What's Wrong with Annuity Markets?*

Stéphane Verani[†] Pei Cheng Yu[‡]

First version: August 2020; this version: August 2023

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 8 percentage points. The contribution of interest rate risk to annuity markups sharply increased after the Global 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, and Santiago Carbó Valverde as well as seminar and conference participants at Academia Sinica, the 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, the European Winter Finance Conference 2022, and the Western Economics Association International Conference 2023. Sam Dreith provided exceptional research assistance. Pei Cheng Yu would like to thank the Institute of Economics at Academia Sinica for its hospitality during the COVID-19 pandemic. 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, 20th & C Streets, NW, Washington, DC 20551.

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

Introduction

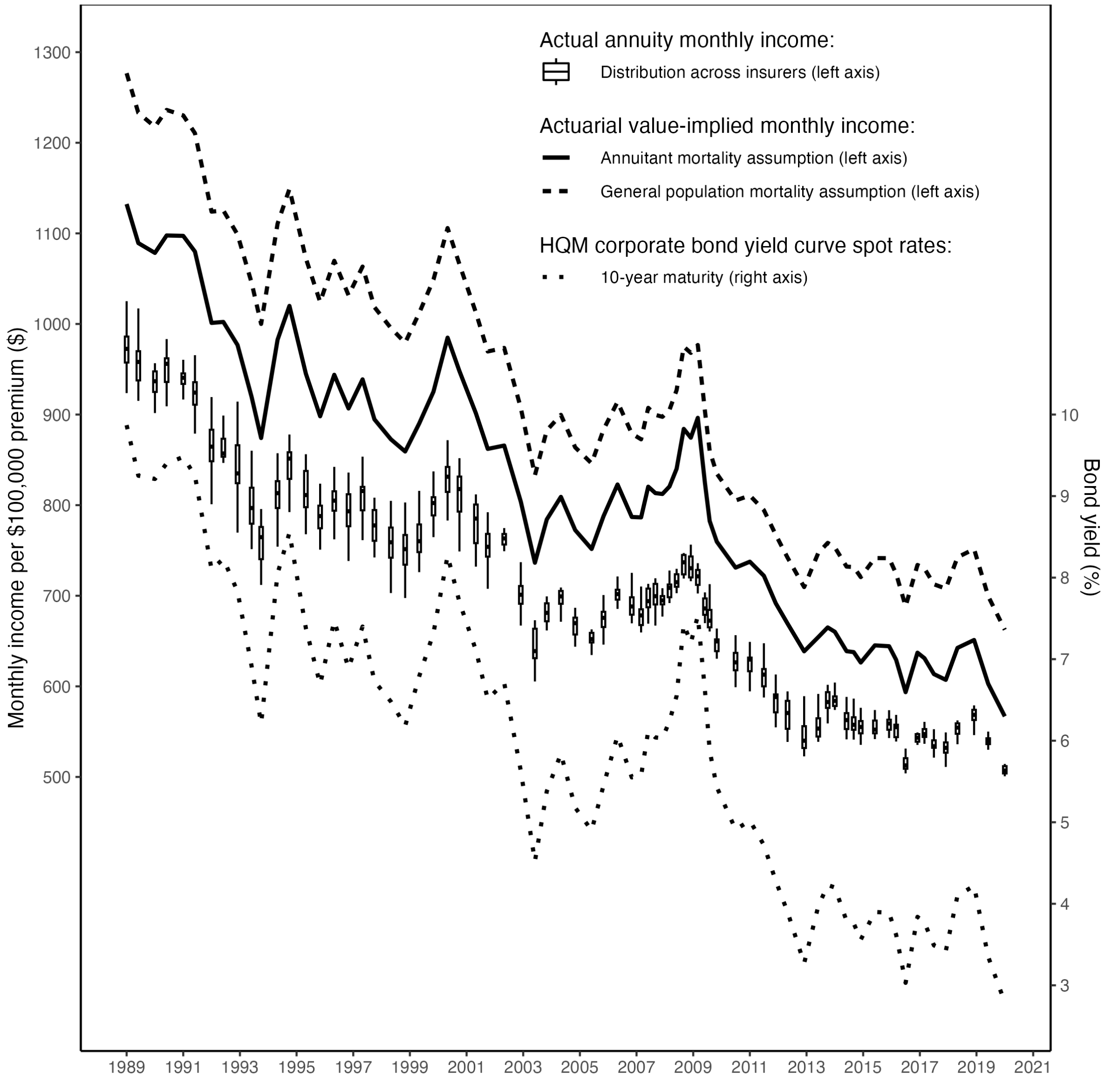
The fundamental risk for retirement is uncertain longevity. Life annuities offer a unique risk transfer solution to retirees wishing to shed the risk of outliving their financial wealth (Yaari 1965; Mitchell et al. 1999; Davidoff et al. 2005). By purchasing a life annuity, retirees transfer their idiosyncratic 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 et al. 2023). 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 and Gruber 1996; Golosov and Tsyvinski 2007). Examining how interest rate risk affects the supply of annuities requires identifying the sources of market inefficiencies that influence longevity insurance markets.

We develop an algorithm for life annuity valuation that decomposes the contributions of demand- and supply-side frictions in life annuity markups observed from 1989 to 2019. The small vertical boxplots in Figure 1 represent the distribution of actual monthly payments offered by US life insurers to a 65-year-old male purchasing a \$100,000 single premium immediate 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 their wealth accumulation methods such as, bank deposits, mutual funds, employer-sponsored defined contribution plans, deferred fixed annuities, or variable annuities.[1] The monthly income offered on new life annuity contracts is positively correlated with the average yield on investment-grade corporate bonds of comparable duration, which is represented by the dotted line below the vertical boxplots.

The short-dashed and solid lines above the boxplots represent the monthly payments implied by the annuity contract's actuarial value calculated using general population mortality and annuitant mortality, respectively. The gap between these two lines is a measure of the industry's average *adverse selection pricing*. This well-known source of demand-side inefficiency stems from life insurers' inability to observe the mortality risk of the individuals seeking longevity insurance, leading to adverse selection (Eichenbaum and

[1]See Appendix A for more details. Note that the price of a life annuity is inversely related to the monthly income it generates.

Figure 1: Life annuity income and investment-grade corporate bond yields This figure plots the insurer distribution (boxplots) of actual monthly dollar income for a nominal $100,000 SPIA offered to a 65-year-old male between 1989 and 2019—it is inversely related to the annuity price. The dotted line below the boxplots represents the 10-year High Quality Market (HQM) corporate bond zero-coupon yield calculated by the U.S. Treasury. The solid and short-dashed lines above the boxplots represent the monthly income implied by the actuarial values of this SPIA using the annuitant and general population mortality estimates, respectively.



Peled 1987; Finkelstein and Poterba 2004). The difference between the actual monthly payment offered (represented by the boxplots) and the solid line is the *adverse selection-adjusted annuity price markup* (henceforth, *AS-adjusted markup*). This portion of the overall annuity markup cannot be explained by differences in survival rates in the annuitant pool and the general population. 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 in present value term during the 1989–2019 period.[2]

In this paper, we study the effect of interest rate risk on life annuity markups. We show that the cost of managing the interest rate risk associated with selling life annuities accounts for around half of the AS-adjusted markup, or 8 percentage points. In other words, besides 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 capital market frictions affects life insurers' product design and capital structure decisions. Additionally, we show that the contribution of interest rate risk to the AS-adjusted markup sharply increased after the Global 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 three reasons. The first reason is that adverse selection alone cannot account for the relatively high annuity markups observed in the data. This is a robust finding in the adverse selection literature, as shown by Brown (2001) and Finkelstein and Poterba (2004), among others, and is visually confirmed by Figure 1. This finding led past researchers to conclude that the AS-adjusted markup reflects mostly "administrative costs"—a catchall term broadly defined as marketing costs, corporate overhead, income taxes, contingency reserves, and profits (Mitchell et al. 1999). Our analysis demonstrates that annuity markets are distorted not only by asymmetric information, which is common in insurance markets, but also frictions in capital markets. The second reason, which is related to the first, is that our results have significant implications for reforms in retirement systems, especially reforms relying on private markets to provide retirement income. Our paper shows that the supply of private life annuities is constrained by interest rate risk. Therefore, reform proposals that do not take this supply-side inefficiency into account may not be welfare maximizing. The third reason is that our results show that the profitability of the entire financial sector, including insurers

[2]Our life annuity valuation framework resembles those of Koijen and Yogo (2015) and Poterba and Solomon (2021). The main difference is the choice of discount rate to value the annuity payment stream. Koijen and Yogo (2015) assume life insurers' discount their annuity liabilities at the same rate as the US Treasury, whereas Poterba and 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 difference means that our AS-adjusted markup is close to the baseline estimate of Poterba and Solomon (2021) and higher than Koijen and Yogo (2015). Section 3.1 discusses this important issue in detail.

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

and banks, can be threatened by large interest rate shocks. For instance, US life insurers struggled to cope with whole-life policy surrenders when the Federal Reserve under Chairman Volcker fought inflation in the late 1970s and early 1980s (National Association of Insurance Commissioners 2013; Brunetti et al. 2023), and US regional banks experienced depositor runs when the Federal Reserve under Chairman Powell rapidly raised rates to fight inflation in 2022. Our paper is a step toward understanding how financial intermediaries' vulnerability to interest rate shocks affects the broader economy.

To understand the effects of interest rate risk management on annuity markups and derive testable hypotheses, we provide a model of annuity pricing with adverse selection and interest rate risk. In the model, limited liability life insurers invest their total annuity premiums—i.e., the total lump sums given by individuals to a life insurer in exchange for life-contingent monthly payments—in corporate bonds issued by nonfinancial firms. Markets are incomplete, so only fixed-rate bonds—i.e., non-contingent debt securities—are traded. Interest rate risk arises because there is aggregate uncertainty over future interest rates. When the supply of long-term corporate bonds is efficient, we show that life insurers optimally manage interest rate risk by investing in a unique portfolio of corporate bonds with varying maturities that perfectly matches the duration of their annuity liabilities. Insurers' capital structure is unique and features zero net worth—i.e., the present value of insurers' assets equals the present value of their annuity liabilities. In this case, the annuity price reflects only the cost of adverse selection. When the long-term bond supply is inefficient and the return on long-term bonds is relatively low, life insurers can no longer perfectly match the duration of their annuity liabilities with corporate bonds.[4] In this case, we show that insurers can close the negative duration gap between their assets and annuity liabilities by adding a positive level of net worth to their capital structure. Net worth equalizes the duration of insurers' corporate bond holdings and total liabilities (annuity liabilities plus net worth), which acts as precautionary savings to avoid insolvency under low realizations of interest rates. We show that if insurers are competitive—even imperfectly so—the cost of managing interest rate risk with net worth is passed on to annuitants in the form of a positive AS-adjusted markup.

Our model offers three important insights to identify the effect of interest rate risk

[4] A large finance literature investigates the frictions causing the supply of long-term corporate bonds to be inefficient (Bolton and Scharfstein 1990, 1996; Hart and Moore 1994, 1998; Huang et al. 2019). We follow this literature and endogenize the long-term corporate bond supply using the framework of Greenwood et al. (2010).

on annuity prices. First, the model establishes the causal channel through which changes in long-term corporate bond returns affect the relative cost of hedging interest rate risk with net worth. Second, the model establishes a direct link between life insurer annuity markups, which are observable at a high frequency, and their net worth, which is typically unobservable. Third, the model demonstrates that the cost of interest rate risk management is a function of insurers' investment and capital structure decisions, which means that the asset and liability sides of insurers' balance sheets are endogenous.

We identify the effect of interest rate risk management on annuity markups using annuity prices offered by U.S. life insurers from 1989 to 2019. Identifying this effect is difficult because demand- and supply-side frictions jointly determine annuity markups. Moreover, life insurers are obviously exposed to interest rate risk in several ways–through different product offerings, their risk-management objectives, and their regulatory environment. We overcome this identification challenge by studying the pricing of life annuities on the margin—as real-life insurers do—and by exploiting the interaction of shocks to contract-level regulatory capital requirements and shocks to long-term corporate bond market conditions. The contract-level regulatory capital requirement shocks create exogenous variations in the relative cost of supplying annuity contracts with different durations. Under the null hypothesis of costless risk management, changes in the long-term corporate bond spread have no effect on markups, as insurers can always hedge the interest rate risk with bonds. Under the alternative hypothesis, changes in long-term bond spreads have a disproportionate effect on the annuity markup of contracts with relatively high reserve requirements. The reason is that changes in long-term bond spreads disproportionately change the relative cost of hedging the annuity duration with net worth for these contracts.

We find that insurers raise their AS-adjusted markups when the relative cost of offering annuities increases as a result of an exogenous increase in the annuity contract-level regulatory reserve requirement, because insurers must create larger reserves backed by more bonds for their annuity contracts. However, this effect is substantially smaller when the return on long-term bonds exogenously increases, as the relative cost of hedging interest rate risk with long-term corporate bonds decreases when long-term corporate bond spreads increase, lowering their reliance on net worth. This difference-in-differences type of result identifies the effect of interest rate risk management on markups without relying on strong assumptions about unobserved changes in annuity demand. In addition, 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 interest rate risk management could account for almost all of the AS-adjusted markup after factoring in the operating expenses reported by the industry.

Our model features a nontrivial feedback mechanism between interest rate risk management and adverse selection. The cost of managing interest rate risk exacerbates the adverse selection problem by raising annuity prices. This development, in turn, deteriorates the annuitant pool and extends the duration of annuity liabilities, further increasing the cost of managing interest rate risk. We provide empirical evidence of this feedback by examining pricing differences in life annuities with different term guarantees. Adverse selection is more pronounced in contracts with period certain guarantees, as prospective buyers perceive a higher mortality risk during the guarantee term. Consistent with this prediction, we find that an exogenous increase in the contract-level regulatory reserve requirements ratio leads to a disproportionate increase in the adverse selection pricing of contracts with longer guarantee terms. This result highlights the effect of the interaction of frictions in insurance and capital markets on insurance pricing.

Lastly, we confirm our baseline empirical findings with a third set of empirical tests that exploit the heterogeneity in insurers' interest rate derivatives positions. Although not as widely used in the industry due to regulatory frictions, large and sophisticated life insurers can manage their balance sheets' negative duration gap by adding positive duration with interest rate swaps. We construct a novel data set on the universe of individual contract-level interest rate swap derivatives data between the end of 2009 and 2015. These data allow us to calculate the aggregate net duration each insurer in our sample adds to its balance sheet using interest rate swaps at any point in time.[5] Focusing on the cross-sectional variation of insurers' interest rate swap portfolios, we find 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 positions) disproportionately increase their AS-adjusted markups. 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. Then, focusing

[5]This approach is different from using aggregate swap gross notional amounts, which is not informative about the direction and size of the hedge.

on the within-insurer variation in interest rate swap portfolios, we use a quantile fixed-effect regression to show that the least competitive insurers who are the most beneficially affected by interest rate shocks (due to their hedging programs) cut their AS-adjusted markups the most—i.e., price their annuities more competitively.

Related literature

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 and Finkelstein (2011)—and the finance literature on risk management of financial institutions—e.g., Froot and Stein (1998). The former tends to focus on frictions that affect supply and demand in insurance markets while typically abstracting from frictions affecting insurers' capital market decisions. The latter tends to focus on frictions affecting supply and demand in capital markets while typically abstracting from frictions in insurance markets. We propose a novel theory of life annuity pricing based on optimal risk management and adverse selection.[6] Our theory emphasizes the nontrivial interaction between frictions in insurance and capital markets.

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, capital structure, and investment decisions—e.g., Koijen and Yogo (2015), Knox and Sørensen (2020), and Ge (2022). We differ from these papers by identifying the specific source of financial frictions—interest rate risk—that affects the supply of longevity insurance. Moreover, we show that supply-side and demand-side frictions are typically not orthogonal, as annuity markups depend on the interaction of frictions on insurance and capital markets. Earlier contributions studying the effect of interest rate risk on insurers' capital structure and investment decisions include Foley-Fisher et al. (2016) and Domanski et al. (2017). In contrast to these studies, insurers in our model price their life annuities by choosing a capital structure and an asset portfolio that are consistent with interest rate risk hedging and adverse selection. This specification also contrasts with recent studies following the seminal work of Koijen and

[6] 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 example, Ellul et al. (2021) and Koijen and Yogo (2022).

Yogo (2015) that abstract from insurers' endogenous capital structure, their asset purchasing decisions, and the interaction of both capital structure and purchasing decisions with adverse selection.[7] Other, more distant studies seek to measure life insurers' *residual* exposure to interest rate risk—e.g., Hartley et al. (2016), Ozdagli and Wang (2019), Sen (2021), Huber (2022), and Brunetti et al. (2023)—that is, insurers' ex post exposure to interest rates after implementing their interest rate risk hedging strategies. Brunetti et al. (2023) find that, consistent with our model's prediction, life insurers are hedged most of the time using high-frequency insurer stock price data. Our focus is different, as we study the effect of life insurers' ex ante 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 modeled in a large class of life-cycle models following the tradition of Yaari (1965) and Blanchard (1985). Most studies do not find that publicly provided annuities, such as Social Security, lead to significant welfare gains—e.g., Hong and Ríos-Rull (2007) and 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 largely results from the assumption that life insurers operate in frictionless capital markets where they can costlessly hedge interest rate risk. In contrast, we identify the cost of hedging interest rate risk as a key financial 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

[7] Recent work by Giambona et al. (2021) studies the effect of interest rate risk management on insurance pricing. Consistent with our study of interest rate swaps, they find that insurers that use relatively more derivatives, measured by higher derivative gross notional amounts, tend to post more competitive prices. That said, the authors focus on posted prices rather than price markups and abstract from insurers' capital structure decisions and adverse selection.

they are alive.

1.1 The U.S. life annuity market

Our paper focuses on the pricing of life annuities that allow retirees to decumulate wealth during retirement and are a type of *immediate annuity*. Roughly half of the U.S. life insurance industry's $600 billion aggregate income in 2018 came from annuity premiums. The other half is roughly split between life and health insurance premiums.[8] Total annuity income includes premiums and considerations related to both deferred annuities, which are pre-retirement savings vehicles, and life annuities, which allow retirees to annuitize their accumulated wealth. For this reason, 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 premiums and considerations in regulatory filings.

Nevertheless, we provide two novel estimates 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 National Association of Insurance Commissioners (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 and above. Second, using the same data, we calculate that in 2018, the U.S. life insurance industry's total payments to annuitants were about 3.5 percent of the total payments made by the U.S. Social Security Administration. These two estimates are consistent with the view in the literature that the market for life annuities in the U.S. is relatively small (Mitchell et al. 1999).

1.2 The life annuity business model

Life insurers' overall business model consists of earning a spread between the returns on their assets and the returns they owe to their policyholders, which is known in the industry as the net investment spread. 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 premiums primarily in fixed-income securities in

[8] See the American Council of Life Insurers' *2018 Life Insurers Fact Book*, https://www.acli.com/posting/rp18-007.

an effort to match their asset and liability cash flows. The illiquidity of life insurance liabilities allows insurers to invest premiums in relatively illiquid fixed-income assets, such as corporate bonds, asset-backed securities, and real estate loans. This approach offers a competitive return to policyholders and compensates 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).[9] By comparison, the rest of life insurers' general account assets include 8 percent in U.S. government securities and 14 percent in mortgage-backed securities, including those backed by the U.S. government.[10]

1.3 Life insurers' 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 in the U.S. 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 pay a fixed rate and are noncallable.[11] This maturity structure contrasts with the duration of a life annuity, which is approximately 10 years when offered to a 65-year-old individual. 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 premium, as discussed in studies such as Krishnamurthy and Vissing-Jorgensen (2012)

[9] General account assets back life insurers' insurance liabilities.

[10] The life insurance industry's relatively low holdings of U.S. government securities, which have relatively lower yields, reflect 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). Moreover, Foley-Fisher et al. (2022) show that a large fraction of the relatively small Treasury holdings of U.S. life insurers is used in connection with derivatives and securities financing transactions.

[11] There is a large literature attempting to explain this phenomenon–for example, through the lens of contracting frictions, as in Bolton and Scharfstein (1990, 1996), Hart and Moore (1994), Hart and Moore (1998), Barclay and Smith Jr (1995), Huang et al. (2019), and Greenwood et al. (2010).

and 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. Treasury securities.

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—roughly the average between 2000 and 2019. For simplicity, assume further that wealth is annuitized only 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 dollars 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, noncallable corporate bonds with maturity over 10 years issued by U.S. firms over the same period. This calculation implies that, although the life annuity market is small, it is larger than the total supply of fixed-rate long-term corporate bonds in the U.S., which has the largest corporate bond market in the world.

The negative duration gap between insurers' assets and their insurance liabilities means that life insurers are exposed to interest rate risk. A decrease in 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. Moreover, the financial condition of life insurers generally deteriorates when interest rates stay low, because periods of low interest rates are typically associated with lower bond coupon rates. These lower coupon rates depress insurers' net investment spread on new business. They also force insurers to reinvest the cash flow from maturing bonds into bonds paying lower coupon rates, which further depresses their net investment spread since the returns they promised policyholders are fixed—this possibility is known as reinvestment risk. Because the prospect of insolvency is incompatible with the sale of life annuities, interest risk management (henceforth, IRM) is at the heart of 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.[12] Net worth acts as precautionary savings and helps cushion the effect of interest rate changes that disproportionately

[12] 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 7.

affect the value of life insurers' insurance liabilities.[13] However, this approach is costly because the primary source of life insurer capital is accumulated retained earnings, which is directly related to insurers' net investment spread. 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 annuities that assume frictionless financial markets—e.g., Yaari (1965), Davidoff et al. (2005), and Hosseini (2015). We formalize the relationship between the cost of IRM and annuity prices in the next section.

2 Pricing with adverse selection and interest rate risk

In this section, we present an annuity pricing model with adverse selection and interest rate risk. We introduce two frictions, which are absent from the adverse selection literature. The first friction lies in the bond market. Due to market incompleteness, insurers are limited to investing in fixed rate—i.e., noncontingent—corporate bonds of varying maturities to hedge interest rate risk. However, corporate bond issuers are financially constrained, resulting in an inefficiently low supply of long-term corporate bonds. 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 operate under limited liability: The owners of a life insurer are not liable for losses beyond the value of their assets. Together, these financial frictions mean that insurers may fail to honor annuity payments if they become insolvent along certain interest rate paths.

The model offers three main insights. First, it characterizes the effect of long-term corporate bond returns on the tradeoff between hedging interest rate risk with capital structure (i.e., net worth) and hedging such risk with the asset portfolio (i.e., long-term bond holdings). Second, the model provides a link between life insurers' unobservable capital structure and their observable annuity pricing decisions. Third, the model endogenizes the assets and liabilities sides of insurers' balance sheets, highlighting the challenges in identifying the sources of inefficiencies within the annuity market. Proofs for this section's results can be found in Appendix D.

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

2.1 Economic environment

We consider a three-period model with heterogeneous retirees, competitive life insurers offering life annuities, and a representative nonfinancial firm issuing corporate bonds. 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 probability α, which is drawn at the beginning of $t = 0$ from cumulative distribution function $G(\alpha)$ with support $[\underline{\alpha}, \overline{\alpha}] \subset [0, 1]$ and probability density function $g(\alpha)$. The survival probability α is an individual's private information. Every individual is deceased at the end of $t = 2$.

Life insurers offer life annuities to individuals at $t = 0$. An annuity contract provides one unit of consumption in each period the contract holder is alive in exchange for a single investment q in $t = 0$.[14] With this normalization, the annuity premium q can be interpreted as the annuity price. The market is competitive, and insurers compete on prices.[15]

Nonfinancial firms issue one- and two-period zero-coupon corporate bonds. The one-period corporate bond return follows an exogenous process: One unit of the one-period bond returns $R_1 \geq 1$ in $t = 1$ and R_2 in $t = 2$, where $R_2 \in \left[1, \overline{R}\right]$ is a shock realized in $t = 1$. We refer to the two-period corporate bond as the long-term bond. The long-term bond return is determined endogenously and denoted by R_l. Closely following the modeling approach of Greenwood et al. (2010), we constrain the supply of long-term bonds because nonfinancial firms face refinancing risk. These firms target a long-term bond supply of T to hedge against rising interest rates. In Appendix E, we show that if T is at a specific threshold value, the bond market is *unconstrained*, and $\frac{1}{R_l} = \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$. That is, corporate bond pricing is consistent with the expectations hypothesis. Note that the long-term corporate bond is not risk free because of the interest rate shock R_2. Conversely, if T is below the threshold value, then the bond market is *constrained*, and $\frac{1}{R_l} > \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$. In this case, the long-term bond return is inefficiently low. Even in a

[14] We assume that insurers offer a single contract to individuals, consistent with actual practices of offering the same SPIA contract based on age and gender without further screening.

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

constrained bond market, we focus on the empirically relevant case where T is such that $R_l \geq R_1$.[16] For the rest of our analysis, we focus on the constrained bond market case.

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 the following: (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: Individuals with higher survival risk invest more of their wealth in annuities. The first condition also requires annuity demand to decrease with annuity price. The second condition requires strictly positive annuity demand even when insurers break even on contracts offered to individuals with the highest survival probability and under the lowest interest rate realization ($R_2 = 1$). This requirement ensures the existence of a market for annuities and a well-defined equilibrium price. The third condition requires annuity demand to be zero for insurers at risk of insolvency. This assumption allows a sharper analysis of interest rate risk on annuity markups. A stark interpretation of the third condition is that annuity contracts are worthless if there is a positive probability the insurer may not fulfill its obligations, as individuals may be unwilling to trade off longevity risk for insurer default risk. Alternatively, this assumption may reflect unmodeled regulatory or credit rating agency requirements for insurers to hold a minimum level of capital to prevent insolvency. Finally, note that $a(\alpha, q)$ represents annuity demand at $t = 0$, so it implicitly depends on the returns of other assets at that time, $\{R_l, R_1, \mathbb{E}(R_2)\}$, but not on the realized value of R_2.[17]

We now turn to life insurers' decisions and introduce their investments and balance sheet equations. We normalize life insurers' preexisting balance sheet (assets and total

[16] If $R_l < R_1$, then the long-term bond is too expensive, and it is more economical for insurers to manage interest rate risk with only one-period bonds.

[17] Specifically, our reduced-form demand assumption implies a negative relationship between interest rates, R_1 and $\mathbb{E}(R_2)$, and annuity sales. That is, individuals buy fewer annuities when interest rates (or spreads) are low, which is a movement along the demand curve rather than a shift in the demand curve.

liabilities) to zero in $t = 0$ and focus our analysis on the contribution of new annuity business to their balance sheet.[18] In $t = 0$, an insurer invests its total annuity premiums, which are the total lump sums received from individuals signing up for an annuity, in a portfolio of corporate bonds $(b_1, l_2) : b_1 + l_2 = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q) g(\alpha) d\alpha$, where b_t and l_2 denote investments in one-period and long-term bonds, respectively. The insurer's balance sheet at $t = 0$ is given by

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

where the first term on the right-hand side of equation (1) is the present value of the insurer's annuity liability and NW_t is the insurer's net worth in t. Note that there is no residual risk in insurers after they have implemented the optimal interest rate risk hedging strategy in our simple model. This feature of the model means that investors in insurers do not demand additional compensation to bear the insurer's residual risk.[19] Therefore, insurers' cost of capital (i.e., its discount rate) is aligned with the one-period bond rate.[20] Evidently, real-world insurers do have residual risk since they are rated around A, which is reflected in their cost of capital. A formal treatment of residual risk would considerably increase the complexity of the model without changing its main insights. Nevertheless, our empirical analysis is fully consistent with the presence of insurers' residual risk, and we discuss how we bring the model to the data in Section 3.

After the one-period bond return R_2 is realized at the beginning of $t = 1$, the insurer's balance sheet becomes $b_2(R_2) = \frac{1}{R_2} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q) g(\alpha) d\alpha + NW_1(R_2)$. This means that an insurer finances the purchase of one-period bonds b_2 using the proceeds from its initial bond holdings (b_1, l_2) net of the annuity payments made to the surviving individuals in $t = 1$: $b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q) g(\alpha) d\alpha$.

These equations show that an insurer is at risk of becoming insolvent in $t = 1$ if the present value of its liabilities exceeds the present value of its assets for certain R_2 realizations. Due to limited liability, insolvent insurers are not liable for losses exceeding

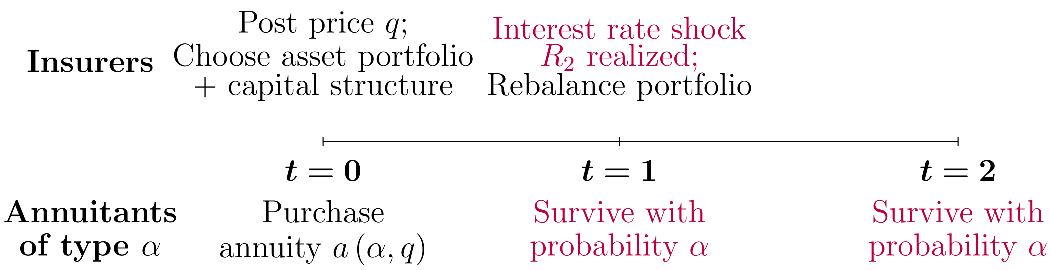
[18]This normalization is valid with perfect price competition. As will be clear in Section 2.3, a life insurer with a higher level of net worth at $t = 0$ could, in principle, cut its annuity price. However, financing the greater $t = 0$ net worth necessarily implies posting uncompetitive prices in previous periods, which is not an equilibrium. That said, we investigate this issue by exploiting insurer hedging strategy heterogeneity in Section 7.

[19]Investors in insurers are either the policyholders, in the case of a mutual insurance company, or shareholders, in the case of a stock insurance company.

[20]See, for example, Allen (1993).

the value of their assets and, therefore, can default on some or all of their promised annuity payments. Under Assumption 1, individuals avoid purchasing annuities from insurers with a nonzero insolvency probability. Therefore, limited liability, together with our assumption on annuity demand, means that life insurers have an incentive to effectively manage interest rate risk to remain solvent. Figure 2 summarizes the timing of the model.

Figure 2: Model timeline. The top and bottom sections outline the sequence of events for insurers and annuitants, respectively, with black text for decisions and red text for random shocks.



2.2 Optimal interest rate risk management strategy

The optimal IRM strategy ensures the present values of an insurer's assets and total liabilities—i.e., annuity liabilities plus net worth—change at the same rate after the interest rate shock R_2 is realized. Theorem 1 characterizes the unique optimal IRM strategy for a given annuity price q, which we examine in the next subsection, and long-term bond return R_l.[21]

Theorem 1 *The unique optimal IRM strategy requires a higher level of net worth when the bond market is constrained. Specifically, for a given annuity price q and long-term bond return R_l, the unique optimal IRM strategy requires an asset allocation and a capital structure satisfying the following: (i) Asset portfolio:* $l_2 = \frac{1}{R_l}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha$, $b_1 = \frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha$, *and* $b_2(R_2) = \frac{1}{R_2}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha$, *and (ii) Capital structure:*

$$NW_0 = \begin{cases} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[\frac{1}{R_l} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha & \text{if } \frac{1}{R_l} > \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right) \\ 0 & \text{if } \frac{1}{R_l} = \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right) \end{cases},$$

and $NW_1(R_2) = 0$ *for all* R_2.

[21] The optimal IRM strategy is unique in a perfectly competitive annuity market. In Appendix F.2, we show that insurers with market power also manage interest rate risk with an asset portfolio and a capital structure satisfying Theorem 1 at a minimum.

The main insight of Theorem 1 is that the long-term bond return determines insurers' relative cost of hedging interest rate risk with net worth. Insurers can close the duration gap between their assets and annuity liabilities by (1) increasing their net worth, as net worth has a lower duration than annuity liabilities; (2) increasing their long-term bond investment; or (3) some combination of the two. Theorem 1 shows that insurers prefer hedging the interest rate risk by investing premiums in long-term bonds. To understand this result, note that when the long-term bond return is high relative to the one-period bond return—i.e., the bond market is unconstrained—insurers manage interest rate risk by investing in an optimal portfolio of one-period and long-term bonds, without relying on net worth. This optimal bond portfolio replicates the features of state-contingent bonds in an incomplete market.[22] Note further that, contrary to the law of demand, insurers' bond demand is inversely related to bond returns, as hedging the annuity duration requires more bonds when bond returns are low. Therefore, when the long-term bond return is sufficiently low relative to the one-period bond return—i.e., the bond market is constrained—it becomes cost effective to hedge interest rate risk with relatively more net worth. However, the substitution is imperfect, and insurers never exclusively hedge interest rate risk with net worth when long-term bonds are available.[23]

2.3 Equilibrium annuity pricing

The equilibrium annuity price is determined by Bertrand competition. Competitive life insurers implementing the optimal IRM strategy of Theorem 1 set their annuity price so that total annuity premiums in $t = 0$ are equal to the optimal total bond demand $b_1 + l_2$, ensuring zero profits. Theoretical details of annuity pricing can be found in Appendix D.

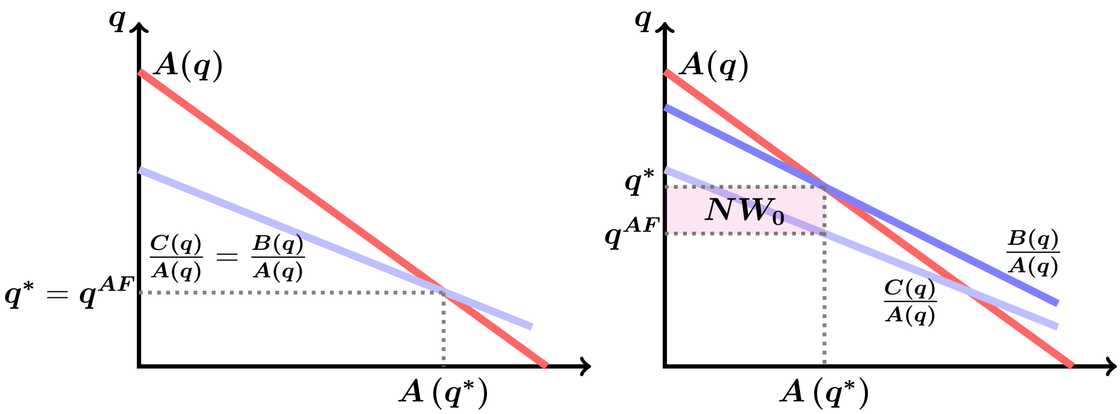
Figure 3 provides a graphical illustration of equilibrium annuity pricing with adverse selection and interest rate risk. We drew Figure 3 in a way that emphasizes the differences between our approach and the textbook model of adverse selection (Einav and Finkelstein

[22] This result echoes the analysis of Angeletos (2002) showing that a rich enough non-state contingent bond maturity structure can replicate the allocation obtained with a full set of state contingent securities if certain conditions are met.

[23] In Appendix F.1, we extend the model to include long-term zero-coupon "government bonds." The return on the long-term government bond is lower than the return on the long-term corporate bond because of a liquidity premium or convenience yield. We show that Theorem 1 also applies in this case and insurers do not use long-term government bonds for IRM, which is consistent with the low share of government securities in life insurers' asset portfolios, mentioned in Section 1.2.

2011). The vertical axis represents annuity price q and the horizontal axis represents the quantity of annuities sold. The competitive equilibrium price q^*, is determined by the intersection of the downward-sloping aggregate demand curve $A(q)$ and insurers' average bond demand curve $B(q)/A(q)$, which represents the amount of bonds of all maturities insurers demand per unit of annuity sold.

Figure 3: Equilibrium annuity price in two different bond markets. Panels (a) and (b) depict the equilibrium annuity price in unconstrained and constrained bond markets, respectively, as functions of quantity.



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

When the bond market is unconstrained, competitive insurers hedge interest rate risk by investing in an optimal corporate bond portfolio. Equation (1) and Figure 3a show that the average bond demand curve $B(q)/A(q)$ is equal to the average cost curve $C(q)/A(q)$. The average cost curve represents the cost per unit of annuity sold and is downward sloping due to adverse selection. Therefore, the equilibrium annuity price q^* is equal to the *risk-adjusted actuarially fair price* q^{AF} when the bond market is unconstrained, which corresponds to the equilibrium price in the textbook model of Einav and Finkelstein (2011).[24]

When the bond market is constrained, Theorem 1 shows that insurers increase their net worth NW_0 to hedge the interest rate risk. Equation (1) and Figure 3b show that, in this case, the average bond demand curve $B(q)/A(q)$ exceeds the average cost curve $C(q)/A(q)$ by the average net worth $NW_0/A(q)$ at the equilibrium price q^*. This set-up

[24] Unlike in Einav and Finkelstein (2011), an insurer's asset portfolio and its capital structure matter in our model, even when the bond market is unconstrained. 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 and 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.

means that insurers finance their net worth by charging a strictly positive AS-adjusted markup ($q^* - q^{AF} > 0$). The following equilibrium relationship links optimal net worth, which is generally unobservable, to annuity price markups, which are observable at a high frequency:

$$NW_0 = \left(q^* - q^{AF}\right) \int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha. \tag{2}$$

This 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 3b.

To understand the channel though which IRM affects annuity markups, we examine how the equilibrium annuity price changes in response to a marginal change in the long-term bond return. Let $\psi = \frac{1}{R_l} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, so the bond market is more constrained when ψ increases. To proceed, we make an additional assumption about annuity demand.

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

In addition to requiring that individuals with high α—i.e., high longevity risk—buy more annuities (Assumption 1), Assumption 2 requires that these individuals are not overly sensitive to price. Theorem 2 demonstrates that when Assumption 2 holds, the AS-adjusted markup is higher when the bond market is more constrained. Importantly, it implies that the average optimal net worth $NW_0\left(q\right)/A\left(q\right)$ increases with q. Thus, if insurers need higher average net worth for IRM, they must finance it by raising the annuity price.

Theorem 2 *The AS-adjusted markup $q^* - q^{AF}$ is higher when the bond market is more constrained: $\frac{\partial q^*}{\partial \psi} - \frac{\partial q^{AF}}{\partial \psi} > 0$. Furthermore, when the bond market is unconstrained, the AS-adjusted markup is zero: $q^* = q^{AF}$.*

In general, our model shows that the effect of frictions in capital markets and in annuity markets is not orthogonal, because the average survival probability of the annuitant pool increases when insurers charge higher annuity prices to hedge the interest rate risk. We postpone a more in-depth analysis of this interaction to Section 6. The model underscores the challenges in identifying the sources of market inefficiencies influencing annuity prices, as net worth is jointly determined by the average cost curve $C\left(q\right)/A\left(q\right)$ and average bond demand curve $B\left(q\right)/A\left(q\right)$. In the next section, we discuss how we measure annuity markups in the data and how we can identify the IRM channel using exogenous shifters of the average cost curve $C(q)/A(q)$ and average bond demand curve $B(q)/A(q)$.

3 Identification of the IRM channel

Testing for the interest rate risk channel requires identifying the effect of changes in long-term bond market conditions on 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.[25] 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 nor the allocation of net worth to support the issuance of different types of insurance liabilities. Second, different types of supply-side frictions may lead to observationally equivalent annuity markups.[26] Third, demand- and supply-side frictions are likely to have nontrivial interactions. In Section 6, we show theoretically and empirically that adverse selection in annuity markets depends on the severity of frictions in the corporate bond market, since a higher level of net worth exacerbates adverse selection by increasing annuity prices.

We overcome this identification challenge by exploiting long-term corporate bond market shocks that differentially affect the relative cost of hedging interest rate risk with net worth of *different annuity contracts* offered by the *same insurer*. Specifically, we identify the IRM channel by comparing the changes in AS-adjusted markup, $q^* - q^{AF}$, for annuity contracts offered by the same insurer resulting from an exogenous increase in contract-level relative cost when long-term corporate bond market conditions are favorable with the changes when long-term corporate bond market conditions are relatively less favorable. In the remainder of this section, we discuss how we measure $q^* - q^{AF}$, and the two sources of variation that exogenously shock insurers' relative cost of hedging.

[25] 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 and Wang (2019) propose an indirect measure of the duration gap based on insurers' stock prices.

[26] For example, Appendix F.2 shows that annuity markups could also be the outcome of monopolistic competition.

3.1 Life annuity price markups measurement

Life insurers reprice their annuities frequently in response to changes in market conditions—they are especially attentive to changes in the long end of the yield curve. Our model focuses on the marginal pricing decision of a life insurer since we normalize the $t = 0$ balance sheet to zero. Using this interpretation, a life insurer creates a new block of business at date t, which is added to its existing block of insurance liabilities. Therefore, the first step in our identification strategy is to evaluate insurers' marginal pricing decisions conditional on bond market conditions.

Two inputs are needed to price new insurance liabilities. We discuss our choices for these two inputs in detail later, as they have important implications for annuity valuation (Poterba and 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 investor in this insurer is willing to commit capital to support the issuance of *illiquid* long-term fixed-rate liabilities backed by illiquid assets.[27] Almost all life insurers offering annuities in the U.S. have an Insurance Financial Strength (IFS) rating around A.[28] This rating means that life insurers are not riskless firms and investors demand additional compensation to bear residual risk. The compensation for

[27]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 owner and that of 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.

[28]Most U.S. life insurers have an IFS rating from a credit rating agency. For example, Moody's states that "Moody's Insurance Financial Strength Ratings are opinions of the ability of insurance companies to repay punctually senior policyholder claims and obligations." Life insurers may also have long-term senior and junior unsecured credit ratings that determine their cost of issuing senior and junior debt, respectively. Moody's uses a comparable scale for all its long-term ratings: "Moody's rating symbols for Insurance Financial Strength Ratings are identical to those used to indicate the credit quality of long-term obligations. These rating gradations provide investors with a system for measuring an insurance company's ability to meet its senior policyholder claims and obligations." Because insurance obligations in the general account of a life insurer are senior to all debt claims from that insurer, an insurer IFS is typically higher than its long-term senior debt rating (if it has one).

bearing the insurer's residual risk is reflected in the insurer's cost of capital—i.e., its discount rate. 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.[29] The HQM yield curve is calculated daily using AAA-, AA-, and A-rated U.S. corporate bonds and is heavily weighted toward A-rated bonds, consistent with their large market share.[30] Consistent with our choice of discount rate, Huber (2022) finds that 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 depend only on the date of birth and gender of the individual. We use three different types of mortality assumptions. 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.[31]

We also use two different versions of the Individual Annuitant Mortality (IAM) table produced by the Society of Actuaries (SOA) in collaboration with the NAIC. The first version of the IAM table is the "basic" one, which is estimated by the SOA from the actual mortality experience of a large pool of annuitants from multiple insurers over a long period. Before 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. In 2015, the industry transitioned to a generational (dynamic) mortality table, and we carefully analyzed individual state legislation to follow its staggered implementation during 2015–16. In addition to the basic IAM table, the SOA produces a "loaded" IAM table, which adds a loading factor to the basic IAM

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

[30] 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.

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

table mortality estimate. The loaded table is used by state insurance regulators to set regulatory reserves. See Appendix G 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.[32] 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.[33]

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

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

where $M \geq 0$ is the number of years the life annuity pays a guaranteed fixed income, $p_{S,n}^k$ is the one-year survival probability for an individual of gender S at 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 the period m cash flow evaluated at time t using the HQM yield curve (r = HQM) or the regulatory reference rate (r = NAIC), which we will explain later.

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

[32] 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.

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

average measure of adverse selection pricing (AS pricing):

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

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 Contract-level variation in 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 and Yogo (2015), exogenous time-series variation in reserve requirements *across* contract maturities 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 insurers' average cost curve (Koijen and Yogo 2015).

Before 2018, state insurance regulations required that insurers calculate their annuity reserves—i.e., the value of 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 Moody's Investors Service, Inc."[34] 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. Starting in 2018, state insurance regulators adopted a new but related

[34] https://content.naic.org/sites/default/files/inline-files/committees_ex_pbr_implementation_tf_related_820.doc

methodology, which we discuss in Appendix G.

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 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 maturities—i.e., those sold to older individuals. Moreover, the difference between regulatory reserves and insurer reserves fluctuates exogenously over time across annuity contracts with different maturities. The reason is that the regulatory interest rate resets infrequently—once a year before 2018 and once a quarter from that year onward—whereas the yield curve used by insurers to price their insurance liabilities changes daily.

Figure 4: Exogenous variation in annuity contract–level regulatory reserve ratio. This figure plots the regulatory reserve ratio for an SPIA sold to a 65-year-old male (top line) and 70-year-old male (bottom line). The shaded area between the two lines varies exogenously because the regulatory discount rate used to calculate the regulatory reserve ratio is the same for all SPIA maturities and resets infrequently. The partially overlapping green and blue areas represent the regulatory ratio for states that adopted the 2012 mortality table in 2015 and 2016, respectively.

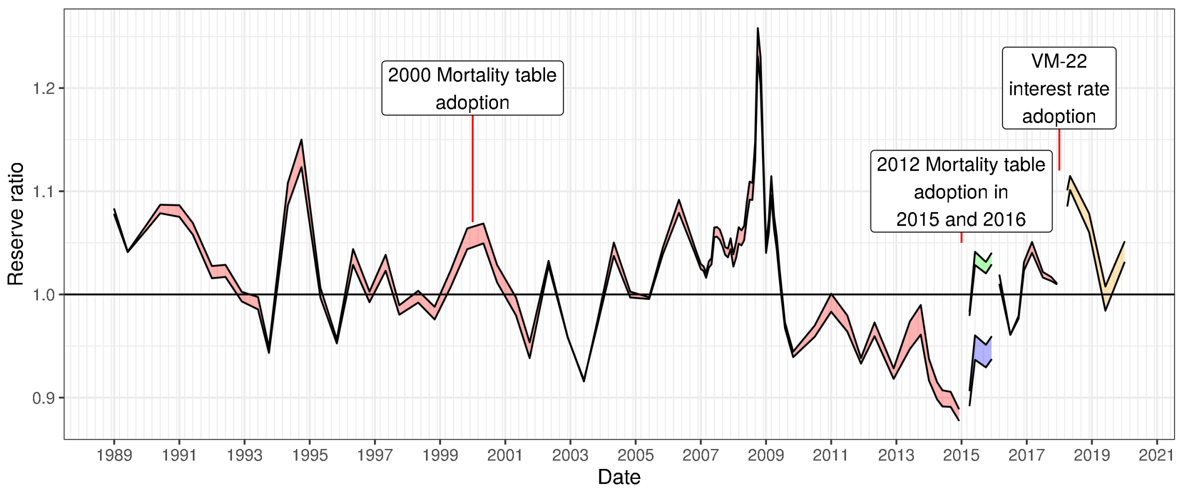


Figure 4 illustrates this source of exogenous variation by plotting the reserve dollars an insurer needs to set aside for each dollar of annuity sold on day t to 65- and 70-year-old males only—in our empirical analysis, we use the full set of price quotes for male and female individuals aged between 50 and 90 years with 5-year intervals. We denote the regulatory interest rate by $r =$ 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 must create a reserve that is greater than the insurance liability warranted by the insurer's yield curve–based actuarial calculation. Conversely, a ratio below 1 indicates that the reserve requirement is *nonbinding* 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 insurer's discount rate are aligned on average. Notice how this distance exogenously fluctuates over time because the flat regulatory interest rate resets infrequently. Figure 4 also depicts additional sources of variation arising from not only U.S. states' staggered adoption of new regulatory mortality assumptions between 2015 and 2016, but also the 2018 adoption of the new methodology to calculate the regulatory reference interest rate. These developments are represented by the partially overlapping green and blue shaded areas, which indicate the variation in relative reserve requirements for early and late adoption states, respectively.

3.3 Variation in long-term investment-grade bond yield spreads

The second source of exogenous variation comes from the variation in long-term investment-grade bond yield spreads relative to insurers' cost of funding. Life insurers' solvency depends on their net worth, which is a stock variable. Life insurers' profitability derives from the spread between the returns they earn on their asset portfolio and the returns they owe their policyholders, which is known as the net investment spread and is a flow variable. The two concepts are intimately related because life insurers finance net worth with retained earnings, which depend on their net investment spread.

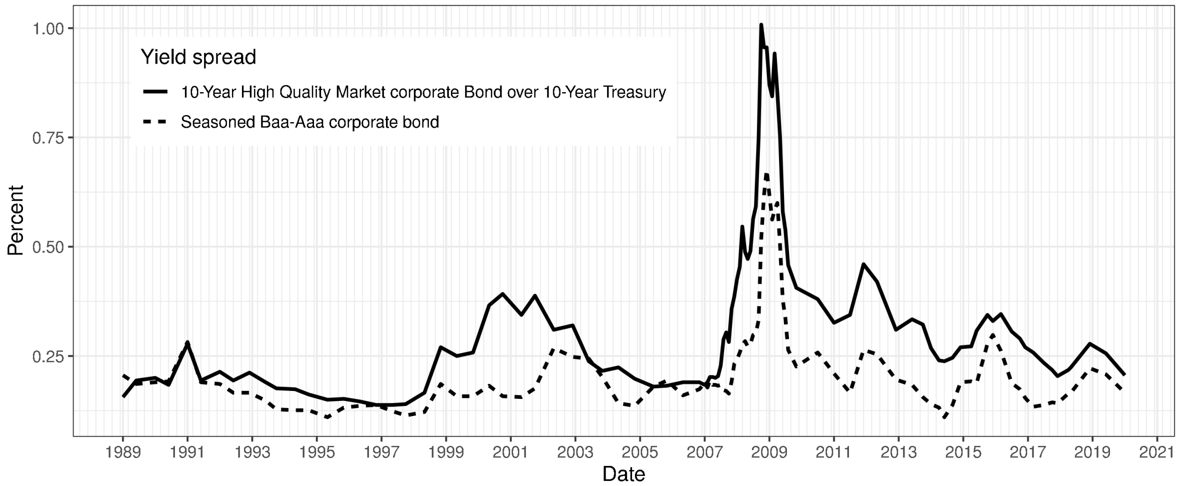
When life insurers cannot perfectly match their assets' and insurance liabilities' durations, changes in interest rates affect insurers' net worth and profitability. For example, an unexpected decline in interest rates depresses insurers' net worth by increasing the present value of insurance liabilities faster than the present value of the assets backing the insurance liabilities. If the decline in interest rates persists or interest rates remain low, life insurers' net worth deteriorates further, as low interest rates are typically associated with lower coupon rates. Therefore, persistently low interest rates force insurers to reinvest the proceeds from their maturing bonds into new bonds paying lower coupon rates, which further depresses their overall net investment spread.

Our empirical analysis of annuity markups focuses on new business pricing—that is, we focus on flow variables. The equilibrium relationship between marginal net worth and the AS-adjusted markup is summarized by equation (2) in our model (Section 2). From Theorem 1, the long-term bond return determines insurers' relative cost of hedging interest rate risk with net worth. However, this relative cost is irrelevant for insurers if they can perfectly match the duration of their annuity liabilities with bonds. Therefore, changes in long-term bond returns have an effect on the AS-adjusted markup only when the bond market is constrained. In this case, changes in long-term bond returns affect the relative cost of hedging interest rate risk with net worth, which is reflected in markups.

We measure shocks to long-term corporate bond returns using exogenous variation in the spread between the yield on Moody's Baa-rated and Moody's Aaa-rated corporate bonds with at least 20 years of maturity. Note that this is the yield spread on long-duration investment grade bonds, not the yield on the bonds themselves. Note further that the yield on Aaa-rated corporate bonds with at least 20 years of maturity is a quasi-risk-free benchmark for this spread. Under state insurance regulations, corporate bonds rated above Moody's Baa are designated as NAIC 1 and uniformly attract the lowest statutory risk-based capital charge in our sample period. Therefore, life insurers benefit from widening long-term investment-grade bond spreads *relative* to their cost of funding when the supply of long-term bonds is inefficient because it corresponds to higher coupon rates for a given bond rating, maturity, and risk-based capital charge. Therefore, changes in long-term investment-grade bond spreads relative to insurers' cost of funding act as exogenous shifters to the relative cost of hedging interest rate risk with the long-term bonds.

Figure 5 illustrates this exogenous variation by plotting the Baa–Aaa spread for seasoned corporate bonds and the spread between the 10-year HQM yield and the 10-year U.S. Treasury yield in percentage points. The 10-year HQM yield spread is our proxy for insurers' average cost of funding, because life annuities offered to a 65-year-old have an initial duration of about 10 years and we use the HQM yield curve as 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, which decreases the relative cost of hedging interest rate risk with long-term bonds.

Figure 5: Long-duration investment-grade bond spreads and insurers' cost of funding. The dashed line represents 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. The solid line is a proxy for insurers' average cost of funding, calculated as the spread between the 10-year HQM yield and the 10-year U.S. Treasury spread in percentage points. Life insurers can generate more yield per dollar of annuity sold when the long-duration Baa–Aaa spread increases *more than* the 10-year HQM yield spread.



3.4 Empirical test of the IRM channel

We now discuss our empirical test of the IRM channel. Under the null hypothesis of costless IRM, the effect of an increase in the reserve requirement is not affected by changes in the long-term investment-grade bond spread. Under the alternative hypothesis of costly IRM, the effect of an increase in the relative reserve requirement on markups is offset by an increase in the long-duration bond yield spread, which is unique to the IRM channel. Therefore, 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.

The empirical 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 long-term investment-grade bond spread is the "treatment" that differentially affects annuity contracts from the same insurer with exogenously varying contract-level relative 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 long-term investment-grade bond spread is high and periods in which it is low.

Lastly, note that the variation in the long-term investment-grade bond 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 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 long-term bond market conditions on the AS-adjusted markup is similar across the different annuity contracts. Under this assumption, the double difference nets out the potential effects of changes in annuity demand associated with changes in corporate bond market conditions. While this identifying assumption is not directly testable, we provide supportive evidence that it is valid using term annuity markups that do not depend on an individual's age or gender.

4 Data and variable definitions

We focus our analysis on SPIAs without term certain guarantees and with 10- and 20-year term certain guarantees. SPIAs with term certain guarantees ensure a payment to a beneficiary during the term period, regardless 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*.[35] Table 1 reports the summary statistics for the variables used in our analysis.

Our main dependent variable $Annuity\_markup_{ijt}$ is the normalized AS-adjusted markup for product j sold by insurer i at date t. It is 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, on average, just under 16 percent and consistently above 10 percent during our sample period. The variable $Reserve\_Ratio_{jt}$ is the ratio of reserve

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

Table 1: Summary statistics

Variables	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

dollars insurers need to set aside for each dollar of annuity j sold on day t. It is 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 2001–19 from S&P Global Market Intelligence. We measure insurer size as the log of insurers' general account assets and leverage as the ratio between 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 rate minus 3-month Treasury constant-maturity rate from the St. Louis Fed's FRED database. We proxy for insurers' cost of funding by calculating the spread between the 10-year HQM yield, and the 10-year Treasury constant-maturity yield. 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 discuss the $Net\_swap\_duration_{it}$ variable in Section 7.

5 Main empirical analysis and results

Regression (3) below implements our empirical test of the IRM channel in a linear regression framework:

$$
\begin{aligned}
Annuity\_markup_{ijt} = & \beta_1 Baa\text{-}Aaa\_spread_t + \beta_2 Reserve\_Ratio_{jt} \\
& + \beta_3 Reserve\_Ratio_{jt} \times Baa\text{-}Aaa\_spread_t \\
& + \beta_4 10\_HQM\_spread_t + \beta_5 10\_HQM\_spread_t \times Baa\text{-}Aaa\_spread_t \\
& + \mathbf{z}'_{it}\boldsymbol{\gamma_1} + Baa\text{-}Aaa\_spread_t \times \mathbf{z}'_{it}\boldsymbol{\gamma_2} \\
& + \alpha_1^i + \alpha_2^j + \epsilon_{ijt}.
\end{aligned} \tag{3}
$$

The unit of observation is an individual annuity contract j offered by insurer i at date t. The sample of observations extends from 1989 to 2019. We focus on within-insurer variation, and regression (3) includes insurer fixed effects 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 behavior.

The first difference in regression (3) is between annuity contracts j offered by insurer i with *relatively* high reserve requirements, measured by the variable $Reserve\_Ratio_{jt}$, and annuity contracts $-j$ with relatively low reserve requirements. The contract-level reserve requirement shocks create a within-insurer random assignment of annuity contract relative cost that varies over time. As we explained in Section 3, this variable helps identify the effect of supply-side frictions by exogenously shifting the annuity supply cost while holding annuity demand fixed (Koijen and Yogo 2015). Regression (3) also includes a contract age-gender-guarantee fixed effect α_2^j to absorb the effect of unobserved fixed demand characteristics that may influence annuity pricing. The second difference in regression (3) is between periods with relatively high and low $Baa\text{-}Aaa\_spread_t$, respectively. The variable $Baa\text{-}Aaa\_spread_t$ is an aggregate shock that differently affects annuity contracts with exogenously varying relative reserve requirements. As we explained in Section 3, this variable allows us to identify the type of supply-side friction among plausible alternatives by studying its effect as a function of changes in $Reserve\_Ratio_{jt}$.

Under the null hypothesis of costless IRM, the effect of an increase in the relative reserve requirement is not a function of changes in bond market conditions. Under the alternative hypothesis, the increase in AS-adjusted markup is lower when the long-duration investment-grade bond spread is higher, which is unique to the IRM channel. The coefficient β_3 is the difference-in-differences estimator that allows us to test the null hypothesis. It is negative under the alternative hypothesis because the cost of IRM is lower when $Baa\text{-}Aaa\_spread_t$ is relatively high.

The main control variable in regression (3) is the average insurer's cost of funding that we proxy with the variable $10\_HQM\_spread_t$. Regression (3) allows for an interaction between $10\_HQM\_spread_t$ and $Baa\text{-}Aaa\_spread_t$, because the two variables are correlated and a change in insurer funding conditions could potentially differently affect markups in times of high and low $Baa\text{-}Aaa\_spread_t$. The vector $\boldsymbol{z}'_{it}$ contains quarterly insurers' general account assets (logged) and leverage, defined as general account assets to policyholders' surplus, which are available only in a subset of our sample starting at the beginning of 2000. We use these additional controls in our robustness tests and also allow for an interaction with $Baa\text{-}Aaa\_spread_t$.

To estimate the effect of IRM on the annuity markup, we analyze the effect of an increase in the contract-level reserve requirement on annuities' AS-adjusted markup, *conditional* on the cost of funding. This effect is captured by the term $\beta_2 + \beta_3 Baa\text{-}Aaa\_spread_t$, and its estimated value should be evaluated at different points within the $Baa\text{-}Aaa\_spread_t$ distribution. According to our model, the effect of an exogenous increase in the relative reserve requirement is smaller during periods with favorable bond market conditions.

5.1 Baseline results

Table 2 summarizes our main results. Column 1 reports our baseline results using the full sample and insurer-clustered robust standard errors throughout. Using the coefficient estimates in Column 1 to evaluate the estimated value of $\beta_2 + \beta_3 Baa\text{-}Aaa\_spread_t$, we find 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 almost 1 full percentage point (0.87) when $Baa\text{-}Aaa\_spread_t$ is at its median level (0.9). Importantly, we estimate that the effect of a one standard deviation increase in $Reserve\_Ratio_{jt}$ on the AS-adjusted markup is about 43 percent lower in periods when $Baa\text{-}Aaa\_spread_t$ is in the third quartile of its distribution relative to periods when $Baa\text{-}Aaa\_spread_t$ is in the first quartile of its distribution.[36]

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, while 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)	(4)	(5)
$Reserve\_Ratio_{jt}$	41.98***	41.98***	33.76***	41.01***	32.70***
	(5.98)	(9.95)	(10.85)	(9.87)	(10.81)
$Reserve\_Ratio_{jt} \times Baa\text{-}Aaa\_spread_t$	-29.18***	-29.18***	-22.53**	-28.90***	-22.05**
	(5.23)	(9.75)	(10.95)	(9.71)	(10.92)
$Baa\text{-}Aaa\_spread_t$	25.61***	25.61***	17.83*	25.29***	17.40*
	(5.07)	(8.71)	(9.74)	(8.64)	(9.68)
$10\_HQM\_spread_t$	1.64***	1.64*	2.06*	1.50*	1.98*
	(0.52)	(0.85)	(1.06)	(0.87)	(1.06)
$10\_HQM\_spread_t \times Baa\text{-}Aaa\_spread_t$	1.09***	1.09*	0.98	1.13**	0.98
	(0.22)	(0.55)	(0.66)	(0.55)	(0.66)
$Log\_total\_assets_{it}$			-0.21		-0.18
			(0.69)		(0.67)
$Log\_total\_assets_{it} \times Baa\text{-}Aaa\_spread_t$			0.50*		0.49*
			(0.25)		(0.25)
$Leverage\_ratio_{it}$			0.02		0.01
			(0.06)		(0.05)
$Leverage\_ratio_{it} \times Baa\text{-}Aaa\_spread_t$			-0.04		-0.04
			(0.05)		(0.05)
Fixed effects:					
Contract characteristics (j)	Y	Y	Y	N	N
Insurer (i)	Y	Y	Y	N	N
Insurer (i) × Contract characteristics (j)	N	N	N	Y	Y
SE Clustering	Insurer	Insurer/Date	Insurer/Date	Insurer/Date	Insurer/Date
Observations	40,790	40,790	29,462	40,790	29,462
Adjusted R^2	0.54	0.54	0.57	0.62	0.64

This baseline 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.[37] This result is consistent with the IRM strategy of life insurers in a constrained bond market. The result is also consistent with recent work by Ozdagli and Wang (2019), who find that when interest rates decline, life insurers rebalance their portfolios toward higher-yielding bonds by increasing the duration, rather than the credit risk, of their portfolios.[38]

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

[37]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, at the height of the European debt crisis in 2012–13, and around the 2014–16 oil shock.

[38]Ozdagli and Wang (2019) do not analyze the effects of IRM on life insurers' product pricing. Rather,

5.2 Robustness tests

The rest of Table 2 investigates the robustness of the baseline results to potential threats to identification. In Column 2, we investigate 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 almost twice as large as those reported in Column 1. Nevertheless, our difference-in-differences coefficient estimate remains significant at below the 1 percent significance level.[39]

In Column 3, we investigate the robustness of our baseline results to potential time-varying heterogeneity in insurer characteristics. We estimate regression (3) on a subset of the sample starting in 2000, as we have access to additional time-varying insurer-level financial controls that are available. We focus on two insurer controls: 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 liabilities minus statutory surplus; statutory surplus is correlated to our definition of net worth in the model in Section 2. Although we observe these financial variables only from 2001 onward, the coefficient estimates in Column 3 obtained on this reduced sample are very similar to those obtained on the full sample in Column 1.[40]

In Columns 4 and 5, we investigate the robustness of our baseline results to potentially unobserved insurer-contract heterogeneity. This situation could happen if, for example, the variation in insurers' business mix affects the pricing sensitivity of different annuity contracts to investment-grade spread and reserve requirement shock when pricing their life annuities, or if customers have preferences for certain insurers' offerings beyond contract characteristics and pricing. We control for heterogeneity at the insurer-contract level

the authors focus on the effect of changes in an indirect measure of life insurers' duration gap on life insurers' bond holdings.

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

[40]In Appendix J, we show that our baseline results are not driven by the variations in the 2007–09 period by estimating regression (3) in the pre- and post-crisis periods.

by adding a contract × insurer fixed effect to regression (3). This specification requires estimating a large number of fixed-effect parameters relative to our baseline specification. Nevertheless, we obtain nearly exactly the same results as in Column 1. The effect of a one standard deviation increase in $Reserve\_Ratio_{jt}$ on the AS-adjusted markup is about 44 percent lower in periods when $Baa\text{-}Aaa\_spread_t$ is in the third quartile of its distribution relative to periods when $Baa\text{-}Aaa\_spread_t$ is in the first quartile of its distribution.

In Appendix I, we investigate yet another potential threat to identification arising from an age-specific correlation between bond market conditions and annuity demand. That is, although our regression test can cope with correlation between (unobservable) annuity demand and bond market conditions, one of our identifying assumptions is that this correlation is constant across customer age and gender. Our estimated effect could be biased if, for example, young and old annuity shoppers' annuity demand heterogeneously changes as bond market conditions change. This identifying assumption is not directly testable, as we do not observe the contract-level quantity of annuities sold—only posted prices. Nevertheless, we can gauge the likelihood of this bias by estimating a version of regression (3) using term certain annuities markups. The pricing of term certain annuities is not age or gender dependent, as an insurer makes fixed regular payments for a fixed number of years irrespective of the contract holder's gender or survival. Therefore, it is very unlikely that there is an age-specific correlation between bond market conditions and term annuity demand. We find that an exogenous increase in the relative reserve requirement has a relatively higher effect on the markup of term annuities with shorter maturities and that this effect is smaller in periods when the $Baa\text{-}Aaa\_spread_t$ is larger for lower-duration term certain annuities. Therefore, an age-specific correlation between bond market conditions and annuity demand is unlikely to bias our baseline results.

5.3 Measuring the contribution of IRM in life annuity markups

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. As previously explained, 5-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. Moreover, 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.[41]

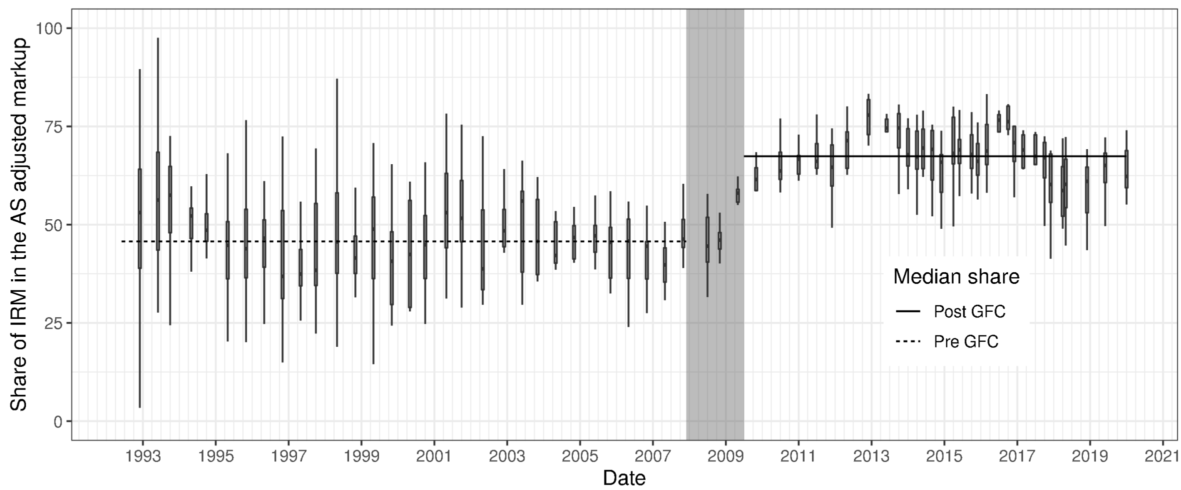


Figure 6: Contribution of IRM cost in the AS-adjusted markup for an SPIA offered to a 65-year-old male

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 actuarial value. This estimate suggests that if insurers' business expenses are indeed 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

[41]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 bank certificates of deposits, while life insurers do not face competition from banks for their life annuity offerings. That said, Figure 10 in Appendix C shows that average competition for the entire fixed annuity market is high.

after the 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 et al. 2023).

6 Feedback between IRM and adverse selection

This section extends the analysis of the model presented in Section 2 to emphasize the nontrivial interaction between shocks from the corporate bond market and adverse selection in the annuity market. Additionally, we conduct further empirical tests of this feedback mechanism.

It is well known that higher annuity prices are associated with more severe adverse selection (Rothschild and Stiglitz 1976). This effect is also present in our model. The unique feature of our model is the enforcement between adverse selection and the cost of managing interest rate risk. As the pool of annuitants deteriorates with higher annuity prices, the overall duration of the annuity liability lengthens. The latter requires insurers to raise their markups, leading to a further deterioration of the annuitant pool.

We focus on the case when the long-term bond supply is constrained ($R_1/R_l > \mathbb{E}(1/R_2)$) so that the optimal net worth at $t = 0$ is strictly positive. We let $z = R_1/R_l$, where a higher z means a more constrained bond market. Using the model, we can decompose the effect of a change in z on the equilibrium annuity price q^* into a *risk management effect* and an *adverse selection effect* by implicitly differentiating insurers' zero-profit condition (see Appendix D) as follows:

$$\frac{\partial q^*}{\partial z} = \underbrace{\frac{\frac{1}{R_1}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q^*)\, g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*)\, g(\alpha)\, d\alpha}}_{\text{Risk management effect}} + \underbrace{\frac{\partial q^*}{\partial z}\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} e(\alpha, q^*)\left[1 - \frac{\frac{\alpha}{R_1}(1+\alpha z)}{q^*}\right] a(\alpha, q^*)\, g(\alpha)\, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*)\, g(\alpha)\, d\alpha}}_{\text{Adverse selection effect}}, \quad (4)$$

where $e(\alpha, q^*) = -\frac{\partial a(\alpha,q^*)}{\partial q^*}\frac{q^*}{a(\alpha,q^*)}$ is the price elasticity of annuity demand. Since each component is normalized by the total amount of annuity supplied, $\frac{\partial q^*}{\partial z}$ represents the average marginal effect of IRM on the annuity price.

An increase in the cost of IRM leads to a higher equilibrium annuity price because insurers must use the annuity markup to finance a greater level of average net worth at $t = 0$. Given the optimal net worth in Theorem 1, the risk-management effect is the change in average net worth $NW_0(q)/A(q)$ in response to a marginal change in the long-term bond return R_l for a fixed annuity price. Intuitively, similar to Theorem 2, a more constrained bond market requires a higher average net worth for IRM, which necessitates a higher equilibrium annuity price. A rise in IRM cost can amplify the effect of adverse selection in the annuity market because the average survival probability of individuals purchasing annuities increases with annuity price. The adverse selection effect in equation (4) is primarily determined by the price elasticity of annuity demand $e(\alpha, q^*)$. When demand is more elastic for individuals with lower survival probability α, the adverse selection effect is more severe since the insurer loses more high-risk individuals than low-risk individuals when prices increase. This development worsens the adverse selection problem and potentially triggers a death spiral.[42] This theoretical result establishes a nontrivial link between the supply- and demand-side frictions, connected by the IRM channel.

6.1 Regression results

We look for evidence of the interaction between IRM and adverse selection by exploiting pricing differences in insurers' life annuities with different types of period certain guarantees. Specifically, individuals choosing a life annuity with a 10- or 20-year period certain guarantee think they are at a higher risk of dying within the next 10 or 20 years (Finkelstein and Poterba 2004, 2006). For these contracts, we follow Section 5 and measure *adverse selection pricing* as the difference between the total annuity markup and the AS-adjusted markup:

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

We then test the hypothesis that the average adverse selection pricing, $AS\_pricing_{ijt}$, of life annuity contracts with longer guarantee periods increases relatively more when regu-

[42] To see how a death spiral can occur, first note that, among annuity purchasers, 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, insurers make a profit off of mortality types $\alpha < \tilde{\alpha}$ and incur a loss from types with $\alpha > \tilde{\alpha}$. If demand is more elastic for agents with low α, then insurers lose more of these agents than high-α agents from an increase in annuity price, causing the insurer to raise prices further to compensate for the more severe adverse selection, potentially triggering a death spiral.

latory 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 without a 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:

$$\begin{aligned} AS\_pricing_{ijt} =& \beta_3 10yr\_guarantee\_period + \beta_4 20yr\_guarantee\_period \\ &+ \beta_5 10yr\_guarantee\_period \times Reserve\_Ratio_{jt} \\ &+ \beta_6 20yr\_guarantee\_period \times Reserve\_Ratio_{jt} \\ &+ \beta_7 Reserve\_Ratio_{jt} + Reserve\_Ratio_{jt} \times \mathbf{z}'_{it}\boldsymbol{\gamma_1} + \mathbf{z}'_{it}\boldsymbol{\gamma_2} \\ &+ \alpha_1^i + \alpha_2^j + \epsilon_{ijt}, \end{aligned} \tag{5}$$

where $10yr\_guarantee\_period$ and $20yr\_guarantee\_period$ are binary variables indicating the guarantee period length. As with our main specification in Section 5, we focus on within-insurer variation using insurer fixed effects. We condition our results on the vector of controls $\mathbf{z}'_{it}$, which includes $Baa\text{-}Aaa\_spread_t$ and the average cost of funding of the insurer, proxied with $10\_HQM\_spread_t$ in our baseline specification. As before, we allow an interaction between the variables in $\mathbf{z}'_{it}$ and the treatment variable, which in this case is the contract-level reserve requirement ratio, $Reserve\_Ratio_{jt}$. As a robustness test, we include insurer-level time-varying financial variables, such as insurer log asset size and leverage, in $\mathbf{z}'_{it}$. The coefficients β_5 and β_6 on the interaction terms are measured relative to the effect on life annuities without a guarantee period, which is the third type of life annuity contract in our sample and is omitted from this regression.

Table 3 summarizes the results of regression (5). We report two-way insurer- and date-clustered robust standard errors. The coefficients in Column 1 show that an exogenous increase in relative 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 one standard deviation increase in the reserve ratio *decreases* the AS pricing of life annuities without a guarantee period by 1.07 percentage point. In contrast, the AS pricing of life annuities with 10- and 20-year guarantees *increases* by 0.33 percentage point and 0.42 percentage point, respectively, in response to the same shock.

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. These results show that changes in corporate bond market conditions have a direct effect on adverse selection in annuity markets. The reason is that individuals choosing life annuities with period certain guarantees think they are at a higher risk of dying within a few years after signing up for a life annuity contract, which is reflected, at least partially, in annuity prices.

Table 3: 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-clustered robust standard errors are reported in parentheses in Columns 1 and 2, respectively. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Dependent variable	$AS\_pricing_{ijt}$	
	(1)	(2)
$Reserve\_Ratio_{jt}$	-19.26***	-21.10***
	(3.26)	(4.22)
10yr_Guarantee	-29.97***	-27.74***
	(3.43)	(3.77)
$10yr\_Guarantee \times Reserve\_Ratio_{jt}$	25.25***	23.15***
	(3.35)	(3.75)
20yr_Guarantee	-34.83***	-34.33***
	(3.69)	(4.45)
$20yr\_Guarantee \times Reserve\_Ratio_{jt}$	26.83***	26.49***
	(3.59)	(4.43)
$Baa\text{-}Aaa\_spread_t$	-4.67	-0.14
	(4.60)	(4.08)
$Baa\text{-}Aaa\_spread_t \times Reserve\_Ratio_{jt}$	5.14	0.76
	(4.56)	(4.02)
$10\_HQM\_spread_t$	1.16	-2.31
	(3.20)	(2.80)
$10\_HQM\_spread_t \times Reserve\_Ratio_{jt}$	-1.89	1.46
	(3.15)	(2.75)
$Leverage\_ratio_{it}$		0.44**
		(0.19)
$Leverage\_ratio_{it} \times Reserve\_Ratio_{jt}$		-0.43**
		(0.18)
$Log\_total\_assets_{it}$		-2.24***
		(0.74)
$Log\_total\_assets_{it} \times Reserve\_Ratio_{jt}$		2.75***
		(0.70)
Insurer FE	Y	Y
Observations	40,790	29,462
Adjusted R^2	0.70	0.68

7 Further evidence from interest rate derivatives

We emphasized the creation and preservation of net worth as life insurers' primary tools to hedge interest rate risk. That is, life insurers build net worth *physically* by investing their annuity premiums in a bond portfolio whose present value exceeds the present value of the annuity liabilities at origination. Although not as widely used in the industry, large and sophisticated life insurers also incorporate interest rate swaps into their interest rate risk hedging programs. These sophisticated life insurers can add positive duration *synthetically* to their balance sheet by entering into a long-term *fixed-for-float* interest rate swap with a counterparty—usually a commercial bank. That said, the regulatory treatment of derivatives means that this is a costly alternative.[43]

Issuing a fixed-for-float interest rate swap is economically equivalent to financing a relatively long-duration fixed-maturity bond with short-term floating-rate debt and, thus, is a form of leverage. The duration of a fixed-for-float swap contract is the difference between the (hypothetical) underlying fixed-rate instrument—e.g., a U.S. Treasury bond—and the duration of the floating-rate liability that finances the fixed-rate instrument—e.g., three-month LIBOR (prior to January 2022). The present value of a swap is zero at inception and becomes either positive (an asset) or negative (a liability) depending on the movements in the reference rates. For example, the present value of an interest rate swap with positive duration becomes positive after a decrease in the term premium and no change in the short rate. The increase in asset present value due to the change in swap valuation is mechanically matched by an increase in net worth, thereby providing downside protection against interest rate risk.

In this section, we exploit the heterogeneity in large and sophisticated life insurers' ex ante exposure to interest rate risk arising from interest rate swaps to study their reaction to interest rate shocks. We construct a proxy for the aggregate net duration added by each life insurer's interest rate swap portfolio using individual swap 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 relatively more positive net duration with swaps ex ante is relatively more hedged against a flattening yield curve that is driven by decreasing term premiums ex

[43] In Appendix G, we discuss in greater detail the mechanics of interest rate swaps in the context of life insurer interest rate risk hedging, how swaps work, and the frictions in the U.S. life insurance industry leading to their relatively low use.

post, and vice versa. Although insurers' swap position is an ex ante endogenous variable, variations in the shape of the yield curve act as an exogenous shock to insurers' net worth ex post via their effect on the swap portfolio value.

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.[44] Therefore, we can compare the AS-adjusted markups of insurers that are beneficially, or at least less adversely, affected by changes in the term premium *ex post* because of their *ex ante* hedging program with the AS-adjusted markups of insurers that are more adversely affected by the interest rate shock.

7.1 Interest rate swaps data

We use position-level interest rate swap data to calculate a novel estimate of the net duration added by swaps as a fraction of an insurer's general account assets.[45] Our position-level swap data come from Schedule DB in the NAIC statutory filing obtained from A.M. Best. Schedule DB provides detailed information on insurers' derivative contracts, including a description of each contract's terms and notional amount. We carefully parsed the text of more than 82,000 individual contract-year observations from 44 U.S. life insurers from 2009 to 2015 and extracted the receiving leg, notional amount, and residual maturity of each contract. The 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 an estimate of 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

[44] 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.

[45] See Appendix G for details about our swap data construction.

insurer's total general account assets to obtain an estimate of the amount of net duration added by 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.

7.2 Cross-sectional regression results

We begin by implementing a cross-sectional test. The test consists of regressing the *AS-adjusted markup* on the interaction of $Net\_swap\_duration_{it}$ and $10Y\text{-}3M\_Treasury\_spread_t$ and a date fixed effect α_3^t using the following framework:

$$\begin{aligned} Annuity\_markup_{ijt} =& \beta_1 Net\_swap\_duration_{it} + \beta_2 Net\_swap\_duration_{it} \times 10Y\text{-}3M\_Treasury\_spread_t \\ &+ \mathbf{z}'_{it}\boldsymbol{\gamma_1} + 10Y\text{-}3M\_Treasury\_spread_t \times \mathbf{z}'_{it}\boldsymbol{\gamma_2} \\ &+ \alpha_1^i + \alpha_2^j + \alpha_3^t + \epsilon_{ijt}. \end{aligned} \tag{6}$$

Regression (7) also includes an insurer fixed effects α_1^i, a product fixed effects α_2^j and a vector of quarterly contract and insurer controls, $\mathbf{z}'_{it}$. The control vector $\mathbf{z}'_{it}$ contains the contract-level reserve requirement ratio, $Reserve\_Ratio_{jt}$—which, as we know from our earlier analysis, significantly affects annuity markups—as well as logged asset size and leverage. We continue to allow for an interaction between each of the control variables and the yield curve shock $10Y\text{-}3M\_Treasury\_spread_t$ and report two-way insurer- and date-clustered robust standard errors as our benchmark. That said, we obtain very similar standard errors using insurer-clustered robust standard errors.

The coefficient estimate on the main interaction term in Column 1 of Table 4 suggests that an insurer with a median-level $Net\_swap\_duration_{it}$ decreases its AS-adjusted markup by about 0.033 percentage point in response to an unexpected flattening of the yield curve—a one standard deviation decrease in $10Y\text{-}3M\_Treasury\_spread_t$. The relatively small economic magnitude of this average effect is consistent with the observation that the median $Net\_swap\_duration_{it}$ is close to zero. However, this effect is almost 12 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. That is, insurers adding relatively

Table 4: Cross-sectional evidence of the risk-management channel The unit of observation is an insurer-product-quarter. The sample of observations extends from 2009 to 2015, which covers the period of the *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-effect regressions implemented using the penalized fixed-effect estimation method proposed by Koenker (2004). Each column contains the estimated coefficients for a particular quantile (τ). The intercept coefficient estimates for each quantile are omitted for legibility. Clustered bootstrapped standard errors with 2,999 replications are implemented using the generalized bootstrap of Chatterjee and 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	
$Annuity\_markup_{ijt}$		$\tau = 0.25$	$\tau = 0.5$	$\tau = 0.75$
$Net\_swap\_duration_{it} \times$	5.04**	6.78***	4.62***	3.94***
$10Y\text{-}3M\_Treasury\_spread_t$	(2.32)	(0.56)	(0.35)	(0.35)
$Net\_swap\_duration_{it}$	-8.85	-14.98***	-9.34***	-8.46***
	(5.97)	(1.59)	(1.09)	(1.03)
$10Y\text{-}3M\_Treasury\_spread_t$		7.79*	9.34***	7.96**
		(3.07)	(2.81)	(2.61)
$Reserve\_Ratio_{jt}$	58.64***	39.64***	53.39******	62.31***
	(17.14)	(8.52)	(7.1)	(6.18)
$Reserve\_Ratio_{jt} \times$	-12.95***	-9.86**	-12.88***	-14.75***
$10Y\text{-}3M\_Treasury\_spread_t$	(4.34)	(3.4)	(2.94)	(2.6)
$Baa\text{-}Aaa\_spread_t$		-13.54***	-16.27***	-11.97***
		(2.39)	(2.13)	(2.08)
$Baa\text{-}Aaa\_spread_t \times$		6.06***	6.92***	4.32***
$10Y\text{-}3M\_Treasury\_spread_t$		(1.02)	(0.94)	(0.87)
$10\_HQM\_spread_t$		10.16***	11.9***	5.77***
		(1.5)	(1.32)	(1.35)
$10\_HQM\_spread_t \times$		-3.66***	-4.07***	-0.88
$10Y\text{-}3M\_Treasury\_spread_t$		(0.65)	(0.61)	(0.6)
$Leverage\_ratio_{it}$	0.02	-0.39***	-0.23***	-0.04
	(0.35)	(0.09)	(0.07)	(0.07)
$Leverage\_ratio_{it} \times$	0.01	0.1**	0.04	-0.05
$10Y\text{-}3M\_Treasury\_spread_t$	(0.14)	(0.03)	(0.03)	(0.03)
$Log\_total\_assets_{it}$	-2.72	0.58	-0.59*	-1.56***
	(1.70)	(0.32)	(0.26)	(0.25)
$Log\_total\_assets_{it} \times$	0.66	-0.02	0.5***	0.92***
$10Y\text{-}3M\_Treasury\_spread_t$	(0.57)	(0.13)	(0.11)	(0.11)
Fixed effects:				
Product char. (j)	Y		Y	
Insurer (i)	Y		Y	
Date (t)	Y		N	
Observations	9,149		9,149	
Adjusted R^2	0.67		χ^2_1-test	26.31***

more positive duration with swaps decrease their AS-adjusted markup by almost one-third of 1 percentage point—1.7 percent of the average AS-adjusted markup—in response to a flattening yield curve.

7.3 Quantile fixed-effect regression results

Columns 2 to 4 of Table 4 delve deeper by estimating a quantile regression with insurer fixed effects. This specification exploits within-insurer variation to investigate how a flattening yield curve affects insurers at different points of the markup distribution with different interest rate swap portfolios. 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

$$\begin{aligned} Q_{Annuity\_markup_{ijt}}(\tau|\mathbf{x}'_{ijt}) =& \beta_1(\tau)Net\_swap\_duration_{it} + \beta_2(\tau)10Y\text{-}3M\_Treasury\_spread_t \\ &+ \beta_3(\tau)10Y\text{-}3M\_Treasury\_spread_t \times Net\_swap\_duration_{it} \\ &+ \alpha_1^i + \alpha_2^j + 10Y\text{-}3M\_Treasury\_spread_t \times \mathbf{z}'_{it}\boldsymbol{\gamma}(\tau)\ , \end{aligned} \tag{7}$$

where $\tau \in \{0.25, 0.5, 0.75\}$ are three quartiles of interest, the vector $\mathbf{x}'_{ijt}$ contains all the covariates, and $\mathbf{z}'_{it}\boldsymbol{\gamma}(\boldsymbol{\tau})$ is a vector of contract-, insurer-, and time-varying controls. In addition to the contract-level reserve requirement ratio, insurer-level logged asset size, and leverage, the control vector $\mathbf{z}'_{ijt}$ also includes time-varying long-term investment-grade spreads, $Baa\text{-}Aaa\_spread_t$, and the insurer funding-cost proxy, $10\_HQM\_spread_t$. We continue to allow for an interaction between our control variables and the yield curve shock $10Y\text{-}3M\_Treasury\_spread_t$. Finally, α_1^i and α_2^j are the insurer and contract fixed effects, respectively.

The estimated coefficients of interest are $\hat{\beta}_2(\tau) + \hat{\beta}_3(\tau) \times Net\_swap\_duration_{it}$, reported in Columns 2 to 4 of Table 4. The bottom row of Column 4 reports the value of a Wald test statistic rejecting the null hypothesis that the 25th and 75th percentile coefficients on the interaction term are equal at below the 1 percent significance level. The estimated coefficient values suggest insurers with a better hedge against a flattening yield curve (a value of $Net\_swap\_duration_{it}$ in the top quartile of the distribution) cut their markups by about 4 percent, on average, after a one standard deviation *decrease* in $10Y\text{-}3M\_Treasury\_spread_t$. In contrast, insurers with a $Net\_swap\_duration_{it}$ in the bottom quartile of the distribution do not significantly cut their markups in response to a flattening yield curve. Moreover, the counterfactual *decrease* in AS-adjusted markup in response to a flattening of the yield curve (i.e., a one standard deviation *decrease* in $10Y\text{-}3M\_Treasury\_spread_t$) is about 22 percent larger for an insurer moving from the bottom to the top of the AS-adjusted markup distribution. This result suggests that insurers post more competitive annuity prices when they experience a relatively positive, or at least less adverse, net worth shock.

8 Conclusion

In this paper, we showed that a large share of the notoriously high life annuity price markups can be explained by the cost of managing interest rate risk. We proposed a novel theory of annuity pricing that reflects frictions in both the insurance and capital markets. A key insight of the theory is that the cost of interest rate risk management is a function of insurers' investment and capital structure decisions. This fact means that it is generally not possible to analyze the effect of capital market frictions on annuity prices without also taking into account the effect of frictions in insurance markets, and vice versa. We identified the IRM channel by using annuity markup data covering a 30-year period and by exploiting long-term corporate bond market shocks and the U.S. insurance regulatory framework. We found 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 periods of overall financial market stress, as annuity prices tend to be lower when long-term investment-grade bond spreads are higher!

Our results have important implications for the macroeconomic literature studying insurance markets. For example, a robust result in the literature studying the welfare effects of social insurance programs using life-cycle models is that social insurance crowds out private insurance—e.g., Cutler and 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 and Tsyvinski 2007). However, this policy conclusion is largely the outcome of assuming that life insurers operate in frictionless capital markets. Under this assumption, life insurers costlessly hedge interest rate risk. Contrary to the premise in these studies, we showed that the supply of private life annuities is constrained by interest rate risk, which is the outcome of a nontrivial interaction between frictions in insurance and capital markets. Another area of interest involves questions surrounding the shrinking U.S. long-term care insurance market—e.g., Ameriks et al. (2018) and Braun et al. (2019). Long-term care insurance is another type of long-duration insurance product that exposes insurers to interest rate risk, in addition to uncertainties about future health-care and demographic costs. Here, too, studying the effect of interest rate risk management could shed some light. We leave these important questions to future research.

References

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

Allen, F. (1993). Estimating Divisional Cost of Capital for Insurance Companies. In Cummins, J. D. and Lamm-Tennant, J., editors, *Financial Management of Life Insurance Companies*, pages 101–23. Springer Netherlands, Dordrecht.

Ameriks, J., Briggs, J., Caplin, A., Shapiro, M. D., and Tonetti, C. (2018). The Long-Term-Care Insurance Puzzle: Modeling and Measurement. *NBER Working Paper 22726.*

Angeletos, G.-M. (2002). Fiscal policy with noncontingent debt and the optimal maturity structure. *The Quarterly Journal of Economics*, 117(3):1105–1131.

Bailey, A. H. (1862). On the Principles on Which the Funds of Life Assurance Societies Should Be Invested. *Assurance Magazine, and Journal of the Institute of Actuaries*, 10(3):142–47.

Barclay, M. J. and Smith Jr, C. W. (1995). The Maturity Structure of Corporate Debt. *Journal of Finance*, 50(2):609–31.

Berends, K., King, T. B., et al. (2015). Derivatives and Collateral at US Life Insurers. *Economic Perspectives*, 39(1):21–38.

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

Blanchard, O. J. (1985). Debt, Deficits, and Finite Horizons. *Journal of Political Economy*, 93(2):223–47.

Bolton, P. and Scharfstein, D. S. (1990). A Theory of Predation Based on Agency Problems in Financial Contracting. *American Economic Review*, pages 93–106.

Bolton, P. and 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., and 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.

Brunetti, C., Foley-Fisher, N., and Verani, S. (2023). Measuring Interest Rate Risk Management by Financial Institutions. *mimeo*.

Chatterjee, S. and Bose, A. (2005). Generalized Bootstrap for Estimating Equations. *Annals of Statistics*, 33(1):414–36.

Cutler, D. M. (1996). Why Don't Markets Insure Long-Term Risk? *Working paper*.

Cutler, D. M. and Gruber, J. (1996). Does Public Insurance Crowd Out Private Insurance? *Quarterly Journal of Economics*, 111(2):391–430.

Davidoff, T., Brown, J., and Diamond, P. (2005). Annuities and Individual Welfare. *American Economic Review*, 95(5):1573–89.

Domanski, D., Shin, H. S., and Sushko, V. (2017). The Hunt for Duration: Not Waving but Drowning? *IMF Economic Review*, 65(1):113–53.

Eichenbaum, M. and Peled, D. (1987). Capital Accumulation and Annuities in an Adverse Selection Economy. *Journal of Political Economy*, 95(2):334–54.

Einav, L. and 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., and Wagner, W. (2021). Insurers As Asset Managers and Systemic Risk. *Kelley School of Business Research Paper*.

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

Foley-Fisher, N., Heinrich, N., and Verani, S. (2022). How Do US Life Insurers Manage Liquidity in Times of Stress? *Available at SSRN*.

Foley-Fisher, N., Heinrich, N., and Verani, S. (2023). Are Life Insurers the New Shadow Banks? *Available at SSRN*.

Foley-Fisher, N., Narajabad, B., and Verani, S. (2016). *Securities Lending as Wholesale Funding: Evidence from the US Life Insurance Industry*. Technical report, National Bureau of Economic Research.

Froot, K. A. and 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. *Journal of Finance*, 77(1):449–503.

Giambona, E., Kumar, A., and Phillips, G. M. (2021). *Hedging and Competition*. Technical report, National Bureau of Economic Research.

Golosov, M. and Tsyvinski, A. (2007). Optimal Taxation with Endogenous Insurance Markets. *Quarterly Journal of Economics*, 122(2):487–534.

Greenwood, R., Hanson, S., and Stein, J. (2010). A Gap-Filling Theory of Corporate Debt Maturity Choice. *Journal of Finance*, 65:993–1028.

Hart, O. and Moore, J. (1994). A Theory of Debt Based on the Inalienability of Human Capital. *Quarterly Journal of Economics*, 109(4):841–79.

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

Hartley, D., Paulson, A., and Rosen, R. J. (2016). Measuring Interest Rate Risk in the Life Insurance Sector: The US and the UK. *The Economics, Regulation, and Systemic Risk of Insurance Markets*, page 124.

Hong, J. H. and Ríos-Rull, J.-V. (2007). Social Security, Life Insurance and Annuities for Families. *Journal of Monetary Economics*, 54(1):118–40.

Hosseini, R. (2015). Adverse Selection in the Annuity Market and the Role for Social Security. *Journal of Political Economy*, 123(4):941–84.

Huang, C., Oehmke, M., and Zhong, H. (2019). A Theory of Multiperiod Debt Structure. *Review of Financial Studies*, 32(11):4447–500.

Huber, M. (2022). Regulation-Induced Interest Rate Risk Exposure. *Working paper*.

Knox, B. and 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. and Yogo, M. (2022). The Fragility of Market Risk Insurance. *Journal of Finance*, forthcoming.

Koijen, R. S. and Yogo, M. (2015). The Cost of Financial Frictions for Life Insurers. *American Economic Review*, 105(1):445–75.

Krishnamurthy, A. and Vissing-Jorgensen, A. (2012). The Aggregate Demand for Treasury Debt. *Journal of Political Economy*, 120(2):233–67.

Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., and Brown, J. R. (1999). New Evidence on the Money's Worth of Individual Annuities. *American Economic Review*, 89(5):1299–1318.

Modigliani, F. and Miller, M. H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review*, 48(3):261–97.

National Association of Insurance Commissioners (2013). Historical Evolution of Life Insurance. In CIPR, editor, *State of the Life Insurance Industry: Implications of Industry Trends*. Center for Insurance Policy and Research, Kansas City, Mo., 540.

Ozdagli, A. K. and Wang, Z. K. (2019). Interest Rates and Insurance Company Investment Behavior. *Available at SSRN 3479663*.

Poterba, J. M. and Solomon, A. (2021). Discount Rates, Mortality Projections, and Money's Worth Calculations for US Individual Annuities. *NBER Working Paper 28557*.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics*, 90(4):629–49.

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., and 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):137–50.

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.

A.1 Pre- and post- retirement annuity contracts

Our paper focuses on life annuities that allow retirees to decumulate wealth during retirement. 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 sometimes discussed together in the context of the different phases of an annuity contract in advertising materials, which can be confusing.[46] 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 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-Global Financial Crisis life insurers offered different types of guaranteed minimum benefits in an effort to compete and differentiate their VAs.[47] Therefore, VAs are essentially a mutual fund with an insurance wrapper. The VA assets are segregated from insurers balance sheet and remain the exclusive property of the VA contract holders. The only connection between VA contracts and the balance

[46]See Black, K., Jr., Skipper, H. D. Black, K., III (2015) for a detailed discussion of the various types of annuities.

[47]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.

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 and Yogo 2022), so aggressive market-based minimum guarantees 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 disbursement is sometimes referred to as an annuity "payout phase." Section 1035 of the U.S. tax code allows individuals to exchange an existing VA contract for a new annuity contract without paying any tax on the income and investment gains in their current VA 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.

B Size of the U.S. life annuity market

We provide two estimates of the size of the U.S. life annuity market using company-level data on the number of annuity contracts and account balances reported in the National Association of Insurance Commissioners (NAIC) 2018 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 dollar value corresponds to roughly \$42,500 per American aged 50 to 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 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.[48]

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.

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

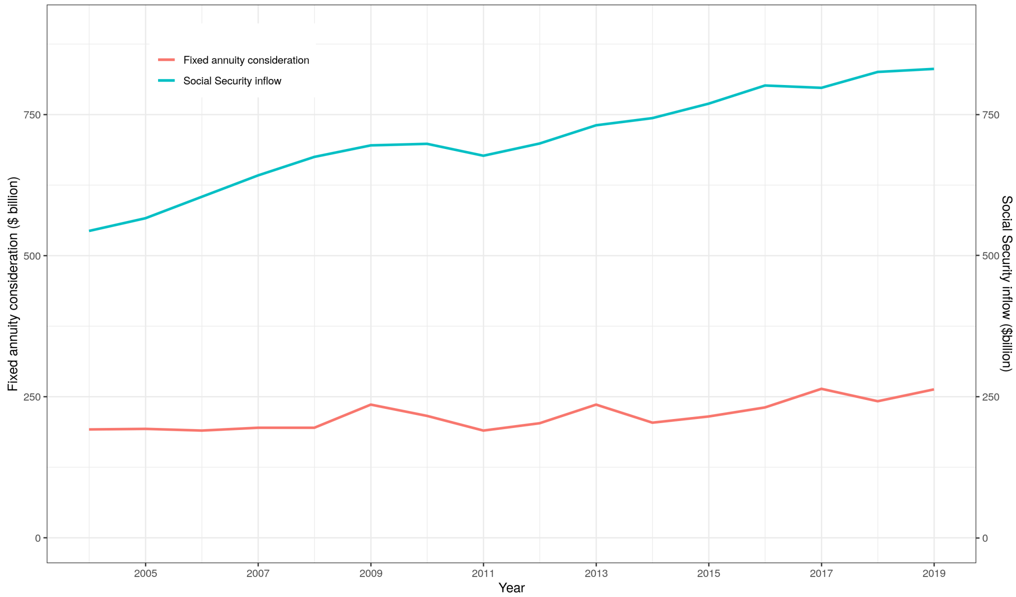
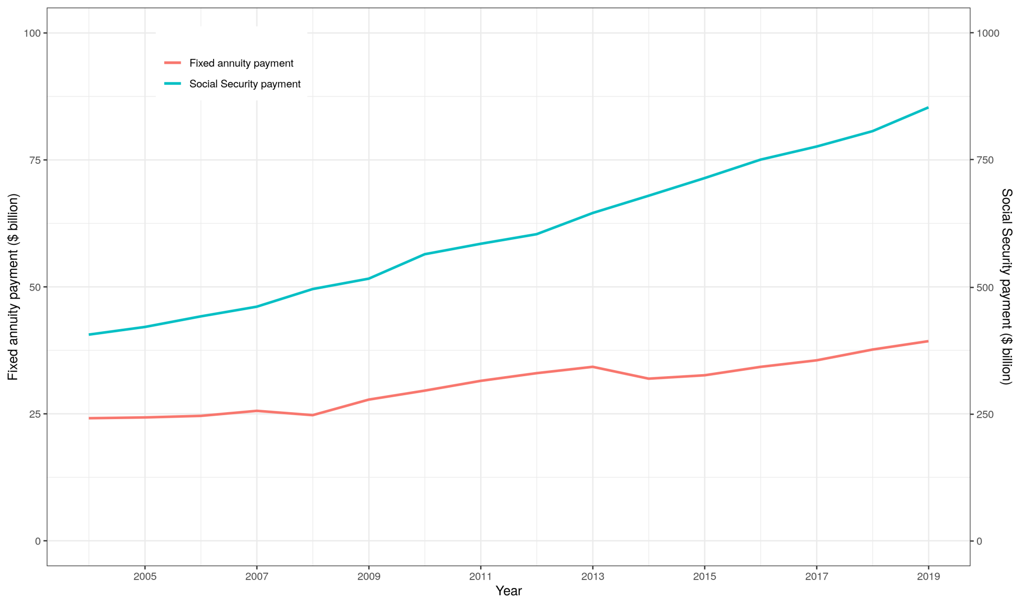


Figure 8: U.S. life insurers' payout on fixed annuity and Social Security payout



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

To put these numbers in perspective, assume that there are 3.5 million new 65-year-old individuals in the U.S. in a given year—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 per 3.5 million new 65-year-old individuals.

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 cover 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.

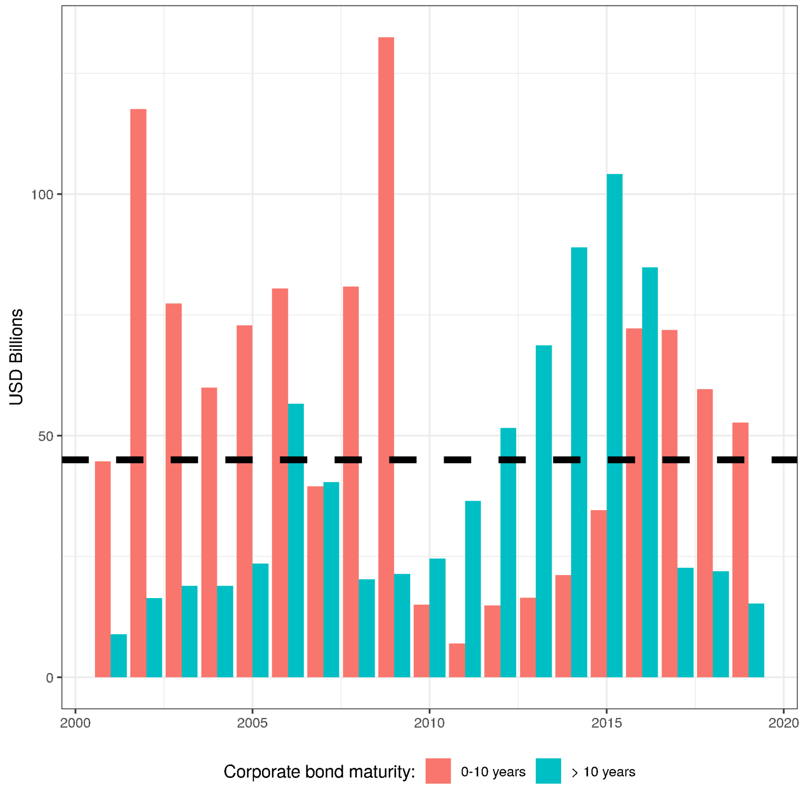


Figure 9: Annual wealth annuitization with life insurers and fixed-rate corporate bond issuance

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 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, noncallable 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 U.S. 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 funds in and out of the U.S.

C Competition in the fixed annuity markets

This appendix presents some evidence of competition in the fixed annuity marketplace. 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 2019:Q2 earnings call provides 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 U.S. fixed annuity market concentration is consistently below 8 percent. Figure 10 confirms industry commentaries that the fixed

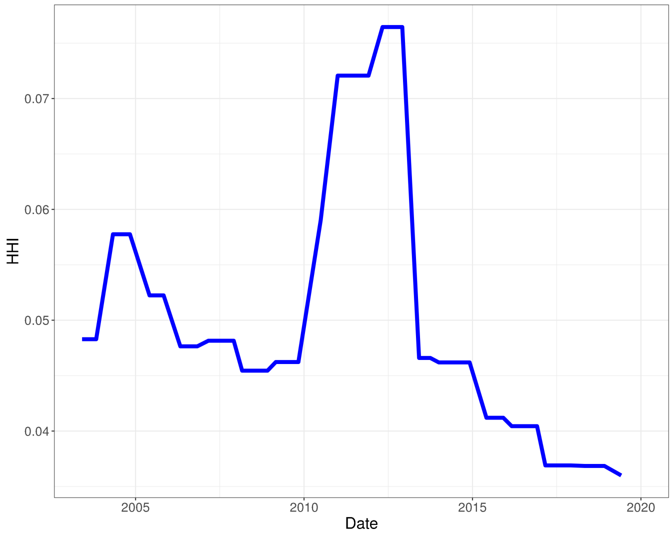


Figure 10: Fixed annuity market Herfindahl-Hirschman Index

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 GFC. This effect is consistent with the decrease in the cross-sectional variance of markups we noted in Section 5. Foley-Fisher et al. (2023) explain 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 charges, these PE-backed insurers, such as Athene, can offer a higher yield on their new annuity liability, thereby lowering prices.

D Main theoretical results and proofs for Section 2

In this appendix, we first prove Theorem 1 in Appendix D.1. Then we provide the details for equilibrium annuity pricing and prove Theorem 2 in Appendix D.2. Finally, we derive equation (4) in Appendix D.3.

D.1 Optimal IRM strategy

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 interest risk management (IRM) so that $NW_1(R_2) \geq 0$ for any R_2. Specifically, competitive insurers finance just enough net worth so that $NW_1(1) = 0$ and $NW_1(R_2) \geq 0$ for $R_2 > 1$. As a result, from the balance sheet equation at $t = 1$,

$$b_2(R_2) = \frac{1}{R_2}\int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha + NW_1(R_2), \tag{8}$$

we get $b_2(1) = \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha$. Hence, by the insurer's budget constraint in $t = 1$,

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha, \tag{9}$$

we obtain the following demand for one-period bonds $b_1(l_2)$ as a function of long-term corporate bonds:

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

Substituting the demand $b_1(l_2)$ into (1) yields

$$NW_0 + \left(\frac{R_l}{R_1} - 1\right) l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1}\left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{11}$$

As a result, when the bond market is constrained, insurers prefer to hold a positive level of net worth to manage interest rate risk, because the yield on long-term bonds is low. To determine the optimal composition of net worth and long-term bonds, consider the following three cases: (i) $NW_1(R_2) = 0$ for all R_2, (ii) $NW_0 = 0$, and (iii) $NW_0 > 0$ with $NW_1(R_2) > 0$ for some R_2.

For case (i), we have $b_2(R_2) = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_2} a(\alpha, q)\, g(\alpha)\, d\alpha$ when $NW_1(R_2) = 0$ for all R_2 from (8). By (9) and (10), we get

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

Therefore, when the bond market is constrained, equation (11) yields

$$NW_0 = \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[\frac{1}{R_l} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha > 0,$$

and $NW_0 = 0$ when the bond market is unconstrained due to limited liability.

For case (ii), when $NW_0 = 0$, equation (11) gives us

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

Therefore, by (8), (9), and (10),

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

which is strictly negative when $\frac{1}{R_l} > \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$. Therefore, $NW_0 = 0$ is not optimal when the bond market is constrained.

For case (iii), if $NW_1\left(\tilde{R}_2\right) > 0$ for some $\tilde{R}_2 > 1$, then

$$R_1 b_1 + \frac{R_l l_2}{\tilde{R}_2} > \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left(1 + \frac{\alpha}{\tilde{R}_2}\right) a(\alpha, q)\, g(\alpha)\, d\alpha$$

by (8) and (9). Consider a new asset allocation $\left(\tilde{b}_1, \tilde{l}_2\right) = \left(b_1 - \epsilon, l_2 + \frac{R_1}{R_l}\epsilon\right)$, so that $NW_1(1) = 0$ under the new allocation and $R_1\tilde{b}_1 + \frac{R_l\tilde{l}_2}{\tilde{R}_2} = R_1 b_1 + \frac{R_l l_2}{\tilde{R}_2} - R_1\left(1 - \frac{1}{\tilde{R}_2}\right)\epsilon$. Hence, for sufficiently small ϵ, $NW_1\left(\tilde{R}_2\right) > 0$. However, $\tilde{b}_1 + \tilde{l}_2 = b_1 + l_2 - \left(1 - \frac{R_1}{R_l}\right)\epsilon$, so by the budget constraint in $t = 0$,

$$b_1 + l_2 = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q)\, g(\alpha)\, d\alpha, \tag{12}$$

the insurer can lower the annuity price under the new asset portfolio since $R_l \geq R_1$. Therefore, case (iii) is inconsistent with a competitive annuity market.

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 (12), NW_0 is uniquely pinned down by (1), $NW_1(R_2)$ is uniquely pinned down by (8), and $b_2(R_2)$ is uniquely pinned down by (9). Therefore, to show uniqueness, it is sufficient to show that at the optimum, $l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$. Taking the original annuity price q as given, first suppose an insurer deviates and chooses $\hat{l}_2 = l_2 - \epsilon$, where $\epsilon \in (0, l_2]$. Then, by (12), 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 $\hat{b}_2(R_2) = b_2(R_2) + \frac{R_1}{R_2}\left(R_2 - \frac{R_l}{R_1}\right)\epsilon$ by (9). Hence, by (8), the new net worth at $t = 1$ is $\hat{NW}_1(R_2) = \frac{R_1}{R_2}\left(R_2 - \frac{R_l}{R_1}\right)\epsilon$. If returns are such that $\hat{NW}_1(1) < 0$, then the new allocation is not optimal since insurers can become insolvent. If returns are such that $\hat{NW}_1(R_2) \geq 0$ for all R_2, then NW_1 is strictly positive for large R_2. This allocation is also not optimal since the insurer can perform IRM with less net worth and charge a lower price. Also, by the same argument as before, the insurer can become insolvent if it deviates and chooses $\hat{l}_2 = l_2 + \epsilon$, where $\epsilon > 0$. Hence, it is optimal for $l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_l} a(\alpha, q) g(\alpha) d\alpha$. This outcome proves uniqueness. ■

D.2 Life annuity pricing

Here, we provide some details on the Bertrand equilibrium in our environment in three separate sections. First, we show that competitive insurers implement the optimal IRM strategy of Theorem 1 in a Bertrand equilibrium. Second, we characterize the equilibrium annuity price in an unconstrained bond market setting. Finally, we prove Theorem 2.

D.2.1 Properties of Bertrand competition

In our setting, life insurers compete over prices, à la Bertrand. Lemma 1 characterizes the basic properties of the Bertrand equilibrium of our model.

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. ■

D.2.2 Pricing with adverse selection in an unconstrained bond market

To see how adverse selection contributes to the annuity price markup, consider first the case when the bond market is unconstrained. By Theorem 1, competitive insurers optimally choose zero net worth. Therefore, equations (12) and (1) show that the equilibrium annuity price in an unconstrained bond market is given by

$$q^{AF} \int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a\left(\alpha, q^{AF}\right) g\left(\alpha\right) d\alpha, \tag{13}$$

where q^{AF} is the risk-adjusted actuarially fair price. The price q^{AF} 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 q^{AF} is higher than the average full-information actuarially fair price.

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

Proof 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

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

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

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

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

D.2.3 Pricing with adverse selection in a constrained bond market

To see how IRM affects the annuity markup, first note that with $\psi = \frac{1}{R_l} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, the bond market is constrained when $\psi > 0$ and unconstrained when $\psi = 0$. Next, insurers' profit $\Pi(q, \psi)$ is given by the difference between their annuity sales revenue and total bond demand:

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

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

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

This definition means that the risk-adjusted actuarially fair price is evaluated using the pool of annuitants purchasing annuities 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^*.

Finally, to prove Theorem 2 (AS-adjusted markup increases with ψ), 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* $\psi > 0$.

Proof First, note that at the equilibrium price q^*, $\frac{\partial \Pi(q^*,\psi)}{\partial q^*} \geq 0$. The reason is that if $\frac{\partial \Pi(q^*,\psi)}{\partial q^*} < 0$, 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 is

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

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

$$\begin{aligned}
&\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha \qquad (14)\\
&+ \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha} a\left(\alpha, q^*\right) - \frac{\partial a\left(\alpha, q^*\right)}{\partial q^*}\right] g\left(\alpha\right) d\alpha \\
&+ \psi \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha} a\left(\alpha, q^*\right) - \frac{\partial a\left(\alpha, q^*\right)}{\partial q^*}\right] g\left(\alpha\right) d\alpha \geq 0.
\end{aligned}$$

Next, we show that Assumption 2 implies that the last term in inequality (14) is strictly negative when $\psi > 0$. 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 of those two inequalities 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 uses the definition of covariance. By rearranging the terms of the second inequality, we find that the last term in inequality (14) is strictly negative if $\psi > 0$.

Therefore, we have

$$\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1+\alpha\mathbb{E}\left(\frac{1}{R_2}\right)\right]\left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha,q^*)}{\partial q^*} g\left(\alpha\right) d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha} a\left(\alpha, q^*\right) - \frac{\partial a\left(\alpha, q^*\right)}{\partial q^*}\right] g\left(\alpha\right) d\alpha > -\int_{\underline{\alpha}}^{\overline{\alpha}} a\left(\alpha, q^*\right) g\left(\alpha\right) d\alpha. \qquad (15)$$

By the definition of q^{AF} and dividing both sides of inequality (15) 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 $\psi > 0$. ■

The actuarially fair price q^{AF} is determined by the equilibrium price q^*, which means that if q^{AF} increases quickly as q^* rises, the AS-adjusted markup could decrease. Lemma 2 proves that this situation cannot occur. Therefore, we can now prove Theorem 2.

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

When $\psi = 0$, then $NW_0 = 0$ by Theorem 1. Hence, by the definition of q^{AF} and q^*, we have $q^* = q^{AF}$ when $\psi = 0$.

Next, we show that q^* increases as $\psi > 0$ increases. Through implicit differentiation,

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

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

Finally, for $\psi > 0$, we show that $q^* - q^{AF}$ increases with ψ. Notice that $\frac{\partial q^*}{\partial \psi} - \frac{\partial q^{AF}}{\partial \psi} = \frac{\partial q^*}{\partial \psi}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. When $\psi > 0$, we have $\frac{\partial q^*}{\partial \psi} > 0$ and $1 > \frac{\partial q^{AF}}{\partial q^*}$ by Lemma 2, which yields $\frac{\partial q^*}{\partial \psi} - \frac{\partial q^{AF}}{\partial \psi} > 0$. ■

D.3 Feedback between IRM and adverse selection

Here, we derive equation (4). With $z = \frac{R_1}{R_l}$, insurers' profit can be rewritten as

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

Implicitly differentiating insurers' zero-profit condition $\Pi(q^*,z) = 0$ yields

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

which gives us equation (4). By Theorem 1, when the bond market is constrained, the optimal amount of net worth at $t = 0$ is a strictly increasing function of z for any given price q:

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

Thus, it follows that the risk-management effect is the additional amount of average net worth needed for IRM under a fixed annuity price.

E Endogenous corporate bond supply for Section 2

This section presents a simple approach to model the corporate long-term bond supply. This section shows how the equilibrium long-term bond return, together with insurers bond demand, is determined and provides a bound on an exogenous parameter that delivers a constrained bond market.

We adopt the model of corporate bond issuers of Greenwood et al. (2010). Corporate bond issuers have an optimal capital structure that involves issuing a target T of long-term bonds l_2. A deviation from the target incurs a quadratic cost of $\frac{1}{2}(l_2 - T)^2$. These costs could reflect firms' exposure to interest rate risk that arises from refinancing risk. For example, firms may be financing long-term projects with relatively shorter-maturity debt. These firms may be targeting an amount of long-term bonds to hedge interest rate rises, and straying from this target in either direction is costly. As in Greenwood et al. (2010), we can interpret the quadratic cost function as a reduced-form way to model the firms' tightening financial constraints when it deviates from its optimal long-term bond supply T.

Taking the return R_l as given, we specify that the representative firm's objective is to solve the following optimization problem:

$$\min_{l_2} R_l l_2 + \frac{1}{2}(l_2 - T)^2.$$

Continuing the interest risk management interpretation of the firm's quadratic cost, we note that a firm is minimizing the interest rate cost and the cost of being exposed to interest rate risk. The solution to the firm's problem yields the long-term bond supply:

$$l_2(R_l) = T - R_l,$$

which is inversely related to the bond return.

The equilibrium return R_l^* is determined by the market-clearing condition for the long-term bond—bond supply and bond demand are equated. For simplicity, we assume that only life insurers purchase long-term corporate bonds.[49] Then, by the long-term bond demand given in Theorem 1, the market-clearing condition for the long-term bond is

$$R_l^{*2} - TR_l^* + L = 0,$$

where $L = \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha$ is the annuity liability in $t = 2$.

Lemma 3 *For any given L, (i) if $T < 2\sqrt{L}$, then long-term corporate bonds are not traded in the economy; (ii) if $2\sqrt{L} \le T < \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, then $\psi > 0$; and, (iii) if $T = \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, then $\psi = 0$.*

Proof By the market-clearing condition,

$$R_l^* = \frac{T \pm \sqrt{T^2 - 4L}}{2}.$$

Therefore, if $T < 2\sqrt{L}$, then there are no long-term corporate bonds in the economy.

Next, we focus on the case with $T \ge 2\sqrt{L}$. Note that the solution $R_l^* = \frac{T - \sqrt{T^2 - 4L}}{2}$ is not plausible, because higher bond supply or higher T would induce lower R_l. Therefore, the equilibrium long-term bond return is

$$R_l^* = \frac{T + \sqrt{T^2 - 4L}}{2}.$$

As a result, by definition, a bond market is constrained if $R_l < \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)}$. As a result, given the equilibrium bond return R_l^*, we have $\frac{1}{R_l} > \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$ when $T < \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, and $\frac{1}{R_l} = \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$ when $T = \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$.

[49] We can enrich the model by considering long-term bond demand from agents and arbitrageurs without changing the main message of this section.

Finally, we need to check that the bounds are valid, i.e., $2\sqrt{L} \leq \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$. It is easy to verify that for any positive value of $\frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, it is always the case that $2\sqrt{L} \leq \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, and it holds with equality only when $\frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right) = \frac{1}{\sqrt{L}}$. This verification completes the proof. ■

Lemma 3 provides the conditions for when the bond market would be constrained. Intuitively, when the targeted bond supply T of the corporate bond issuers is sufficiently low (relative to future annuity liability L), then the bond market would be constrained. Furthermore, our analysis is focused on a constrained bond market with $R_l \geq R_1$, so the relevant bounds of T for our analysis is $\left[\max\left\{2\sqrt{T}, \frac{L+R_1^2}{R_1}\right\}, \frac{R_1}{\mathbb{E}\left(\frac{1}{R_2}\right)} + \frac{L}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)\right)$.

F Extensions to the benchmark model in Section 2

We evaluate the robustness of our theoretical results through two extensions. First, we examine an environment with unlimited long-term government bond supply and demonstrate that our Section 2 results still hold if long-term government bond returns are lower than those of corporate bonds. This result is consistent with government bonds composing a small fraction of life insurers' assets—see Foley-Fisher et al. (2022) for an estimate. Lastly, we consider a monopolistic competition scenario, showing that life insurers continue to adjust their markups in response to interest rate risk management costs, even outside a perfectly competitive setting.

F.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 the corporate bond market is constrained. In this appendix, we consider a model with an unlimited supply of long-term government bonds. We find that unless the long-term government bond provides a higher return than the long-term 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 corporate bond market is constrained.

In addition to the economic environment of Section 2, there is an unlimited supply of zero-coupon long-term government bonds g_2. One unit of government bond purchased at

$t = 0$ returns R_g at $t = 2$. For simplicity, we assume that R_g is exogenously determined.

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

$$b_1 + l_2 + g_2 = q \int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{16}$$

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

$$b_1 + l_2 + g_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1} \left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha + NW_0. \tag{17}$$

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

$$b_2(R_2) = \frac{1}{R_2} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha + NW_1(R_2), \tag{18}$$

where

$$b_2(R_2) = R_1 b_1 + \frac{R_l l_2}{R_2} + \frac{R_g g_2}{R_2} - \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha. \tag{19}$$

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, the unique optimal IRM strategy when $R_g > R_l$ requires an asset allocation and a capital structure that satisfies the following: (i) Asset portfolio: $b_1 = \frac{1}{R_1} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha a(\alpha, q)\, g(\alpha)\, d\alpha$, $l_2 = 0$, $g_2 = \frac{1}{R_g} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha$, and $b_2(R_2) = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_2} a(\alpha, q)\, g(\alpha)\, d\alpha$, and (ii) Capital structure:*

$$NW_0 = \begin{cases} \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[\frac{1}{R_g} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)\right] a(\alpha, q)\, g(\alpha)\, d\alpha & \text{if } \frac{1}{R_g} > \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right) \\ 0 & \text{if } \frac{1}{R_g} = \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right) \end{cases},$$

and $NW_1(R_2) = 0$ for all R_2.

When $R_g \leq R_l$, the optimal IRM has $g_2 = 0$ and requires the same asset portfolio and capital structure as in 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 (18), $b_2(1) = \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 a(\alpha, q)\, g(\alpha)\, d\alpha$. Hence, by equation (19), 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:

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

Substituting the demand $b_1\left(l_2, g_2\right)$ into (17) yields

$$NW_0 + \left(\frac{R_g}{R_1} - 1\right) g_2 + \left(\frac{R_l}{R_1} - 1\right) l_2 = \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha^2}{R_1}\left[1 - \mathbb{E}\left(\frac{1}{R_2}\right)\right] a\left(\alpha, q\right) g\left(\alpha\right) d\alpha. \tag{20}$$

By (20), the long-term government bond and the long-term corporate bond are perfect substitutes. Therefore, competitive insurers purchase the bond with the highest returns to minimize net worth. When $R_l \geq R_g$, insurers purchase only corporate bonds, and the result from Theorem 1 applies. When $R_g > R_l$, insurers replace long-term corporate bonds with long-term government bonds, and the result follows from the proof for Theorem 1. ■

Proposition 2 shows how insurers use long-term government bonds to manage interest rate risk. Long-term government bonds and long-term corporate bonds of the same maturity are perfect substitutes. Competitive insurers seek to minimize the cost of IRM. Therefore, when the yield on long-term government bonds is too low ($R_g \leq R_l$), their availability has no effect on 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 long-term government bonds. When the yield on long-term government bonds is higher ($R_g > R_l$), insurers replace all of their long-term corporate bond demand and 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 $\frac{1}{R_g} > \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$. In this case, net worth at $t = 0$ remains positive.

Typically, the returns from the long-term government bond are less than those from the long-term corporate bond: $R_g < R_l$. The lower government bond yield could reflect, for example, a convenience yield due to the government bond's high liquidity and safe-haven status. 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 and Vissing-Jorgensen 2012). Therefore, Proposition 2 is consistent with the relatively low share of government securities in life insurers' asset portfolio, as noted in Section 1.2.

By Proposition 2, it is almost immediate that the competitive annuity price remains above the risk-adjusted actuarially fair price when the bond market is constrained, which (in the presence of long-term government bonds) is redefined as $\psi_g = \min\left\{\frac{1}{R_l}, \frac{1}{R_g}\right\} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right) > 0$. Similarly, analogous to Theorem 2, the AS-adjusted markup is strictly positive when the bond market is constrained ($\psi_g > 0$).

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

Proof The result follows from the proof of Theorem 2. ■

F.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 $\gamma\left(q_X - q_Y\right)$ 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 are at least consistent with Theorem 1. Specifically, the demand for one-period and long-term bonds and the net worth in each period have to be weakly greater than the amounts 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^*. Following Appendix 6, we can redefine the degree of bond market constraint as $z = \frac{R_1}{R_l}$, so that z is positively related to ψ. Our analysis will focus on the case with a constrained bond market, so $z > \mathbb{E}\left(\frac{1}{R_2}\right)$. 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.[50] This operation 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 aforementioned equilibrium condition, Theorem 3, presented next, characterizes the relationship 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 for* $z > \mathbb{E}\left(\frac{1}{R_2}\right)$: $\frac{\partial q^*}{\partial z} - \frac{\partial q^{AF}}{\partial z} > 0$, *and it 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 $\frac{\partial W}{\partial z} > 0$ for $z > \mathbb{E}\left(\frac{1}{R_2}\right)$. From implicit differentiation, $\frac{\partial q^*}{\partial z} = -\frac{\partial W}{\partial z} / \frac{\partial W}{\partial q^*} > 0$, and $\frac{\partial q^*}{\partial \gamma} = -\frac{\partial W}{\partial \gamma} / \frac{\partial W}{\partial q^*} < 0$.

[50] 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 insurers are monopolists.

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

$$q^* > \frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}(1+\alpha z) a(\alpha, q^*) g(\alpha) \, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) \, d\alpha}. \tag{21}$$

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 preceding inequality holds with equality when $\gamma = 0$. Using the definitions of $z = \frac{R_1}{R_l}$ and $\psi = \frac{1}{R_l} - \frac{1}{R_1}\mathbb{E}\left(\frac{1}{R_2}\right)$, we can rewrite that inequality as

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

where $\psi > 0$. Since $\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) \, d\alpha < 0$, we can replace q^* in the foregoing inequality with

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

which is strictly smaller than q^* by (21) and yields

$$\begin{aligned} &\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) \, d\alpha \\ &+ \int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\alpha}{R_1}\left[1 + \alpha \mathbb{E}\left(\frac{1}{R_2}\right)\right] \left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) \, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) \, d\alpha} a(\alpha, q^*) - \frac{\partial a(\alpha, q^*)}{\partial q^*}\right] g(\alpha) \, d\alpha \\ &+ \psi \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^2 \left[\frac{\int_{\underline{\alpha}}^{\overline{\alpha}} \frac{\partial a(\alpha, q^*)}{\partial q^*} g(\alpha) \, d\alpha}{\int_{\underline{\alpha}}^{\overline{\alpha}} a(\alpha, q^*) g(\alpha) \, d\alpha} a(\alpha, q^*) - \frac{\partial a(\alpha, q^*)}{\partial q^*}\right] g(\alpha) \, d\alpha > 0. \end{aligned} \tag{22}$$

Notice (22) resembles inequality (14) 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 $\psi > 0$ if Assumption 2 holds.

Finally, notice that $\frac{\partial q^*}{\partial z} - \frac{\partial q^{AF}}{\partial z} = \frac{\partial q^*}{\partial z}\left(1 - \frac{\partial q^{AF}}{\partial q^*}\right)$. When the bond market is constrained, $\frac{\partial q^*}{\partial z} > 0$ and Lemma 2 applies, so 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 $\psi > 0$, 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 $NW_0 = 0$ when $\psi = 0$, so inequality (22) becomes

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

By the definition of q^{AF} and dividing both sides of the aforementioned 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 $\psi \leq 0$. Therefore, for any ψ, 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 because of IRM or market power. Insurers increase their AS-adjusted markup either 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—$\psi = 0$ or $z = \mathbb{E}\left(\frac{1}{R_2}\right)$—can generate an AS-adjusted markup that is observationally equivalent to that generated by 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—$\mathbb{E}\left(\frac{1}{R_2}\right) < z$. This fact 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.

G Additional details about variable construction

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

G.1 Regulatory interest rate to discount annuity liabilities

Before 2018, state insurance regulations 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."[51] 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.

In 2018, state insurance regulators adopted a new methodology to calculate the single reference interest rate used in regulatory reserve regulations. With the new methodology, the reference interest rate is the sum of a weighted average U.S. Treasury yield plus a credit spread and an expected default cost. The spread over the reference Treasury rate is calculated by the NAIC using the public bond portion of an average U.S. life insurer's asset portfolio. The new reference interest rate varies by type of annuity contract guarantee period and is reset once a quarter (for retail annuity contracts). For example, the reference rate for a single premium immediate annuity issued on March 2, 2018, without a guarantee period to a 68-year-old was 3.25 percent, which is about 75 basis points (0.75 percentage point) higher than the reference Treasury rate used in the reference rate calculation.[52] By comparison, Moody's seasoned Aaa and Baa corporate bond yields on the same day are 3.9 percent and 4.58 percent, respectively.

G.2 Mortality assumption

The Society of Actuaries (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 reserves per dollar of annuity sold. Second, statutory regulations did not require insurers to apply the SOA generational mortality improvement factor to the static loaded mortality table for their

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

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

reserve calculations before 2015, when the 2012 Individual Annuitant Mortality Table was adopted in most states. As a consequence, regulatory reserves before 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 before 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 in 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 Hedging interest rate risk with interest rate swaps

In our stylized model, insurers would be indifferent between managing interest rate risk with net worth or with interest rate swaps absent any additional frictions. To see this finding, consider a straightforward extension of the model. At $t = 0$, insurers have the option to purchase a fixed-for-float swap from some (unmodeled) counterparty at price p^s. The insurer would pay the one-period interest rate and receive the two-period interest rate. The price of the swap is a spread over the one-period bond rate, which is determined in equilibrium in the swap market. The present value of the swap at inception at $t = 0$ is 0. If the interest rate realization at $t = 1$ is low, then the present value of the fixed-for-float swap becomes positive (the swaps become an asset), thereby offsetting the disproportionate increase in the present value of the annuity liabilities. Therefore, insurers in this model extension can hedge the interest rate risk with fixed-for-float swaps, and the cost of the derivative position would be passed on to annuity consumers through the annuity markup.

H.1 Statutory accounting treatment of swaps

In practice, however, there are a number of frictions making interest risk management with swaps relatively unpopular in the U.S. life insurance industry As a consequence of these frictions, only the largest and sophisticated life insurers include interest rate swaps in their hedging programs.

One reason is that state insurance regulations require each open swap position to be tied to a specific asset of liabilities to be treated under hedge-effective accounting. If the insurer cannot continuously prove to its regulator the hedge effectiveness of the swap, then the swap needs to be carried at fair value. Therefore, an insurer "fixing" the duration of its balance sheet with swaps will carry swaps at fair value. Although this distinction is irrelevant in our simple model, it matters for actual insurers in the U.S. for at least two reasons.

First, recurring fair value measurement requires an insurer to invest in information and control systems to assess relevant economic conditions and estimate fair values of swaps quarterly. This assessment is required to provide the regulator with a detailed report about the hedge effectiveness of each swap. This is unlike statutory accounting for assets and liabilities that are not marked to market and whose value evolves according to predetermined amortized cost measurement. Only the largest, most sophisticated insurers in the U.S. have this back-office technology to monitor their swap value.

Second, swaps increase the insurer cash flow and risk-based capital variability. The fair value of a swap becomes either an asset or a liability, depending on the term of the contract and the change in interest rate. Collateral must be posted with (received from) the counterparty when the fair value of the swap exceeds a predetermined threshold, as it increases credit risk. For example, if interest rate were to increase a lot rapidly, then insurers that entered into fixed-for-float swaps to close their negative duration gap would have to post collateral with their counterparty, typically a bank. Insurers typically post Treasury securities or a cash equivalent as collateral, but the question, of course, is how the insurer will finance the collateral. Moreover, the change in the swap fair value changes the insurer risk-based capital charge. As the fair value of the swap increases, so does the capital charge that is related to the time-varying rating of the swap counterparty. Therefore, the insurer needs to finance more capital to maintain its regulatory capital level.

H.2 Measuring the net 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 $\times$ the residual maturity of the contract and that the interest rate reset on the floating leg of the swap occurs every three months. The factor 0.75 is a commonly used rule of thumb when the actual swap curve is unavailable. Although it is quite crude, this assumption is reasonable to study the variation in average swap duration across insurers in our setting. Moreover, assuming that the interest rate reset on the floating leg of the swap occurs every three months is consistent with the widely used three-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 insurers' total general account assets to obtain the amount of duration added by swaps as a fraction of 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 Evidence from term annuity markups

As a robustness test, we implement our main regression test using data on term annuities markups. Term annuities pay the policyholder a stream of income for a fixed period. Unlike bank certificate of deposits, the principal payment (or premium) in a term annuity is fully amortized during the contract. Moreover, unlike life annuities, term annuities do not pay a mortality-contingent stream of cash flows, and their price does not depend on the customer's age or gender. Therefore, the markup of term annuities is not affected by adverse selection, and the demand for term annuities is unlikely to be systematically driven by age or gender.

The unit of observation for this test is a term annuity contract j offered by insurer i at date t. The sample of observation extends from 1989:Q1 to 2019:Q4. We estimate Equation (23) as follows:

$$
\begin{aligned}
& Term\_Annuity\_markup_{ijt} = \\
& +\beta_1 Reserve\_Ratio_{jt} + \beta_2 10\_HQM\_spread_t + \mathbf{z}'_{it}\boldsymbol{\gamma_1} + \beta_3 Baa\text{-}Aaa\_spread_t \\
& + (\beta_4 Reserve\_Ratio_{jt} + \beta_5 10\_HQM\_spread_t + \mathbf{z}'_{it}\boldsymbol{\gamma_2}) \times Baa\text{-}Aaa\_spread_t \\
& + \sum_{k \neq 5\text{yr}} \mathbf{1}_{\{\text{Term}=k\}} \times \{ \\
& \quad \beta_{1,k} Reserve\_Ratio_{jt} + \beta_{2,k} 10\_HQM\_spread_t + \mathbf{z}'_{it}\boldsymbol{\gamma_{1,k}} + \beta_{3,k} Baa\text{-}Aaa\_spread_t \\
& + (\beta_{4,k} Reserve\_Ratio_{jt} + \beta_{5,k} 10\_HQM\_spread_t + \mathbf{z}'_{it}\boldsymbol{\gamma_{2,k}}) \times Baa\text{-}Aaa\_spread_t\} \\
& + \alpha^i + \epsilon_{ijt}.
\end{aligned}
\tag{23}
$$

Equation (23) uses Term $= 5yr$ as the benchmark effect. The variable $\mathbf{1}_{\{\text{Term}=k\}}$ takes the value 1 if the maturity of the term annuity is equal to $k \in \{10yr, 15yr, 20yr, 25yr, 30yr\}$. Equation (23) includes an insurer fixed effect α^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 behavior. We allow for an interaction between all the control variables and the aggregate shock variable $Baa\text{-}Aaa\_spread_t$. In our reduced sample starting in 2000:Q1, we include the vector $\boldsymbol{z}'_{it}$ containing insurer-level time-varying asset size and leverage. As in the main text, we report insurer-clustered robust standard errors throughout as our baseline and two-way insurer- and date-clustered robust standard errors as a robustness test.

Table 5 reports the regression results, omitting the many interactions with control variables for legibility. We are interested in estimating the effect of a change in the contract-level reserve requirement on this contract's markup during times of relatively wide and tight $Baa\text{-}Aaa\_spread_t$ *conditional* on insurers' cost of funding, $10\_HQM\_spread_t$. For our benchmark case with five-year term annuities, the effect is estimated by $\hat{\beta}_1 + \hat{\beta}_4 \times Q_{Baa\text{-}Aaa\_spread_t}(p)$, where $Q_{Baa\text{-}Aaa\_spread_t}(p)$ is the p-th percentile of $Baa\text{-}Aaa\_spread_t$. The effect for the other annuity maturities adds the relevant coefficients on the interaction terms with the binary variable $\mathbf{1}_{\{\text{Term}=k\}}$. We evaluate the estimated effect using the first, second, and third quartiles of the distribution of $Baa\text{-}Aa\_spread_t$, as we are interested in evaluating how the effect varies across periods in which the long-duration investment-grade spread is relatively tight, neutral, and wide, respectively.

Table 5: The effect of long-duration investment-grade corporate bond yield spread on term annuity markups. The unit of observation is a life insurer-product-quarter. The sample of observation extends from 1989:Q1 to 2019:Q4. The dependent variable *Term annuity* $markup_{ijt}$ is term annuity markups for product j sold by insurer i in quarter t. The baseline coefficients are estimated for the five-year term annuity contract. The intercept coefficients and the full set of interacted controls are not reported for legibility. Insurer-clustered standard errors are reported in parentheses in Columns 1 and 3, and two-way insurer- and date-clustered standard errors are reported in parentheses in Columns 2 and 4. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Dependent variable $Term\_annuity\_markup_{ijt}$	(1)	(2)	(3)	(4)
$Reserve\_Ratio_{ijt}$	106.21***	106.21***	133.60***	133.60***
	(14.24)	(18.21)	(44.85)	(46.87)
$Reserve\_Ratio_{ijt} \times 10yr\_maturity$	-59.44***	-59.44***	-81.66**	-81.66**
	(13.43)	(20.40)	(32.19)	(33.67)
$Reserve\_Ratio_{ijt} \times 15yr\_maturity$	-84.44***	-84.44***	-113.54**	-113.54**
	(20.53)	(24.50)	(44.15)	(49.89)
$Reserve\_Ratio_{ijt} \times 20yr\_maturity$	-92.34***	-92.34***	-122.43**	-122.43**
	(25.62)	(30.62)	(52.60)	(58.00)
$Reserve\_Ratio_{ijt} \times 25yr\_maturity$	-110.06***	-110.06***	-131.92**	-131.92**
	(30.29)	(34.21)	(62.16)	(63.49)
$Reserve\_Ratio_{ijt} \times 30yr\_maturity$	-129.89***	-129.89***	-145.19**	-145.19**
	(34.50)	(36.36)	(66.59)	(66.75)
$Reserve\_Ratio_{ijt} \times Baa\text{-}Aa\_spread_t$	-78.12***	-78.12***	-112.76***	-112.76***
	(15.21)	(22.02)	(38.54)	(40.01)
$Reserve\_Ratio_{ijt} \times Baa\text{-}Aa\_spread_t \times 10yr\_maturity$	58.60***	58.60**	80.45***	80.45**
	(16.51)	(24.63)	(28.59)	(29.73)
$Reserve\_Ratio_{ijt} \times Baa\text{-}Aa\_spread_t \times 15yr\_maturity$	74.11***	74.11***	111.59***	111.59**
	(21.54)	(27.31)	(39.56)	(46.28)
$Reserve\_Ratio_{ijt} \times Baa\text{-}Aa\_spread_t \times 20yr\_maturity$	82.36***	82.36**	121.02**	121.02**
	(26.17)	(33.97)	(47.83)	(54.98)
$Reserve\_Ratio_{ijt} \times Baa\text{-}Aa\_spread_t \times 25yr\_maturity$	99.78***	99.78***	130.17**	130.17**
	(30.10)	(37.20)	(57.76)	(60.24)
$Reserve\_Ratio_{ijt} \times Baa\text{-}Aa\_spread_t \times 30yr\_maturity$	118.32***	118.32***	144.37**	144.37**
	(34.51)	(38.91)	(62.80)	(63.93)
SE clustering	Insurer	Insurer & Date	Insurer	Insurer & Date
Fixed effects	Insurer	Insurer	Insurer	Insurer
Quarterly log asset and leverage controls	N	N	Y	Y
Observations	4,576	4,576	3,026	3,026
Adjusted R^2	0.60	0.60	0.65	0.65

Table 6 summarizes the estimated economic magnitudes for each term annuity contract using the three quartiles of the $Baa\text{-}Aaa\_spread_t$ variable. The last two columns of Table 5 report the p-value of a one-sided statistical test of the null hypotheses that the first-quartile economic effect is lower than or equal to the third-quartile effect using insurer-clustered and the more conservative two-way insurer- and date-clustered robust standard errors, respectively.

The main takeaway from Table 6 is that when $Baa\text{-}Aaa\_spread_t$ is at its median level, an exogenous increase in the reserve requirement has a relatively lower effect on the markup of term annuities with longer-maturity. The reason is that, as we explained in the main text, the early part of the yield curve has less weight in the present value calculation of longer-maturity term annuities. However, this effect is smaller in periods when the $Baa\text{-}Aaa\_spread_t$ is larger for lower-duration annuities, such as 5-year and 10-year term annuities.

Table 6: The effect of investment-grade corporate bond yield spread on term annuity markups. This table summarizes the economic effect of a one standard deviation increase in $Reserve\_Ratio_{ijt}$ on $Annuity\_markup_{ijt}$ for different levels of $Baa\text{-}Aaa\_spread_t$ and different term annuity maturities.

	$Baa\text{-}Aaa\_spread_t$ quartiles			p-values	p-values
Annuity term	1st	2nd	3rd	Insurer-clustered SEs	Insurer- and date-clustered SEs
5yr	2.081	1.328	0.375	0.002	0.002
10yr	1.204	0.988	0.715	0.037	0.053
15yr	0.852	0.844	0.835	0.466	0.482
20yr	0.78	0.835	0.905	0.326	0.394
25yr	0.671	0.787	0.934	0.262	0.303
30yr	0.567	0.778	1.045	0.163	0.19

J Baseline 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 7 reproduces our main result on samples before and after the 2007–09 financial crisis, respectively. The results in Table 7 are broadly consistent with the results in Table 2. For example, focusing on Column 2, we note that 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 1.45 percentage point when $Baa\text{-}Aaa\_spread_t$ is at its median level (0.8). This effect is about 25 percent lower in periods when $Baa\text{-}Aaa\_spread_t$ is in the third quartile of its distribution relative to periods when $Baa\text{-}Aaa\_spread_t$ is in the first quartile of its distribution.

Table 7: 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 and 2 and from 2010 to 2019 in Columns 3 and 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)	(4)
$Baa\text{-}Aaa\_spread_t \times Reserve\_Ratio_{jt}$	-35.73***	-35.73**	27.37***	27.37*
	(8.92)	(13.59)	(6.08)	(14.42)
$Reserve\_Ratio_{jt}$	56.91***	56.91***	-10.67	-10.67
	(7.93)	(12.20)	(6.55)	(13.36)
$Baa\text{-}Aaa\_spread_t$	40.11***	40.11**	-28.79***	-28.79*
	(9.12)	(15.08)	(6.91)	(15.41)
$10\_HQM\_spread_t$	4.90***	4.90*	8.83***	8.83**
	(1.69)	(2.47)	(1.87)	(4.04)
$Baa\text{-}Aaa\_spread_t \times 10\_HQM\_spread_t$	-5.17***	-5.17	-1.83	-1.83
	(1.79)	(3.10)	(1.58)	(3.09)
SE clustering	Insurer	Insurer/Date	Insurer	Insurer/Date
Observations	16,767	16,767	21,502	21,502
Adjusted R^2	0.57	0.57	0.61	0.61