July 13, 2021

Abstract

We show that the supply of life annuities in the U.S. is constrained by interest rate risk. We identify this effect using annuity prices offered by U.S. life insurers from 1989 to 2019 and exogenous variations in contract-level regulatory capital requirements. The cost of interest rate risk management accounts for at least half of the average life annuity markups or eight percentage points. The contribution of interest rate risk to annuity markups sharply increased after the great financial crisis, suggesting new retirees' opportunities to transfer their longevity risk are unlikely to improve in a persistently low interest rate environment.

JEL Codes: G10; G22; G32

Keywords: life insurance; annuities; corporate bond market; retirement; interest rate risk

*We would like to thank, without implication, Toni Braun, Jeff Brown, Celso Brunetti, Dean Corbae, Nathan Foley-Fisher, Stefan Gissler, Jim Hines, Richard Holden, Ivan Ivanov, Narayana Kocherlakota, Olivia Mitchell, Andrew Melnyk, Borghan Narajabad, Ali Ozdagli, Radek Paluszynski, David Rahman, Coco Ramirez, Adam Solomon, as well as seminar participants at Academia Sinica, Atlanta FRB, National Tsing Hua University, and the University of Michigan. Sam Dreith provided exceptional research assistance. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

[†]stephane.h.verani@frb.gov, +1-202-912-7972, 20^{th} & C Streets, NW, Washington, D.C. 20551.

[‡]pei-cheng.yu@unsw.edu.au, +61-2-9385-3704, West Lobby Level 4. UNSW Business School building. Kensington Campus UNSW Australia

Introduction

The fundamental risk for retirement is unknown longevity. Life annuities offer a unique risk transfer solution to new retirees wishing to shed the risk of outliving their financial wealth (Yaari 1965, Mitchell, Poterba, Warshawsky & Brown 1999, Davidoff, Brown & Diamond 2005). However, falling long-term interest rates from the late 1980s have eroded the profitability of the life annuity business (Foley-Fisher, Narajabad & Verani 2020). A natural question is, how will historically low interest rates affect new retirees' opportunities to manage their longevity risk? Answering this question is crucial for policymakers, as the provision of social insurance depends on the conditions in private longevity insurance markets (Cutler & Gruber 1996, Golosov & Tsyvinski 2007). Examining how interest rate risk affects the supply of annuities requires identifying the sources of market inefficiencies that influence longevity insurance markets.

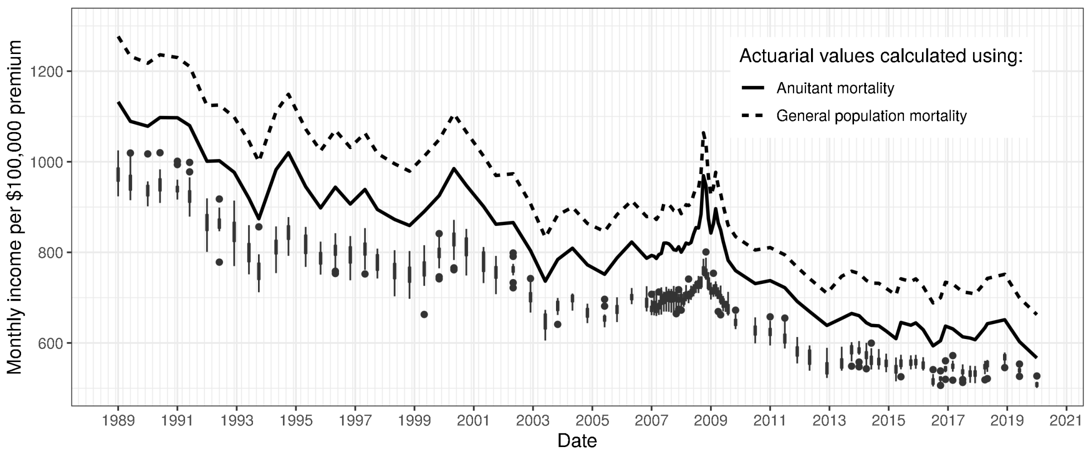


Figure 1: Actual and actuarial monthly payment for a nominal $100,000 SPIA offered to a 65-year-old male.

We develop an algorithm for annuity valuation that decomposes the contribution of demand- and supply-side frictions in annuity markups observed from 1989 to 2019. Figure 1 plots the distribution of actual monthly payments offered by U.S. life insurers to a 65-year-old male purchasing a $100,000 single premium immediate life annuity (SPIA) (the vertical box plots) together with the monthly payments implied by the actuarial value of the annuity contract calculated using the general population mortality (the dashed line) and annuitant mortality (the solid line). The difference between the dashed and solid lines is a measure of the industry's average *adverse selection pricing*. This well-known source of demand-side inefficiency arises because life insurers do not observe the mortality risk of individuals seeking longevity insurance, leading to adverse selection (Eichenbaum & Peled 1987, Finkelstein & Poterba 2004). The difference between the actual monthly payment offered and the solid line is the *adverse selection adjusted annuity price markup* (henceforth, *AS-adjusted markup*). The average *AS-adjusted markup* is substantial and around 16 percent.

We show that the cost of managing the interest rate risk associated with selling life annuities accounts for at least half of the *AS-adjusted markup* or eight percentage points. That is, in addition to the well-known cost of adverse selection, the supply of private longevity insurance is constrained by life insurers' own vulnerability to uninsurable aggregate shocks.[1] This supply-side inefficiency that arises from financial frictions affects life insurers' product design and capital structure decisions. Moreover, we show that the contribution of interest rate risk to the *AS-adjusted markup* sharply increased after the great financial crisis (GFC), in the aftermath of unprecedented actions by central banks around the world that accelerated the decrease of long-term interest rates.

Our results are important for two reasons. The first reason is that the risk management channel provides a novel explanation for the low wealth annuitization in the U.S. A large literature attributes low annuity demand to high markups (Mitchell et al. 1999). Our results show that a substantial fraction of annuity markups reflects insurers' cost of managing interest rate risk. The effect of interest rate risk, a supply-side friction, is likely to add to the adverse effects of other noted demand-side frictions on annuity demand including, for example, bequest motives (Friedman & Warshawsky 1990, Bernheim 1991, Lockwood 2011, 2018), behavioural biases (Brown et al. 2008, 2017, Gottlieb 2018), and pre-existing annuitization, such as social security (Bernheim 1991, Dushi & Webb 2004).

The second reason our results are important is that the risk management chan-

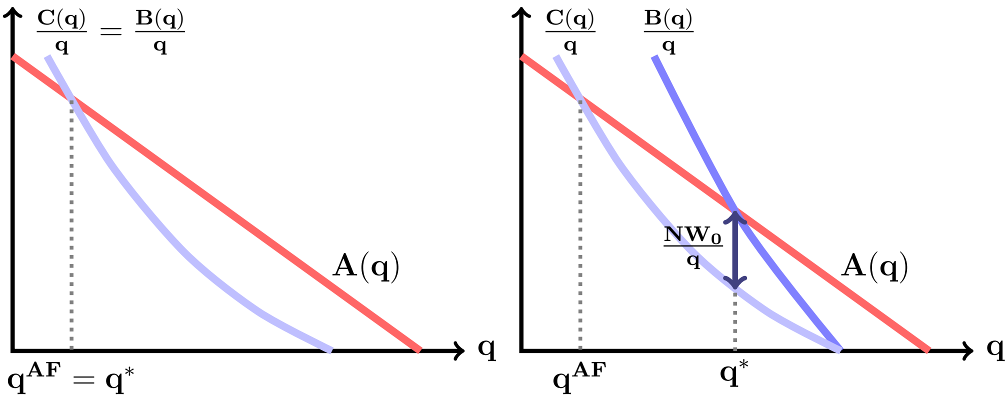
[1] Cutler (1996) makes a similar point focusing on long-term care insurance.

nel has significant implications for the macroeconomic literature studying the welfare effects of social insurance programs using life-cycle models—e.g., Hong & Ríos-Rull (2007), Hosseini (2015), Hosseini & Shourideh (2019). This literature typically models life insurers as profit-maximizing firms operating in frictionless financial markets. Under this assumption, life insurers costlessly hedge interest rate risk and their capital structure is irrelevant. In contrast, we identify the cost of hedging interest rate risk as a key friction shaping the private supply of longevity insurance. Therefore, the ubiquitous result from this literature that social insurance does not significantly improve welfare may not hold when taking into account the cost of managing interest rate risk.

We provide a model of annuity pricing in a market with adverse selection—a demand-side friction—and interest rate risk—a supply-side friction. In the model, interest rate risk arises because there is aggregate uncertainty over future interest rates, and corporate debt maturity is constrained to be relatively short (Bolton & Scharfstein 1990, 1996, Hart & Moore 1994, 1998, Huang et al. 2019). As a result, the duration of the bonds life insurers use to fund their annuities (the asset side of the insurers' balance sheet) is less than the duration of their annuity liabilities (Domanski, Shin & Sushko 2017). This negative duration gap means that the present value of life insurers' annuity liabilities increases faster than the present value of their bond holdings when long-term interest rates decrease relative to short-term interest rates (i.e., the yield curve flattens), which could lead to insolvency.

Figures 2a and 2b summarize our main theoretical results. The horizontal axis indexes the annuity price q. The vertical axis indexes the market demand for annuities, which is denoted by $A(q)$. The insurers' asset portfolio and capital structure matter in determining the equilibrium annuity price because financial markets are inefficient. The competitive equilibrium price q^* is determined by the intersection of the demand curve $A(q)$ and the insurers' average bond demand curve $B(q)/q$, which is the amount of bonds insurers demand per dollar of annuities sold.

When the bond market is unconstrained (Figure 2a), life insurers invest their



(a) Unconstrained bond market **(b)** Constrained bond market

Figure 2: Equilibrium annuity price

annuity sales revenue in an optimal bond portfolio that perfectly hedges the interest rate risk. As a result, the insurers' average cost curve $C(q)/q$, which is the cost per dollar of annuities sold and is downward sloping due to adverse selection, is equal to the average bond demand curve $B(q)/q$. The equilibrium annuity price q^* is such that the *AS-adjusted markup* is zero, i.e., $q^* - q^{AF} = 0$, where q^{AF} is the actuarially fair price for the average annuitant. On the other hand, when the supply of long-term bonds is constrained (Figure 2b), insurers hedge the interest rate risk with a positive level of net worth per dollar of annuities sold ($NW_0/q > 0$) to cushion the adverse effects of future interest rate shocks and prevent insolvency. In this case, the average bond demand curve $B(q)/q$ is higher than the average cost curve $C(q)/q$ because insurers finance their net worth by charging a positive *AS-adjusted markup*, i.e., $q^* - q^{AF} > 0$.

We identify the effect of interest rate risk management on annuity markups using annuity prices offered by U.S. life insurers from 1989 to 2019 and exogenous variations in contract-level regulatory capital requirements. Identifying this effect is difficult because demand- and supply-side frictions jointly contribute to relatively high annuity markups. We address this identification challenge by exploiting shocks that differentially affect the average cost curve $C(q)/q$ and the average bond demand curve $B(q)/q$ of *different annuity contracts* offered by the *same insurer*. Put differently, our identification strategy exploits shocks to the spread between what life insurers can earn by investing new annuity money and what they credit as interest on new annuity contracts, which is known in the industry as the net investment spread.

We find that insurers raise their AS-adjusted markups when the net investment spread on new annuity business decreases as a result of an exogenous increase in regulatory reserve requirement. However, this effect is substantially smaller when the insurers' potential yield from investing new annuity money exogenously increases. This difference-in-difference result identifies life insurers managing interest rate risk with net worth, as the marginal cost of interest rate risk management decreases when the net investment spread on new annuity business increases. Moreover, by exploiting the difference between five-year term certain annuity markups and life annuity markups offered by the *same* insurer at the *same* time, we find that the average marginal cost of risk management accounts for almost all of the *AS-adjusted markup* after factoring in standard operating expenses.

We confirm the economic mechanism by conducting several additional robustness tests of our findings, each relying on a different set of identifying assumptions. First, we focus on the cross-sectional variation of insurers' interest rate hedging programs. Using the universe position-level interest rate swap derivatives data between 2009 and 2015, we show that life insurers that are relatively more adversely affected by an unexpected change in the shape of the yield curve due to their ex-ante interest rate derivative position disproportionately increase their AS-adjusted markup. This identification strategy exploits the unusual zero lower bound period from 2009 to 2015 during which all the movements in the yield curve came from fluctuations in the long end of the curve. Second, we use a quantile fixed-effect regression to show that the least competitive insurers that are the most beneficially affected by interest rate shocks (due to their hedging programs) cut their AS-adjusted markups the most.

Our paper contributes to several strands of literature. First, we bridge the gap between the economic literature on adverse selection in insurance markets and the finance literature on risk management of financial institutions—e.g., Froot &

Stein (1998), Foley-Fisher, Narajabad & Verani (2016), Ozdagli & Wang (2019). Our theory of annuity pricing differs from the textbook model of adverse selection in insurance markets—e.g., Einav & Finkelstein (2011)—because insurers must credibly show to annuity shoppers that they are managing interest rate risks with a unique ex-ante capital structure. Second, we contribute to the recent literature identifying the effect of supply-side frictions in insurance markets. Previous research has identified the general effect of financial constraints on insurer product design and capital structure decisions—e.g., Koijen & Yogo (2015), Ge (2019). We differ from these papers by identifying the specific financial friction—interest rate risk—that affects the supply of insurance contracts.

1 Selling and managing fixed annuities

In this section, we provide some background about the U.S. annuity market. New retirees can manage the risk of outliving their financial wealth by purchasing a life annuity from a life insurer either directly or through their employer's pension plan. An individual purchasing a life annuity contract transfers its idiosyncratic longevity risk to the life insurer by surrendering his or her wealth in exchange for a stream of payments while he or she is alive. In the remainder of this section, we briefly discuss the U.S. life annuity market.

Roughly half of the U.S. life insurance industry's $600 billion aggregate income in 2018 came from annuity considerations.[2] About half of the industry's annuity considerations relate to fixed annuity sales, which are annuity contracts for which the principal is backed by the life insurer's general account.[3]

Estimating the size of the U.S. life annuity market is difficult. The vast majority of fixed annuities sold in the U.S. are deferred fixed annuities. Deferred fixed annuities are purchased by individuals before their retirement age to accumulate

[2]The other half is roughly split between life and health insurance premiums. See the ACLI's 2018 Life Insurers Fact Book https://www.acli.com/posting/rp18-007.

[3]Another large type of annuities are variable annuities for which the underlying assets are segregated from the life insurer's general account and remain the property of the variable annuity contract holders—see, for example, Koijen & Yogo (2018).

wealth on a tax-deferred basis during the contract period. At the end of the deferred fixed annuity contract period and after reaching the age of 59.5 years of age, contract holders have the option of receiving their accumulated wealth as a lump sum, a term annuity, or a life annuity.[4] It is generally difficult to precisely separate life annuity sales from deferred annuity sales in regulatory filings.

Using insurer-level data on the number of annuity contracts and account balances reported in the 2018 NAIC Statutory Fillings of over 800 life insurers, we estimate that Americans annuitize about $625 billion of their wealth with life insurers. This amount corresponds to approximately $12,700 per person aged 65 years and above or about 3.5 percent of the total payments made by the US Social Security Administration in 2018.[5] These two estimates are consistent with the view in the literature that the market for immediate annuities in the U.S. is small (Mitchell et al. 1999).[6]

1.1 The fixed annuity business model

Life insurers' overall business model consists of earning a spread between the yield they promise on their insurance liabilities and the yield they earn on the assets backing these liabilities. Life annuities and life insurance contracts are fixed-rate liabilities that are illiquid, as they are not transferable from one individual to another. Consequently, life insurers tend to invest their annuity considerations and premiums primarily in fixed-income securities in an effort to match their asset and liability cash flows. The illiquidity of life insurance liabilities allows insurers to invest considerations and premiums in relatively illiquid fixed income, such as corporate bonds, asset-backed securities and real estate loans to offer a competitive

[4]Note that variable annuities are a type of deferred annuity and an alternative to fixed deferred annuities. As with fixed deferred annuities, an individual must chose a payout option for his or her accumulated wealth at the end of the variable annuity contract period, which includes signing up for a life annuity.

[5]See Appendix A for more details about these calculations.

[6]The literature has proposed several theories to account for the low wealth annuitization demand in the US. This includes, for example, bequest motives (Friedman & Warshawsky 1990, Bernheim 1991, Lockwood 2011, 2018), behavioural biases (Brown et al. 2008, 2017, Gottlieb 2018), and pre-existing annuitization such as social security (Bernheim 1991, Dushi & Webb 2004). Our focus in this paper is to empirically identify a specific friction that shapes the supply of annuities, taking as given any of these plausible demand-side frictions.

return to policyholders and compensate them for bearing the insurance contract's illiquidity.

U.S. life insurers have been the largest institutional investor in corporate bonds issued by U.S. corporations since the 1930s. At the end of 2017, U.S. life insurers held about $2.1 trillion of corporate bonds in their general account, about half of their general account assets and roughly one-third of the total corporate bond amount outstanding in the U.S. (ACLI 2018).[7] By comparison, the rest of the life insurers' general account assets includes 8 percent in U.S. government securities and 14 percent in mortgage-backed securities, including those backed by the U.S. government.[8]

1.2 Interest rate risk management

The duration of life insurers' assets is typically less than the duration of their insurance liabilities because the maturity of corporate debt is typically much shorter than the duration of life insurance liabilities. For example, corporate bonds have a median initial maturity of about 5 years, and over 90 percent have an initial maturity that is 10 years or less. This maturity structure contrasts with the duration of a life annuity offered to a 65-year-old individual, which is approximately 10 years.[9] Moreover, long-duration U.S. government securities are unattractive to life insurers because they carry a substantial liquidity premium. This negative duration gap means that a decrease in the interest rates increases the present value of a life insurer's fixed-rate liabilities faster than the present value of its fixed-income assets, which could lead to insolvency. Because the prospect of insolvency is incompatible with the sale of life annuities, interest risk management (henceforth,

[7]General account assets back a life insurer's insurance liabilities.

[8]The life insurance industry's relatively low holding of U.S. government securities, which have relatively lower yields, reflects their substantial liquidity premium. This liquidity premium means that investing annuity considerations in U.S. government securities is unprofitable because life insurers must compensate annuity contract holders for the illiquidity they bear when signing up for a life annuity. Moreover, backing long-term insurance liabilities with government securities creates additional problems for life insurers during times of overall market stress, as the market value of government securities typically moves in the opposite direction of the market value of the insurer's liabilities (Bailey 1862).

[9]Note that the duration of a fixed-income instrument is less than or equal to its maturity.

IRM) is at the heart of the modern insurer's annuity business model.

Life insurers primarily manage interest rate risk by maintaining a suitable level of net worth, which is also referred to as surplus in the industry.[10] Net worth helps cushion the effect of interest rate changes that disproportionately affect the value of the life insurer's insurance liabilities. Because accumulated retained earnings are the primary source of life insurer capital, the cost of preserving net worth is reflected in annuity prices. That said, the effect of IRM on annuity pricing is typically absent from economic models with life annuities that assume frictionless financial markets—e.g., Yaari (1965), Davidoff et al. (2005), Hosseini (2015). We formalize the relationship between IRM and annuity prices in the next section.

2 Pricing with adverse selection and interest rate risk

In this section, we show how life insurers set prices in an annuity market with adverse selection and interest rate risk. We introduce two frictions, which are absent from the adverse selection literature. The first friction is that life insurers are protected by limited liability, which means that the owners of a life insurer are not liable for corporate losses that exceed the value of the insurer's assets. The second friction is that the bond market is constrained, such that debt maturity is relatively short. Competitive life insurers have an incentive to manage interest rate risk with net worth because annuity shoppers can rationally anticipate the prospect of insolvency associated with certain portfolio and capital structure choices. The cost of net worth is reflected in annuity prices that are higher than the contract's actuarial value.

2.1 Economic environment

The economy is populated by a continuum of new retirees with wealth normalized to 1 and lasts for three periods: $t = 0, 1, 2$. Each individual survives from period to period with survival probability α, which is drawn at the beginning of

[10] To a lesser extent, larger and more sophisticated life insurers use derivatives in conjunction with net worth to hedge interest rate risk (Berends et al. 2015), which we analyze in Section 6.

$t = 0$ from c.d.f $G(\alpha)$ with support $[\underline{\alpha}, \overline{\alpha}] \subset [0,1]$ and p.d.f $g(\alpha)$. The survival probability α is the individuals' private information. Every individual is deceased at the end of $t = 2$.

There are two types financial instruments in the economy that can be used to transfer wealth across periods. First are annuity contracts offered by life insurers in $t = 0$. An annuity contract pays one unit of consumption in each period the contract holder is alive in exchange for a lump sum payment q in $t = 0$.[11] The annuity market is competitive and life insurers compete over annuity prices.[12] Second are one- and two-period zero-coupon corporate bonds issued by non-financial firms, which we do not model explicitly. Corporate bonds only differ by their maturity and, therefore, we do not keep track of the face value of each individual bond's principal outstanding. One unit of the one-period bond returns $R_1 \geq 1$ in $t = 1$ and R_2 in $t = 2$, where $R_2 \in \left[1, \overline{R}\right]$ with mean $\mathbb{E}(R_2)$ and is realized in $t = 1$. The two-period bond, which is the long-term bond in our model, is priced efficiently in $t = 0$, such that R_l satisfies $\frac{1}{R_l} = \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$.

We do not make explicit assumptions about the individuals' consumption and investment decisions. Instead, we require that the annuity demand $a(\alpha, q)$ of individuals with survival probability α satisfies Assumption 1.

Assumption 1 *The individual annuity demand $a(\alpha, q)$ satisfies: (i) $a(\alpha, q)$ is differentiable in α and q, with $\frac{\partial a}{\partial \alpha} > 0$ and $\frac{\partial a}{\partial q} < 0$; (ii) there exists $\alpha \in (\underline{\alpha}, \overline{\alpha})$ such that $a(\alpha, q) > 0$ when $q = \frac{\overline{\alpha}}{R_1}(1 + \overline{\alpha})$; and (iii) $a(\alpha, q) = 0$ for all α and q if there is a positive probability that the insurer is insolvent in period $t \geq 1$ and $a(\alpha, q) \geq 0$ otherwise.*

The first condition of Assumption 1 follows from the adverse selection literature. It requires individuals with higher survival risk to have a higher demand for

[11] We do not consider the effects of screening through the offering of multiple contracts. Instead, we focus on life insurers offering a single contract. This assumption is consistent with actual life insurers offering the same SPIA contract to individuals of the same age and gender—i.e., life insurers do not screen the annuitants beyond age and gender.

[12] Unlike variable annuities for which life insurers compete over prices and product characteristics (Koijen & Yogo 2018), fixed annuities are standardized products and insurers compete over prices. Nevertheless, we consider an extension of our benchmark environment with monopolistic competition in Appendix F.

annuities and that the demand for annuities to be downward sloping. The second condition requires the annuity demand to be strictly positive even when insurers break even on individuals with the highest survival probability and interest rate R_2 is at its lowest level—i.e., $R_2 = 1$. This condition ensures that there is a market for annuities and an equilibrium price exists. The third condition requires that the demand for annuities from insurers with a strictly positive probability of becoming insolvent to be zero. A stark interpretation of this condition is that an annuity contract is worthless to individuals if there is a positive probability that the insurer may not honor its contractual obligations. A more nuanced interpretation is that an un-modelled insurance regulator or a rating agency requires insurers to hold minimum annuity reserves to prevent insolvency. As a reminder, $a(\alpha, q)$ is the demand for annuities at $t = 0$, so it is implicitly a function of R_1 and $\mathbb{E}(R_2)$, but not of the realized value of R_2. Figure 3 summarizes the timing of the model.

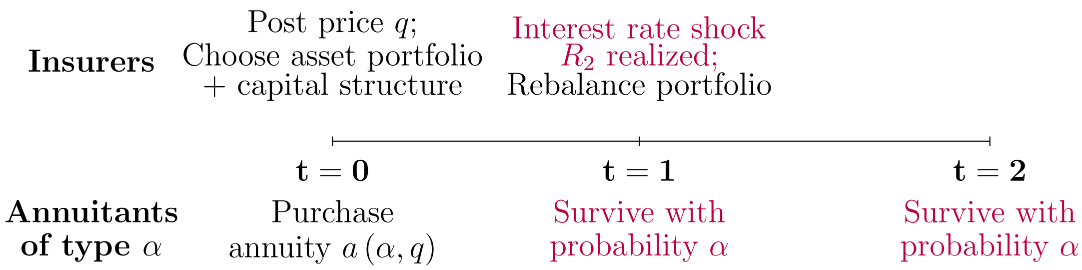


Figure 3: Model timeline

2.2 Interest rate risk management

An insurer is insolvent in $t = 1$ if the present value of its liabilities exceeds the present value of its assets once the aggregate shock R_2 is realized. Life insurers are concerned about their solvency because they operate under limited liability and rationally anticipate that they will not sell annuities at $t = 0$ if there is a chance they become insolvent at $t = 1$.

The dynamics of a life insurer's balance sheet is as follows. In $t = 0$, the insurer invests its annuity considerations (the revenue from the annuity sales in $t = 0$) in a portfolio of bonds (b_1, l_2) :

$$b_1 + l_2 = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q)\, g(\alpha)\, d\alpha, \tag{1}$$

where b_t and l_2 denote the insurer's investment in one- and two-period bonds, respectively. The insurer's balance sheet at $t = 0$ equates the insurer's assets (b_1, l_2) with its annuity liabilities and net worth:

$$b_1 + l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha + NW_0. \tag{2}$$

The first term on the right-hand side of equation (2) is the present value of the insurer's annuity liability and $NW_t \geq 0$ is the insurer's net worth in t. The weakly positive net worth reflects the insurer's limited liability.

After the aggregate shock R_2 is realized at the beginning of $t = 1$, the insurer's balance sheet becomes:

$$b_2(R_2) = \frac{1}{R_2} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha + NW_1(R_2), \tag{3}$$

where

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{4}$$

Equation (4) shows that the insurer finances the purchase of one-period bonds $b_2(R_2)$ in $t = 1$ using the proceeds from its initial bond holdings (b_1, l_2) net of the annuity payments made to the surviving individuals in $t = 1$. Collectively, equations (3) and (4) show that the insurer risks becoming insolvent in $t = 1$ if the present value of its liabilities exceeds the present value of its assets for certain realizations of R_2. Under Assumption 1, individuals do not purchase annuities from an insurer that has a non-zero probability of becoming insolvent in $t = 1$. Therefore, life insurers have an incentive to manage interest rate risk by choosing an asset portfolio, an annuity price, and a capital structure, such that $NW_1(R_2) \geq 0$ for all R_2.

Optimal IRM requires that the present value of an insurer's assets and liabilities (including the insurer's net worth) change at the same rate for any change in the interest rate. To see this, define the duration D of an asset or liability as the elasticity of its present value PV with respect to the interest rate: $D = -\frac{\partial PV}{\partial R}\frac{R}{PV}$. When the duration of the insurer's liabilities is greater than the duration of its assets, the present value of an insurer's liabilities increases more rapidly than the present value of its assets when the interest rate decreases, which may lead to insolvency.

To see how insurers invest their annuity considerations in corporate bonds to perform IRM, consider first an economy in which insurers can purchase as many units of the two-period bond as they need in $t = 0$ to perfectly hedge their interest rate risk. We refer to this special case as the economy with an unconstrained bond market. Insurers choose their bond holdings such that they remain solvent for all realizations of R_2. Specifically, insurers purchase bonds such that their net worth is always weakly positive in $t \geq 1$ and, due to competition, zero when $R_2 = 1$. In essence, competitive pressure forces life insurers to minimize their annuity price and hold the minimal level of net worth necessary to hedge the interest rate risk. In fact, when the bond market is unconstrained, we can show that if

$$l_2 = \frac{1}{R_l}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a\left(\alpha, q\right) g\left(\alpha\right) d\alpha \text{ and } b_1 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} a\left(\alpha, q\right) g\left(\alpha\right) d\alpha,$$

then by equations (2), (3), and (4), the optimal capital structure is such that $NW_0 = NW_1\left(R_2\right) = 0$ for any realization of R_2. Therefore, insurers can perfectly hedge their interest rate risk by investing in a suitable portfolio of one- and two-period bonds when the bond market is unconstrained without maintaining a strictly positive net worth.

We now consider the case when the bond market is constrained. Let $\zeta \in [0, 1]$ index the limit on the two-period bond supply. Specifically, insurers can purchase at most $l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a\left(\alpha, q\right) g\left(\alpha\right) d\alpha$ units of the two-period bond. The bond market is constrained when $\zeta \in [0, 1)$, which implies that insurers can no longer perfectly hedge the interest rate risk without net worth. Theorem 1 characterizes the optimal unique IRM strategy for any ζ. Importantly, since annuity liabilities do not change with respect to ζ, Theorem 1 implies that insurers' total bond demand increases to finance a higher level of net worth in a more constrained bond market.

Theorem 1 *Under the unique optimal IRM strategy, insurers require a higher level of net worth when the bond market is more constrained (ζ is lower), which increases insurers' total demand for bonds $b_1 + l_2$.*

Specifically, for a given annuity price q and $\zeta \in [0,1]$, the unique optimal IRM strategy requires an asset allocation and a capital structure that satisfies:

i. Asset portfolio:

$$l_2 = \frac{\zeta}{R_l} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha,$$

$$b_1 = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left[1 + (1-\zeta)\alpha\right] a(\alpha, q)\, g(\alpha)\, d\alpha,$$

$$b_2(R_2) = \left(1 - \zeta + \frac{\zeta}{R_2}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha.$$

ii. Capital structure:

$$NW_0 = \frac{1-\zeta}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha,$$

$$NW_1(R_2) = (1-\zeta)\left(1 - \frac{1}{R_2}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha.$$

When there is an insufficient supply of two-period bonds ($\zeta < 1$), the duration of an insurer's assets is strictly lower than the duration of its annuity liabilities.[13] Using the IRM strategy shown in Theorem 1, an insurer eliminates its duration gap by decreasing the duration of its total liabilities by maintaining a strictly positive

[13] In Appendix E, we consider an extension of the model with $\zeta < 1$ and an unlimited supply of two-period zero-coupon "government bonds" available in $t = 0$. The government bond has a lower yield than the two-period corporate bond, which captures its superior liquidity in a reduced form. We show that the lower the yield on the two-period government bond, the larger the net worth necessary to hedge interest rate risk. Thus, as long as the yield on the two-period government bond is strictly less than R_l, the value of the interest rate hedge provided by the two-period government bond is limited and the main message of Theorem 1 applies. This result is consistent with the relatively low share of government securities in life insurers' asset portfolio noted in Section 1.1.

level of net worth (recalling that net worth has zero duration).[14] Furthermore, the IRM strategy presented in Theorem 1 is unique when the market for annuity is perfectly competitive.[15]

2.3 Annuity pricing

The equilibrium annuity price is determined by Bertrand competition. In Appendix B, Lemma 1 characterizes the basic properties of the Bertrand equilibrium. We focus our analysis on the decision of life insurers implementing the optimal IRM strategy of Theorem 1. Crucially, by equation (1), competitive insurers choose an annuity price such that total annuity sales revenue is equal to total bond demand $b_1 + l_2$ under the optimal IRM—the zero-profit condition in our model.

We start by considering how adverse selection contributes to the annuity price markup. To do so, consider first the case when the bond market is unconstrained ($\zeta = 1$). By Theorem 1, competitive insurers optimally choose zero net worth. Therefore, equations (1) and (2) show that the equilibrium annuity price in an unconstrained bond market q^{AF} is given by:

$$q^{AF} \int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha. \quad (5)$$

We refer to q^{AF} as the *risk-adjusted actuarially fair price*, which accounts for adverse selection in the annuity market when the bond market is unconstrained.

Next, consider a complete information economy with an unconstrained bond market. Let $q^{CI}\left(\alpha\right)$ denote the equilibrium price in an economy where insurers can observe individual survival types α. The full information actuarially fair price is given by $q^{CI}\left(\alpha\right) = \frac{\alpha}{R_1}\left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right]$. Theorem 2 establishes the classic adverse selection result that the *risk-adjusted actuarially fair price* is higher than the

[14]Duration is a local measure of interest rate sensitivity. In practice, insurers are also concerned about large interest rate fluctuations, which is better measured by convexity. Convexity measures how duration responds to changes in the interest rate. In Appendix C, we show that optimal IRM in the model not only matches the insurer's asset and liability duration, but also convexity.

[15]In Appendix F, we show that insurers with market power also need to engage in IRM and build an asset portfolio and capital structure that satisfies Theorem 1 at the minimum.

average full-information actuarially fair price:

Theorem 2 $q^{AF} > \int_{\underline{\alpha}}^{\overline{\alpha}} q^{CI}(\alpha)\, g(\alpha)\, d\alpha$.

Finally, to see how IRM affects the annuity markup, we characterize how the equilibrium annuity price is affected by a marginal change in the supply of two-period bonds in a constrained bond market ($\zeta \in [0,1)$). The insurers' profit $\Pi(q,\zeta)$ is given by the difference between their annuity sales revenue and total bond demand:

$$q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q)\, g(\alpha)\, d\alpha - \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left[1 + \alpha \left(1 - \zeta + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha,$$

where total bond demand is given by equation (2) and Theorem 1. Let q^* be the equilibrium annuity price—the lowest positive annuity price such that $\Pi(q^*,\zeta) = 0$. Theorem 3 shows that the equilibrium annuity price q^* is higher than the risk-adjusted actuarially fair price q^{AF}.

Theorem 3 *The competitive annuity price q^* is higher when the bond market is more constrained (ζ is lower):* $\frac{\partial q^*}{\partial \zeta} < 0$.

Theorem 3 demonstrates how IRM affects the *AS-adjusted markup* defined as: $q^* - q^{AF}$. When the bond market is unconstrained ($\zeta = 1$), the AS-adjusted markup is zero, as $q^* = q^{AF}$. The AS-adjusted markup increases when the bond market becomes more constrained, as indexed by a lower level of $\zeta < 1$.

2.4 Discussion of the model's properties

Figures 2a and 2b presented in the Introduction depict how insurers finance their net worth by charging a markup over the annuity contract's actuarial value in a constrained bond market. These figures are adapted from the textbook model of adverse selection in Einav & Finkelstein (2011) to highlight the unique features of our annuity pricing model.

In the textbook model, financial markets are efficient and the competitive annuity price is determined by the intersection of the demand curve $A(q)$ and the downward-sloping insurer average cost curve $C(q)/q$. Because insurers can only offer a single annuity price to heterogeneous individuals when their mortality risk is not observable, perfect competition drives the equilibrium annuity price to be equal to the insurers' average cost.

Unlike the textbook model, an insurer's asset portfolio and its capital structure matter in our model, even when the bond market is unconstrained ($\zeta = 1$).[16] Equation (1) shows how our competitive equilibrium price q^* is determined by the intersection of the demand curve $A(q)$ and the insurers' average bond demand curve $B(q)/q$, which is the amount of bonds insurers demand per dollar of annuities sold. The intersection of $A(q)$ and the average cost curve $C(q)/q$ pins down the *risk-adjusted actuarially fair price* q^{AF}.

When the bond market is unconstrained, competitive insurers construct an optimal bond portfolio that perfectly hedges the interest rate risk without net worth. As shown in equation (2) and Figure 2a, the average cost curve $C(q)/q$ is equal to the bond demand curve $B(q)/q$. Therefore, the equilibrium price q^* in our model corresponds to the equilibrium price in the textbook model when the bond market is unconstrained, such that $q^* = q^{AF}$.

Limited liability leads to a different outcome when the supply of long-term bonds is constrained ($\zeta < 1$). By Theorem 1, insurers finance a positive level of net worth per dollar of annuities sold ($NW_0/q > 0$) to cushion the effect of future interest rate shocks and prevent insolvency. As shown in equation (2) and Figure 2b, the average cost curve $C(q)/q$ and average bond demand curve $B(q)/q$ are no longer equal when the long-term bond supply is constrained. Specifically, the total bond demand is equal to the sum of total annuity liabilities and the positive level of net worth. Theorem 3 shows how insurers finance their net worth by charging a higher annuity price—the equilibrium price q^* is higher than the *risk-adjusted actuarially fair price* q^{AF}.

[16] The insurers' capital structure is irrelevant in the textbook model because financial markets are efficient. Due to limited liability, the environment in our model violates the conditions for the Modigliani-Miller theorem to hold (Modigliani & Miller 1958). Limited liability implies that insurers must credibly show to annuity shoppers that they are managing risk, which pins down a unique ex-ante capital structure even when the bond market is unconstrained.

This discussion shows how the size of the AS-adjusted markup, $q^* - q^{AF}$, reflects the severity of financial frictions associated with interest rate risk. We can derive the equilibrium relationship between insurers' optimal net worth, which is generally not observable, and annuity price markups, which are observable at a high frequency, as follows:

$$NW_0 = \int_{\underline{\alpha}}^{\overline{\alpha}} \left(q^* - q^{CI}(\alpha)\right) a(\alpha, q^*)\, d\alpha - \int_{\underline{\alpha}}^{\overline{\alpha}} \left(q^{AF} - q^{CI}(\alpha)\right) a\left(\alpha, q^{AF}\right) d\alpha.$$

This expression shows how changes in an insurer's net worth are reflected in annuity markups.

Finally, Appendix D extends the present discussion by showing that life insurers' IRM amplifies the effect of adverse selection on annuity markups. This is because the average survival probability of individuals purchasing annuities increases as insurers charge a higher annuity price to finance their net worth. This additional theoretical result shows that the effects of supply-side and demand-side frictions on annuity markups are not orthogonal, and highlights the great difficulty in disentangling the sources of market inefficiencies that may affect annuity markets. In the next section, we discuss how we measure annuity markups in the data and how we identify the risk management channel using exogenous shifters of the average cost curve $C(q)/q$ and average bond demand curve $B(q)/q$.

3 Identification

The main empirical challenge to identifying the risk management channel in annuity price data is threefold. First, it is not possible to directly measure the duration gap between U.S. insurers' assets and insurance liabilities.[17] The reason is that the actual discount rate used by life insurers to value their insurance liabilities is not observable and insurance liabilities are not reported at the contract level in

[17]For this reason, most of the literature seeking to estimate life insurers' interest rate risk has proposed an indirect measure of life insurers' duration gap. For example, Hartley et al. (2016) and Ozdagli & Wang (2019) proposes an indirect measure of the duration gap based on insurers' stock price.

statutory fillings. Therefore, it is also not possible to observe life insurers' actual net worth position at a high frequency. Second, different types of "supply-side" frictions may lead to observably equivalent annuity markups. For example, annuity markups could be the outcome of monopolistic competition. Third, demand- and supply-side frictions could have non-trivial interactions. In Appendix D, we show theoretically and empirically that adverse selection in the annuity market depends on the severity of frictions in the corporate bond market, as higher level of net worth exacerbates adverse selection by increasing annuity prices.

We overcome these challenges by exploiting shocks that differentially affect the average cost curve $C(q)/q$ and the average bond demand curve $B(q)/q$ of *different annuity contracts* offered by the *same insurer*. Our identification strategy exploits the property that, under the assumptions of our model, the $C(q)/q$ and $B(q)/q$ curves determine the AS-adjusted markup when the bond market is constrained ($\zeta < 1$). On the one hand, an exogenous increase in $C(q)/q$ means that an insurer faces a larger insurance liability for a given pair of annuity price and quantity. Because the interest rate risk is unchanged, competitive insurers finance additional bonds per unit of annuity sold by increasing q^* to maintain the optimal net worth straegy $NW(q)/q$. In essence, $B(q)/q$ increases to fund the higher average cost along with maintaining the original net worth strategy, which raises the AS-adjusted markup $q^* - q^{AF}$. On the other hand, an exogenous decrease in $B(q)/q$ means that the insurer can fund a block of new annuity business with fewer bonds. Competitive insurers respond by offering a higher annuity yield—i.e., a lower price q^*—which increases the average cost along the $C(q)/q$ curve and $q^* - q^{AF}$ decreases.

The combination of exogenous shocks to the average cost curve $C(q)/q$ and average bond demand curve $B(q)/q$ is necessary to identify the risk management channel. We identify the risk management channel by comparing the change in AS-adjusted markup $q^* - q^{AF}$ for annuity contracts offered by the same insurer with higher average cost to those with lower average cost as a response to common shocks to the average bond demand curve $B(q)/q$. In the remainder of this section, we discuss how we measure $q^* - q^{AF}$, and the two sources of variation that exogenously shift the average cost curve and the bond demand curve.

3.1 Annuity price markups measurement

Life insurers reprice their annuities frequently in response to changes in market conditions. It is straightforward to interpret our model as the marginal pricing decision of a life insurer. Using this interpretation, a life insurer creates a new block of business at date t, which is added to its existing block of annuities. Therefore, the first step in our identification strategy is to evaluate the insurers' marginal pricing decisions conditional on bond market conditions.

Two inputs are needed to price new insurance liabilities. The first input in valuing annuity cash flows is a discount rate. We follow our theory closely and value new annuity cash flows from the perspective of the *owner* of a life insurer operating under limited liability. As discussed in Section 1, annuity contracts are illiquid fixed-rate liabilities, and life insurers invest their annuity considerations primarily in relatively illiquid fixed-income securities in an effort to match their asset and liability cash flows and offer a competitive return to annuitants. Therefore, our choice of cash flow discount rate needs to be consistent with the yield at which the marginal shareholder of this insurer is willing to commit capital to support the issuance of *illiquid* long-term fixed rate liabilities.[18] Almost all life insurers offering annuities in the U.S. have a rating around A and invest their annuity considerations in a portfolio of A-rated or equivalent fixed-income instruments, on average. Therefore, the discount rate of an average insurer's marginal shareholder should be close to the duration-matched yield on A-rated debt securities.

We proxy for the unobserved discount rate of the marginal life insurer shareholder using the zero-coupon High Quality Market (HQM) yield curve produced

[18] Note that the discount rate of an annuity shopper is likely very different from the discount rate of the owner of a life insurer. An annuity shopper seeking a safe longevity insurance contract may *perceive* an annuity contract to be relatively "safe" because of the existence, for example, of a state insurance guarantee fund. Consequently, the payoff structure of a limited liability life insurer's shareholders and the annuity contract holders are vastly different in the event the life insurer is placed in receivership by its state insurance regulator. Using a default free discount rate to value annuity contracts may be appropriate for the latter, but not for the former.

by the U.S. Treasury.[19] The HQM yield curve is calculated daily using AAA, AA, and A-rated U.S. corporate bonds and is heavily weighted towards A-rated bonds, consistent with their large market share.[20]

The second input to valuing annuity cash flows is an assumption about individuals' mortality. Virtually none of the fixed annuities sold by U.S. life insurers are underwritten, which means they require no medical exam and their terms only depend on the date of birth and gender of the individual. We use three different types of mortality assumptions. First, we use a "general" population period mortality table produced by the U.S. Internal Revenue Service that is updated annually with the mortality experience of the entire U.S. general population.[21] Second, we use two different versions of the Individual Annuitant Mortality table produced by the Society of Actuaries (SOA) in collaboration with the National Association of Insurance Commissioners (NAIC). In addition to a "basic" annuitant mortality table, which is estimated from the actual mortality experience of a large pool of annuitants from multiple insurers, the SOA produces a "loaded" annuitant mortality table, which is used by state insurance regulators to set regulatory reserves.[22]

The actuarial value of a life annuity contract with an M-year guarantee term per dollar using mortality assumption $k \in \{\text{General}, \text{Basic}, \text{Loaded}\}$ is defined as

$$V_t^k(n,S,M,r) = \underbrace{\sum_{m=1}^{M} \frac{1}{R_t(m,r)^m}}_{\text{M-year term certain annuity}} + \underbrace{\sum_{m=M+1}^{N_S^k - n} \frac{\Pi_{l=0}^{m-1} p_{S,n+l}^k}{R_t(m,r)^m}}_{\text{Life annuity from year } M+1},$$

where $M \geq 0$ is the number of years the life annuity pays a guaranteed fixed income, $p_{S,n}^k$ is the one-year survival probability for an individual of gender S at

[19] The HQM yield curve data is available at https://www.treasury.gov/resource-center/economic-policy/corp-bond-yield/Pages/Corp-Yield-Bond-Curve-Papers.aspx.

[20] For example, the sample of bonds used to calculate the HQM yield curve on August 31, 2011 includes 12 commercial papers, 42 AAA bonds, 299 AA bonds, and 1,345 A bonds. For more information, see https://www.treasury.gov/resource-center/economic-policy/corp-bond-yield/Documents/ycp_oct2011.pdf

[21] The general population mortality tables are available at https://www.irs.gov/retirement-plans/actuarial-tables.

[22] Appendix H contains more details about mortality assumptions.

age n from the k-th mortality table, N_S^k is the maximum attainable age for this gender in the k-th mortality table, and $1/R_t(m,r)^m$ is the reference discount factor for period m cash flow evaluated at time t using the HQM yield curve ($r =$ HQM), or the regulatory reference rate ($r =$ NAIC), which we will explain below.

Let $P_t(n,S,M)$ be the price of an M-year guaranteed life annuity offered to an individual of gender S and age n at date t. We decompose the total annuity price markup into an insurer-contract-level AS-adjusted markup and an industry-contract average measure of adverse selection pricing (AS pricing):

$$P_t(n,S,M) - V_t^{\text{General}}(n,S,M,r) = \underbrace{\Big(P_t(n,S,M) - V_t^{\text{Basic}}(n,S,M,r)\Big)}_{\text{Adverse selection adjusted markup}} + \underbrace{\Big(V_t^{\text{Basic}}(n,S,M,r) - V_t^{\text{General}}(n,S,M,r)\Big)}_{\text{Average adverse selection pricing}} \,,$$

where r is the HQM yield curve. It follows that the insurer-contract-level variable $P_t(n,S,M) - V_t^{\text{Basic}}(n,S,M,r)$ is the counterpart of the AS-adjusted markup $q^* - q^{AF}$ in our model.

Figure 1, which we discussed in the introduction, plots the distribution of actual monthly payments offered to a 65-year-old male for a \$100,000 SPIA from a sample of U.S. life insurers against the monthly payment implied by the different actuarial values.

3.2 Regulatory reserve requirements

The first source of exogenous variation comes from annuity contract-level time series variation in the regulatory reserves that insurers are required to set aside for each dollar of annuity they sell. As explained by Koijen & Yogo (2015), exogenous time-series variation in reserve requirements across contract maturity arises because regulatory reserves are calculated using a single interest rate that is reset infrequently. We denote this regulatory interest rate by $r =$ NAIC and use it in the calculation of $V_t^k(n,S,M,r)$.

By construction, the regulatory reference interest rate is close to the average of the longer end of the HQM yield curve. When the actual yield curve is upward sloping, the actuarial value of a life annuity calculated using the average of the long end of a yield curve is mechanically smaller than the corresponding actuarial value calculated using the entire yield curve. Moreover, this difference is greater for life annuities with shorter expected maturity—i.e., sold to older individuals. Because the regulatory interest rate is reset infrequently—once a year prior to 2018 and once a quarter from 2018 instead of daily—the difference between the ratio of reserves required to the reserve calculated using the daily yield curve exogenously fluctuates over time across annuity contract maturity.

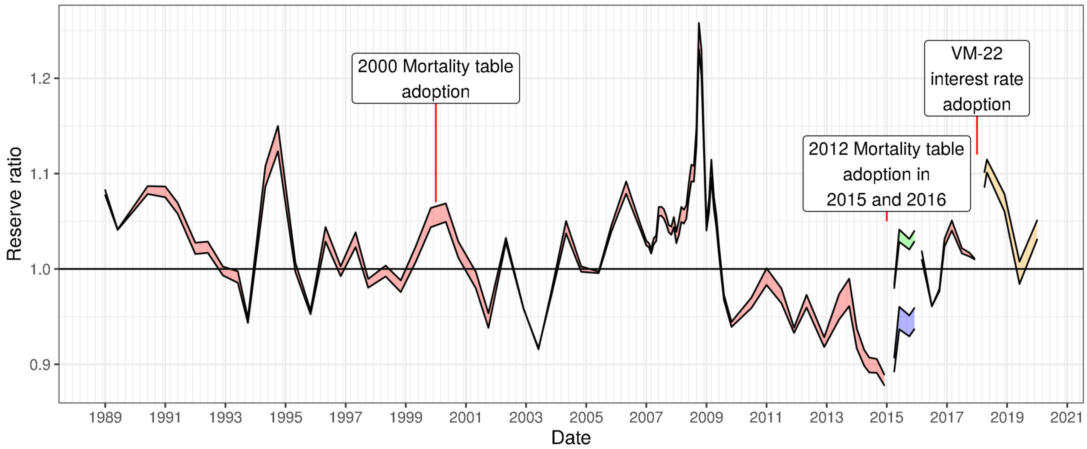


Figure 4: Reserve ratio for SPIA sold to a 65-year-old male (top line) and 70-year-old male (bottom line)

Figure 4 illustrates this source of exogenous variation by plotting the reserve dollars an insurer needs to set aside for each dollar of annuity sold on day t to 65- and 70-year-old males only. We calculate the regulatory reserve ratio as $V_t^{\text{Loaded}}(n, S, M, r = \text{NAIC})/V_t^{\text{Basic}}(n, S, M, r = \text{HQM})$. A ratio above 1 indicates that the reserve requirement is *binding*, as the insurer must create a reserve that is greater than the insurance liability warranted by the insurers' yield curve-based actuarial calculation. Conversely, a ratio below 1 indicates that the reserve requirement is *non-binding* because the required reserve is below the insurer's own actuarial calculation. The *distance* between the two lines—depicted by the colored shaded area—measures the relative cost of each contract and exogenously fluctuates overtime because the flat regulatory interest rate resets infrequently.

Figure 4 shows that there are additional sources of variation arising from U.S. states' staggered adoption of new regulatory mortality assumptions between 2015 and 2016, and the 2018 adoption of the new methodology to calculate the regulatory reference interest rate.[23]

Mapping this data to Figure 2b, an exogenous rise in reserve requirements increases the average cost curve $C(q)/q$ because the insurer must create a larger insurance liability for a given pair of annuity price and quantity. The insurer matches this larger insurance liability by financing additional bonds per unit of annuity sold, which requires increasing q^* to maintain the optimal net worth strategy $NW(q)/q$ and offset some of the increase in $C(q)/q$. As a result, $B(q)/q$ increases and q^* rises. However, even when the reserve ratio is binding, the risk-adjusted actuarially fair price q^{AF} remains the same because it is meant to capture the cost of providing annuities from the insurers' perspective, independent of the regulator. As a result, a tightening of the reserve requirement increases the AS-adjusted markup—i.e., $q^* - q^{AF}$ rises.[24]

3.3 Yield spreads on long-duration investment grade bonds

The second source of exogenous variation comes from aggregate time-series variation in the spread between the yield on Moody's Baa-rated and Moody's Aaa-rated corporate bonds that have at least 20 years of maturity—i.e., the yield spread on long-duration investment grade bonds. As we explained above, the regulatory interest rate used prior to 2018 is a weighted average of the corporate bond yields rated above Baa that have at least 20 years of maturity. Under state insurance regulation, corporate bonds rated above Moody's Baa are designated as NAIC 1 and attract the lowest statutory risk-based capital charge. This risk-based capital charge reflects the fact that, historically, the increase in credit risk for a firm moving from an A to a Aaa rating is negligible—this is because firms are usually downgraded before they become insolvent. As explained in Section 1,

[23] We provide more details about the regulatory discount rate in Appendix H.

[24] Appendix G provides a formal treatment of the effect of a reserve requirement shock on the optimal IRM strategy and annuity pricing in our model.

the life insurer's business model consists of earning a spread between the yield on the assets purchased with premiums and considerations and the yield credited on insurance liabilities. Therefore, a widening in Baa-Aaa spread for long-duration corporate bonds *conditional* on the insurer's cost of funding corresponds to higher-yielding investment opportunities for new annuity money.

Figure 5 illustrates this exogenous variation by plotting the Baa-Aaa spread for seasoned corporate bonds and the 10-year HQM yield spread over 10-year U.S. Treasury in percentage points. Life insurers can generate more yield per dollar of annuity sold when the Baa-Aaa spread increases more than the 10-year HQM yield spread, which is our proxy for the insurers' average cost of funding, since life annuities have an original duration of about 10 years.

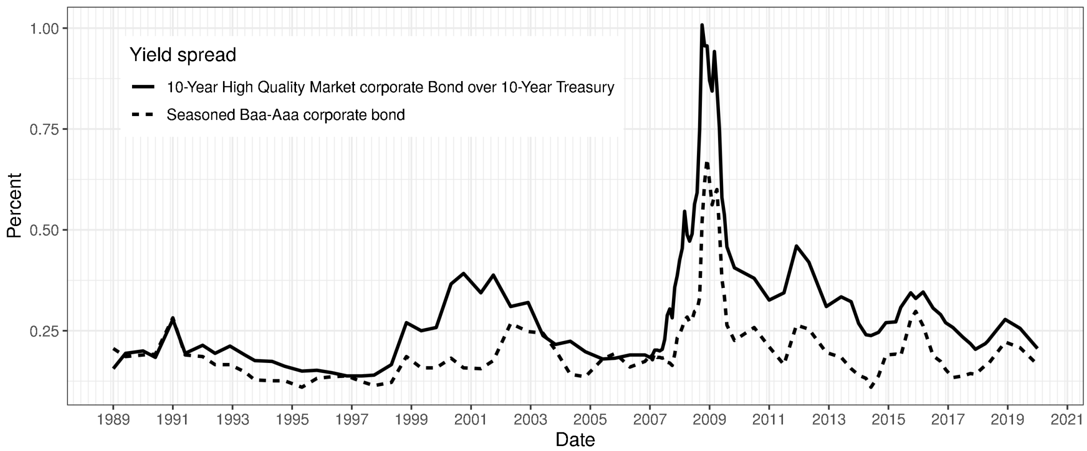


Figure 5: Baa-Aaa spread and insurers' average cost of funding

Mapping this data to Figure 2b once more, an exogenous increase in long-duration bond yield decreases $B(q)/q$ because the insurer can fund a block of new annuity business with fewer bonds. Other things being equal, the decrease in $B(q)/q$ implies that insurers no longer need to hold as much net worth—i.e., the original level of $NW(q)/q$ is now higher than the new optimal level. Competitive insurers respond by lowering the price of their annuity q^*—increasing the yield on their annuities. This price adjustment corresponds to a movement up along $C(q)/q$ and results in a lower AS-adjusted markup $q^* - q^{AF}$.[25]

[25] Appendix G contains a formal treatment of the effect of long-term bond spread shock on the optimal IRM strategy and annuity pricing. A caveat with this simplification is that our model

4 Data and variable definitions

We focus our analysis on Single Premium Immediate Annuities (SPIA) and SPIA with 10 and 20 year term certain guarantees. SPIA with 10- and 20-year term certain guarantees promise a payment to a beneficiary during a term period irrespective of the annuitant's survival. Our sample includes quotes from 99 life insurers, with about 20 life insurers per reporting dates. Price quotes are typically reported for male and female individuals aged between 50 and 90 years with 5-year intervals. Annuity prices are collected from the 1989-2019 issues of the *Annuity Shopper Buyer's Guide*.[26] Table 1 reports the summary statistics for the variables used in our analysis.

Table 1: Summary statistics

	Obs.	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
Number of insurers by period		19.9	6.19	16	19	23.2
Number of contracts by period		634	374	294	686	914
Life annuity contract (binary):						
Life only	40,790	0.40				
10 year guarantee	40,790	0.33				
20 year guarantee	40,790	0.27				
55 years old	40,790	0.11				
60 years old	40,790	0.14				
65 years old	40,790	0.15				
70 years old	40,790	0.15				
75 years old	40,790	0.14				
80 years old	40,790	0.10				
85 years old	40,790	0.09				
90 years old	40,790	0.03				
Male	40,790	0.50				
Female	40,790	0.50				
$Annuity\_markup_{ijt}$ (%)	40,790	15.84	4.76	12.71	15.43	18.44
$Reserve\_Ratio_{jt}$	40,790	1.01	0.06	0.97	1.01	1.05
$10Y\text{-}3M\_Treasury\_spread_t$	40,790	1.69	1.08	0.80	1.73	2.54
$Baa\text{-}Aaa\_spread_t$	40,790	0.95	0.33	0.74	0.90	1.04
$10\_HQM\_spread_t$	40,790	1.46	0.57	1.12	1.35	1.72
$Log\_total\_assets_{it}$ (from 2001)	29,462	2.69	1.58	1.76	2.84	3.78
$Leverage\_ratio_{it}$ (from 2001)	29,462	10.58	5.28	6.98	10.08	13.57
$Net\_swap\_duration_{it}$ (from 2009 to 2015)	9,149	0.09	0.16	0.002	0.01	0.11

Our main dependent variable $Annuity\_markup_{ijt}$ is the normalized AS-adjusted

in Section 2 assumes the annuity demand curve is fixed. In reality, annuity demand and bond market conditions could be correlated, as both financial instruments allow individuals to transfer wealth across periods. In Section 5, we explain how we overcome this issue by implementing our test of the IRM channel in a difference-in-differences regression framework that compares the change in AS-adjusted markup for annuity contracts with higher relative reserve requirement to those with lower relative reserve requirement in response to the same shock to the Baa-Aaa spread.

[26] Koijen & Yogo (2015) use a smaller sample of the same data extending from 1989 to 2011.

markup for product j sold by insurer i, defined as

$$Annuity\_markup_{ijt} = \frac{P_{ijt}(n, S, M)}{V_{jt}^{\text{Basic}}(n, S, M, r = \text{HQM})} - 1 \ .$$

The AS-adjusted markup is just under 16 percent on average and consistently above 10 percent during our sample period. The variable $Reserve\_Ratio_{jt}$ is the ratio of reserve dollars insurers need to set aside for each dollar of annuity j sold on day t, defined as

$$Reserve\_Ratio_{jt} = \frac{V_{jt}^{\text{Loaded}}(n, S, M, r = \text{NAIC})}{V_{jt}^{\text{Basic}}(n, S, M, r = \text{HQM})} \ .$$

The reserve ratio fluctuates around 1, confirming that the regulatory discount rate and the insurers' discount rate are aligned on average. We also obtain time-varying insurer characteristics data from NAIC statutory filings for 2001-2019. We measure insurer size as the log of insurers' general account assets and leverage as the ratio between the insurers' general account assets and general account liabilities minus statutory accounting surplus.[27]

We obtain Moody's Seasoned Aaa and Baa corporate bond yields, the 10-year Treasury constant maturity rate, and 10-year Treasury constant maturity minus 3-month Treasury constant maturity from the St. Louis Fed's FRED database. We proxy for the insurers' cost of funding by calculating the spread between the 10-year high quality market and the 10-year Treasury constant maturity rate. For all of our regressions, we retain the last set of prices observed in a quarter. Our final data set contains 40,790 insurer-contract-quarter observations with an average of 634 insurer-contract observations per reporting period.

5 Main empirical analysis and results

Adopting a difference-in-differences approach, we test the hypothesis that the AS-adjusted markups on annuity contracts facing higher regulatory reserve re-

[27] We discuss the $Net\_swap\_duration_{it}$ variable in Section 6.

quirements are relatively higher in periods when the Baa-Aaa spread is lower. In the main specification, the first difference is between annuity contracts j offered by insurer i with relatively high reserve requirements and different annuity contracts $-j$ offered by the *same* insurer with relatively low reserve requirements. The second difference is between periods in which the Baa-Aaa spread is high and periods in which the Baa-Aaa spread is low. As discussed in Section 3, we condition all our tests on the average cost of funding of the insurer, which we proxy using the 10-year HQM zero coupon yield over the 10-year U.S. Treasury spread.

We implement our test in a linear regression framework. The unit of observation is a life insurer-product-time. The sample of observation extends from 1989 to 2019. The coefficient β_3 on the interaction between $Reserve\_Ratio_{jt}$ and the $Baa\text{-}Aaa\_spread_t$ in the following linear model captures the difference-in-differences estimate of the reduction in reserve requirement on AS-adjusted markup during times of increasing Baa-Aaa spreads *conditional* on the level of the $10\_HQM\_spread_t$:

$$\begin{aligned} Annuity\_markup_{ijt} = & \alpha_1^i + \alpha_2^j + \beta_1 Baa\text{-}Aaa\_spread_t + \beta_2 Reserve\_Ratio_{jt} \\ & + \beta_3 Baa\text{-}Aaa\_spread_t \times Reserve\_Ratio_{jt} \\ & + \beta_4 10\_HQM\_spread_t + \mathbf{z}_{it}'\boldsymbol{\gamma} + \epsilon_{ijt} \ . \end{aligned} \tag{6}$$

Equation (6) includes an insurer fixed effect α_1^i to absorb the effects of potentially unobserved fixed insurer characteristics—e.g., differences in state regulations and insurer ratings—that may directly affect life insurers' pricing behaviour. We also include a complete set of product fixed effects α_2^j—age, gender, and annuity guarantee type—to absorb the effect of fixed demand characteristics that may influence pricing. The vector $\boldsymbol{z}_{it}'$ includes other insurer-level time varying financial variables, such as insurer size and leverage. We report insurer clustered robust standard errors throughout as our baseline.

Table 2 summarizes our main result. The coefficient estimate on the interaction term suggests that, conditional on insurers' average cost of funding, a one standard deviation increase in $Reserve\_Ratio_{jt}$ (0.056) raises the AS-adjusted markup by 1

Table 2: The effect of investment-grade corporate-bond yield spread on life annuity markups The unit of observation is a life insurer-product-year. The sample of observation extends from 1989 to 2019. The dependent variable $Annuity\ markup_{ijt}$ is the adverse selection adjusted markup for life annuity j sold by insurer i at date t. Column 1 reports insurer clustered robust standard errors in parentheses and Column 2 and 3 report two-way insurer and date clustered robust standard errors in parentheses. *** p < 0.01; ** p<0.05; * p<0.1.

Dependent variable:		$Annuity\_markup_{ijt}$	
	(1)	(2)	(3)
$Baa\text{-}Aaa\_spread_t \times Reserve\ Ratio_{jt}$	-12.14***	-12.14**	-11.73**
	(3.34)	(5.29)	(5.32)
$Reserve\_Ratio_{jt}$	28.98***	28.98***	28.06***
	(4.17)	(6.18)	(6.20)
$Baa\text{-}Aaa\_spread_t$	10.55***	10.55*	9.87*
	(3.36)	(5.56)	(5.62)
$10\_HQM\_spread_t$	2.93***	2.93***	3.15***
	(0.40)	(0.71)	(0.84)
$Log\_total\_assets_{it}$			0.30
			(0.68)
$Leverage\_ratio_{it}$			-0.04
			(0.03)
Fixed effects:			
Contract characteristics (j)	Y	Y	Y
Insurer (i)	Y	Y	Y
SE Clustering	Insurer	Insurer/Date	Insurer/Date
Observations	40,790	40,790	29,462
Adjusted R^2	0.41	0.41	0.41

percentage point when $Baa\text{-}Aaa\_spread_t$ is at its average level (0.95). This effect is about 18 percent lower in periods when $Baa\text{-}Aaa\_spread_t$ is in the 3rd quartile of its distribution relative to periods when $Baa\text{-}Aaa\_spread_t$ is in the first quartile of its distribution.[28] This means that, conditional on their average cost of funding, insurers tend to raise their AS-adjusted markups when the reserve requirement becomes binding—when the $Reserve\_Ratio_{jt}$ is higher—but do so significantly less when $Baa\text{-}Aaa\_spread_t$ is wider.

This first result shows that insurers decrease their AS-adjusted markup when the cost of IRM decreases on the margin. That is, conditional on an insurer's cost of funding, a widening in Baa-Aaa spread for long-duration corporate bonds

[28]This difference is statistically significant at the less than 1 percent level.

corresponds to higher yielding investment opportunities for new annuity money and, therefore, a lower AS-adjusted markup.[29] This result is consistent with the IRM strategy of life insurers in a constrained bond market. In Appendix G, when the bond market is unconstrained, we formally show that $q^{*} - q^{AF}$ is unaffected by changes in long-term bond yield spreads in an extension of our Section 2 model with perfect competition and monopolistic competition, respectively. This result is also fully consistent with recent work by Ozdagli & Wang (2019), who find that when interest rates decline, life insurers re-balance their portfolios toward higher-yielding bonds by increasing the duration, rather than the credit risk, of their portfolios.[30]

Column 2 investigates the robustness of our inference by reporting two-way insurer and date clustered robust standard errors that allow for arbitrary types of within-insurer correlation as well as contemporaneous correlation of the errors across different insurer clusters. Although onerous in terms of degrees of freedom, allowing for cross-insurer cluster correlation could be important given that insurers reprice their annuity products in response to aggregate bond market shocks. Consistent with this prior, Column 2 shows that the two-way clustered robust standard errors are about twice as large as those reported in Column 1. Nevertheless, our difference-in-differences coefficient estimate remains significant at below the 5 percent significance level.[31] Column 3 controls for time-varying insurer size, measured as the log of the insurer's general account assets, and insurer leverage, measured as the ratio of the insurer's general account assets to liability minus statutory surplus (statutory surplus is correlated to our definition of net worth in the model in Section 2). Although we only observe these financial variables from 2001, the coefficient estimates in Column 3 are almost identical to those obtained

[29]Although we do not observe actual annuity sales on a per contract basis, Figure 7 in Appendix A shows that aggregate fixed annuity sales sharply increase whenever the Baa-Aaa yield spread increases. This effect is apparent during the 2008-09 financial crisis, the height of the European debt crisis in 2012-13, and around the 2014-16 oil shock.

[30]Ozdagli & Wang (2019) do not analyze the effects of IRM on life insurers' product pricing. Rather, the authors focus on the effect of changes in an indirect measure of life insurers' duration gap on life insurers' bond holdings.

[31]We investigate the robustness of our inference to different clustering assumptions by calculating block bootstrap standard errors and wild bootstrap standard errors and find no evidence of bias—the results are available on request.

with the full sample in Column 2.

We conclude this section by estimating the contribution of IRM to the life annuity AS-adjusted markup. Although formally estimating the effect of IRM on markups with a structural model is outside the scope of this paper, we can nevertheless obtain a rough estimate using the markup on 5-year term certain annuities offered by the same insurer at the same time as a benchmark. Five-year term certain annuities are not affected by adverse selection, as the insurer makes fixed regular payments for 5 years irrespective of the contract holder's survival. Life insurers can easily match the duration and illiquidity profile of 5-year term annuities, as roughly half of corporate bonds issued have an initial maturity ranging from 5 to 10 years. Therefore, we expect the 5-year term annuity markup to largely reflect insurers' expenses associated with issuing these types of liabilities. Indeed, we find that this markup is around zero after netting the industry reported a 3 to 5 percent issuance and maintenance expense in 2019. Assuming the expenses associated with issuing 5-year term annuities are not greater than those associated with issuing a life annuity and that competition for each product is similar, the insurer-level difference between the life annuity's AS-adjusted markup and the 5-year term annuity markup is an upper bound estimate of the cost of IRM.[32]

Figure 6 plots the distribution of this markup difference calculated for each date and for each insurer offering both contracts simultaneously. The shaded region indicates the 2007–09 U.S. recession. Figure 6 shows that the cost of IRM accounts for at most 50 to 70 percent of the AS-adjusted markup, or about 8 to 11 percent of the life annuity's actuarial value. This estimate suggests that if the insurers' business expense are around 3 to 5 percent, IRM accounts for almost all of the average AS-adjusted markup. Figure 6 also shows that the share of IRM in markup significantly increased after the Global Financial Crisis (GFC) and that its cross-sectional variance decreased significantly. These two observations are

[32]This is a rough estimate in the sense that there could be material differences in market structure across the two products that could bias this calculation. For example, 5-year term annuities are an imperfect substitute for banks' certificate of deposits (CDs), while life insurers do not face competition from banks for their life annuity offerings. That said, Figure 9 in Appendix F shows that average competition for the entire fixed annuity market is high.

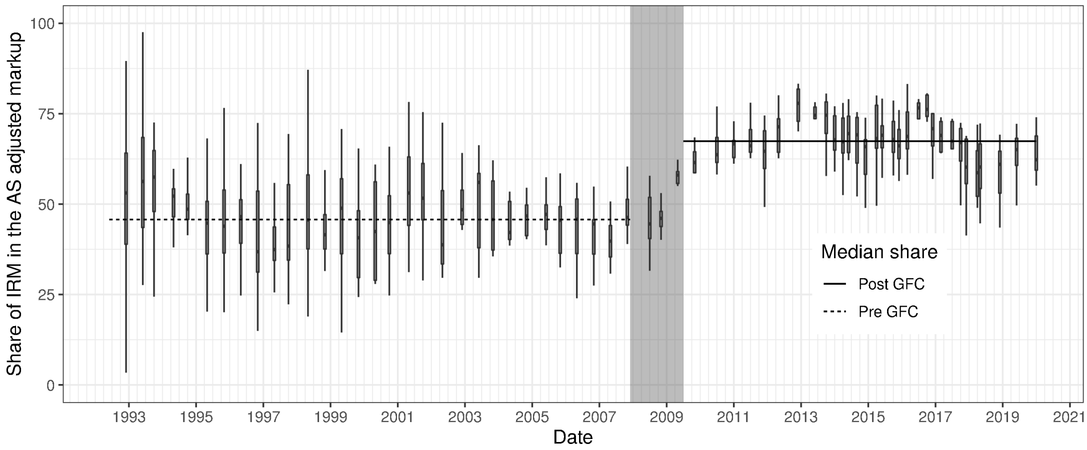


Figure 6: Contribution of IRM cost in the AS adjusted markup for SPIA offered to a 65-year-old male.

consistent with the adverse effect of lower long-term rates and spread compression on the life annuity business model and the increase in competition in the annuity market space (Foley-Fisher, Heinrich & Verani 2020).

6 Evidence from interest rate derivatives

This paper so far has emphasized building a cushion of net worth to manage future interest rate shocks. Another important aspect of life insurers' risk management is the preservation of net worth, as a long enough sequence of bad shocks could lead to insolvency. This motive is absent from our model because the life insurers only need to manage a one-time interest rate shock. In this section, we venture outside the predictions of our simple model and dig deeper into the mechanism by which interest rate risk affects insurers' annuity pricing decisions by studying how life insurers' interest rate derivative positions affect their annuity pricing decisions.

Life insurers can add positive duration to their balance sheet by entering into a long-term *fixed-for-float* interest rate swap with a counterparty. Issuing a fixed-for-float interest rate swap is economically equivalent to financing a fixed-maturity bond with short-term floating rate debt. The duration of a fixed-for-float swap contract is then the difference between the (hypothetical) underlying fixed-rate instrument (usually a U.S. Treasury bond) and the duration of the floating rate liability that finances the fixed rate instrument—usually 3-month LIBOR. An insurer adding more net positive duration with interest rate swaps is more likely to manage the risks associated with a widening negative duration gap between its assets and insurance liabilities.

Although it is not possible to measure a U.S. life insurer's duration gap, we can construct a proxy for the *marginal* net duration added by a life insurer's interest rate swaps. We can then measure how different hedging programs perform facing the *same* sequence of aggregate interest rate shocks and trace out the effect on annuity prices. For example, an insurer adding positive net duration with swaps is relatively more hedged against a flattening yield curve and vice versa. Although an insurer's swap position is an ex-ante endogenous variable, variations in the 10-year over 3-month Treasury spread during the period of the *zero lower bound* act as an exogenous shifter of the swap portfolio value ex-post. Therefore, we can compare the AS-adjusted markups of insurers that are favorably affected by the aggregate interest rate shock *ex-post* because of their *ex-ante* hedging program relative to those that are adversely affected by the shock. We focus on the period of the *zero lower bound* from 2009 to 2015, during which all the variations in the yield curve were driven by movements in long rates.[33]

6.1 Interest rate swaps data

We use position-level interest rate swap data to calculate a novel estimate of the net duration added by the swaps as a fraction of an insurer's general account asset portfolio.[34] Our position-level swap data comes from Schedule DB in the NAIC statutory filing. Schedule DB provides detailed information on each insurer's position-level derivative contracts, including a description of the contract term and its notional amount. We carefully parsed the text of more than 82,000 individual contract-year observations from 44 U.S. life insurers from 2009 to 2015

[33] Outside of the zero lower bound period, the value of an insurer's swap portfolio may respond differently to a steepening of the yield curve that is driven by lower short rates and higher long rates, which would greatly complicate the analysis.

[34] See Appendix H for details.

and extracted the receiving leg, notional amount, and residual maturity of the contracts. Life insurers in our sample have on average 1,416 open interest rate swap contracts at year's end with a standard deviation of 978. The average notional amount of a swap contract is \$45 million with a standard deviation of \$83 million.

We first calculate the quarter-end individual swap position using each contract's residual maturity. At every quarter-end, we normalize an individual swap contract's duration using the duration of a reference 10-year fixed-for-float swap contract, and then we multiply by the original contract's notional amount. This number is the dollar amount of duration, which can be positive or negative, contributed by an individual swap contract. We then sum over an insurer's entire swap portfolio to obtain the aggregate dollar amount of duration added by the swaps. Finally, we divide by the insurer's total general account assets to obtain the amount of net duration added by the swaps expressed as a fraction of the insurer's asset portfolio, which we denote by $Net\_swap\_duration_{it}$. A value of zero indicates that the insurer is not adding positive or negative duration using swaps. A value of 0.5 indicates that the insurer is adding net positive duration that is 50 percent of its size.

6.2 Cross-sectional regression results

We implement our cross-sectional test by interacting the $Net\_swap\_duration_{it}$ variable with $10Y\text{-}3M\_Treasury\_spread_t$ in the following equation:

$$\begin{aligned} Annuity\_markup_{ijt} = & \alpha_1^i + \alpha_2^j + \alpha_3^t + \beta_1 Net\_swap\_duration_{it} + \beta_2 Reserve\_Ratio_{jt} \\ & + \beta_3 10Y\text{-}3M\_Treasury\_spread_t \times Net\_swap\_duration_{it} \\ & + \mathbf{z}_{it}'\boldsymbol{\gamma} + \epsilon_{ijt} \ . \end{aligned}$$

We focus on the cross-sectional variation in life insurer $Net\_swap\_duration_{it}$ by including date fixed effect α_3^t. We continue to include an insurer fixed effect α_1^i, a complete set of product fixed effects α_2^j, and time-varying insurer-level $\mathbf{z}_{it}'$, which includes insurer size and leverage. We report two-way insurer and date clustered robust standard errors as our benchmark, although the addition of a date fixed effect means that we obtain very similar standard errors using insurer clustered robust standard errors.

Table 3: Cross-sectional evidence of the risk management channel The unit of observation is an insurer-product-year. The sample of observation extends from 2009 to 2015, which covers the period of *zero lower bound*. The dependent variable $Annuity\ markup_{ijt}$ is the life annuity markup for product j sold by insurer i in year t. Column 1 is a fixed-effect regression with two-way insurer and date clustered robust standard errors reported in parentheses. Columns 2 to 4 are quantile fixed-effects regressions implemented using the penalized fixed-effects estimation method proposed by Koenker (2004). Percentiles are indicated in the square parenthesis. Clustered bootstrapped standard errors (1,000 replications) are implemented using the generalized bootstrap of Chatterjee & Bose (2005) with unit exponential weights sampled for insurer-contract observations and reported in parentheses. *** p < 0.01; ** p<0.05; * p<0.1.

Dependent variable:	(1)	(2)	(3)	(4)
			Quantiles	
$Annuity\_markup_{ijt}$		$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
$Net\_swap\_duration_{it} \times$	5.63**	7.06***	4.98***	3.99***
$10Y\text{-}3M\_Treasury\_spread_t$	(2.33)	(0.55)	(0.34)	(0.33)
$Net\_swap\_duration_{it}$	-11.49**	-15.67***	-10.25***	-9.02***
	(5.44)	(1.52)	(1.01)	(0.91)
$10Y\text{-}3M\_Treasury\_spread_t$		-0.77***	-0.92***	-0.76***
		(0.13)	(0.13)	(0.12)
$Reserve\_Ratio_{jt}$	43.12***	15.19***	21.06***	27.45***
	(8.59)	(1.36)	(1.97)	(2.05)
$Baa\text{-}Aaa\_spread_t$		0.68*	-0.23	-1.83***
		(0.28)	(0.4)	(0.45)
$10\_HQM\_spread_t$		1.58***	2.35***	3.92***
		(0.21)	(0.28)	(0.32)
$Leverage\_ratio_{it}$	0.10	-0.12***	-0.13***	-0.12***
	(0.11)	(0.03)	(0.03)	(0.02)
$Log\_total\_assets_{it}$	-1.35	0.43***	0.55***	0.51***
	(1.57)	(0.08)	(0.08)	(0.08)
Fixed effects:				
Product char. (j)	Y		Y	
Insurer (i)	Y		Y	
Date (t)	Y		N	
Observations	9,149		9,149	
Adjusted R^2	0.52		t-test	17.4***

Column 1 of Table 3 summarizes our cross-sectional results. The coefficient estimate on the interaction term suggests that an insurer with the median level of $Net\_swap\_duration_{it}$ decreases its AS-adjusted markup by about 0.034 percentage point as a response to an unexpected flattening of the yield curve (a one standard deviation decrease in $10Y\text{-}3M\_Treasury\_spread_t$). The small economic magnitude of this average effect suggests that insurers' hedging strategy is effective on average. However, this effect is almost 50 times larger for insurers in the top quartile of the $Net\_swap\_duration_{it}$ distribution relative to those in the bottom quartile of the distribution. Insurers in the top quartile of the $Net\_swap\_duration_{it}$ distribution decrease their AS-adjusted markup by almost 1/3 percentage point (1.8 percent of the average AS-adjusted markup) in response to a flatter yield curve.

6.3 Quantile fixed-effect regression results

Columns 2 to 4 of Table 3 delve deeper by estimating a quantile regression with fixed effects. We estimate the conditional quantile functions $Q_{Annuity\_markup_{ijt}}(\tau|x_{ijt})$ of the response of the t-th observation on the j-th annuity contract offered by the i-th insurer's $Annuity\_markup_{ijt}$ given by

$$\begin{aligned}
Q_{Annuity_markup_{ijt}}(\tau|\mathbf{x}'_{ijt}) =& \beta_3(\tau)10Y\text{-}3M_Treasury_spread_t \times Net_swap_duration_{it} \\
&+ \beta_1(\tau)Net_swap_duration_{it} + \beta_2(\tau)10Y\text{-}3M_Treasury_spread_t \\
&+ \beta_3(\tau)Reserve_Ratio_{jt} + \beta_4(\tau)Baa\text{-}Aaa_spread_t \\
&+ \beta_5(\tau)10_HQM_spread_t + \alpha_1^i + \alpha_2^j + \mathbf{z}'_{it}\boldsymbol{\gamma}(\tau)\ ,
\end{aligned}$$

for quantile $\tau \in \{0.25, 0.5, 0.75\}$, where $\mathbf{x}'_{ijt}$ is the vector of covariates, $\mathbf{z}'_{it}\boldsymbol{\gamma}(\boldsymbol{\tau})$ is a vector of insurer-level time-varying controls, and α_1^i and α_2^j are the insurer and contract fixed effects, respectively. We also control for the effect of corporate bond market shocks on $Reserve\_Ratio_{jt}$, $Baa\text{-}Aaa\_spread_t$, and $10\_HQM\_spread_t$ that we discussed in the previous section.

The coefficients on the interaction terms in Columns 2 to 4 suggest that the least competitive insurers (those with relatively high markups) that are beneficially affected by the interest rate shocks as a result of their hedging programs disproportionately cut their AS-adjusted markups. The bottom row of Column 4 reports that the t test statistic rejects the null hypothesis that the 25^{th} and 75^{th} percentile coefficients on the interaction term are equal at below the 1 percent significance level. The coefficient estimates suggest that the counterfactual decrease in AS-adjusted markup in response to a flattening of the yield curve (i.e, a one standard deviation decrease in $10Y\text{-}3M\_Treasury\_spread_t$) would have been about twice as large for an insurer moving from the bottom to the top of the AS-adjusted markup distribution. Moreover, within the least competitive insurers that are at the top of the AS-adjusted markup distribution, the quantile fixed-effect regression allows us to estimate the counterfactual response of insurers that have a better hedge because they are in the top of the $Net\_swap\_duration_{it}$ distribution to those with a worse hedge. We find that among the least competitive insurers, those with a better hedge cut their markup by about 2 percent after a one standard deviation decrease in $10Y\text{-}3M\_Treasury\_spread_t$. In contrast, the least competitive insurers at the bottom of the $Net\_swap\_duration_{it}$ do not significantly cut their markup.

7 Conclusion

In this paper, we show that a large share of the notoriously high life annuity price markups can be explained by the cost of managing interest rate risk. We propose a novel theory of insurance pricing that reflects both informational frictions and interest rate risk. We develop an algorithm for annuity valuation to decompose the contribution of demand- and supply-side frictions in annuity markups using over 30 years of annuity price data and a novel identification strategy that exploits bond market shocks and the U.S. insurance regulatory framework. Our main result is that interest rate risk significantly constrains the supply of life annuities. A corollary is that the best time to sign up for a life annuity is during a time of overall financial market stress, as annuity prices are lower when investment grade bond spreads are higher!

This result has important implications for the literature seeking to assess the welfare effects of social insurance programs using life-cycle models. A robust result in this literature is that social insurance crowds out private insurance—e.g., Cutler & Gruber (1996) and Hosseini (2015). This result holds even when there are informational asymmetries in insurance and labor markets, as contracts can be designed to overcome this friction (Golosov & Tsyvinski 2007). It is unclear if this seminal result holds when considering the problem of life insurers managing interest rate risk. Moreover, studying life insurers' IRM may also shed light on important puzzles surrounding the shrinking U.S. long-term care insurance market in the "low-for-long" interest rate environment—e.g., Ameriks, Briggs, Caplin, Shapiro & Tonetti (2016), Braun, Kopecky & Koreshkova (2019). We leave these important questions to future research.

References

ACLI (2018), Life Insurers Fact Book 2018, Technical report, American Council of Life Insurers.

Ameriks, J., Briggs, J., Caplin, A., Shapiro, M. D. & Tonetti, C. (2016), Late-in-life risks and the under-insurance puzzle, Technical report.

Bailey, A. H. (1862), 'On the principles on which the funds of life assurance societies should be invested', *The Assurance Magazine, and Journal of the Institute of Actuaries* **10**(3), 142–147.

Berends, K., King, T. B. et al. (2015), 'Derivatives and collateral at US life insurers', *Economic Perspectives* **39**(1), 21–38.

Bernheim, D. (1991), 'How Strong Are Bequest Motives? Evidence Based on Estimates of the Demand for Life Insurance and Annuities', *Journal of Political Economy* **99**(5), 899–927.

Bolton, P. & Scharfstein, D. S. (1990), 'A theory of predation based on agency problems in financial contracting', *American Economic Review* pp. 93–106.

Bolton, P. & Scharfstein, D. S. (1996), 'Optimal debt structure and the number of creditors', *Journal of Political Economy* **104**(1), 1–25.

Braun, R. A., Kopecky, K. A. & Koreshkova, T. (2019), 'Old, Frail, and Uninsured: Accounting for Features of the US Long-Term Care Insurance Market', *Econometrica* **87**(3), 981–1019.

Brown, J., Kapteyn, A., Luttmer, E. F. & Mitchell, O. (2017), 'Cognitive Constraints on Valuing Annuities', *Journal of European Economic Association*

15(2), 429–462.

Brown, J., Kling, J., Mullainathan, S. & Wrobel, M. (2008), 'Why Don't People Insure Late Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle', *American Economic Review, Papers and Proceedings* **98**(2), 304–309.

Chatterjee, S. & Bose, A. (2005), 'Generalized Bootstrap for Estimating Equations', *The Annals of Statistics* **33**(1), 414–436.

Cutler, D. M. (1996), Why don't markets insure long-term risk?, Technical report, Working Paper, Harvard University and National Bureau Economic Research.

Cutler, D. M. & Gruber, J. (1996), 'Does public insurance crowd out private insurance?', *The Quarterly Journal of Economics* **111**(2), 391–430.

Davidoff, T., Brown, J. & Diamond, P. (2005), 'Annuities and Individual Welfare', *American Economic Review* **95**(5), 1573–1589.

Domanski, D., Shin, H. S. & Sushko, V. (2017), 'The hunt for duration: not waving but drowning?', *IMF Economic Review* **65**(1), 113–153.

Dushi, I. & Webb, A. (2004), 'Household annuitization decisions: simulations and empirical analyses', *Journal of Pension Economics and Finance* **3**(2).

Eichenbaum, M. & Peled, D. (1987), 'Capital accumulation and annuities in an adverse selection economy', *Journal of Political Economy* **95**(2), 334–354.

Einav, L. & Finkelstein, A. (2011), 'Selection in insurance markets: Theory and empirics in pictures', *Journal of Economic Perspectives* **25**(1), 115–38.

Finkelstein, A. & Poterba, J. (2004), 'Adverse selection in insurance markets: Policyholder evidence from the U.K. annuity market', *Journal of Political Economy* **112**(1), 183–208.

Finkelstein, A. & Poterba, J. (2006), 'Testing for Adverse Selection with 'Unused Observables'', *NBER Working Paper 12112* .

Foley-Fisher, N., Heinrich, N. & Verani, S. (2020), 'Capturing the Illiquidity Premium', *Available at SSRN* .

Foley-Fisher, N., Narajabad, B. & Verani, S. (2016), Securities lending as wholesale funding: Evidence from the us life insurance industry, Technical report, National Bureau of Economic Research.

Foley-Fisher, N., Narajabad, B. & Verani, S. (2020), 'Self-fulfilling runs: Evidence from the us life insurance industry', *Journal of Political Economy* **128**(9), 3520–3569.

Friedman, B. & Warshawsky, M. (1990), 'The cost of annuities: Implications for saving behavior and bequests', *The Quarterly Journal of Economics* **105**(1), 135–154.

Froot, K. A. & Stein, J. C. (1998), 'Risk management, capital budgeting, and capital structure policy for financial institutions: an integrated approach', *Journal of Financial Economics* **47**(1), 55–82.

Ge, S. (2019), Internal Capital Markets and Product Pricing: Evidence from Weather and Life Insurance Premiums, Technical report, Working Paper.

Golosov, M. & Tsyvinski, A. (2007), 'Optimal taxation with endogenous insurance markets', *The Quarterly Journal of Economics* **122**(2), 487–534.

Gottlieb, D. (2018), 'Prospect theory, life insurance, and annuities', *Working Paper* .

Hart, O. & Moore, J. (1994), 'A theory of debt based on the inalienability of human capital', *The Quarterly Journal of Economics* **109**(4), 841–879.

Hart, O. & Moore, J. (1998), 'Default and Renegotiation: A Dynamic Model of Debt', *The Quarterly Journal of Economics* **113**(1), 1–41.

Hartley, D., Paulson, A. & Rosen, R. J. (2016), 'Measuring interest rate risk in the life insurance sector', *The economics, regulation, and systemic risk of insurance markets* p. 124.

Hong, J. H. & Ríos-Rull, J.-V. (2007), 'Social security, life insurance and annuities for families', *Journal of Monetary Economics* **54**(1), 118–140.

Hosseini, R. (2015), 'Adverse selection in the annuity market and the role for social security', *Journal of Political Economy* **123**(4), 941–984.

Hosseini, R. & Shourideh, A. (2019), 'Retirement financing: An optimal reform approach', *Econometrica* **87**(4), 1205–1265.

Huang, C., Oehmke, M. & Zhong, H. (2019), 'A Theory of Multi-Period Debt Structure', *Review of Financial Studies* .

Koenker, R. (2004), 'Quantile Regression for Longitudinal Data', *Journal of Multivariate Analysis* **91**(1), 74–89.

Koijen, R. S. & Yogo, M. (2015), 'The cost of financial frictions for life insurers', *American Economic Review* **105**(1), 445–75.

Koijen, R. & Yogo, M. (2018), The fragility of market risk insurance, Technical report, National Bureau of Economic Research.

Lockwood, L. M. (2011), 'Bequest motives and the annuity puzzle', *Review of Economic Dynamics* **15**(2), 226–243.

Lockwood, L. M. (2018), 'Incidental bequests and the choice to self-insure late-life risks', *American Economic Review* **108**(9), 2513–50.

Mitchell, O. S., Poterba, J. M., Warshawsky, M. J. & Brown, J. R. (1999), 'New evidence on the money's worth of individual annuities', *American economic review* **89**(5), 1299–1318.

Modigliani, F. & Miller, M. H. (1958), 'The cost of capital, corporation finance and the theory of investment', *American Economic Review* **48**(3), 261–297.

Ozdagli, A. K. & Wang, Z. K. (2019), 'Interest rates and insurance company investment behavior', *Available at SSRN 3479663* .

Rothschild, M. & Stiglitz, J. (1976), 'Equilibrium in competitive insurance markets: An essay on the economics of imperfect information', *The Quarterly Journal of Economics* **90**(4), 629–649.

Yaari, M. E. (1965), 'Uncertain lifetime, life insurance, and the theory of the consumer', *Review of Economic Studies* **32**(2), 137–150.

What's Wrong with Annuity Markets?

Stéphane Verani and Pei Cheng Yu

APPENDIX FOR ONLINE PUBLICATION ONLY

A Sizing up the U.S. life annuity market

We provide two estimates of the size of the U.S. life annuity market using company-level data on the number of annuity contracts and account balances reported in the 2018 NAIC Statutory Fillings of over 800 life insurers. First, we calculate that individuals in the U.S. accumulated about \$2.5 trillion in the form of deferred fixed annuities. This corresponds to roughly \$42,500 per American aged between 50 and 65 years. By contrast, a back-of-the-envelope calculation using the aggregate payment from life insurers to life annuity contract holders, assuming a 6 percent average yield, suggests that Americans annuitize only about \$625 billion of their wealth with life insurers, or approximately \$12,700 per person aged 65 years and above. This first estimate suggests that new retirees annuitize a relatively small share of their wealth with a life insurer. Second, using the same data, we calculate that the U.S life insurance industry's total payments to annuitants is about 3.5 percent of the total payments made by the U.S. Social Security Administration in 2018. Figure 7 and 8 plots the time series of income and payout for the U.S. life insurance industry and the U.S. Social Security Administration, respectively.

B Theoretical results and proofs

Proof of Theorem 1: By Assumption 1, insurers must remain solvent in order to sell annuities. Thus, limited-liability insurers must engage in IRM so that $NW_1(R_2) \geq 0$ for any R_2. The optimal IRM strategy comprises of an optimal asset portfolio and an optimal capital structure, which can be derived following the arguments in Section 2.

Figure 7: U.S. life insurers income from fixed annuity sales and Social Security income

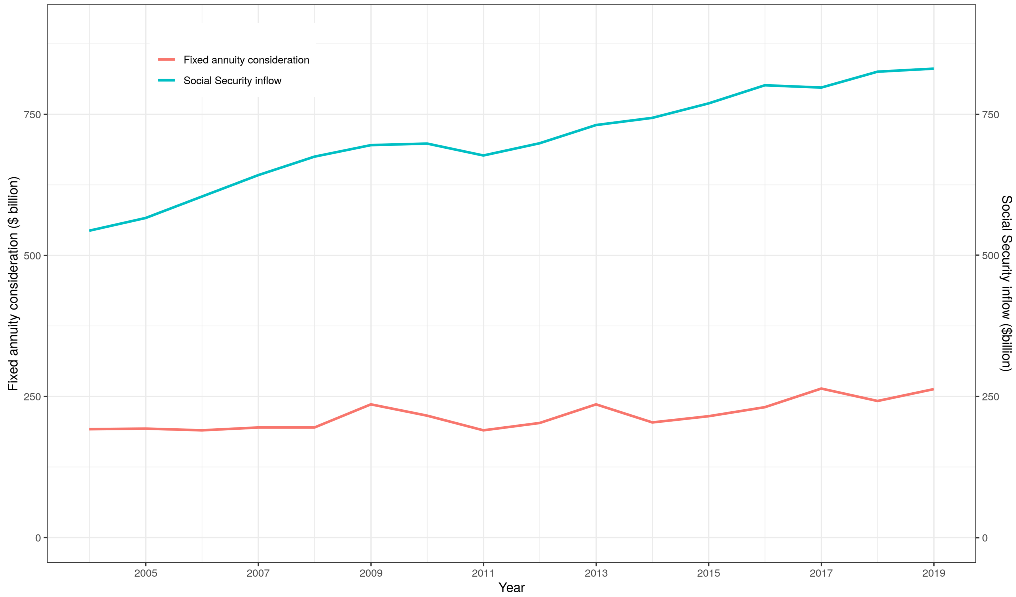
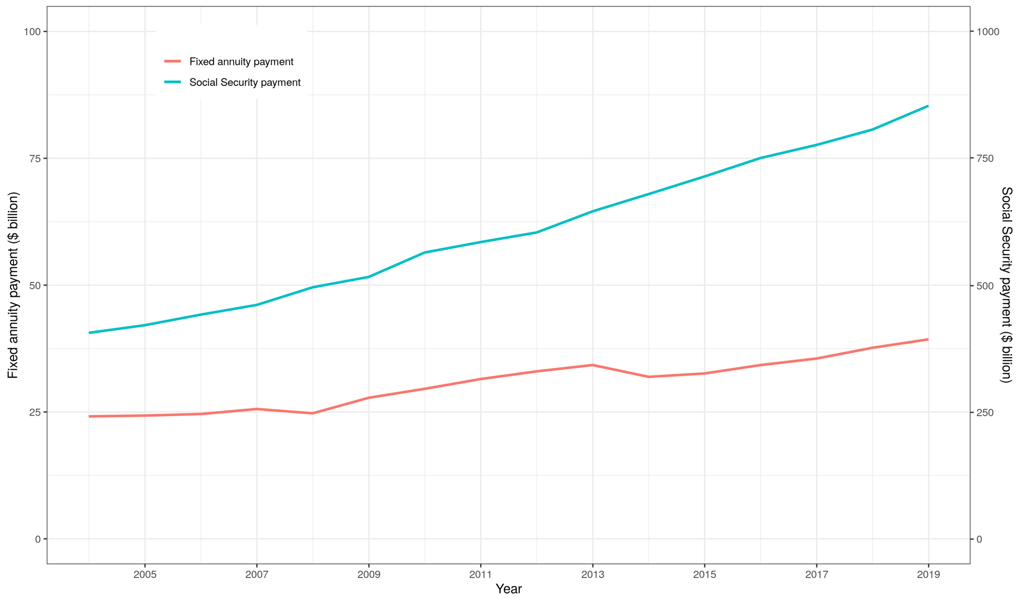


Figure 8: U.S. life insurers payout on fixed annuity and Social Security payout



Finally, we show that the optimal IRM is also unique in a competitive equilibrium. Let $\{b_1, l_2, b_2(R_2), NW_0, NW_1(R_2)\}$ denote the optimal asset portfolio and capital structure for a given market price q. Notice that b_1 is uniquely pinned down by (1), NW_0 is uniquely pinned down by (2), $NW_1(R_2)$ is uniquely pinned down by (3), and $b_2(R_2)$ is uniquely pinned down by (4). Therefore, to show uniqueness, it is sufficient to show that at the optimum $l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$. There are two cases to consider: (i) when $\zeta < 1$ and (ii) when $\zeta = 1$. First, we will examine case (i). When $\zeta < 1$, taking the original price q as given, suppose an insurer deviates and chooses $\hat{l}_2 = l_2 - \epsilon$, where $\epsilon \in (0, l_2]$. Then, by (1), the new short-term bond demand at $t = 0$ is $\hat{b}_1 = b_1 + \epsilon$. This implies that the new short-term bond demand at $t = 1$ is $\hat{b}_2(R_2) = b_2(R_2) + \frac{R_1}{R_2}\left(R_2 - \frac{1}{\mathbb{E}\left(\frac{1}{R_2}\right)}\right)\epsilon$ by (4). Hence, by (3), the new net worth at $t = 1$ is $\hat{NW}_1(1) < 0$ when $R_2 = 1$. As a result, insurers become insolvent if their long-term bond demand is smaller than $\zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$. Finally, we will analyze case (ii). When $\zeta = 1$, taking the original price q as given, suppose an insurer deviates and chooses $\hat{l}_2 = l_2 + \epsilon$, where $\epsilon > 0$. The new short-term bond demand is $\hat{b}_1 = b_1 - \epsilon$ and $\hat{b}_2(R_2) = b_2(R_2) - \frac{R_1}{R_2}\left(R_2 - \frac{1}{\mathbb{E}\left(\frac{1}{R_2}\right)}\right)\epsilon$. Therefore, the new net worth at $t = 1$ is $\hat{NW}_1(1) > 0$ when $R_2 = 1$. This is not optimal since, to be competitive in a Bertrand setting, the insurer can perform IRM with less net worth and charge a lower price. Also, by the same argument as above, the insurer will violate the limited liability constraint if it deviates and chooses $\hat{l}_2 = l_2 - \epsilon$, where $\epsilon \in (0, l_2]$. Hence, it is optimal for $l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$ for any $\zeta \in [0, 1]$. This proves uniqueness. ■

Lemma 1 *Under Bertrand competition, no insurer earns a strictly positive profit and at least two insurers implement the optimal IRM strategy.*

Proof The first part of Lemma 1 follows from a standard Bertrand competition argument. To see why the equilibrium features at least two insurers managing interest rate risk, suppose that instead no insurers manage interest rate risk according to the strategy in Theorem 1. In this case, an insurer can earn a strictly positive profit by choosing a price q and implementing the hedging strategy in

Theorem 1, which is a contradiction. ■

Proof of Theorem 2: Rewrite

$$q^{AF}\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha = \frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha.$$

as

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \left[q^{AF} - q^{CI}\left(\alpha\right)\right] a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha = 0.$$

There exists $\alpha^* \in (\underline{\alpha}, \overline{\alpha})$ such that $a\left(\alpha^*, q^{AF}\right) > 0$, and $q^{AF} > q^{CI}\left(\alpha\right)$ for any $\alpha < \alpha^*$ and $q^{AF} < q^{CI}\left(\alpha\right)$ for any $\alpha > \alpha^*$. This yields

$$0 = \int_{\underline{\alpha}}^{\alpha^*} \left[q^{AF} - q^{CI}\left(\alpha\right)\right] a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha + \int_{\alpha^*}^{\overline{\alpha}} \left[q^{AF} - q^{CI}\left(\alpha\right)\right] a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha$$
$$< a\left(\alpha^*, q^{AF}\right)\int_{\underline{\alpha}}^{\overline{\alpha}} \left[q^{AF} - q^{CI}\left(\alpha\right)\right] g\left(\alpha\right) d\alpha.$$

The result follows as $a\left(\alpha^*, q^{AF}\right) > 0$. ■

Proof of Theorem 3: Let $z = 1 - \zeta + \zeta\mathbb{E}\left(\frac{1}{R_2}\right)$ so insurers' profit can be equivalently expressed as

$$\Pi\left(q, z\right) = q\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q\right) g\left(\alpha\right) d\alpha - \frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha\left(1+\alpha z\right) a\left(\alpha, q\right) g\left(\alpha\right) d\alpha,$$

where $z \in \left[\mathbb{E}\left(\frac{1}{R_2}\right), 1\right]$. First, we show that there exists a $q^* = \min\left\{q | \Pi\left(q, z\right) = 0\right\}$ for any z. By Assumption 1, when $q = \frac{\overline{\alpha}}{R_1}\left(1+\epsilon\overline{\alpha}\right)$, then $\Pi\left(q, z\right) > 0$. Also, when $q = 0$, then $\Pi\left(q, z\right) < 0$. Since $\Pi\left(q, z\right)$ is continuous in q, the intermediate value theorem implies that there exists q such that $\Pi\left(q, z\right) = 0$. Therefore, the set $\left\{q | \Pi\left(q, z\right) = 0\right\}$ is non-empty. Also, $\left\{q | \Pi\left(q, z\right) = 0\right\}$ is closed, because $\{0\}$ is closed and Π is continuous in q so $\Pi^{-1}\left(\{0\}, z\right)$ is closed. Furthermore, $\left\{q | \Pi\left(q, z\right) = 0\right\}$ is bounded below by zero. Hence, a minimum for $\left\{q | \Pi\left(q, z\right) = 0\right\}$ exists. Next, through implicit differentiation,

$$\frac{\partial q^*}{\partial z} = \frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha}{\frac{\partial \Pi\left(q^*, z\right)}{\partial q^*}}.$$

Immediately, notice the numerator is strictly positive. Suppose the denominator, $\frac{\partial \Pi\left(q^*, z\right)}{\partial q^*}$, is strictly negative. This then implies an insurer can deviate by lowering the price to capture the whole market and earn strictly positive profits. However, this contradicts the fact that $q^* = \min\{q|\Pi(q,z)=0\}$. Hence, $\frac{\partial q^*}{\partial z} > 0$ and since z is inversely related to ζ, we have $\frac{\partial q^*}{\partial \zeta} < 0$. ■

C Duration and convexity matching with optimal interest rate risk management

In this appendix, we show that insurers achieve both duration and convexity matching under the optimal IRM strategy outlined in Theorem 1. Recall that, the duration of an asset or liability is the elasticity of its present value PV with respect to the interest rate $R_2 : D = -\frac{\partial PV}{\partial R_2}\frac{R_2}{PV}$. Let $D(L)$ and $D(A)$ denote the duration of the liabilities and assets, respectively. The present value of annuity liabilities as a function of the realized value of R_2 is $\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left(1+\frac{\alpha}{R_2}\right) a(\alpha,q)\, g(\alpha)\, d\alpha$. By Theorem 1, for a given ζ, the present value of net worth NW_0 as a function of the realized value of R_2 is $\left(\frac{1-\zeta}{R_1}\right)\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2\left(1-\frac{1}{R_2}\right) a(\alpha,q)\, g(\alpha)\, d\alpha$. Therefore, the duration of insurance liabilities at $t=0$ (before the realization of R_2) under the optimal IRM is given by

$$D(L) = \frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1 R_2} a(\alpha,q)\, g(\alpha)\, d\alpha - \left(\frac{1-\zeta}{R_1 R_2}\right)\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q)\, g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left(1+\frac{\alpha}{R_2}\right) a(\alpha,q)\, g(\alpha)\, d\alpha + \left(\frac{1-\zeta}{R_1}\right)\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2\left(1-\frac{1}{R_2}\right) a(\alpha,q)\, g(\alpha)\, d\alpha}.$$

To calculate the insurer's asset duration, first note that the one-period bond has 0 duration. Moreover, by Theorem 1, the present value of two-period bonds as a function of the realized value of R_2 is $\frac{\zeta}{R_1 R_2}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q)\, g(\alpha)\, d\alpha$. Therefore, the duration of the insurer's assets under the optimal IRM is

$$D(A) = \frac{\frac{\zeta}{R_1 R_2}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q)\, g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left(1+(1-\zeta)\alpha\right) a(\alpha,q)\, g(\alpha)\, d\alpha + \frac{\zeta}{R_1 R_2}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q)\, g(\alpha)\, d\alpha},$$

which is equal to $D(L)$ and shows that optimal IRM strategy in Theorem 1 closes the duration gap.

Finally, define convexity as $C = \tilde{D}^2 - \frac{\partial \tilde{D}}{\partial R_2}$, where $\tilde{D} = D/R_2$, which is known as the modified duration. Since $D(L) = D(A)$ for any R_2, the convexity of the assets and liabilities under the optimal IRM strategy are also matched: $C(L) = C(A)$.

Therefore, the optimal IRM strategy also achieves convexity matching.

D Interaction between supply- and demand-side frictions

In this appendix, we derive a relationship between shocks originating from the corporate bond market and adverse selection in the annuity market from the model presented in Section 2. We then test this relationship using variation in life annuity markups arising from the length of the contract guarantee period.

It is well known that higher annuity prices are associated with more severe adverse selection (Rothschild & Stiglitz 1976). We can redefine the limit on the supply of long-term bonds as $z = 1 - \zeta + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)$, so z is inversely related to ζ. Using our model, we can decompose the effect of a change in the supply of the long-term bond z on the equilibrium annuity price into a *risk management effect* and an *adverse selection effect*:

$$\frac{\partial q^*}{\partial z} = \underbrace{\frac{\frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q^*) g(\alpha) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) d\alpha}}_{\text{Risk management effect}} + \underbrace{\frac{\partial q^*}{\partial z} \frac{\int_{\underline{\alpha}}^{\overline{\alpha}} e(\alpha, q^*) \left[1 - \frac{\frac{\alpha}{R_1}(1+\alpha z)}{q^*}\right] a(\alpha, q^*) g(\alpha) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) d\alpha}}_{\text{Adverse selection effect}}, \tag{7}$$

where $e(\alpha, q^*)$ is the price elasticity of annuity demand: $e(\alpha, q^*) = -\frac{\partial a(\alpha, q^*)}{\partial q^*} \frac{q^*}{a(\alpha, q^*)}$. Since a marginal decrease in two-period bond supply necessitates an increase in net worth, each component of $\frac{\partial q^*}{\partial z}$ represents the average marginal effect of IRM on annuity price because each component is normalized by the total amount of annuity supplied.

A marginal increase in IRM raises the equilibrium annuity price because the insurer must finance greater net worth in $t = 0$. By Theorem 1, we can express the optimal amount of net worth at $t = 0$ as a function of z:

$$NW_0(z) = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[z - \mathrm{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q^*) g(\alpha) d\alpha.$$

It follows that the risk management effect in equation (7) can be written as

$$\text{Risk management effect} = \frac{\frac{\partial NW_0(z)}{\partial z}}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha},$$

showing that higher net worth for IRM requires insurers to charge a higher equilibrium annuity price.

Life insurers' IRM amplifies the effect of adverse selection in the annuity market because the average survival probability of individuals purchasing annuities increases as insurers charge a higher annuity price to finance their net worth. Equation (7) shows that the main determinant of the adverse selection effect is the price elasticity of annuity demand. The adverse selection effect amplifies the price increase if demand is more elastic for agents with lower survival probability $\alpha : e\left(\alpha', q^*\right) > e\left(\alpha'', q^*\right)$ when $\alpha' < \alpha''$. This is because the insurer would lose more agents of high mortality risk than those with lower risk when price increases, which worsens the adverse selection problem and precipitates further increases in price. This phenomenon is often referred to as a death spiral. To see this, first note that among agents purchasing annuities (all α such that $a\left(\alpha, q^*\right) > 0$), there exists a $\tilde{\alpha}$ such that $q^* < \frac{\alpha}{R_1}\left(1 + \alpha z\right)$ for any $\alpha > \tilde{\alpha}$ and $q^* > \frac{\alpha}{R_1}\left(1 + \alpha z\right)$ for any $\alpha < \tilde{\alpha}$. Due to competition,

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \left(1 - \frac{\frac{\alpha}{R_1}\left(1 + \alpha z\right)}{q^*}\right) a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha = 0,$$

which implies the existence of $\tilde{\alpha}$.

Next, notice that $\frac{\alpha}{R_1}\left(1 + \alpha z\right)$ corresponds to the *full information* actuarially fair price when the constraint on the supply of long-term bonds is *binding*. This means insurers make a profit off of mortality types $\alpha < \tilde{\alpha}$, which equates to the loss from types with $\alpha > \tilde{\alpha}$ due to competition. If demand is more elastic for agents with low α, then insurers lose more agents with mortality type $\alpha < \tilde{\alpha}$ than mortality type $\alpha > \tilde{\alpha}$ from an increase in annuity price. This creates a net loss, so the insurer must further raise prices to compensate for more severe adverse selection. This theoretical result establishes a link between the supply-side and demand-side frictions, connected by the IRM channel.

D.1 Regression results

The above discussion suggests that corporate bond market shocks may have a direct effect on adverse selection. We look for evidence of this effect by exploiting differences in life insurers' pricing of SPIA with different types of "period certain" guarantees and by measuring the relative change in AS pricing for these products. Note that individuals choosing a life annuity with 10- or 20-year guarantee think they are at a higher risk of dying within the next 10 or 20 years (Finkelstein & Poterba 2004, 2006).

We measure AS pricing as the difference between the total annuity markup and the AS-adjusted markup as follows:

$$AS\_pricing_{ijt} = \frac{P_{ijt}(n,S,M)}{V_{jt}^{\text{General}}(n,S,M,r=\text{HQM})} - \frac{P_{ijt}(n,S,M)}{V_{jt}^{\text{Basic}}(n,S,M,r=\text{HQM})}.$$

We then test the hypothesis that $AS\_pricing_{ijt}$ for annuity contracts with longer guarantee periods increase more when regulatory reserve requirements increase in a difference-in-differences framework. In this test, the first difference is between annuity contracts j offered by insurer i with a long guarantee period and annuity contracts $-j$ offered by the same insurer with no guarantee period. The second difference is between periods in which reserve requirements are more binding and periods in which reserve requirements are less binding. We implement our test in a linear regression framework as follows:

$$\begin{aligned} AS\_pricing_{ijt} = & \alpha_1^i + \alpha_2^j + \beta_1 Baa\text{-}Aaa\_spread_t + \beta_2 10\_HQM\_spread_t \\ & + \beta_3 10yr\_guarantee\_period + \beta_4 20yr\_guarantee\_period \\ & + \beta_5 10yr\_guarantee\_period \times Reserve\_Ratio_{jt} \\ & + \beta_6 20yr\_guarantee\_period \times Reserve\_Ratio_{jt} \\ & + \beta_7 Reserve\_Ratio_{jt} + \mathbf{z}_{it}'\boldsymbol{\gamma} + \epsilon_{ijt}, \end{aligned} \tag{8}$$

where $10yr\_guarantee\_period$ and $20yr\_guarantee\_period$ are binary variables indicating the guarantee period length. The coefficients β_5 and β_6 on the interaction terms are measured relative to the effect on SPIA without guarantee period, which is the third type of life annuity contract in our sample and is omitted from this regression. As with our main specification in Section 5, we focus on within-insurer variation using insurer fixed effects and we condition our test on $Baa\text{-}Aaa\_spread_t$ and the average cost of funding of the insurer proxied with $10\_HQM\_spread_t$. The vector $\mathbf{z}'_{it}$ includes other insurer-level time-varying financial variables, such as insurer size and leverage.

Table 4 summarizes the results of regression (8). The coefficients in Column 1 show that an exogenous increase in statutory reserve requirements disproportionately increases the AS pricing in life annuities with 10- and 20-year guarantees relative to life annuities without guarantees. For example, a standard deviation increase in the reserve ratio *decreases* the AS pricing of life annuities without a guarantee period by 0.84 percentage point. In contrast, the AS pricing of life annuities with 10- and 20-year guarantees *increase* by 0.5 and 0.6 percentage point, respectively, as a response to the same shock. We continue to report two-way insurer and date clustered robust standard errors. The results in Column 2 are broadly similar when the same specification is estimated on a shorter sample period with time-varying insurer-level financial controls. Because individuals choosing life annuities with period-certain guarantees think they are at a higher risk of dying within the next few years, this implies that changes in corporate bond market conditions have a direct effect on adverse selection in annuity markets, which is reflected, at least partly, in annuity prices.

E Unlimited long-term government bonds

In the paper, we showed how insurers have to accumulate a positive net worth to hedge against the interest rate risk when there is a shortage of efficiently-priced long-term (two-period) corporate bonds. In this appendix, we consider a model with an unlimited supply of two-period government bonds. We find that unless the two-period government bond is priced efficiently, insurers still require a positive net worth and competitive annuity prices would be strictly higher than the risk-adjusted actuarially fair price when the supply of long-term corporate bonds is constrained ($\zeta < 1$).

Table 4: The effect of corporate bond market shocks on adverse selection The unit of observation is a life insurer-product-year. The dependent variable *Adverse selection pricing*$_{ijt}$ is the difference between the markup computed using the general population mortality table and the corresponding markup computed using the annuitant pool mortality table for annuity j sold by insurer i in year t. Two-way insurer and date cluster robust standard errors are reported in parentheses in Columns (1) and (2), respectively. *** p < 0.01; ** p<0.05; * p<0.1.

Dep. variable:	$AS\_pricing_{ijt}$	
	(1)	(2)
$10yr\_Guarantee \times Reserve\_Ratio_{jt}$	25.46***	23.29***
	(3.33)	(3.75)
$20yr\_Guarantee \times Reserve\_Ratio_{jt}$	26.97***	26.34***
	(3.62)	(4.42)
$Reserve\_Ratio_{jt}$	-17.17***	-14.67***
	(2.62)	(3.00)
$10yr\_Guarantee$	-30.18***	-27.89***
	(3.41)	(3.78)
$20yr\_Guarantee$	-34.99***	-34.19***
	(3.71)	(4.44)
$Baa\text{-}Aaa\_spread_t$	0.64***	0.72***
	(0.23)	(0.24)
$10\_HQM\_spread_t$	-0.76***	-0.84***
	(0.15)	(0.17)
$Leverage\_ratio_{it}$		-0.01
		(0.01)
$Log\_total\_assets_{it}$		0.58**
		(0.26)
Insurer FE	Y	Y
Observations	40,790	29,462
Adjusted R^2	0.70	0.68

In addition to the economic environment of Section 2, there is an unlimited supply of zero-coupon two-period government bonds g_2. One unit of government bond purchased at $t = 0$ returns R_g at $t = 2$. We assume that the returns from the long-term government bond is less than the long-term corporate bond: $R_g < R_l$. If not, the corporate bond market would disappear. For simplicity, we do not model the economic forces, such as risk and liquidity, that determine the difference between R_l and R_g.

At $t = 0$, an insurer invests its annuity considerations in a portfolio of corporate bonds and long-term government bonds:

$$b_1 + l_2 + g_2 = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{9}$$

The insurer's balance sheet at $t = 0$ is given by

$$b_1 + l_2 + g_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha + NW_0. \tag{10}$$

At $t = 1$, the aggregate shock R_2 is realized, and the insurers' balance sheet becomes:

$$b_2(R_2) = \frac{1}{R_2} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha + NW_1(R_2), \tag{11}$$

where

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{R_2} + \frac{R_g g_2}{R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{12}$$

Since $R_l > R_g$, we know that when $\zeta < 1$ the insurers will purchase

$$l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q)\, g(\alpha)\, d\alpha.$$

The following theorem shows how insurers construct their asset portfolio and capital structure to manage interest rate risk when government bonds are available.

Theorem 4 *For a given annuity price q and $\zeta < 1$, the unique optimal IRM strategy when $R_g \in (R_1, R_l)$ requires an asset allocation and a capital structure that satisfies:*

i. Asset portfolio:

$$b_1 = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q) g(\alpha) d\alpha,$$
$$l_2 = \frac{\zeta}{R_l} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q) g(\alpha) d\alpha,$$
$$g_2 = \frac{1-\zeta}{R_g} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q) g(\alpha) d\alpha,$$
$$b_2(R_2) = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_2} a(\alpha, q) g(\alpha) d\alpha.$$

ii. Capital structure:

$$NW_0 = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left(\frac{1}{R_g} - \frac{1}{R_l}\right) a(\alpha, q) g(\alpha) d\alpha,$$
$$NW_1(R_2) = 0 \textit{ for all } R_2.$$

When $R_g \leq R_1$, the optimal IRM has $g_2 = 0$ and requires the same asset portfolio and capital structure as the environment without government bonds.

Proof The competitive insurers accumulate just enough net worth so that $NW_1(R_2) = 0$ when $R_2 = 1$ and $NW_1(R_2) \geq 0$ when $R_2 > 1$. As a result, from (11), $b_2(1) = \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q) g(\alpha) d\alpha$. Hence, by (12), we have the following demand for one-period bonds $b_1(l_2, g_2)$ as a function of long-term corporate and government bonds:

$$b_1(l_2, g_2) = \frac{1}{R_1}\left[\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha(1+\alpha) a(\alpha, q) g(\alpha) d\alpha - R_l l_2 - R_g g_2\right].$$

Substituting the demand $b_1(l_2, g_2)$ into (10) yields

$$NW_0 + \left(\frac{R_g}{R_1} - 1\right) g_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1}\left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) g(\alpha) d\alpha - \left(\frac{R_l}{R_1} - 1\right) l_2. \tag{13}$$

Since $\zeta < 1$ and $R_g < R_l$, the equilibrium demand for long-term corporate bonds is

$$l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha.$$

Therefore, (13) can be rewritten as

$$NW_0 + \left(\frac{R_g}{R_1} - 1\right) g_2 = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1}\left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) g(\alpha) d\alpha. \tag{14}$$

The right-hand side of the equation above is total corporate assets minus insurer liabilities. As a result, when $R_g \leq R_1$, insurers prefer to accumulate net worth to manage the interest rate risk since the return to long-term government bonds is too low.

Finally, when $R_g \in (R_1, R_l)$, net worth and long-term government bonds are imperfect substitutes. Due to competition, insurers choose g_2 such that

$$\min \{NW_0, NW_1(R_2)\} \geq 0$$

for any R_2. We will first focus on the case with $NW_1(R_2) = 0$ and then show that $NW_0 = 0$ is not optimal when $R_g < R_l$.

From (11) and (12), we have

$$NW_1(R_2) = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left(1 - \frac{1}{R_2}\right) a(\alpha, q) g(\alpha) d\alpha - R_g \left(1 - \frac{1}{R_2}\right) g_2.$$

Suppose that $g_2 = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_g} a(\alpha, q) g(\alpha) d\alpha$ so $NW_1(R_2) = 0$ for all R_2. Then, we have the following net worth at $t = 0$:

$$NW_0 = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left(\frac{1}{R_g} - \frac{1}{R_l}\right) a(\alpha, q) g(\alpha) d\alpha,$$

which is strictly positive when $R_g < R_l$.

Next, we show that $NW_0 = 0$ is not optimal. To see this, suppose $NW_0 = 0$, then (14) implies

$$g_2 = \frac{1-\zeta}{R_g - R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) g(\alpha) d\alpha.$$

From (12),

$$\begin{aligned} b_2(R_2) = &\left(1 - \zeta + \frac{\zeta}{R_2}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q) g(\alpha) d\alpha \\ &- \left(1 - \frac{1}{R_2}\right) \left(\frac{(1-\zeta) R_g}{R_g - R_1}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) g(\alpha) d\alpha, \end{aligned}$$

so by (11) the net worth at $t = 1$ is

$$NW_1(R_2) = (1-\zeta) \left(1 - \frac{1}{R_2}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[1 - \frac{R_g \left(1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right)}{R_g - R_1}\right] a(\alpha, q) g(\alpha) d\alpha,$$

which is strictly negative when $R_g < R_l$. Hence, at the optimum, $NW_0 > 0$ and $NW_1(R_2) = 0$ for all R_2.

Given the optimal capital structure when long-term government bonds are available, we can derive the optimal asset portfolio. Since $g_2 = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_g} a(\alpha, q) g(\alpha) d\alpha$

and $l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$, the short-term bond demand is

$$b_1 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} a(\alpha, q) g(\alpha) d\alpha,$$

and from (12),

$$b_2(R_2) = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_2} a(\alpha, q) g(\alpha) d\alpha.$$

The proof for uniqueness follows a similar argument as the one presented in the proof for Theorem 1 in Appendix B. This completes the proof. ■

Theorem 4 shows how insurers use long-term government bonds to manage interest rate risk. When the returns of long-term government bonds are too low ($R_g \leq R_1$), then their presence does not change the insurers' IRM. When the returns of long-term government bonds are higher, then insurers can substitute some of the net worth at $t = 0$ with long-term government bonds to perform IRM. However, it is not a perfect substitute when the returns of government bonds are such that $R_l < R_g$, which could reflect a convenience yield due to their high liquidity and safe-haven status. In this case, net worth at $t = 0$ is remains positive.

To formally show that the competitive annuity price remains above the risk-adjusted actuarially fair price in this setting, Theorem 4 implies that the equilibrium annuity price is characterized by

$$q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q) g(\alpha) d\alpha = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left[1 + \alpha \left(\frac{R_1}{R_g} - \zeta \frac{R_1}{R_g} + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)\right)\right] a(\alpha, q) g(\alpha) d\alpha.$$

Similar to Section 2, when the corporate bond market is unconstrained ($\zeta = 1$), the competitive annuity price is equal to the risk-adjusted actuarially fair price. On the other hand, in contrast to Section 2, the annuity liabilities are lower when unlimited long-term government bonds are available but long-term corporate bonds are not ($\zeta < 1$). This shows how the AS-adjusted markup decreases when insurers can use long-term government bonds to manage interest rate risk, but it is still strictly positive when the corporate bond market is constrained.

Let $\hat{z} = \frac{R_1}{R_g} - \zeta \frac{R_1}{R_g} + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)$ and define $\Pi(q, \hat{z})$ as the insurers' profit, such that

$$\Pi(q,\hat{z}) = q\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q)\,g(\alpha)\,d\alpha - \frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha(1+\alpha\hat{z})\,a(\alpha,q)\,g(\alpha)\,d\alpha,$$

where $\hat{z} \in \left[\mathbb{E}\left(\frac{1}{R_2}\right), \frac{R_1}{R_g}\right]$. Let q^* be the equilibrium annuity price—the lowest positive annuity price such that $\Pi(q^*,\hat{z}) = 0$. Then, we can derive a theorem similar to Theorem 3, which shows that q^* is higher than the risk-adjusted actuarially fair price q^{AF} when $\zeta < 1$, even when unlimited long-term government bonds are available.

Theorem 5 *When unlimited long-term government bonds are available, the competitive annuity price increases as the bond market becomes more constrained (higher $\hat{z}$): $\frac{\partial q^*}{\partial \hat{z}} > 0$.*

Proof The proof is similar to the proof of Theorem 3 ■

F Monopolistic competition

In this section, we explore how insurer IRM is affected by market competition. To model imperfect competition, we consider a market with two insurers: $\{1,2\}$. Each insurer is matched with a continuum of individuals of measure 1 with identical survival probability distributions. An insurer can lower the price to capture a portion of its competitor's market: If Insurer 2 sets price q_2, then Insurer 1 can seize $\gamma(q_2 - q_1)$ of individuals that were matched with Insurer 2 when it sets $q_1 < q_2$. For simplicity, the additional individuals that an insurer captures when it lowers its price are independent of the individual's type. Under this assumption about market structure, insurers are monopolists when $\gamma = 0$ and the model collapses to our baseline specification with Bertrand competition when $\gamma \to \infty$.

We restrict our attention to symmetric Nash equilibria in which both life insurers charge the same price q^*. Crucially, prices are affected by how insurers manage the interest rate risk. Though insurers with market power accumulate net worth in the form of monopoly profits, the net worth from exercising market power may prove to be inadequate. This is because insurance regulators or ratings agencies would require insurers to hold minimum annuity reserves to guarantee solvency for all interest rate realizations. Therefore, at the bare minimum, even insurers with market power must at least acquire an asset portfolio and capital structure outlined in Theorem 1. As a result, the cost of selling an annuity is determined by the expected present value of liabilities and the optimal net worth the insurer must hold.

Suppose Insurer 1 deviates to price $\hat{q} < q^*$ while Insurer 2 sets the equilibrium price q^*. We can redefine the limit on the supply of long-term bonds as $z = 1 - \zeta + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)$, so z is inversely related to ζ. Insurer 1 chooses $\hat{q}$ to maximize profit:

$$[1 + \gamma (q^* - \hat{q})] \int_{\underline{\alpha}}^{\overline{\alpha}} \left[\hat{q} - \frac{\alpha}{R_1}(1 + \alpha z)\right] a(\alpha, \hat{q}) \, g(\alpha) \, d\alpha.$$

To solve for $\hat{q}$, we take the first-order condition of the profit maximization problem and then use the fact that in a symmetric Nash equilibrium, $\hat{q} = q^*$ at the optimum.[35] This yields the following equilibrium condition for price q^* :

$$\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) \, g(\alpha) \, d\alpha + \int_{\underline{\alpha}}^{\overline{\alpha}} \left[q^* - \frac{\alpha}{R_1}(1 + \alpha z)\right] \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) \, d\alpha \\ - \gamma \int_{\underline{\alpha}}^{\overline{\alpha}} \left[q^* - \frac{\alpha}{R_1}(1 + \alpha z)\right] a(\alpha, q^*) \, g(\alpha) \, d\alpha = 0.$$

Using this equilibrium condition, Theorem 6 characterizes the relationship between the market structure γ and the AS-adjusted markup.

Theorem 6 *The annuity price increases as the bond market becomes more constrained (higher z): $\frac{\partial q^*}{\partial z} > 0$, and decreases as the annuity market becomes more competitive: $\frac{\partial q^*}{\partial \gamma} < 0$.*

Proof Rewrite the first-order condition as $W(q^*, z, \gamma) = 0$, where

$$W(q^*, z, \gamma) = \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) \, g(\alpha) \, d\alpha + \int_{\underline{\alpha}}^{\overline{\alpha}} \left[q^* - \frac{\alpha}{R_1}(1 + \alpha z)\right] \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) \, d\alpha \\ - \gamma \int_{\underline{\alpha}}^{\overline{\alpha}} \left[q^* - \frac{\alpha}{R_1}(1 + \alpha z)\right] a(\alpha, q^*) \, g(\alpha) \, d\alpha.$$

[35] A sufficient condition for the second order condition to hold is to assume that $\int_{\underline{\alpha}}^{\overline{\alpha}} \left[q - \frac{\alpha}{R_1}(1 + \alpha z)\right] a(\alpha, q) \, g(\alpha) \, d\alpha$ is strictly concave. This assumption is equivalent to requiring a unique optimum to exist when the insurers are monopolists.

Note that $\frac{\partial W}{\partial q^*} < 0$ from the second-order condition, and $\frac{\partial W}{\partial \gamma} < 0$, and $\frac{\partial W}{\partial z} > 0$. From implicit differentiation, $\frac{\partial q^*}{\partial z} = -\frac{\partial W}{\partial z} / \frac{\partial W}{\partial q^*} > 0$ and $\frac{\partial q^*}{\partial \gamma} = -\frac{\partial W}{\partial \gamma} / \frac{\partial W}{\partial q^*} < 0$. ■

Theorem 6 shows that the AS-adjusted markup can increase due to IRM or market power. In other words, insurers either increase annuity prices to fund the net worth needed for IRM or to limit the quantity sold in the market and exercise market power. In essence, Theorem 6 implies that monopolistic competition with an unconstrained bond market—$z = \mathbb{E}\left(\frac{1}{R_2}\right)$—can generate observationally equivalent AS-adjusted markup as a perfectly competitive annuity market facing a bond market that is constrained. Furthermore, Theorem 6 shows that for any given market structure γ, the AS-adjusted markup is strictly positive when the bond market is constrained—$\mathbb{E}\left(\frac{1}{R_2}\right) < z \leq 1$. This implies that even insurers with varying degrees of market power raise their annuity prices to manage the interest rate risk. In the main text, we explain how our difference-in-differences strategy is not sensitive to the annuity market structure.

We investigate competition in the fixed annuity market by calculating a Herfindahl-Hirschman Index (HHI) for the industry. Figure 9 calculates the HHI using insurer-level data on fixed annuity premiums and considerations extracted from about 800 NAIC Statutory fillings—unfortunately, our collection of statutory fillings only starts in 2003. The solid line represents the HHI and shows that the US fixed annuity market concentration is consistently below 8 percent. Figure 9 confirms industry commentaries that the fixed annuity market is very competitive and justifies using perfect competition as a benchmark for our theoretical analysis. Moreover, in addition to starting from an already high level, competition has only intensified further in the aftermath of the 2008–09 Global Financial Crisis. This effect is consistent with the decrease in the cross-sectional variance of markups we noted in Section 5. Foley-Fisher, Heinrich & Verani (2020) explains that this increase in competition coincides with the arrival of private equity firms in the industry that purchased large blocks of legacy annuity business and invested in relatively more liquid assets. By adding more illiquidity on their asset side, these

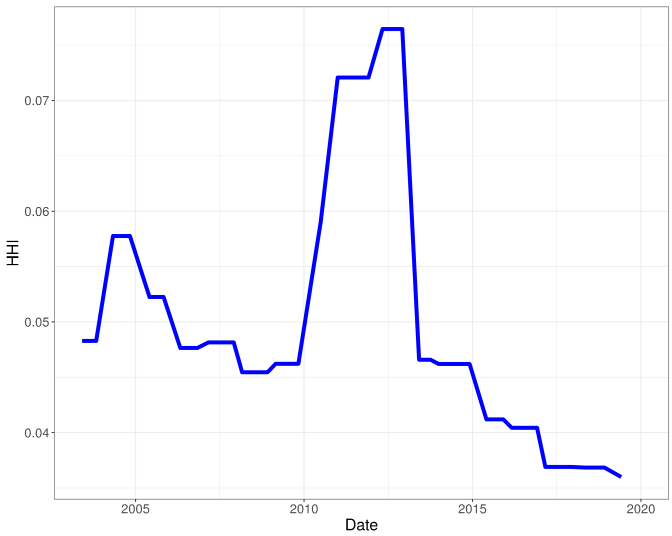


Figure 9: Fixed annuity market Herfindahl-Hirschman Index

PE-backed insurers can offer a higher yield on their new annuity liability, thereby lowering prices.

G Reserve requirements and bond spread shocks

In this appendix, we provide a simple and tractable extension of our model to analyze how exogenous changes in reserve requirements and two-period bond yield spread affect insurers' risk management and pricing decisions.

G.1 Model with regulatory reserve requirements

Let τ denote the reserve ratio as defined in Section 3.2. The variable τ captures both the regulatory interest rate and the "loaded" mortality assumption used to calculate regulatory reserves. We focus on the case when the reserve ratio is binding ($\tau > 1$), since our main results in Section 2 are unchanged if it were slack ($\tau = 1$).

At $t = 0$, the insurer's balance sheet is

$$b_1 + l_2 = \tau \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha + NW_0, \tag{15}$$

and the asset portfolio and capital structure are determined by (1), (3), (4), and

(15). The asset portfolio and capital structure at $t = 1$ remain the same as in Theorem 1 because reserve requirements do not change the risk management problem at $t = 1$. From the insurer's perspective, τ is never strictly below 1 even when the regulatory reserve requirements are relaxed, since it must be able to fulfill the expected present value of liabilities irrespective of the requirements.

Following the derivation in Section 2,

$$b_1(l_2) = \frac{1}{R_1}\left[\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha(1+\alpha)\, a(\alpha,q)\, g(\alpha)\, d\alpha - R_l l_2\right].$$

Equation (15) can be rewritten as

$$\left(\frac{1}{\mathbb{E}\left(\frac{1}{R_2}\right)} - 1\right) l_2 + NW_0 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1}\left[1 - \tau\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha - (\tau-1)\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} a(\alpha,q)\, g(\alpha)\, d\alpha. \quad (16)$$

Equation (16) shows that insurers stop selling annuities when the right-hand side of (16) is negative, i.e., if the reserve ratio is sufficiently high. Hence, we focus on values of τ that are above but sufficiently close to 1 so the insurers choose to sell annuities despite the higher cost. By (16) and compared to Theorem 1, $\tau > 1$ means that either the demand for long-term bonds or the net worth at $t = 0$ must decrease. Because real world insurers' financial strength ratings and regulatory risk-based capital ratios are tied to the insurers' net worth, we can rule out the case when the insurers change their net worth strategy for $t = 0$ as a response to a change in τ and focus on the change in two-period bond holdings.

While maintaining the same net worth strategy for $t = 0$ as in Theorem 1 ($NW_0 = \frac{1-\zeta}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2\left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha$), the optimal $t = 0$ asset portfolio with $\tau > 1$ is

$$l_2 = \zeta\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha,q)\, g(\alpha)\, d\alpha - \frac{(\tau-1)\,\mathbb{E}\left(\frac{1}{R_2}\right)}{1-\mathbb{E}\left(\frac{1}{R_2}\right)}\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1 + \alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha,$$

$$b_1 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1 + (1-\zeta)\alpha\right] a(\alpha,q)\, g(\alpha)\, d\alpha + \frac{\tau-1}{1-\mathbb{E}\left(\frac{1}{R_2}\right)}\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1 + \alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha.$$

It follows that when τ increases, the two-period bond demand decreases; however, aggregate bond demand increases due to a relatively larger demand for the one-period bonds. Furthermore, the optimal asset portfolio is the same as Theorem 1 when $\tau = 1$.

To see how a tightening of reserve requirements affects the equilibrium annuity price, define $\Pi(q,\tau)$ as the insurers' profit:

$$\Pi(q,\tau) = q\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q)\, g(\alpha)\, d\alpha - \tau \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha - (1-\zeta)\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1}\left[1-\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\, g(\alpha)\, d\alpha.$$

As before, the competitive price q^* is determined by the zero profit condition: $\Pi(q^*,\tau) = 0$. By implicit differentiation, we find that $\frac{\partial q^*}{\partial \tau}$ is strictly positive if $\Pi(q,\tau)$ is increasing at q^*, which must be true at the equilibrium. To see this, suppose that $\Pi(q,\tau)$ is strictly decreasing at q^*, then an insurer that deviates from q^* by lowering their annuity price would capture the entire market and earn strictly positive profits, contradicting the fact that q^* is an equilibrium price.

Finally, to see how a change in τ affects the AS-adjusted markup, we first need to clarify the difference between the risk-adjusted actuarially fair price q^{AF} and the regulator's risk-adjusted actuarially fair price $\hat{q}^{AF}$. The risk-adjusted actuarially fair price q^{AF} is defined by the competitive annuity price in an unconstrained bond market from the *insurers'* perspective:

$$q^{AF}\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha,q^{AF}\right) g(\alpha)\, d\alpha = \frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a\left(\alpha,q^{AF}\right) g(\alpha)\, d\alpha.$$

Meanwhile, the regulator's risk-adjusted actuarially fair price $\hat{q}^{AF}$ is defined by the competitive annuity price in a compete bond market from the *regulator's* perspective:

$$\hat{q}^{AF}\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha,\hat{q}^{AF}\right) g(\alpha)\, d\alpha = \frac{\tau}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a\left(\alpha,\hat{q}^{AF}\right) g(\alpha)\, d\alpha.$$

As we explained in Section 3, the regulatory actuarial values $\hat{q}^{AF}$ is different from the insurer's own actuarial valuation q^{AF} because regulators prescribe a discount rate to value annuity liabilities that resets infrequently and also impose a flat loading factor on the annuitant mortality probabilities. The effect of these actuarial assumptions on the regulatory actuarial value $\hat{q}^{AF}$ is captured by the parameter τ. Therefore, $\hat{q}^{AF}$ captures how changes in the reserve ratio affects the cost of providing annuities, while q^{AF} does not. The AS-adjusted annuity markup is defined as the difference between the competitive annuity price q^* and the risk-adjusted actuarially fair price q^{AF}. This definition captures the fact that insurers use a fraction of their annuity markup to finance the binding reserve requirement. As a result, since q^{AF} does not vary with τ and q^* increases with τ, the AS-adjusted markup unambiguously increases with τ.

G.2 Model with two-period yield spread shocks

We map the effect of long-term Baa-Aaa bond yield spread shocks to our model by considering the effect of a change in the expected two-period bond return—the spread of the two-period bond in the model relative to the cash instrument. Let ρ denote the shock on the long-term bond returns such that the return is $\frac{R_l}{\rho}$ in $t = 2$. Our baseline model in Section 2 corresponds to the case when when $\rho = 1$. Because we do not explicitly model bond corporate bond issuers' credit risk, the parameter ρ captures the yield spread on long-duration investment grade bonds in a reduced form. When ρ decreases (increases), the yield on long-term bonds increases (decreases), which corresponds to a widening (narrowing) of the Baa-Aaa spread in the data. Following the exposition in Section 3.3, we assume that ρ is orthogonal to the insurers' discount rate.

Specifically, the insurers' balance sheet is characterized by (1), (2), and (3). However, since the return on long-term bonds is now affected by ρ, (4) becomes

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{\rho R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha.$$

The above equation shows that the insurers' optimal IRM strategy changes with different values of ρ, as ρ affects the cost of managing interest rate risk.

Following the derivation in Section 2, the optimal one-period bond holding in $t = 0$ is given by:

$$b_1(l_2) = \frac{1}{R_1}\left[\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha(1+\alpha)\, a(\alpha, q)\, g(\alpha)\, d\alpha - \frac{R_l l_2}{\rho}\right].$$

Furthermore, when the two-period bond supply is constrained ($\zeta < 1$) and the increase in the two-period bond spread is small—i.e., the decrease in ρ is sufficiently small—the optimal demand for two-period bonds in $t = 0$ is given by:

$$l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha.$$

On the other hand, if the two-period bond supply is unconstrained ($\zeta = 1$) or if the increase in the two-period bond spread is large, the demand for the two-period bond decreases due to competition. Intuitively, when the two-period bond spread is high, insurers need less two-period bonds to manage the interest rate risk associated with their block of annuity business. Therefore, even in an unconstrained bond market, insurers can fully hedge the interest rate risk when $\rho < 1$ with less two-period bonds than $l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$. In the remainder of this section, we focus on the case when the two-period bond supply is constrained and the increase in the two-period bond spread is relatively small.

With $l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$, by (2), the optimal net worth at $t = 0$ is given by:

$$NW_0 = (1 - \zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) g(\alpha) d\alpha - \zeta \left(\frac{1}{\rho} - 1\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} a(\alpha, q) g(\alpha) d\alpha.$$

It follows that insurers decrease their demand for one-period bonds b_1 and decrease their net worth NW_0 when ρ falls. In essence, when the two-period bond spread increases, insurers respond by decreasing both their net worth at $t = 0$ and their demand for one-period bonds.

To see how an increase in the two-period bond spread affects the competitive price, let $\Pi(q, \rho)$ denote the insurers' profit function, which is

$$\Pi(q, \rho) = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q) g(\alpha) d\alpha - \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha (1 + \alpha z) a(\alpha, q) g(\alpha) d\alpha + \zeta \left(\frac{1}{\rho} - 1\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} a(\alpha, q) g(\alpha) d\alpha, \quad (17)$$

where $z = 1 - \zeta + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)$. The competitive price q^* is determined by the zero profit condition: $\Pi(q^*, \rho) = 0$. By implicit differentiation, $\frac{\partial q^*}{\partial \rho}$ is strictly positive if $\Pi(q, \rho)$ is increasing at q^*, which must be true at the equilibrium price.[36] Since

[36]To see this, note that if $\Pi(q, \rho)$ is strictly decreasing at q^*, then an insurer can deviate from q^* by lowering the price and earn strictly positive profits, contradicting the fact that q^* is

the change in ρ is orthogonal to the insurer's discount rate, the risk-adjusted actuarially fair price q^{AF} is unaffected by a change in ρ. Therefore, the AS-adjusted markup decreases with an increase in the two-period bond yield spread.

An alternative way to see why q^{AF} is unaffected by changes in ρ is to interpret q^{AF} as the competitive equilibrium price when the bond market is unconstrained. When the bond market is unconstrained, insurers can freely adjust their bond demand to ρ.[37] As a result, insurers are able to perfectly hedge the interest rate risk without a strictly positive net worth. More precisely, we can follow the argument above and derive the optimal long-term and short-term bond demands at $t = 0$ in an unconstrained bond market setting as

$$l_2 = \frac{\rho}{R_l}\left[\frac{1-\mathbb{E}\left(\frac{1}{R_2}\right)}{1-\rho\mathbb{E}\left(\frac{1}{R_2}\right)}\right]\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q)\,g(\alpha)\,d\alpha,$$

$$b_1 = \frac{1}{R_1}\left[\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha(1+\alpha)\,a(\alpha,q)\,g(\alpha)\,d\alpha - \left[\frac{1-\mathbb{E}\left(\frac{1}{R_2}\right)}{1-\rho\mathbb{E}\left(\frac{1}{R_2}\right)}\right]\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q)\,g(\alpha)\,d\alpha\right].$$

As a result, the aggregate bond demand b_1+l_2 at $t = 0$ is $\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha,q)\,g(\alpha)\,d\alpha$. Hence, the competitive equilibrium price in an unconstrained bond market q^{AF} is invariant to changes in ρ. This is also true in a setting with monopolistic competition. Therefore, the equilibrium price is invariant to variations in the two-period bond yield spread when the bond market is unconstrained, regardless of market power.

H Additional details about variable construction

This appendix contains details about the regulatory interest rate, mortality table, and our measure of interest rate swap duration.

a Bertrand equilibrium.

[37] This is also why we focused our analysis in this section on the case where the bond market is constrained—$\zeta < 1$—with small changes in ρ.

H.1 Regulatory interest rate to discount annuity liabilities

Prior to 2018, state insurance regulation required that insurers calculate their annuity reserves—i.e., their insurance liabilities—using a single reference interest rate calculated as "the average over a period of twelve (12) months, ending on June 30 of the calendar year of issue or year of purchase, of the monthly average of the composite yield on seasoned corporate bonds, as published by Moody's Investors Service, Inc."[38] The Moody's composite yield on seasoned corporate bonds is a weighted average yield on all investment grade corporate bonds rated between Baa and Aaa with maturity of at least 20 years.

From 2018, state insurance regulators adopted a new methodology to calculate the single reference interest rate used in regulatory reserve regulations. With the new methodology, the reference interest rate is the sum of a weighted average U.S. Treasury yield plus a credit spread and an expected default cost. The spread over the reference Treasury rate is calculated by the NAIC using the public bond portion of an average U.S. life insurer's asset portfolio. The new reference interest rate varies by type of annuity contract guarantee period and is reset once a quarter (for retail annuity contracts). For example, the reference rate for a Single Premium Immediate Annuity issued on March 2, 2018 without a guarantee period to a 68 year-old was 3.25 percent, which is about 75 basis points (0.75 percentage point) higher than the reference Treasury rate used in the reference rate calculation.[39] By comparison, Moody's seasoned Aaa and Baa corporate bond yields on the same day are 3.9 and 4.58 percent, respectively.

H.2 Mortality assumption

The SOA mortality tables are available at https://mort.soa.org/. There are two important differences between the "basic" and the "loaded" annuitant mortality tables. First, the loaded table adds a flat 10 percent loading on the estimated

[38] https://www.naic.org/store/free/MDL-820.pdf

[39] For more details, see https://www.soa.org/globalassets/assets/library/newsletters/financial-reporter/2018/june/fr-2018-iss113-hance-gordon-conrad.pdf.

survival probabilities, which requires insurers to hold more reserve per dollar of annuity sold. Second, statutory regulation did not require insurers to apply the SOA generational mortality improvement factor to the static loaded mortality table for their reserve calculations prior to 2015 when the 2012 Individual Annuitant Mortality Table was adopted in most states. As a consequence, regulatory reserves prior to the adoption of the generational table in 2015-2016 became less conservative over time, as the population mortality naturally improved. This phenomenon led the NAIC to update the loaded table in 2000 to essentially "reset" the loading factor. For all our calculations prior to the adoption of the 2012 SOA generational table, we follow industry practice and apply the SOA generational factor to adjust the mortality estimate from the static basic table to the year of observation.

Roughly half of the states required insurers to use the 2012 Individual Annuitant Mortality Table in 2015 and the other half from 2016. We carefully parse each state insurance regulator's website to identify the year at which a new mortality table is implemented for the purpose of regulatory reserve calculations based on the NAIC standard valuation model law 820-1.

H.3 Measuring duration added by interest rate swaps

We proxy for the duration of each individual swap contract by assuming that the duration of the hypothetical zero coupon fixed rate bond is 0.75× the residual maturity of the contract and that the interest rate reset on the floating leg of the swap occurs every 3 months. The factor 0.75 is a commonly used rule of thumb when the actual swap curve is unavailable. Although it is quite crude, this assumption is reasonable to study the variation in average swap duration across insurers in our setting. Assuming that the interest rate reset on the floating leg of the swap occurs every 3 months is consistent with the widely used 3-month LIBOR benchmark among life insurers. It follows that that the duration of a fixed-for-float swap is given by $\text{Swap duration}_{it}^{\text{Receive Fixed}} = 0.75 \times \text{Contract residual maturity} - 1/4 \times 1/2$. Similarly, we calculate the swap duration of individual float-for-fixed swap contracts as $\text{Swap duration}_{it}^{\text{Receive Float}} =$

$-0.75 \times$ Contract residual maturity $+ 1/4 \times 1/2$.

We then multiply each individual swap contract duration by its respective notional amount and divide this number by the duration of a reference 10-year fixed-for-float swap contract, which is calculated as $0.75 \times 10 - 1/4 \times 1/2$. Taking the average over each individual life insurer's swap portfolio in each year yields how much the insurer buys of the reference 10-year fixed-for-float swap. Finally, we divide by the insurers' total general account assets to obtain the amount of duration added by the swaps as a fraction of the insurer asset portfolio. This ratio is a measure of life insurers' interest rate risk management. A value of zero indicates that the insurer is not adding positive or negative duration to its portfolio using swaps.