What's Wrong with Annuity Markets?*

Stéphane Verani† Pei Cheng Yu‡

First version: August 2020; this version: May 2022

Abstract

We show that the supply of U.S. life annuities is constrained by interest rate risk. We identify this effect using annuity prices offered by life insurers from 1989 to 2019 and exogenous variations in contract-level regulatory capital requirements. The cost of interest rate risk management—conditional on the effect of adverse selection—accounts for about half of 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: D82; E44; G22

KEYWORDS: life annuities; corporate bond market; retirement; interest rate risk; adverse selection

*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, Anastasia Kartasheva, Ben Knox, Narayana Kocherlakota, Olivia Mitchell, Andrew Melnyk, Borghan Narajabad, Ali Ozdagli, Radek Paluszynski, David Rahman, Coco Ramirez, Jason Seligman, Adam Solomon, as well as seminar and conference participants at Academia Sinica, Atlanta Fed, CREST-École Polytechnique, National Taiwan University, National Tsing Hua University, University of Michigan, University of Sydney, University of St. Gallen, the National Tax Association meetings and the European Winter Finance Conference 2022. 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 retirees wishing to shed the risk of outliving their financial wealth (Yaari 1965, Mitchell, Poterba, Warshawsky & Brown 1999, Davidoff, Brown & Diamond 2005). New retirees purchasing a life annuity transfer their idiosyncratic longevity risk to a life insurer by surrendering some of their wealth in exchange for a stream of payments while they are alive. 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 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.

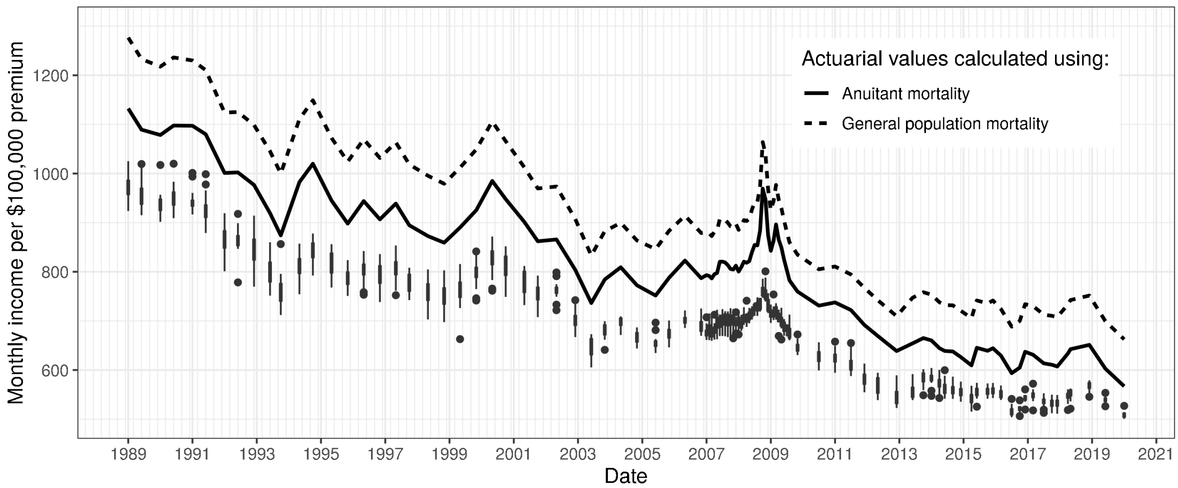


Figure 1: Actual and actuarial monthly payment for a nominal $100,000 SPIA offered to a 65-year-old male.

We develop an algorithm for life annuity valuation that decomposes the contribution of demand- and supply-side frictions in life annuity markups observed from 1989 to 2019. The vertical box plots on Figure 1 represent 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). This is the price that all individuals at retirement age in the US face when choosing how to structure their retirement income, regardless of whether they accumulated wealth through bank deposits, mutual funds, employer-sponsored defined contribution plans, deferred fixed annuities, or variable annuities.[1] The dashed and solid lines are the monthly payments implied by the actuarial value of the annuity contract calculated using the general population mortality and annuitant mortality, respectively. 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 for a $100,000 SPIA offered to a 65-year-old male is substantial and ranges from 10 to 16 percent over the 1989-2019 period.[2]

We focus on a specific supply-side friction, namely interest rate risk, and 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.[3] 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 findings are important for two reasons. The first reason is that adverse selection alone cannot account for high annuity markups. This is a robust finding in the adverse selection literature—e.g., Brown (2001), Finkelstein & Poterba (2004)—and is confirmed

[1]See Section 1.1 for more details.

[2]Our framework for life annuity valuation is similar to the framework used by Koijen & Yogo (2015) and Poterba & Solomon (2021). The main difference is the choice of discount rate to value the annuity payment stream. Koijen & Yogo (2015) assume life insurers' discount their annuity liabilities at the same rate as the US Treasury, whereas Poterba & Solomon (2021) consider money worth calculations from the perspective of prospective annuity shoppers. We differ from these studies by valuing new annuity cash flows from the perspective of the *owner* of a limited liability life insurer contributing capital to support the issuance of *illiquid* fixed rate liabilities. In practice, this means that our AS-adjusted markup is close to the baseline estimate of Poterba & Solomon (2021) and higher than Koijen & Yogo (2015). We discuss this important issue in detail in Section 3.1.

[3]Cutler (1996) makes a similar point focusing on long-term care insurance without identifying the source of aggregate risk.

by Figure 1. This finding led past researchers to conclude that the AS-adjusted markup reflects mostly "administrative costs"—a catch-all term broadly defined as marketing costs, corporate overhead, income taxes, contingency reserves, and profits (Mitchell et al. 1999). We show that the cost of managing interest rate risk accounts for about half of the AS-adjusted markup. This distinction is important because there is a feedback between interest rate risk and adverse selection, which is absent with administrative costs.

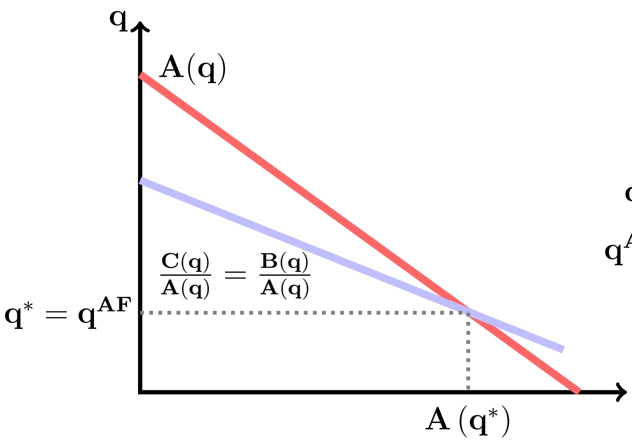
The second reason our results are important is that they bear implications for retirement systems reforms, especially those reforms relying on private markets in providing retirement income. Most studies do not find that publicly provided annuities, such as Social Security, lead to significant welfare gains—e.g., Hong & Ríos-Rull (2007), Hosseini (2015). In these models, social insurance tends to crowd out private insurance markets, because policymakers do not have an advantage over life insurers in assessing individuals' longevity risk. However, this policy conclusion is largely the outcome of assuming that life insurers operate in frictionless financial markets. Under this assumption, life insurers costlessly hedge interest rate risk. Contrary to the premise in these studies, we show that the supply of private life annuity is constrained by interest rate risk. Therefore, reforms that address this type of supply-side inefficiency may be welfare enhancing.

To illustrate the effects of interest rate risk management on annuity markups, we provide a model of annuity pricing in a market with adverse selection and interest rate risk. 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. The assumption that corporate debt maturity is constrained to be short builds on a large finance literature studying this phenomenon—e.g., Bolton & Scharfstein (1990, 1996), Hart & Moore (1994, 1998), Huang et al. (2019).

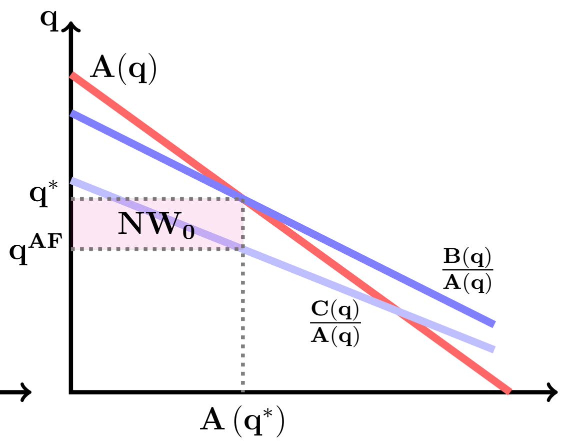
When long-term corporate bonds are relatively scarce, the duration of the bonds life insurers use to fund their annuities (the asset side of the insurers' balance sheet) is less than the duration of their annuity liabilities (Domanski, Shin & Sushko 2017). This creates a negative duration gap, which 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 and could lead to insolvency. The negative duration gap limits insurers' ability to manage their interest rate risk and leads them to hold net worth as precautionary saving to avoid insolvency, which is costly. If insurers are competitive—even imperfectly so—the cost of managing interest rate risk is passed on to annuitants and leads to higher annuity markups.

Moreover, in contrast to a standard adverse selection model with administrative costs, the demand- and supply-side frictions in our model are interdependent. On the one hand, the cost of interest rate risk exacerbates the adverse selection problem by raising annuity prices. On the other hand, more intense adverse selection affects the interest rate risk, as it changes the duration of annuity liabilities.

Figure 2 summarizes our main theoretical results. The vertical axis indexes the annuity price q. The horizontal axis indexes the quantity of annuities. 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 downward-sloping demand curve $A(q)$ and the insurers' average bond demand curve $B(q)/A(q)$, which is the amount of bonds insurers demand per annuity sold.



(a) Unconstrained bond market



(b) Constrained bond market

Figure 2: Equilibrium annuity price

When the bond market is unconstrained (Figure 2a), life insurers invest their annuity sales revenue in an optimal bond portfolio that perfectly hedges the interest rate risk. As a result, the insurers' average cost curve $C(q)/A(q)$, which is the cost per annuity sold and is downward sloping due to adverse selection, is equal to the average bond demand curve $B(q)/A(q)$. 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 by holding some net worth ($NW_0 > 0$, depicted by the shaded area) to cushion the adverse effects of future interest rate shocks and prevent insolvency. In this case, the average bond demand curve $B(q)/A(q)$ is above the average cost curve $C(q)/A(q)$ because insurers charge a positive AS-adjusted markup, i.e., $q^* - q^{AF} > 0$, to finance the net worth.

We identify the effect of interest rate risk management on life annuity markups using life annuity prices offered by U.S. life insurers from 1989 to 2019. Identifying this effect is difficult because demand- and supply-side frictions jointly contribute to relatively high annuity markups. Moreover, life insurers are obviously exposed to interest risk in several ways: through different product offerings, their risk management objective, and their regulatory environment. As shown in Figure 2, identifying the IRM channel requires finding exogenous variations that shift an insurer's average cost curve $C(q)/A(q)$ and average bond demand curve $B(q)/A(q)$. In other words, identification requires shocks to the spread between what a life insurer can earn by investing new annuity money and what it credits as interest on new annuity contracts, which is known in the industry as the net investment spread.

We overcome this identification challenge by studying the pricing of life annuities on the margin—as real-life insurer do—and by exploiting the interaction of shocks to contract-level regulatory capital requirements and shocks to corporate bond market conditions. The contract-level regulatory capital requirements identify the general effect of financial frictions, while the corporate bond shocks identify the type of financial friction. The interaction of these shocks allows us to study how an insurer sets the AS-adjusted markups of different annuity contracts in response to contract-level reserve requirement shocks in favorable and unfavorable bond markets. Specifically, the contract-level reserve requirement shocks shift the average cost curve $C(q)/A(q)$ and the average bond demand curve $B(q)/A(q)$ of *different annuity contracts* offered by the *same insurer* in a similar fashion under both the null hypothesis of costless interests rate risk management and the alternative hypothesis. Meanwhile, the corporate bond market shocks affect the cost of financing net worth under the alternative only, which shifts the average bond demand curve $B(q)/A(q)$. This identification strategy is robust to different assumptions about the annuity market structure—i.e., regardless of insurers' market power.

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 annuity contract-level regulatory reserve requirement. However, this effect is substantially smaller when the insurers' expected return from investing new annuity money exogenously increases. This difference-in-differences 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 cost of risk management could account for almost all of the AS-adjusted markup after factoring in the operating expenses reported by the industry.

Our model derives a life insurer's optimal interest rate hedging strategy conditional on knowing the distribution of interest rate shocks. In reality, life insurers may have different priors about the distribution of interest rate shocks, leading to different ex-ante interest rate hedging positions and ex-post residual exposure to interest rate risk. We exploit this heterogeneity to study the reaction of different insurers in response to a common interest rate shock. Our second set of empirical tests relies on a different set of identifying assumptions, which helps confirm our proposed economic mechanism. First, we focus on the cross-sectional variation of insurers' interest rate hedging programs. We construct a novel data set on the universe of position-level interest rate swap derivatives data between the end of 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—e.g., Einav & Finkelstein (2011)—and the finance literature on risk management of financial institutions—e.g., Froot & Stein (1998). We propose a novel theory of life annuity pricing based on optimal risk management and adverse selection. Importantly, the life annuities we study in this paper allow wealth decumulation during retirement. They should not be confused with *deferred annuities*, which include variable annuities and are tax-deferred savings vehicles that individuals can use to accumulate wealth before retirement—see for e.g., Ellul et al. (2021) and Koijen & Yogo (2022). We show that the feedback between interest rate risk and adverse selection means that it is not possible to study the former without taking into account the latter.

Second, we contribute to the growing literature identifying the effect of supply-side frictions in insurance markets. Previous research has identified the general effects of financial constraints on insurer product design and capital structure decisions—e.g., Koijen & Yogo (2015), Ge (2022), Knox & Sørensen (2020). We differ from these papers by identifying the specific financial friction—interest rate risk—that affects the supply of longevity insurance. Moreover, insurers in our model price their life annuities by optimally choosing a capital structure and an asset portfolio that is consistent with interest rate risk hedging and adverse selection. This is in contrast to the recent studies following the seminal work of Koijen & Yogo (2015) that abstract from the endogenous net worth financing due to interest rate risk and its interaction with adverse selection. Other, more distant studies focus on detecting life insurers' *residual* exposure to interest rate risk—e.g., Hartley et al. (2016), Sen (2021), Ozdagli & Wang (2019), Huber (2022). Our focus is different, as we study the effect of life insurers' interest rate hedging strategy on life annuity prices.

Third, our paper contributes to the literature studying pension reforms. The life annuities we study in this paper are the real-world counterpart to the unique financial contracts modelled in a large class of life-cycle models following the tradition of Yaari (1965) and Blanchard (1985). As mentioned above, the interest rate risk channel is absent from these studies—e.g., Hong & Ríos-Rull (2007) and Hosseini (2015). In contrast, we identify the cost of hedging interest rate risk as a key friction shaping the private supply of longevity insurance.

1 Selling and managing life annuities

In this section, we provide some background about the U.S. life 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. Individuals purchasing a life annuity contract transfer their idiosyncratic longevity risk to the life insurer by surrendering some of their wealth in exchange for a stream of payments while they are alive.

1.1 The U.S. life annuity market

Our paper focuses on the pricing of life annuities. Life annuities allow retirees to decumulate wealth during retirement and are a type of immediate annuity. They should not be confused with deferred annuities, which include variable annuities (VAs) and are tax-deferred savings vehicles that individuals can use to accumulate wealth before retirement. That said, life annuities and deferred annuities are sometime discussed together in the context of the different phases of an annuity contract in advertising materials, which can be confusing.[4] In this subsection, we explain this important distinction, the relationship between life annuities and VAs, and provide two novel estimates of the size of the U.S. life annuity market.

Individuals in the U.S. can save for retirement with deferred annuities, which are, once again, different from the life annuities we focus on in this paper. There are two broad types of deferred annuities: deferred fixed annuities and (deferred) VAs. Deferred fixed annuities offer a guaranteed rate of return over a set time period with tax-deferrals, whereas VAs, as their name suggests, have a rate of return that varies with the return on the stock, bond and money market funds underlying the VA contracts. Although VAs do not offer a guaranteed return, pre-GFC life insurers offered different types of guaranteed minimum benefits in an effort to compete and differentiate their VAs.[5] Therefore, VAs are essentially a mutual fund with an insurance wrapper. The VA assets are segregated from the insurers balance sheet and remain the exclusive property of the VA contract holders. The only connection between VA contracts and the balance sheet of life insurers is through the value of the insurance riders offered with the VAs. These complex guaranteed minimum benefits exposed life insurers to significant equity market risk and caused enormous stress to their balance sheet when the stock market crashed in 2008 (Ellul et al. 2021, Koijen & Yogo 2022), so aggressive market-based minimum guarantees

[4]See Black Jr., Skipper & Black III (2015) for a detailed discussion of the various types of annuities.

[5]For example, some of the more aggressive insurers offered VA policies with both Guaranteed Minimum Death Benefits and Guaranteed Minimum Income Benefits riders that protect policyholders against equity market downturn in case of death or annuitization.

are no longer offered.

At the end of the deferred fixed annuity or VA contract period and after reaching 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. This is sometime refereed to as an annuity "payout phase." Section 1035 of the U.S. tax code allows individuals to exchange an existing variable annuity contract for a new annuity contract without paying any tax on the income and investment gains in their current variable annuity account. Therefore, the price of life annuities that we study in this paper is the price that all individuals at retirement age in the U.S. face when choosing how to structure their retirement income, regardless of whether they accumulated wealth through bank deposits, mutual funds, employer-sponsored defined contribution plans, deferred fixed annuities, or VAs.

Roughly half of the U.S. life insurance industry's $600 billion aggregate income in 2018 came from annuity considerations. The other half is roughly split between life and health insurance premiums.[6] Total annuity income includes considerations related to both deferred annuities and life annuities. Estimating the size of the U.S. life annuity market is difficult, as it is not possible to precisely separate life annuity sales from deferred fixed annuity considerations in regulatory filings. Nevertheless, we provide two novel estimate of the size of the U.S. life annuity markets using individual regulatory filings of insurers—see Appendix A for more details. First, using insurer-level data on the number of general account annuity contracts and account balances reported in the 2018 NAIC Statutory Filings 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. 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. 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).

1.2 The life annuity business model

Life insurers' overall business model consists of earning a spread between the yield they owe on their insurance liabilities and the yield they earn on the assets backing

[6] See the ACLI's 2018 Life Insurers Fact Book https://www.acli.com/posting/rp18-007.

these liabilities. Life annuities 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 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 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, which is 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]

1.3 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 annuity liabilities. For example, corporate bonds have a median initial maturity of about 5 years. Over 90 percent have an initial maturity that is 10 years or less and, among those bonds, only a minority pays fixed rate and is non-callable.[9] This maturity structure contrasts with the duration of a life annuity offered to a 65-year-old individual, for example, which is approximately 10 years—note that the duration of a fixed-income instrument is less than or equal to its maturity. Moreover, long-duration U.S. government securities are unattractive to life insurers because they carry a substantial liquidity

[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]There is a large literature attempting to explain this phenomenon, for example, through the lens of contracting frictions—e.g, Bolton & Scharfstein (1990, 1996), Hart & Moore (1994, 1998), Barclay & Smith Jr (1995), Huang et al. (2019).

premium—see for e.g., Krishnamurthy & Vissing-Jorgensen (2012), van Binsbergen et al. (2021). This liquidity premium means that it is not profitable for private life insurers to fund long-term illiquid liabilities, such as fixed annuities, with highly liquid long-term bonds, such as U.S. Treasuries.

To put these numbers in perspective, assume there are 3.5 million new 65-year old individuals in the U.S. in a given year—this is roughly the average between 2000 and 2019. For simplicity, assume further that wealth is only annuitized by new 65-year-old individuals. Our previous calculation suggests that a cohort of 65-year-old individuals annuitizes about $45 billion in wealth with life insurers in a given year—i.e., $12,700 dollar per 3.5 million individuals. Using data on the universe of corporate bond issuance, Appendix A shows that this amount is about 15 percent larger than the average amount of fixed-rate, non-callable corporate bond with maturity over 10 years issued by U.S. firms over the same period. This implies that, although the life annuity market is small, it is larger than the total supply of fixed-rate long term corporate bonds. Moreover, the U.S. has the largest corporate bond market in the world. And, of course, U.S. life insurers compete for these bonds with other long-term investors, such as pension funds and sovereign wealth funds in and out of the U.S.

The negative duration gap between insurers' assets and their insurance liabilities 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, IRM) is at the heart of the modern insurers' 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 acts as precautionary savings and helps cushion the effect of interest rate changes that disproportionately affect the value of the life insurers' insurance liabilities.[11] This, however, is costly because the primary source of life insurer capital is accumulated retained earnings. Therefore, the cost of building and 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

[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.

[11]Net worth is not to be confused with what the industry calls reserves, which is the value of insurance liabilities.

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 the bond market is constrained, such that corporate debt maturity is relatively short. This friction implies that life insurers are exposed to interest rate risk because the duration of their insurance liabilities is longer than the duration of their assets, which could lead to insolvency. The second 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 in excess of the value of the insurer's assets. This friction implies that insurers may fail to honor their annuity payments if they become insolvent. The rest of this section shows how life insurers manage interest rate risk by financing net worth with the annuity markup.

2.1 Economic environment

The economy is populated by a continuum of new retirees with homogeneous wealth 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 $t = 0$ from c.d.f $G(\alpha)$ with support $[\underline{\alpha}, \overline{\alpha}] \subset [0, 1]$ and p.d.f $g(\alpha)$.[12] The survival probability α is the individuals' private information. Every individual is deceased at the end of $t = 2$.

There are two types of 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 $t = 0$.[13] The annuity market is competitive

[12] Our theoretical results do not change substantively if we assumed that the survival probabilities decrease with age.

[13] 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.

and life insurers compete over annuity prices. In Appendix E.2, we consider an extension of our benchmark environment with monopolistic competition, and in Section 3, we argue that assumptions about the market structure are not critical for identification.[14] 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 an aggregate shock 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 its return 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 $-\infty < \frac{\partial a}{\partial q} < 0$; (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 annuities and that the demand for annuities 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 the demand for annuities from insurers with a strictly positive probability of becoming insolvent to be zero. This simplification allows us to more clearly analyze the effect of interest rate risk on annuity markups. 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. This would be the case if individuals are unwilling to swap their longevity risk with insurer default risk,

[14] Unlike variable annuities for which life insurers compete over prices and product characteristics (Koijen & Yogo 2022), fixed annuities are standardized products and insurers compete over prices. In Appendix F, we provide evidence of competitive fixed annuity markets by calculating a Herfindahl-Hirschman Index for the industry.

for example. A more nuanced interpretation of this assumption that is closer to actual practices in the life insurance industry is that an un-modelled insurance regulator or a credit rating agency requires insurers to hold minimum capital to prevent insolvency. In the U.S., insurers are required to have a certain amount of capital and surplus to establish and continue operations. When an insurer's capital and surplus falls below the minimum standard, it is considered to be legally impaired, which triggers regulatory interventions.

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.

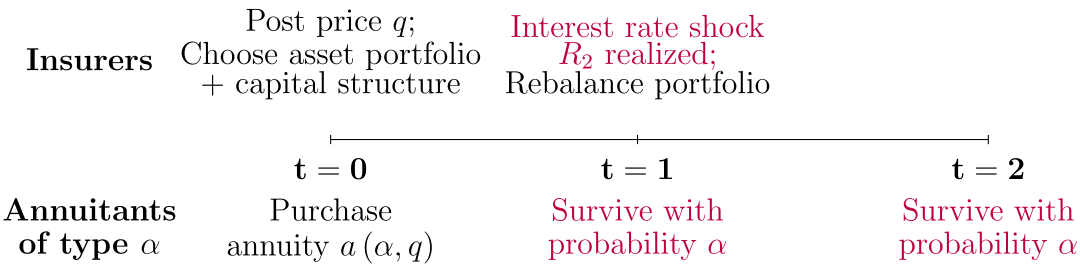


Figure 3: Model timeline

2.2 Interest rate risk management

This section formulates the optimal IRM strategy of the insurer. We begin by examining the dynamics of a life insurer's balance sheet, which 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 is the insurer's net worth in t.

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, the competitive pressure gives life insurers an incentive to minimize their annuity price while maintaining the minimum level of net worth necessary to hedge the interest rate risk. In fact, when the bond market is unconstrained, we can show that if

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

then by equations (2), (3), and (4), the optimal capital structure is such that $NW_0 = NW_1(R_2) = 0$ for any realization of R_2. Therefore, when the bond market is unconstrained, insurers can perfectly hedge their interest rate risk by investing in a suitable portfolio of one- and two-period bonds without maintaining a strictly positive level of 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 $l_2 = \zeta \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q)\, g(\alpha)\, d\alpha$ 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. The proof of all theoretical results in Section 2 can be found in Appendix B.

Theorem 1 *Under the unique optimal IRM strategy, insurers require a higher level of net worth when the bond market is constrained ($\zeta < 1$). 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:

$$\begin{aligned} l_2 &= \frac{\zeta}{R_l} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha, \\ b_1 &= \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left[1 + (1 - \zeta)\alpha\right] a(\alpha, q)\, g(\alpha)\, d\alpha, \\ b_2(R_2) &= \left(1 - \zeta + \frac{\zeta}{R_2}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha. \end{aligned}$$

ii. Capital structure:

$$\begin{aligned} NW_0 &= \frac{1 - \zeta}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha, \\ NW_1(R_2) &= (1 - \zeta)\left(1 - \frac{1}{R_2}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha. \end{aligned}$$

To summarize, 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, which exposes the insurer to interest rate risk.[15] Specifically, an insurer becomes 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 they might become insolvent at $t = 1$. Using the IRM strategy shown in Theorem 1, an insurer maintains a strictly positive level of net worth to decrease the duration of its total liabilities (because net worth has a lower duration than the insurance liabilities) and eliminate its duration gap.[16] Furthermore, the IRM strategy presented in Theorem 1 is unique when the market for annuity is perfectly competitive.[17]

Finally, note that in contrast to the law of demand, the insurers' bond demand is inversely related to bond returns. This is because for a given annuity price, when bond returns are low, life insurers require more bonds to hedge the interest rate risk.

2.3 Life 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

[15]In Appendix E.1, 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 liquidity premium or convenience yield in a reduced form. We show that the lower the yield on the two-period government bond relative to the two-period corporate bond, the larger the level of 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. The intuition for this results is that when there is a liquidity premium on government bonds, insurers backing their annuity liabilities with government bonds pay for the liquidity premium on the asset side. At the same time, they pay an illiquidity premium to the annuitant for holding the illiquid annuity contract. Therefore, the hedging value of long-term government bonds is a decreasing function of their liquidity premium. This result is consistent with the relatively low share of government securities in life insurers' asset portfolio noted in Section 1.2.

[16]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.

[17]In Appendix E.2, we show that insurers with market power also engage in IRM and build an asset portfolio and capital structure that satisfies Theorem 1 at the minimum. This is because insurers with market power have a bond demand of at least b_1 and l_2 in $t = 0$ as specified in Theorem 1 to manage their interest rate risk.

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:

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

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 $q^{CI}(\alpha)$ denote the equilibrium price in an economy where insurers can observe individual survival types α. The full information actuarially fair price is given by $q^{CI}(\alpha) = \frac{\alpha}{R_1}\left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right]$. Proposition 1 establishes the classic adverse selection result that the *risk-adjusted actuarially fair price* is higher than the average full-information actuarially fair price:

Proposition 1 $q^{AF} > \int_{\underline{\alpha}}^{\overline{\alpha}} q^{CI}(\alpha) g(\alpha) d\alpha.$

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:

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

where total bond demand is given by Theorem 1. Let q^* be the equilibrium annuity price—the lowest positive annuity price such that $\Pi(q^*, \zeta) = 0$. Notice that given the equilibrium annuity price q^*, the risk-adjusted actuarially fair price is defined by the average annuity liability at q^* :

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

This means that the risk-adjusted actuarially fair price is evaluated using the pool of annuitants purchasing annuity at the equilibrium price q^*. To streamline notation, we write the risk-adjusted actuarially fair price as q^{AF} with the implicit understanding that it depends on the equilibrium price q^*. Next, we make an additional assumption about the annuity demand curve to rule out pathological cases.

Assumption 2 *For any annuity price q, individuals with higher longevity risk are less responsive to annuity price changes:* $\frac{\partial^2 a(\alpha,q)}{\partial q \partial \alpha} \geq 0$.

Assumption 2 places a restriction on how the price elasticity of demand varies with the survival probability α. In addition to requiring that individuals with high α—i.e., high longevity risk—buy more annuities (Assumption 1), Assumption 2 also requires that these individuals cannot be too sensitive to price changes. Theorem 2 shows that—when Assumption 2 holds—the AS-adjusted markup, defined as $q^* - q^{AF}$, increases when the bond market becomes more constrained, as indexed by a lower level of $\zeta < 1$.

Theorem 2 *The AS-adjusted markup $q^* - q^{AF}$ is higher when the bond market is more constrained (ζ is lower):* $\frac{\partial q^*}{\partial \zeta} - \frac{\partial q^{AF}}{\partial \zeta} < 0$. *Furthermore, when the bond market is unconstrained ($\zeta = 1$), the AS-adjusted markup is zero:* $q^* = q^{AF}$.

Crucially, Assumption 2 implies that the average optimal net worth $NW_0(q)/A(q)$ given in Theorem 1 is increasing in q. Therefore, if insurers require a higher average net worth for IRM, they finance it by raising the annuity price.

2.4 Discussion of the model's properties

Figure 2 presented in the introduction depicts how insurers finance their net worth by charging a markup over the annuity contract's actuarial value in a constrained bond market. This figure highlights the unique features of our annuity pricing model relative to the textbook model of adverse selection in Einav & Finkelstein (2011).

The textbook model abstracts from supply-side frictions by implicitly assuming that financial markets are efficient. The competitive annuity price in the textbook model is determined by the intersection of the aggregate demand curve $A(q)$ and the insurers' average cost curve $C(q)/A(q)$. 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 ($\zeta = 1$).[18] Equation (1) shows how our competitive equilibrium price q^* is determined by the intersection of the aggregate demand curve given by

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

and the insurers' average bond demand curve $B(q)/A(q)$, where the total bond demand $(b_1 + l_2)$ under the unique optimal IRM strategy is

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

so that $q^* = B(q^*)/A(q^*)$. The average cost curve $C(q)/A(q)$ at the equilibrium annuity demand determines the *risk-adjusted actuarially fair price* q^{AF}, with total cost

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

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 $C(q)/A(q)$ is equal to the bond demand curve $B(q)/A(q)$. (Note that the curves in Figure 2 are drawn as functions of quantity, while the objects in the model are functions of annuity price.) 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 NW_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 $C(q)/A(q)$ and average bond demand curve $B(q)/A(q)$ are no longer equal when the long-term bond supply is constrained.

[18] The insurers' capital structure is irrelevant in the textbook model because financial markets are efficient. Due to limited liability, the Modigliani-Miller theorem does not hold in our environment (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.

Specifically, the total bond demand is equal to the sum of total annuity liabilities and the positive level of net worth. Theorem 2 shows how insurers finance the net worth by charging a higher AS-adjusted markup—the equilibrium price q^* is higher than q^{AF}.

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:

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

This equilibrium relationship shows how total revenue generated by the AS-adjusted markup finances the optimal net worth, which is depicted by the shaded area in Figure 2b.

Finally, it should be noted that our model is not equivalent to the textbook model with administrative costs (Einav & Finkelstein 2011). An important difference is that there is no feedback between administrative costs defined earlier and adverse selection, whereas there is a feedback between the cost of IRM and adverse selection. Appendix D extends the present discussion by showing that life insurers' IRM amplifies the effect of adverse selection on annuity markups, while a worsening adverse selection problem would further alter the interest rate risk. 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- 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 prices. 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 $C\left(q\right)/A\left(q\right)$ and average bond demand curve $B\left(q\right)/A\left(q\right)$.

3 Identification

Testing for the interest rate risk channel requires identifying the effect of changes in bond market conditions on the AS-adjusted markups. The main identification challenge is threefold. First, it is not possible to directly measure the duration gap between U.S.

insurers' assets and insurance liabilities.[19] 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 statutory filings. 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 observationally equivalent annuity markups. For example, Appendix E.2 shows that annuity markups could also be the outcome of monopolistic competition. Third, demand- and supply-side frictions are likely to have non-trivial interactions. Appendix D shows theoretically and empirically that adverse selection in annuity markets depends on the severity of frictions in the corporate bond market, as a higher level of net worth exacerbates adverse selection by increasing annuity prices.

We overcome this identification challenge by exploiting corporate bond market shocks that differentially affect the average cost curve $C(q)/A(q)$ and the average bond demand curve $B(q)/A(q)$ of *different annuity contracts* offered by the *same insurer*. The combination of exogenous shocks to the average cost curve $C(q)/A(q)$ and average bond demand curve $B(q)/A(q)$ is necessary to identify the risk management channel in a market with adverse selection. 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 due to an exogenous increase in relative cost when corporate bond market conditions are good to the change when corporate bond market conditions are relatively worse. 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 Life 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.

[19]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) propose an indirect measure of the duration gap based on insurers' stock prices.

Two inputs are needed to price new insurance liabilities. We discuss our choices for these two inputs in detail below, as they have important implications for annuity valuation (Poterba & Solomon 2021). The first input in valuing annuity cash flows is a discount rate. We follow our theory and industry practices 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 backed by illiquid assets.[20] Almost all life insurers offering annuities in the U.S. have a rating around A. Therefore, the discount rate of an average insurer's marginal investor should be close to the duration-matched yield on A-rated illiquid debt securities.

We proxy for the unobserved discount rate of the marginal life insurer investor using the zero-coupon High Quality Market (HQM) yield curve produced by the U.S. Treasury.[21] 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.[22] Consistent with our choice of discount rate, Huber (2022) finds that the insurers' implied average annuity discount rate tracks closely the HQM yield curve on a duration-matched basis.

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.

[20]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.

[21]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.

[22]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

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.[23]

We also use two different versions of the Individual Annuitant Mortality (IAM) table produced by the Society of Actuaries (SOA) in collaboration with the National Association of Insurance Commissioners (NAIC). The first version of the IAM is the "basic" table, which is estimated by the SOA from the actual mortality experience of a large pool of annuitants from multiple insurers over a long period of time. Prior to the 2012 edition, the basic IAM table is static and used in conjunction with a fixed generational improvement factor (Scale G) to adjust for the population's natural mortality improvement. The industry transitioned to a generational (dynamic) mortality table from 2015 and we carefully parse individual state legislation to follow its staggered implementation in 2015-2016. In addition to the basic IAM table, the SOA produces a "loaded" IAM table, which adds a loading factor to the basic IAM table mortality estimate. The loaded table is used by state insurance regulators to set regulatory reserves. See Appendix H for more details about mortality assumptions.

A majority of insurers surveyed by the SOA use the "basic" annuitant mortality table with the Scale G factor to price their annuities.[24] That said, very large insurers with a large annuitant pool may modify these estimates to reflect the mortality experience of their own pool of annuitants. Nevertheless, all insurers must use the loaded IAM to calculate their regulatory annuity reserves and, therefore, their own mortality assumptions cannot deviate too much from the IAM table.[25]

The actuarial value of a life annuity contract with an M-year guarantee term per

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

[24]See the *Report of the Society of Actuaries Mortality Improvement (Annuity) Survey Subcommittee April 2012* available at https://www.soa.org/files/research/exp-study/research-mort-annuity-survey-report.pdf.

[25]Obviously, this unobserved heterogeneity in mortality assumptions could contribute to cross-sectional variation in markups. As will be clear later, this is not a threat to identification in our within-insurer-contract analysis and is orthogonal to our explanatory variables in our cross-sectional analysis.

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 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 ($r = \text{HQM}$), or the regulatory reference rate ($r = \text{NAIC}$), which we will explain below.

Let $P_t(n,S,M)$ be the normalized 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):

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

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 payments implied by the different actuarial values.

3.2 Regulatory reserve requirements

The first source of exogenous variation comes from the effects of changes in corporate bond conditions on the regulatory reserves that insurers are required to set aside for each dollar of annuity they sell. As noted by Koijen & Yogo (2015), exogenous time-series variation in reserve requirements across contract maturity arises because regulatory reserves are calculated using a single regulatory interest rate that resets infrequently. This source of exogenous variation is useful to identify the general effect of financial frictions because it acts as a shifter of the insurers' average cost curve (Koijen & Yogo 2015).

Prior to 2018, state insurance regulation required that insurers calculate their annuity reserves—i.e., their insurance liabilities—using a single reference interest rate defined 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 Moodys Investors Service, Inc."[26] 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 but related methodology, which we discuss in Appendix H for brevity.

By construction, the regulatory reference interest rate is close to the 12-month average of the longer end of the HQM yield curve that we use as a proxy for the insurers' discount rate. When the actual yield curve is upward sloping, the actuarial value of a life annuity calculated using the average of the long end of the yield curve is mechanically smaller than the corresponding actuarial value calculated using the entire yield curve. This difference is greater for life annuities with shorter expected maturity—i.e., sold to older individuals. Moreover, the difference between regulatory reserve and insurer reserves fluctuates exogenously over time across annuity contracts with different maturities. This is because the regulatory interest rate resets infrequently—once a year prior to 2018 and once a quarter from 2018—whereas the yield curve used by insurers to price their insurance liabilities changes daily.

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—in our empirical analysis, we use the full set of price quotes for male and female individuals aged between 50 and 90 years with 5-year intervals. We denote the regulatory interest rate by $r = \text{NAIC}$ and use it in the calculation of $V_t^k(n, S, M, r)$. 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

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

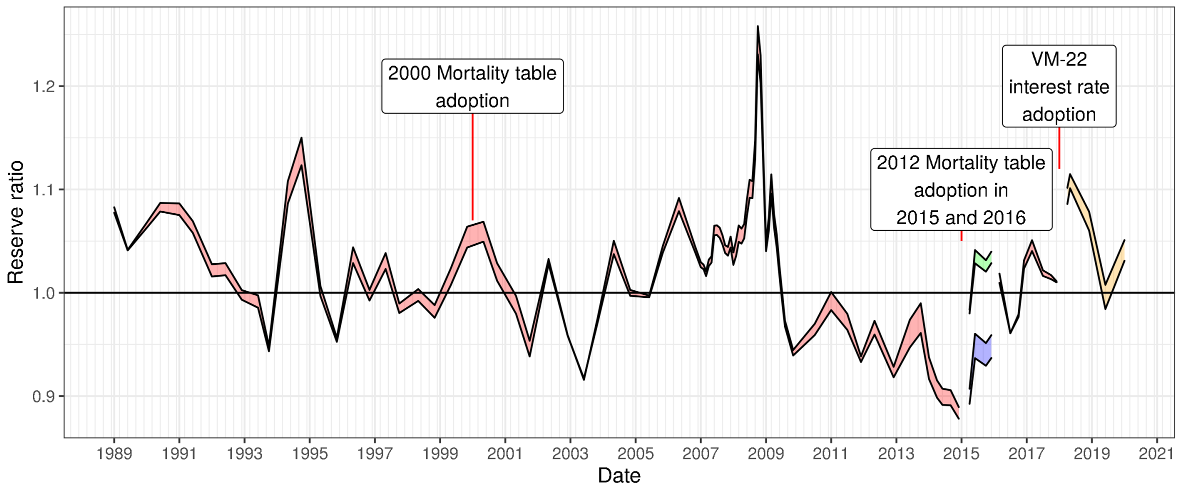


Figure 4: Reserve ratio for SPIA sold to a 65-year-old male (top line) and 70-year-old male (bottom line)

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. The reserve ratio fluctuates around 1, confirming that the regulatory discount rate and the insurers' discount rate are aligned on average. Notice how this distance exogenously fluctuates overtime because the flat regulatory interest rate resets infrequently. Figure 4 also depicts 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.

Mapping this data to Figure 2, an exogenous rise in reserve requirements means that the insurer must create a larger insurance liability for a given pair of annuity price and quantity. A tightening of the reserve requirement creates a wedge between the insurer's own average cost curve and the cost curve implied by the regulatory interest rate and mortality assumption. In essence, when the reserve requirement binds, an insurer's average cost curve $C(q)/A(q)$ is lower than what the regulator thinks it should be. The wedge between the insurers' and regulators' liability valuations is often called *excess reserve* in the industry. The insurer backs the excess reserve by purchasing additional bonds per unit of annuity sold. This means that the insurers' bond demand curve $B(q)/A(q)$ is higher than its average cost curve $C(q)/A(q)$. This effect arises under both the null hypothesis of costless IRM (Figure 2a) and the alternative (Figure 2b). In both cases, the insurer raises the AS-adjusted markup $q^* - q^{AF}$ to finance the excess reserve. The key difference is that, when IRM is costly, the AS-adjusted markup also finances the strictly positive net worth. 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.

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, which is not to be confused with the yield on those bonds. This source of variation allows us to identify the effect of IRM channel among potential alternative explanations.

A widening in Baa-Aaa spread for long-duration corporate bonds *relative* to the insurer's cost of funding corresponds to higher-yielding investment opportunities for new annuity money. As explained in Section 1, the life insurer's business model consists of earning a spread between the yield they receive on the assets purchased with annuity considerations and the yield they owe on their annuity liabilities. Life insurers target bonds rated around A because they match the illiquidity profile of their insurance liabilities. Moreover, under state insurance regulation, corporate bonds rated above Moody's Baa are designated as NAIC 1 and uniformly 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. Therefore, life insurers target the bottom of the investment grade category as it maximizes yield subject to the same risk-based capital charge (Becker & Ivashina 2015).

Figure 5 illustrates this exogenous variation by plotting the Baa-Aaa spread for seasoned corporate bonds and the 10-year HQM yield over 10-year U.S. Treasury spread in percentage points. The 10-year HQM yield spread is our proxy for the insurers' average cost of funding because life annuities offered to a 65-years old have an initial duration of about 10 years and we use the HQM yield curve as the insurers' discount rate. 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. Put differently, a relative widening of the long duration Baa-Aaa spread allows insurers to generate more income by increasing their portfolio returns relative to what they owe on policies.

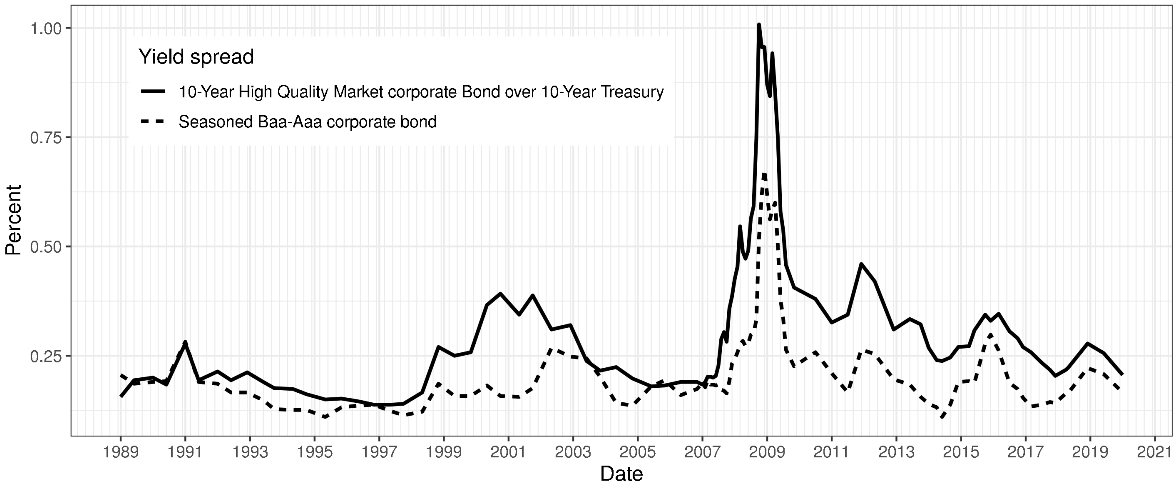


Figure 5: Baa-Aaa spread and insurers' average cost of funding

Mapping this data to Figure 2 once more, an exogenous increase in long-term bond *yield spread* has no impact on annuity markups under the null hypothesis of costless IRM (Figure 2a). This is because insurers can always construct a portfolio of bonds that perfectly hedges the interest rate risk without net worth. Therefore, the average bond demand curve $B(q)/A(q)$ is always equal to the average cost curve $C(q)/A(q)$, which does not move with changes in the long-term bond yield spread. Under the alternative hypothesis (Figure 2b), recall that the insurers' bond demand is inversely related to bond returns. This is because for a given annuity price, when bond returns are low, life insurers require more bonds to hedge the interest rate risk. Therefore, an exogenous increase in long-term bond *yield spread* decreases the average bond demand curve, $B(q)/A(q)$, because the insurer can fund a block of new annuity business with fewer bonds. The decrease in $B(q)/A(q)$ implies that insurers no longer need to hold as much net worth per annuity dollar because the average cost curve is not affected by the shock—i.e., the original level of NW_0 per annuity dollar is higher than the new optimal level. As a result, the AS-adjusted markup $q^* - q^{AF}$ is lower. Appendix G contains a formal treatment of the effect of long-term bond spread shock on the optimal IRM strategy and annuity pricing.

3.4 Testing for the interest risk channel

From the discussion above, the effect of an increase in reserve requirement on markup is offset by an increase in long-term bond yield spread under the hypothesis of costly IRM. The reserve requirement shocks allow us to identify the general effect of financial frictions, while the long-term bond spread shocks allow us to tease out the effect of the costly IRM friction among competing supply-side alternatives.

This effect is unique to our model and does not depend on our assumption about the market structure. For example, Appendix E.2 analyzes the model presented in Section 2 under the assumption of monopolistic competition. Theorem 3 in Appendix E.2 shows that monopolistic competition with an *unconstrained* bond market can generate an AS-adjusted markups that is observationally equivalent to a perfectly competitive annuity market facing a *constrained* bond market. This is intuitive, as insurers either increase their annuity markups to fund net worth or to limit the quantity sold in the market and exercise market power. However, for any given market structure, we show that the effect of an increase in bond spread on the AS-adjusted markup is strictly negative when the bond market is constrained and zero when the bond market is unconstrained.

We build on this theoretical insight to implement an empirical test of the IRM channel. The test consists of estimating the effect of a change in bond market conditions on the AS-adjusted markup using a type of difference-in-differences approach. In our setting, the change in the Baa-Aaa spread is the treatment that differently affects annuity contracts from the same insurer with exogenously varying reserve requirements. The first difference is between annuity contracts j offered by insurer i with relatively high reserve requirements and annuity contracts $-j$ offered by the *same* insurer i with relatively low reserve requirements. The contract-level reserve requirement shocks create a within-insurer random assignment of annuity contract relative cost that varies from one period to the next. The second difference is between periods in which the Baa-Aaa spread is high and periods in which it is low.[27] Under the null hypothesis of costless

[27] Note that the variation in the Baa-Aaa spread is exogenous in the sense that neither the annuity shopper nor the insurer affects it and both are affected by it. We are not assuming that annuity demand is necessarily orthogonal to long-term bond yield spread. Our identifying assumption is that the effect of an (unobserved) change in annuity demand due to a change in bond market conditions on the AS-adjusted markup is proportional across annuity contract types. This is similar to a parallel trend assumption in a more standard difference-in-differences setting. Under this assumption, the difference-in-differences nets out the potential effects of changes in annuity demand associated with changes in corporate bond market conditions.

IRM, the effect of an increase in reserve requirement is not affected by changes in the long-term Baa-Aaa spread. Under the alternative, the increase in markup is lower when the bond spread is higher, which is unique to the IRM channel. As previously explained, we condition 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.

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 date. 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.*[28] 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 markup

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

for product j sold by insurer i at date t, 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

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

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.

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 HQM 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. We relegate the discussion of the $Net\_swap\_duration_{it}$ variable to Section 6.

5 Main empirical analysis and results

We implement our test in a linear regression framework. The unit of observation is an annuity contract j offered by insurer i at date t. The sample of observation extends from 1989 to 2019. The coefficient β_3 on the interaction between $Reserve\_Ratio_{jt}$ and $Baa\text{-}Aaa\_spread_t$ in the following linear model captures the difference-in-differences estimate of the reduction in contract-level reserve requirement on this contract's AS-adjusted markup during times of increasing Baa-Aaa spreads *conditional* on the cost of funding

$10\_HQM\_spread_t$.

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

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: The effect of investment-grade corporate-bond yield spread on life annuity markups The unit of observation is a life insurer-product-quarter. The sample of observation extends from 1989 to 2019. The dependent variable $Annuity\_markup_{ijt}$ is the AS-adjusted markup for life annuity j sold by insurer i at date t. Column 1 reports insurer clustered robust standard errors in parentheses and Columns 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

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 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.[29] 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 $Reserve\_Ratio_{jt}$ 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 corresponds to higher yielding investment opportunities for new annuity money and, therefore, a lower AS-adjusted markup.[30] This result is consistent with the IRM strategy of life insurers in a constrained bond market. 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.[31]

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.[32] Column 3 controls for

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

[30]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.

[31]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.

[32]We investigate the robustness of our inference to different clustering assumptions by calculating

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 with the full sample in Column 2.[33]

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 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—but not necessarily perfectly competitive—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.[34]

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 2008–09 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

block bootstrap standard errors and wild bootstrap standard errors and find no evidence of bias—the results are available on request.

[33]In Appendix I, we show that our baseline results are not driven by variations in the 2007-09 period.

[34]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 10 in Appendix F shows that average competition for the entire fixed annuity market is high.

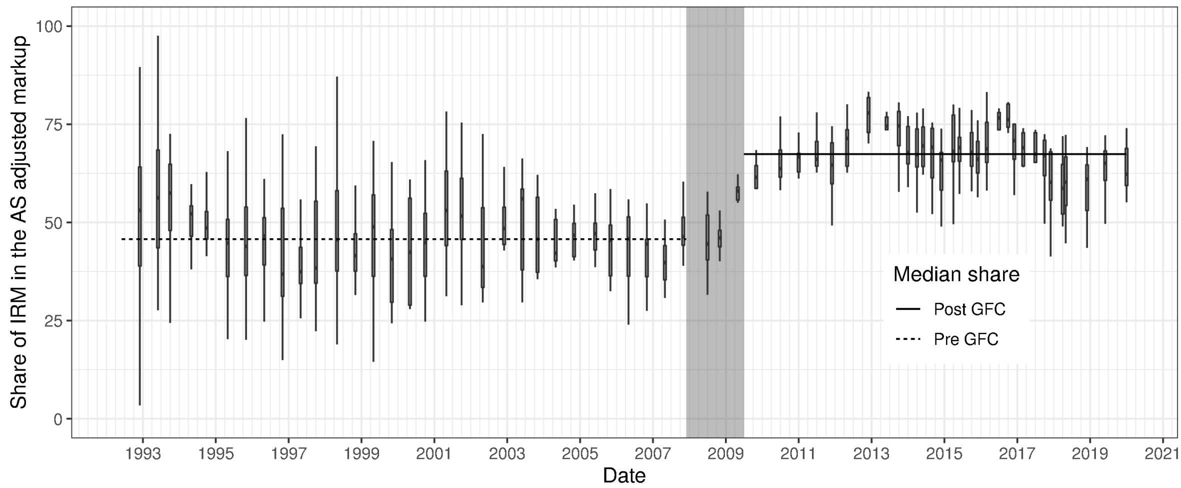


Figure 6: Contribution of IRM cost in the AS-adjusted markup for SPIA offered to a 65-year-old male.

actuarial value. This estimate suggests that if the insurers' business expense are around 3 to 5 percent, IRM could account 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 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

The creation and preservation of net worth are life insurers' primary tools to hedge interest rate risk. Although not widely used by the industry, some large and sophisticated life insurers incorporate interest rate swaps as part of their interest rate hedging program.[35] In this section, we exploit the heterogeneity in ex-ante exposure to interest rate risk that arises from life insurers' interest rate swaps to study the reaction of different insurers in response to a common interest rate shock.

Sophisticated 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—usually a

[35] The cost and complexity of swaps account for the industry's relatively low use. Interest rate swaps are not a panacea for life insurers to manage interest rate risk because they are complicated, require collateral financing, and face margin calls. Moreover, life insurers using swaps need to continuously prove to their state regulators that each swap contract is hedge effective to avoid incurring relatively high capital charges.

commercial bank. 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 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.

We construct a proxy for the aggregate net-duration added by each life insurer's interest rate swaps portfolio using contract-level data. We 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 that is driven by lower term premium and vice versa. Although an insurer's swap position is an ex-ante endogenous variable, variations in the the shape of the yield curve act as exogenous shifters of the swap portfolio value ex-post. We focus on the period of the *zero lower bound* from 2009 to 2015, during which all the variation in the yield curve is driven by movements in the term premium.[36] Therefore, we can compare the AS-adjusted markups of insurers that are favorably affected by a change in the term premium *ex-post* because of their *ex-ante* hedging program relative to those that are adversely affected by the shock.

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 assets.[37] 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 contracts' terms and notional amount. We carefully parsed the text of more than 82,000 individual contract-year observations from 44 U.S. life

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

[37]See Appendix H for details.

insurers from 2009 to 2015 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 multiply this ratio by the original contract's notional amount. This number is the dollar amount of duration contributed by an individual swap contract, which can be positive or negative. 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 this number 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. We denote this variable by $Net\_swap\_duration_{it}$ and report its summary statistics in Table 1. 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 $Net\_swap\_duration_{it}$ 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 financial variables $\mathbf{z}_{it}'$, 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.

Column 1 of Table 3 summarizes our cross-sectional results. The coefficient estimate

Table 3: Cross-sectional evidence of the risk management channel The unit of observation is an insurer-product-quarter. 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 AS-adjusted 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		χ^2_1-test	34.75***

on the interaction term suggests that an insurer with a 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 the insurers' hedging strategies are 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 one third of a percentage point—1.8 percent of the average AS-adjusted markup—in response to a flattening 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 $Q_{Annuity\_markup_{ijt}}(\tau|\mathbf{x}'_{ijt})$ 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 starting 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 value of the Wald 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 ex-ante, those with a better hedge cut their markups 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 markups.

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 annuity 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 studying 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 private contracts can be designed to mitigate this friction (Golosov & Tsyvinski 2007). However, this policy conclusion is largely the outcome of assuming that life insurers operate in frictionless financial markets. Under this assumption, life insurers costlessly hedge interest rate risk. Contrary to the premise in these studies, we show that the supply of private life annuity is constrained by interest rate risk. If private annuity markets are constrained by interest rate risk, the crowding-out effect of social insurance could be overstated and reforms that address this type of supply-side inefficiency could be welfare enhancing.

Lastly, studying life insurers' IRM may also shed light on important puzzles surrounding the shrinking U.S. long-term care insurance market—e.g., Ameriks, Briggs, Caplin, Shapiro & Tonetti (2018), Braun, Kopecky & Koreshkova (2019). Long-term care insurance is another type of long duration insurance product that exposes insurers to interest rate risk in addition to the uncertainty about future healthcare cost. The combination of lower than expected lapse rates, higher than expected morbidity and the low interest rate environment contributed to a rapid deterioration of profitability in this sector. That said, the latter has received less attention in the literature. 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. (2018), 'The Long-Term-Care Insurance Puzzle: Modeling and Measurement', *NBER Working Paper 22726* .

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.

Barclay, M. J. & Smith Jr, C. W. (1995), 'The maturity structure of corporate debt', *The Journal of Finance* **50**(2), 609–631.

Becker, B. & Ivashina, V. (2015), 'Reaching for yield in the bond market', *The Journal of Finance* **70**(5), 1863–1902.

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

Black Jr., K., Skipper, H. D. & Black III, K. (2015), *Life Insurance, 15th Edition*, Lucretian, LLC.

Blanchard, O. J. (1985), 'Debt, deficits, and finite horizons', *Journal of Political Economy* **93**(2), 223–247.

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. (2001), 'Private pensions, mortality risk, and the decision to annuitize', *Journal of Public Economics* **82**, 29–62.

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?', *Working Paper* .

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.

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.

Ellul, A., Jotikasthira, C., Kartasheva, A. V., Lundblad, C. T. & Wagner, W. (2021), 'Insurers as asset managers and systemic risk', *Kelley School of Business Research Paper* .

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. (2020), 'Self-fulfilling runs: Evidence from the us life insurance industry', *Journal of Political Economy* **128**(9), 3520–3569.

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. (2022), 'How Do Financial Constraints Affect Product Pricing? Evidence from Weather and Life Insurance Premiums', *The Journal of Finance* **77**(1), 449–503.

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

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.

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

Huber, M. (2022), 'Regulation-induced interest rate risk exposure', *Working Paper* .

Knox, B. & Sørensen, J. A. (2020), 'Asset-driven insurance pricing', *Working Paper* .

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. (2022), 'The fragility of market risk insurance', *The Journal of Finance* **forthcoming**.

Krishnamurthy, A. & Vissing-Jorgensen, A. (2012), 'The Aggregate Demand for Treasury Debt', *Journal of Political Economy* **120**(2), 233–267.

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* .

Poterba, J. M. & Solomon, A. (2021), 'Discount rates, mortality projections, and money's worth calculations for us individual annuities', *NBER Working Paper 28557* .

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.

Schmidt, K. (2014), 'On inequalities for moments and the covariance of monotone functions', *Insurance: Mathematics and Economics* **55**, 91–95.

Sen, I. (2021), 'Regulatory Limits to Risk Management', *Review of Financial Studies* **forthcoming**.

van Binsbergen, J., Diamond, W. & Grotteria, M. (2021), 'Risk-free interest rates', *Journal of Financial Economics* **143**(1), 1–29.

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

In this appendix, we estimate the size of the private life annuity market and benchmark it against the amount of long-term fixed-rate bonds issued by U.S. corporations.

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 Filings of over 800 life insurers. For each insurance company, we extracted the amount of annual life contingent income payable to individual and group annuities reported in the "Exhibit of Number of Policies, Contracts, Certificates, Income Payable and Account Values in Force for Supplementary Contracts, Annuities, Accident & Health and Other Policies."

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.[38]

Figures 7 and 8 plot the time series of income and payout for the U.S. life insurance industry and the U.S. Social Security Administration, respectively.

To put these numbers in perspective, assume that there are 3.5 million new 65 year old individuals in the US in a given year—this is roughly the average between 2001 and 2019. Our previous calculation suggests that a representative cohort annuitizes almost $45 billion in wealth with life insurers in a given year—i.e., $12,700 dollar per 3.5 million new 65 year old individuals.

[38]The Social Security payout data are available here: https://www.ssa.gov/data/.

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

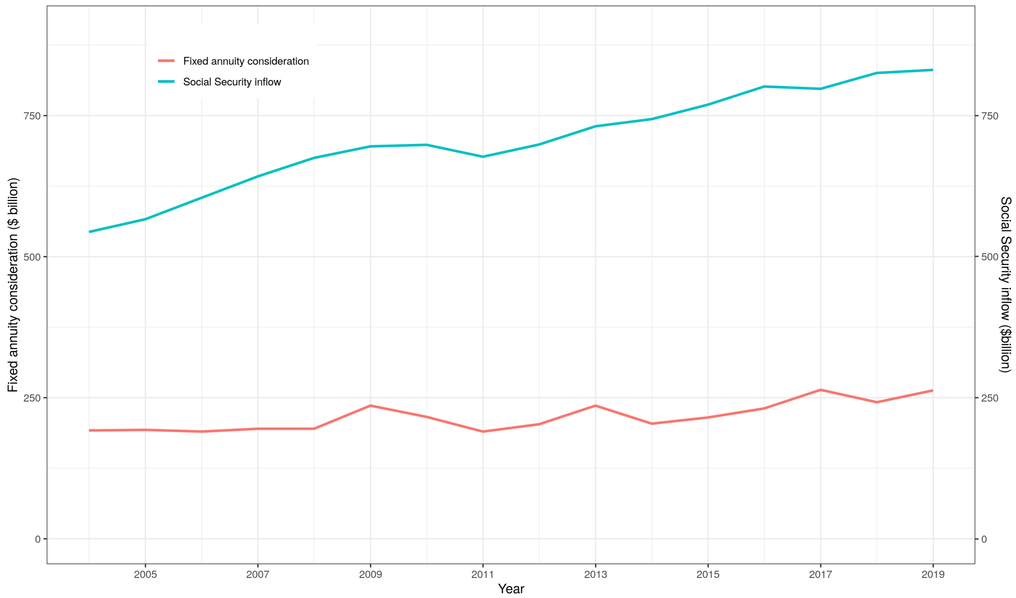
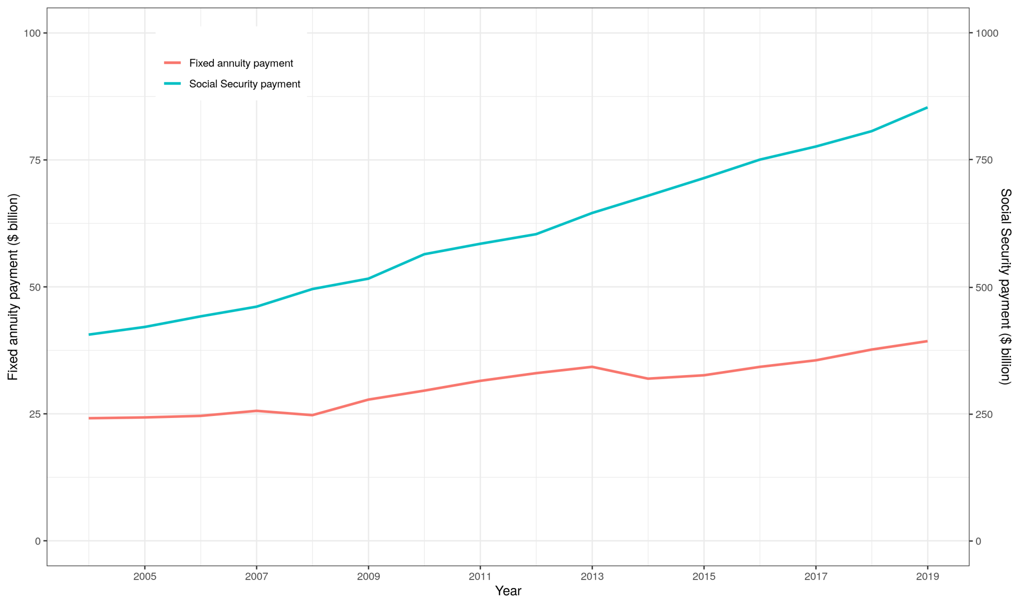


Figure 8: U.S. life insurers payout on fixed annuity and Social Security payout



Next, we show that the average amount of annuitized wealth is greater than the total supply of long-term fixed rate corporate bonds. We use data from Mergent FISD, which covers the universe of corporate bond issuance by U.S. corporations. This database provides information on over hundreds of bond characteristics, including coupon types, call features, ratings, and maturity. We focus on investment grade bonds and exclude callable bonds. Callable bonds are not useful to life insurers issuing life annuities because the issuer usually calls the bond when interest rates fall, which is precisely when insurers need long-term bonds the most.

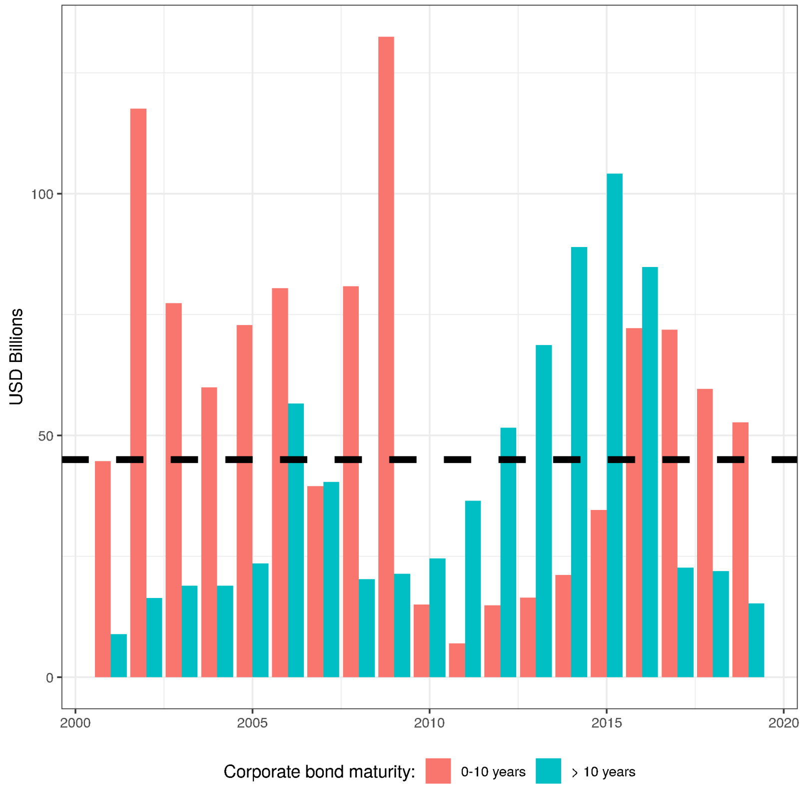


Figure 9: Annual wealth annuitization with life insurers and fixed rate corporate bond issuance

Figure 9 benchmarks the average amount of annuitized wealth against the amount of fixed-rate corporate bonds broken down by initial maturity above and below 10 years. The solid black dashed line represents the aforementioned $45 billion in average annuitized wealth. Taking an average of the blue bars over the sample period shows that the average annual amount of annuitized wealth is about 15 percent larger than the total amount of fixed-rate, non-callable corporate bonds with maturity over 10 years issued by U.S. firms over the same period. This estimate reveals that the private life annuity market is larger than the long-term fixed rate bond market in the US. This finding is striking, as the

U.S. has the largest corporate bond market in the world and, therefore, U.S. life insurers compete for these long-term bonds with other long-term investors, such as pension funds and sovereign wealth fund in and out of the U.S.

B Theoretical results and proofs

Proof of Theorem 1: By Assumption 1, insurers must show that they will 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 outlined in Section 2.

Finally, we show that the optimal IRM is 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 annuity 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) $\zeta < 1$ and (ii) $\zeta = 1$. First, we examine case (i). When $\zeta < 1$, take the original annuity price q as given and 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 analyze case (ii). When $\zeta = 1$, take the original annuity price q as given and 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 can become insolvent 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 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 strictly positive profit by choosing a price q and implementing the hedging strategy in Theorem 1, which is a contradiction. ■

Proof of Proposition 1: 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 $\alpha^* \in \left(\underline{\alpha}, \overline{\alpha}\right)$ 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

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

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

To prove Theorem 2, we use the following lemma, which places an upper bound on the change in the risk-adjusted actuarially fair price q^{AF} with respect to the equilibrium annuity price q^*.

Lemma 2 *If Assumption 2 holds, then* $\frac{\partial q^{AF}}{\partial q^*} < 1$ *for any* $\zeta \in [0, 1)$.

Proof First, note that at the equilibrium annuity price q^*, $\frac{\partial \Pi(q^*, \zeta)}{\partial q^*} \geq 0$. This is because if $\frac{\partial \Pi(q^*, \zeta)}{\partial q^*} < 0$, then insurers can lower the price to capture the entire market and earn strictly higher profit, which is a contradiction. Because the equilibrium annuity price satisfies

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

and $\frac{\partial \Pi(q^*,\zeta)}{\partial q^*} \geq 0$, it follows that

$$\begin{aligned} &\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha \\ &+ \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha} a\left(\alpha, q^*\right) - \frac{\partial a\left(\alpha, q^*\right)}{\partial q^*}\right] g\left(\alpha\right) d\alpha \\ &+ \frac{1-\zeta}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2\left[1-\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha} a\left(\alpha, q^*\right) - \frac{\partial a\left(\alpha, q^*\right)}{\partial q^*}\right] g\left(\alpha\right) d\alpha \geq 0. \end{aligned} \tag{7}$$

Next, we show that Assumption 2 implies that the last term in inequality (7) is strictly negative when $\zeta \in [0, 1)$. Since $\frac{\partial a(\alpha,q)}{\partial q}$ is finite (Assumption 1), Assumption 2 implies that $\text{cov}\left(\alpha^2, \frac{\partial a(\alpha,q)}{\partial q}\right) \geq 0$ for any q (Schmidt 2014). Therefore, we have

$$\begin{aligned} &\frac{\text{cov}\left(\alpha^2, a\left(\alpha, q\right)\right)}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q\right) g\left(\alpha\right) d\alpha} > \frac{\text{cov}\left(\alpha^2, -\frac{\partial a(\alpha,q)}{\partial q}\right)}{\int_{\underline{\alpha}}^{\overline{\alpha}} -\frac{\partial a(\alpha,q)}{\partial q} g\left(\alpha\right) d\alpha} \\ \Longrightarrow &\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a\left(\alpha, q\right) g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q\right) g\left(\alpha\right) d\alpha} > \frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left(-\frac{\partial a(\alpha,q)}{\partial q}\right) g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} -\frac{\partial a(\alpha,q)}{\partial q} g\left(\alpha\right) d\alpha}. \end{aligned}$$

The first inequality above comes from the fact that $\text{cov}\left(\alpha^2, -\frac{\partial a(\alpha,q)}{\partial q}\right) = -\text{cov}\left(\alpha^2, \frac{\partial a(\alpha,q)}{\partial q}\right)$, and Assumption 1 implies $\text{cov}\left(\alpha^2, a\left(\alpha, q\right)\right) > 0$ (Schmidt 2014). The second inequality above uses the definition of covariance. By rearranging the terms of the second inequality, we find that the last term in inequality (7) is strictly negative if $\zeta \in [0, 1)$.

Therefore, we have

$$\begin{aligned} &\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha \\ &+ \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha} a\left(\alpha, q^*\right) - \frac{\partial a\left(\alpha, q^*\right)}{\partial q^*}\right] g\left(\alpha\right) d\alpha > 0. \end{aligned} \tag{8}$$

By the definition of q^{AF} and dividing both sides of inequality (8) by $\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha$, we obtain $1 - \frac{\partial q^{AF}}{\partial q^*} > 0$ when $\zeta \in [0, 1)$. $\blacksquare$

Proof of Theorem 2: First, we show that there exists a $q^* = \min\{q|\Pi(q,\zeta) = 0\}$ for any ζ. By Assumption 1, when $q = \frac{\overline{\alpha}}{R_1}(1 + \epsilon\overline{\alpha})$, then $\Pi(q,\zeta) > 0$. Also, when $q = 0$, then $\Pi(q,\zeta) < 0$. Since $\Pi(q,\zeta)$ is continuous in q, the intermediate value theorem implies that there exists q such that $\Pi(q,\zeta) = 0$. Therefore, the set $\{q|\Pi(q,\zeta) = 0\}$ is non-empty. Also, $\{q|\Pi(q,\zeta) = 0\}$ is closed, because $\{0\}$ is closed and Π is continuous in q so $\Pi^{-1}(\{0\},\zeta)$ is closed. Furthermore, $\{q|\Pi(q,\zeta) = 0\}$ is bounded below by zero. Hence, a minimum for $\{q|\Pi(q,\zeta) = 0\}$ exists.

Next, we show that q^* increases as ζ decreases. Through implicit differentiation,

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

Immediately, notice the numerator is negative. Suppose the denominator, $\frac{\partial \Pi(q^*,\zeta)}{\partial q^*}$, is strictly negative. This would imply that an insurer can deviate by lowering the price to capture the entire market and earn strictly positive profit. However, this contradicts the fact that $q^* = \min\{q|\Pi(q,\zeta) = 0\}$. Hence, we have $\frac{\partial q^*}{\partial \zeta} < 0$.

Finally, we show that $q^* - q^{AF}$ increases as ζ decreases, notice that $\frac{\partial q^*}{\partial \zeta} - \frac{\partial q^{AF}}{\partial \zeta} = \frac{\partial q^*}{\partial \zeta}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. Since $\frac{\partial q^*}{\partial \zeta} < 0$ and by Lemma 2, $1 > \frac{\partial q^{AF}}{\partial q^*}$, we have $\frac{\partial q^*}{\partial \zeta} - \frac{\partial q^{AF}}{\partial \zeta} < 0$. $\blacksquare$

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

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

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 the relationship between shocks originating in 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 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* by implicitly differentiating the insurers' zero-profit condition $\Pi(q,z)=0$:

$$\frac{\partial q^*}{\partial z} = \underbrace{\frac{\frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha,q^*)\,g(\alpha)\,d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\,g(\alpha)\,d\alpha}}_{\text{Risk management effect}} + \underbrace{\frac{\partial q^*}{\partial z}\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} e(\alpha,q^*)\left[1-\frac{\frac{\alpha}{R_1}(1+\alpha z)}{q^*}\right] a(\alpha,q^*)\,g(\alpha)\,d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\,g(\alpha)\,d\alpha}}_{\text{Adverse selection effect}}, \quad (9)$$

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 the cost of IRM raises the equilibrium annuity price because the insurer must finance a greater level of net worth in $t=0$ with the annuity markup. 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 (9) can be written as

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

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 (9) 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 α : $e(\alpha',q^*) > e(\alpha'',q^*)$ 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, potentially triggering a death spiral. To see this, first note that among agents purchasing annuities (all α such that $a(\alpha,q^*) > 0$), there exists a $\tilde{\alpha}$ such that $q^* < \frac{\alpha}{R_1}(1+\alpha z)$ for any $\alpha > \tilde{\alpha}$ and $q^* > \frac{\alpha}{R_1}(1+\alpha z)$ 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. Therefore, the insurer must further raise prices to compensate for more severe adverse selection, which could trigger a death spiral. 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 i 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:

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

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 (10). 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.

Table 4: The effect of corporate bond market shocks on adverse selection The unit of observation is a life insurer-product-quarter. The dependent variable $AS\_pricing_{ijt}$ 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

E Extensions to the benchmark model

E.1 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 provides at least the same return as the two-period corporate bond, 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 ($\zeta < 1$).

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: $R_g < R_l$. This assumption captures the U.S. government bonds' convenience yield. U.S. Treasury bonds have a convenience yield because investors value their liquidity and safety and are willing to accept lower yields to hold them over alternative investments that offer the same cash flows (Krishnamurthy & Vissing-Jorgensen 2012).

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

$$b_1 + l_2 + g_2 = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{11}$$

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

$$b_1 + l_2 + g_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha + NW_0. \tag{12}$$

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

$$b_2(R_2) = \frac{1}{R_2} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha + NW_1(R_2), \tag{13}$$

where

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{R_2} + \frac{R_g g_2}{R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q) g(\alpha) d\alpha. \tag{14}$$

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

Proposition 2 *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:

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

ii. Capital structure:

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

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 Competitive insurers finance 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 (13), $b_2(1) = \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q) g(\alpha) d\alpha$. Hence, by equation (14), we obtain 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 (12) yields

$$NW_0 + \left(\frac{R_g}{R_1} - 1\right) g_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha - \left(\frac{R_l}{R_1} - 1\right) l_2. \quad (15)$$

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, equation (15) can be rewritten as

$$NW_0 + \left(\frac{R_g}{R_1} - 1\right) g_2 = (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha. \quad (16)$$

The right-hand side of the equation above is the total value of corporate bond holdings minus the value of insurance liabilities. As a result, when $R_g \leq R_1$, insurers prefer to hold a positive level of net worth to manage interest rate risk, because the yield on 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 first focus on the case with $NW_1(R_2) = 0$ and then show that $NW_0 = 0$ is not optimal when $R_g < R_l$.

Using equations (13) and (14), we obtain

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

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, net worth at $t = 0$ is given by:

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

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, equation (16) implies that

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

Equation (14) implies that

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

Using equation (13), we obtain the net worth at $t = 1$:

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

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

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

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

and from equation (14),

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

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. ■

Proposition 2 shows how insurers use long-term government bonds to manage interest rate risk. When the yield on the long-term government bonds are too low ($R_g \leq R_1$), their availability has no effect on the insurers' IRM problem. That is, the cost of managing the interest rate risk with net worth is lower than the cost of managing the interest rate risk with low yielding two-period government bonds. When the yield on the long-term government bonds are higher, insurers substitute some of their net worth at $t = 0$ with long-term government bond holdings to perform IRM. However, net worth and long-term government bonds are not perfect substitutes when the yield on the government bonds is such that $R_g < R_l$. The difference between the two yields could reflect, for example, a convenience yield due to the government bond's high liquidity and safe-haven status. In this case, net worth at $t = 0$ remains positive.

We now show that the competitive annuity price remains above the risk-adjusted actuarially fair price in this setting. Proposition 2 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 the arguments in Section 2, when the corporate bond market is unconstrained ($\zeta = 1$), the competitive annuity price is equal to the risk-adjusted actuarially fair price. Next, analogous to Theorem 2, we show how the AS-adjusted markup is strictly positive when the corporate bond market is constrained, even when long-term government bonds are available in unlimited supply.

Proposition 3 *When there is an unlimited supply of long-term government bonds with return $R_g < R_l$, the AS-adjusted markup $q^* - q^{AF}$ is higher when the corporate bond market is more constrained (ζ is lower): $\frac{\partial q^*}{\partial \zeta} - \frac{\partial q^{AF}}{\partial \zeta} < 0$. Furthermore, when the corporate bond market is unconstrained ($\zeta = 1$), the AS-adjusted markup is zero: $q^* = q^{AF}$.*

Proof Following the proof of Lemma 2 closely, it can be shown that $1 > \frac{\partial q^{AF}}{\partial q^*}$ when $\zeta \in [0, 1)$ if Assumption 2 holds in this environment. The rest of the proof is similar to the proof of Theorem 2 ■

E.2 Monopolistic competition

In this section, we explore how life insurers' IRM is affected by market competition. To model imperfect competition, we consider a market with two insurers: $\{X, Y\}$. Each insurer is matched with a continuum of individuals of measure 1 with identical survival probability distributions. An insurer can lower its annuity price to capture a portion of its competitor's market: If Insurer X sets price q_X, then Insurer Y can seize $\gamma(q_X - q_Y)$ of individuals that were matched with Insurer X by setting $q_Y < q_X$. 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 approaches our baseline specification with Bertrand competition as γ increases.

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. Although insurers with market power accumulate net worth in the form of monopoly profit, the net worth from exercising market power may prove to be inadequate for IRM. Therefore, insurers with market power must have an asset portfolio and capital structure that is at least consistent with Theorem 1. Specifically, the demand for one- and two-period bonds and the net worth in each period has to be weakly greater than the amount specified in Theorem 1. As a result, the cost of selling an annuity is determined by the expected present value of the insurance liabilities and the insurers' optimal net worth position.

Suppose Insurer X deviates by setting a price $\hat{q}$ while Insurer Y 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 that z is inversely related to ζ. Insurer X 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 use the fact that $\hat{q} = q^*$ at the optimum in a symmetric Nash equilibrium.[39] This yields the following equilibrium condition for 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 the above equilibrium condition, Theorem 3 below characterizes the relationship

[39] 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 that a unique optimum exists when the insurers are monopolists.

between the annuity market structure γ and the AS-adjusted markup.

Theorem 3 *In a market with monopolistic competition, the AS-adjusted markup increases as the bond market becomes more constrained (higher z): $\frac{\partial q^*}{\partial z} - \frac{\partial q^{AF}}{\partial z} > 0$, and decreases as the annuity market becomes more competitive (higher γ): $\frac{\partial q^*}{\partial \gamma} - \frac{\partial q^{AF}}{\partial \gamma} < 0$.*

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

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

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 $\frac{\partial q^*}{\partial \gamma} = -\frac{\partial W}{\partial \gamma} / \frac{\partial W}{\partial q^*} < 0$.

Next, to see that Lemma 2 holds in this environment, first notice that the insurer's profit is strictly positive at the equilibrium when insurers have market power:

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

which also implies that

$$q^* > \frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}(1+\alpha z) a(\alpha, q^*) g(\alpha) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) d\alpha}. \tag{17}$$

Therefore, by the first-order condition and the fact that $\gamma \geq 0$, it must be the case that

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

where the inequality above holds with equality when $\gamma = 0$. Using the definition of $z = 1 - \zeta + \zeta \mathbb{E}\left(\frac{1}{R_2}\right)$, we can rewrite the inequality above as

$$\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) d\alpha + q^* \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) d\alpha - \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) d\alpha - \frac{1-\zeta}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) d\alpha \geq 0.$$

Since $\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g(\alpha)\, d\alpha < 0$, we can substitute q^* in the inequality above with

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

which is strictly smaller than q^* by (17) and yield

$$\begin{aligned}&\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\, g(\alpha)\, d\alpha \\ &+\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\, g(\alpha)\, d\alpha} a(\alpha,q^*) - \frac{\partial a(\alpha,q^*)}{\partial q^*}\right] g(\alpha)\, d\alpha \\ &+\frac{1-\zeta}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2\left[1-\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\, g(\alpha)\, d\alpha} a(\alpha,q^*) - \frac{\partial a(\alpha,q^*)}{\partial q^*}\right] g(\alpha)\, d\alpha > 0.\end{aligned} \tag{18}$$

Notice (18) resembles inequality (7) from the proof of Lemma 2, but with a strict inequality. Following the proof of Lemma 2, $\frac{\partial q^{AF}}{\partial q^*} < 1$ when $\zeta \in [0,1)$ if Assumption 2 holds.

Finally, notice that $\frac{\partial q^*}{\partial z} - \frac{\partial q^{AF}}{\partial z} = \frac{\partial q^*}{\partial z}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. Since $\frac{\partial q^*}{\partial z} > 0$ and Lemma 2 applies, the AS-adjusted markup increases as z increases: $\frac{\partial q^*}{\partial z} - \frac{\partial q^{AF}}{\partial z} = \frac{\partial q^*}{\partial z}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right) > 0$. Next, note that $\frac{\partial q^*}{\partial \gamma} - \frac{\partial q^{AF}}{\partial \gamma} = \frac{\partial q^*}{\partial \gamma}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. When $\zeta < 1$, Lemma 2 applies and since $\frac{\partial q^*}{\partial \gamma} < 0$, we have $\frac{\partial q^*}{\partial \gamma} - \frac{\partial q^{AF}}{\partial \gamma} = \frac{\partial q^*}{\partial \gamma}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right) < 0$. And, when $\zeta = 1$, inequality (18) can be rewritten as

$$\begin{aligned}&\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\, g(\alpha)\, d\alpha \\ &+\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\, g(\alpha)\, d\alpha} a(\alpha,q^*) - \frac{\partial a(\alpha,q^*)}{\partial q^*}\right] g(\alpha)\, d\alpha > 0.\end{aligned}$$

By the definition of q^{AF} and dividing both sides of the above inequality by the demand $\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha,q^*)\, g(\alpha)\, d\alpha$, we obtain $1 - \frac{\partial q^{AF}}{\partial q^*} > 0$ when $\zeta = 1$. Therefore, for any $\zeta \in [0,1]$, the AS-adjusted markup decreases as γ increases: $\frac{\partial q^*}{\partial \gamma} - \frac{\partial q^{AF}}{\partial \gamma} = \frac{\partial q^*}{\partial \gamma}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right) < 0$. ■

Theorem 3 shows that the AS-adjusted markup can increase due to IRM or market power. Insurers either increase their AS-adjusted markup to finance the net worth needed for IRM or to limit the quantity sold in the market and exercise market power. In essence, Theorem 3 implies that monopolistic competition in an unconstrained bond market—$\zeta = 1$ or $z = \mathbb{E}\left(\frac{1}{R_2}\right)$—can generate observationally equivalent AS-adjusted markup as a perfectly competitive annuity market facing a constrained bond market. Furthermore, Theorem 3 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 can cope with different types of annuity market structure.

F Competition in the fixed annuity markets

This appendix presents some evidence of competition in the fixed annuity market place. Fixed annuities are standardized products that are not underwritten and insurers compete over prices around tight margins. The following quote from Athene's president Bill Wheeler during the 2019Q2 earning call provides an anecdotal evidence that the U.S. fixed annuity market is competitive:

> [I]f you think about the spectrum of companies and how they price new business, we probably are in the [...] top decile in terms of how quickly we reprice. And I suppose that has a lot to do with how we are compensated because we are not compensated on volumes. We're compensated on margin, okay? So that's really important. So there's no interest in trying to keep old pricing out there and try to get some more sales before you're finally forced to move it. [...] So being a first mover is good for margins and good for return on capital. It's not so good necessarily for the competitive environment because you tend to be the price leader downwards, or could be upwards too. But they're–they're downwards in this environment.

We investigate the issue of competition more formally by calculating a Herfindahl-Hirschman Index (HHI) for the industry. Figure 10 calculates the HHI using insurer-level data on fixed annuity premiums and considerations extracted from about 800 NAIC Statutory filings. Our tedious collection of statutory filings starts in 2003. The solid line represents the HHI and shows that the US fixed annuity market concentration is consistently below 8 percent. Figure 10 confirms industry commentaries that the fixed

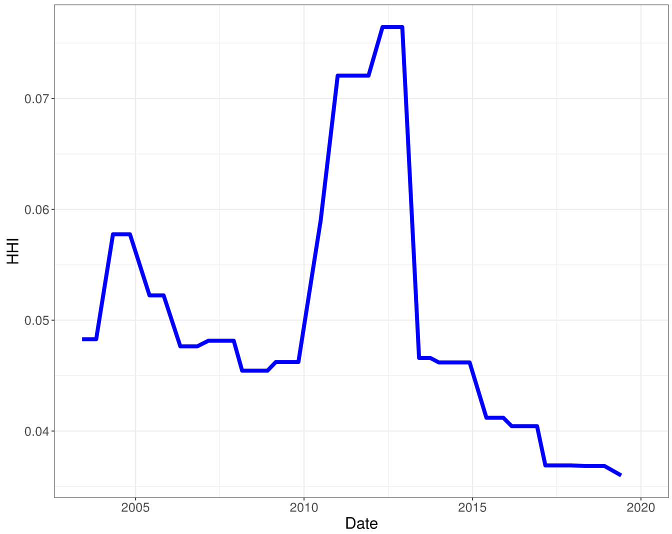


Figure 10: Fixed annuity market Herfindahl-Hirschman Index

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 (PE) firms in the industry. The PE-backed insurers purchased large blocks of legacy annuity business and found innovative ways to invest in relatively more illiquid assets without significantly increasing their regulatory risk based capital charges. By adding more illiquidity to their asset side without incurring a significant increase in risk-based capital charge, these PE-backed insurers, such as Athene, 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. Our goal in this appendix is to bridge the gap between our simple model and our empirical strategy.

First, we introduce the two exogenous shocks—the regulatory reserve requirement and the long-term Baa-Aaa bond yield spread.

- Let τ denote the regulatory 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 ($\tau = 1$). 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 regulatory requirements.

- 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. In this case, the spread of the two-period bond in the model is relative to a riskless cash instrument. Let ρ denote the shock on the long-term bond return such that the return is $\frac{R_l}{\rho}$ in $t = 2$. Our baseline model in Section 2 corresponds to the case when $\rho = 1$. Because we do not explicitly model 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 is analogous to a widening (narrowing) of the long-term Baa-Aaa bond spread in the data. Following the exposition in Section 3.4, we assume that ρ is orthogonal to the insurers' discount rate.

Next, we modify the benchmark model of Section 2 with the two shocks. In this environment, the insurer's balance sheet at $t = 0$—equation (2) of the benchmark model—is rewritten as

$$b_1 + l_2 = \tau \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E} \left(\frac{1}{R_2} \right) \right] a(\alpha, q)\, g(\alpha)\, d\alpha + NW_0. \tag{19}$$

Also, the return on long-term bonds depends on ρ, so equation (4) of the benchmark model becomes

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{\rho R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{20}$$

The asset portfolio and capital structure are determined by equations (1), (3), (19), and (20).

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

is given by:

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

Immediately, we find that equation (19) can be rewritten as

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

$$- (\tau - 1)\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} a(\alpha,q)\, g(\alpha)\, d\alpha. \quad (21)$$

Equation (21) shows that insurers stop selling annuities when the reserve ratio is sufficiently high—i.e., when the right-hand side of equation (21) is negative. Hence, we focus on values of τ that are above but sufficiently close to 1 so that the insurers choose to sell annuities despite their higher cost. By equation (21), $\tau > 1$ means that either the demand for long-term bonds l_2 or the net worth NW_0 must change.

We make two additional assumptions. First, we rule out the case when the insurers change their net worth at $t = 0$ as a response to a change in τ. This is because real world insurers' financial strength ratings and regulatory risk-based capital ratios are tied to (proxies of) the insurers' net worth position or policyholders surplus. In other words, we assume that the net worth in $t = 0$ is invariant to changes in the reserve requirement. Let $\overline{NW}_0(\rho)$ denote the insurers' net worth at $t = 0$ as a function of ρ when $\tau = 1$. Therefore, we focus on how insurers adjust their two-period bond holdings when τ goes above 1.

Second, we assume that the increase in the two-period bond spread is small. 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, in an unconstrained bond market, insurers can fully hedge the interest rate risk with less two-period bonds than $l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha,q)\, g(\alpha)\, d\alpha$ when $\rho < 1$ and $\tau = 1$. Furthermore, when ρ is sufficiently smaller than 1, it is even possible for insurers to fully hedge the interest rate risk in a constrained bond market ($\zeta < 1$). As a result, 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 bond market remains constrained.

Therefore, the optimal demand for two-period bonds in $t = 0$ when $\tau = 1$ is given by:

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

As a result, the optimal net worth at $t = 0$ when ρ is small and $\tau = 1$ is

$$\overline{NW}_0(\rho) = \begin{cases} (1-\zeta) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} \left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) dG(\alpha) - \zeta \left(\frac{1-\rho}{\rho}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1} a(\alpha, q) dG(\alpha), & \zeta < 1 \\ 0, & \zeta = 1. \end{cases}$$

Clearly, for any $\zeta < 1$, the increase in the two-period bond spread has to be sufficiently small for $\overline{NW}_0(\rho)$ to be positive.

To see how a tightening of reserve requirement or an increase in the long-term bond spread affects the equilibrium annuity price q^*, define $\Pi(q, \tau, \rho)$ as the insurers' profit while maintaining the net worth $\overline{NW}_0(\rho)$:

$$\Pi(q, \tau, \rho) = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q) g(\alpha) d\alpha - \tau \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q) g(\alpha) d\alpha - \overline{NW}_0(\rho).$$

The competitive price q^* is determined by the zero profit condition: $\Pi(q^*, \tau, \rho) = 0$. By implicit differentiation, we find that both $\frac{\partial q^*}{\partial \tau}$ and $\frac{\partial q^*}{\partial \rho}$ are strictly positive when the bond market is constrained if $\Pi(q, \tau, \rho)$ is increasing at q^*, which must be true at the equilibrium.[40] Moreover, when the bond market is unconstrained, $\frac{\partial q^*}{\partial \rho}$ is zero while $\frac{\partial q^*}{\partial \tau}$ is strictly positive. This implies that insurers raise the annuity price when the reserve requirement tightens regardless of whether IRM is costly or not, but the long-term bond spread shifts the annuity price only when IRM is costly.

Note that the risk-adjusted actuarially fair price q^{AF} is not directly affected by the two exogenous shocks but is indirectly affected through the equilibrium annuity price q^*. Therefore, we can define q^{AF} as before in Section 2:

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

It is worth emphasizing that the insurer's own actuarial valuation is different from the

[40] Suppose that $\Pi(q, \tau, \rho)$ 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 profit, contradicting the fact that q^* is an equilibrium price.

regulator's, and q^{AF} captures the cost of providing annuities from the insurer's perspective. The regulator and the insurers perspectives match only when the reserve ratio is slack ($\tau = 1$).

Finally, we present a proposition that shows how the AS-adjusted markup responds to changes in the regulatory reserve requirement and the long-term bond spread across different bond market conditions.

Proposition 4 *When the bond market is unconstrained ($\zeta = 1$), the AS-adjusted markup $q^* - q^{AF}$ (i) increases as the reserve requirement tightens ($\tau > 1$ and τ is higher): $\frac{\partial q^*}{\partial \tau} - \frac{\partial q^{AF}}{\partial \tau} > 0$, and (ii) is invariant to changes in the long-term bond spread: $\frac{\partial q^*}{\partial \rho} - \frac{\partial q^{AF}}{\partial \rho} = 0$.*

When the bond market is constrained ($\zeta < 1$), the AS-adjusted markup $q^* - q^{AF}$ (i) increases as the reserve requirement tightens ($\tau > 1$ and τ is higher): $\frac{\partial q^*}{\partial \tau} - \frac{\partial q^{AF}}{\partial \tau} > 0$, and (ii) decreases as the long-term bond spread widens (ρ is lower): $\frac{\partial q^*}{\partial \rho} - \frac{\partial q^{AF}}{\partial \rho} > 0$.

Proof First, following the proof of Lemma 2 closely, it must be the case that $1 > \frac{\partial q^{AF}}{\partial q^*}$ when $\zeta \in [0, 1)$ if Assumption 2 holds in this environment.

Next, we will show that $\frac{\partial q^*}{\partial \tau} - \frac{\partial q^{AF}}{\partial \tau} > 0$ for any $\zeta \in [0, 1]$. Notice that $\frac{\partial q^*}{\partial \tau} - \frac{\partial q^{AF}}{\partial \tau} = \frac{\partial q^*}{\partial \tau}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. In the text, we have shown that $\frac{\partial q^*}{\partial \tau} > 0$ for any $\zeta \in [0, 1]$. Therefore, we only need to show that $1 > \frac{\partial q^{AF}}{\partial q^*}$, which is true when $\zeta < 1$ and Assumption 2 is satisfied. When $\zeta = 1$, we have

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

And since $\frac{\partial \Pi(q^*, \tau, \rho)}{\partial q^*} \geq 0$, we have

$$\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*)\, g(\alpha)\, d\alpha + \tau \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*)\, g(\alpha)\, d\alpha} a(\alpha, q^*) - \frac{\partial a(\alpha, q^*)}{\partial q^*}\right] g(\alpha)\, d\alpha \geq 0.$$

By the definition of q^{AF} and dividing both sides of the inequality above by the demand $\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*)\, g(\alpha)\, d\alpha$, we obtain $1 \geq \tau \frac{\partial q^{AF}}{\partial q^*}$. Since $\tau > 1$, we have $1 - \frac{\partial q^{AF}}{\partial q^*} > 0$. Therefore, we obtain $\frac{\partial q^*}{\partial \tau} - \frac{\partial q^{AF}}{\partial \tau} > 0$ for any $\zeta \in [0, 1]$.

Finally, notice that $\frac{\partial q^*}{\partial \rho} - \frac{\partial q^{AF}}{\partial \rho} = \frac{\partial q^*}{\partial \rho}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. When $\zeta = 1$, we have $\overline{NW}_0(\rho) = 0$

for any ρ and the text above showed that $\frac{\partial q^*}{\partial \rho} = 0$. As a result, we have $\frac{\partial q^*}{\partial \rho} - \frac{\partial q^{AF}}{\partial \rho} = 0$ when $\zeta = 1$. On the other hand, $1 > \frac{\partial q^{AF}}{\partial q^*}$ when $\zeta < 1$ and Assumption 2 is satisfied. Therefore, when $\zeta < 1$, we have $\frac{\partial q^*}{\partial \rho} - \frac{\partial q^{AF}}{\partial \rho} > 0$ since $\frac{\partial q^*}{\partial \rho} > 0$. ■

Proposition 4 shows that the AS-adjusted markup increases when the reserve requirement tightens, regardless of bond market conditions. This is because, independent of the cost of IRM, insurers use a fraction of their annuity markup to finance the excess reserve when the regulatory reserve requirement binds.

The first part of Proposition 4 shows that the AS-adjusted markup is unaffected by changes in the long-term bond spread when the bond market is unconstrained. This is simply due to the fact that the AS-adjusted markup is mainly used to finance net worth and net worth is zero in an unconstrained bond market. More precisely, we can see this by deriving the optimal long-term and short-term bond demands at $t = 0$ in an unconstrained bond market setting:

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

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

The aggregate bond demand $b_1 + l_2$ at $t = 0$ is thus

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

which does not change with the two-period bond spread. It follows that the competitive equilibrium price in an unconstrained bond market is invariant to changes in ρ. It is important to point out that this result also holds in a setting with monopolistic competition. As a result, the equilibrium annuity price is invariant to changes in the two-period bond yield spread when the bond market is unconstrained, regardless of market power.

The second part of Proposition 4 focuses on a constrained bond market, and we show how the AS-adjusted markup decreases when the long-term bond spread widens. Intuitively, insurers finance their net worth with the AS-adjusted markup, and the AS-adjusted markup decreases when the cost of financing net worth falls due to a widening bond spread.

From the discussion above, the effect of an increase in reserve requirement on markup is offset by an increase in long-term bond yield spread under the hypothesis of costly IRM. The reserve requirement shocks allow us to identify the general effect of financial frictions, while the long-term bond spread shocks allow us to tease out the effect of the costly IRM friction among competing supply-side alternatives.

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.

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."[41] 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

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

reference rate calculation.[42] 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 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 using the "basic" mortality tables prior to the adoption of the 2012 SOA generational table, we follow industry practice and apply the SOA generational factor (G2 scale) 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

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

variation in average swap duration across insurers in our setting. Moreover, 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 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 $\text{Swap\_duration}_{it}^{\text{Receive Float}} = -0.75 \times \text{Contract residual maturity} + 1/4 \times 1/2$.

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 insurers' 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.

I Main empirical results excluding the 2007-09 financial crisis

In this appendix, we show that our findings are not driven by variations in the 2007-09 period. Table 5 reproduces our main result on a sample before and after the 2007-09 financial crisis, respectively. The results in Table 5 are broadly consistent with the results in Table 2. For example, focusing on Column 2, the coefficient estimate on the interaction term suggests that, conditional on insurers' average cost of funding, a one standard deviation increase in $Reserve\_Ratio_{jt}$ (0.051) raises the AS-adjusted markup by 1.38 percentage point when $Baa\text{-}Aaa\_spread_t$ is at its median level (0.8). This effect is about 23 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.

Table 5: The effect of investment-grade corporate-bond yield spread on life annuity markups – Robustness excluding the 2007-09 financial crisis The unit of observation is a life insurer-product-quarter. The sample of observation extends from 1989 to 2006 in Columns 1-2 and from 2010 to 2019 in Columns 3-4. The dependent variable $Annuity\_markup_{ijt}$ is the AS-adjusted markup for life annuity j sold by insurer i at date t. Columns 1 and 3 report insurer clustered robust standard errors in parentheses and Columns 2 and 4 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)	(3)
$Baa\text{-}Aaa\_spread_t \times Reserve\_Ratio_{jt}$	-31.68***	-31.68**	29.95***	29.95*
	(9.03)	(14.75)	(5.69)	(15.21)
$Reserve\_Ratio_{jt}$	52.32***	52.32***	-5.71	-5.71
	(8.06)	(12.97)	(6.17)	(13.12)
$Baa\text{-}Aaa\_spread_t$	29.33***	29.33**	-34.55***	-34.55**
	(9.18)	(14.04)	(5.96)	(15.64)
$10\_HQM\_spread_t$	0.90	0.90	7.36***	7.36***
	(0.69)	(0.91)	(0.81)	(1.74)
Fixed effects:				
Contract characteristics (j)	Y	Y	Y	Y
Insurer (i)	Y	Y	Y	Y
SE Clustering	Insurer	Insurer/Date	Insurer	Insurer/Date
Observations	16,767	16,767	21,502	21,502
Adjusted R^2	0.51	0.51	0.43	0.43