Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Measuring Interest Rate Risk Management by Financial Institutions

Celso Brunetti, Nathan Foley-Fisher, Stéphane Verani

2023-067

Please cite this paper as:
Brunetti, Celso, Nathan Foley-Fisher, and Stéphane Verani (2023). "Measuring Interest Rate Risk Management by Financial Institutions," Finance and Economics Discussion Series 2023-067. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2023.067.

Measuring Interest Rate Risk Management by Financial Institutions*

Celso Brunetti[1], Nathan Foley-Fisher[1], and Stéphane Verani[1]

[1]Federal Reserve Board

First version: June 2022; this version: August 2023

Abstract

Financial intermediaries manage myriad interest rate risk exposures. We propose a new method to measure financial intermediaries' residual interest rate risk using high-frequency financial market data. Our method exploits all available high-frequency information and is valid under extremely weak assumptions. Applying the method to U.S. life insurers, we find their interest rate risk management strategies are generally effective. However, life insurers are more sensitive to changes in long-term interest rates than property and casualty insurers. We show that the term premium helps to explain the difference in sensitivities between the two types of insurer.

JEL Codes: G20; C58

Keywords: financial institutions; interest rate risk management; high-frequency financial econometrics; subsampling; life insurers.

*For providing valuable comments, we would like to thank, without implication, Mark Carey, Burcu Duygan-Bump, Peter Hansen, Max Huber, Anastasia Kartasheva, Borghan Narajabad, Andrew Patton, Matt Pritsker, Roberto Renò, Rich Rosen, Oleg Sokolinskiy, Pavel Szerszen and participants in the Society for Economic Measurement Annual Conference 2023, the International Risk Management Conference 2023, and seminars at the European Central Bank, St. Gallen University, and the Federal Reserve Board. We are grateful to Julia Silbert and Renee Garrow for exceptional research assistance. The views in this paper are solely 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.

1 Introduction

Financial intermediaries are exposed to interest rate risk. They have multiple sources of exposure arising from cash flow differences across balance sheet components as well as contractual or embedded options with asymmetric payoff characteristics. Although intermediaries have a wide range of asset and liability management tools available to hedge interest rate risk, they do not fully insulate themselves from all potential changes in interest rates for several reasons.[1] Financial markets may be incomplete, fully hedging may be prohibited by its cost, and carrying interest rate risk may be a source of earnings.[2] Thus, financial intermediaries carry some *residual* exposure to interest rate risk, which could have significant consequences for financial stability and macroeconomic outcomes in bad states of the world (Holmstrom and Tirole, 1997; Brunnermeier and Sannikov, 2014).

In this paper, we propose a new method to measure the time-varying residual interest rate risk exposure of financial intermediaries using minute-

[1]Risk managers at financial institutions are expected to monitor and manage interest rate exposures at prudent levels, but not fully eliminate the risk. Supervisors provide detailed guidance on management practices and coordinate their standards. See, for example, the Office of the Comptroller of the Currency (OCC) Revised Handbook March 2020, the Federal Deposit Insurance Corporation (FDIC) Letter on Financial Institution Management of Interest Rate Risk 2010, the Federal Reserve Board (FRB) Supervisory Manual on Interest Rate Risk, the National Association of Insurance Commissioners (NAIC) Risk-Based Capital for Insurers Model Act, the OCC-FDIC-FRB Joint Policy Statement on Interest Rate Risk 1996, and the Basel Committee on Banking Supervision Guidance on Standards 2014.

[2]Even an established hedging strategy may be exposed to "basis risk"—that is, it might lose its effectiveness.

by-minute financial market data. We calculate the daily realized covariance of high-frequency stock returns for those intermediaries and Treasury security returns. We construct a conditional covariance by projecting out aggregate stock market returns from stock returns and Treasury security returns. We then introduce *realized gamma* as the ratio of the conditional covariance to the daily realized conditional variance of Treasury security returns. *Realized gamma* is a daily estimate of the sensitivity of an individual firm's stock price returns to realized changes in interest rates. We calculate returns at five-minute intervals using every possible five-minute grid point in a trading day, exploiting all available high-frequency information as described in Zhang, Mykland and Aït-Sahalia (2005).

We then propose a new statistical test of the daily residual interest rate risk exposure of financial intermediaries. We conduct statistical inference on the *realized gamma* estimates by calculating asymptotically valid confidence intervals using subsampling (Politis, Romano and Wolf, 1999). The essence of the subsampling method is to approximate the sampling distribution of the daily *realized gamma* with the empirical distribution generated by estimating the *realized gamma* on an exhaustive set of intra-day subsamples.[3] Although computationally intensive, the method of subsampling behaves well under extremely weak, easily satisfied

[3] Our limiting concept is the length of the time interval between two stock price observations going to zero. We provide the main theoretical results for our application in Appendix B.

assumptions.[4] Our approach to statistical inference is crucial because it is by definition impossible to know everything about each financial intermediary's proprietary risk management framework.

Our new method provides a time-varying measure of residual interest rate risk exposure because it is based on financial intermediaries' publicly-traded equity values. Those values reflect intermediaries' exposure to interest rates *after* they have executed their interest rate risk management strategies. The correlation of equity values with interest rates reveals market participants' views on the effectiveness of financial intermediaries' hedging strategies in relation to the changes in the interest rates that occurred. The measure is a reflection of the hedging strategy conditional on the actual changes in interest rates. A measure of zero doesn't necessarily mean that financial intermediaries are fully hedged. That said, intuitively, the stock price of a financial intermediary with fully hedged interest rate risk would be uncorrelated with all possible changes in interest rates (Allen, 1993).

Note that we will not address the question of *why* financial intermediaries bear interest rate risk. Importantly, we are not making any normative statement about how much interest rate risk financial intermediaries could or should carry. In particular, our notion of effectiveness does not imply

[4]By contrast, bootstrapping the confidence intervals would require showing the time-series properties were preserved within samples or impose strong assumptions about the data generating process.

that intermediaries should aim for zero residual interest rate risk. Nor does it imply that market participants think intermediaries should do so. Rather, our measure derives from the compensation for the interest rate risk borne by the ultimate owners of the intermediary, as in Allen (1993). When ownership is obtained through traded equity, the equity market price reflects that compensation.

Monitoring residual interest rate risk exposures is an important component in analysts, policymakers, and supervisors' evaluation of the financial conditions of intermediaries. Interest rate risk exposures are typically included as part of credit rating reports and investment analysis. As part of their financial stability dicussions, central bankers are attuned to the potential effects on their decisions on financial intermediaries, e.g., Brainard (2022). Supervisors of financial institutions expect regular reports concerning interest rate risk management and exposures. Monitoring is required because interest rates can change swiftly and significantly, with large potential effects. The profitability of entire financial sector industries has been threatened by interest rate exposures. For example, the life insurance industry struggled to cope with the sharp rise in interest rates in the late 1970s and early 1980s, when the Federal Reserve under Chairman Volcker fought inflation (NAIC, 2013).

We apply our new method to publicly-listed U.S. life insurers during the period from 2007 to 2022. Interest rate risk management is at the heart of the modern life insurer business model because the duration of life insurers' insurance liabilities, such as life insurance policies and annuity contracts, is typically much longer than the duration of the assets available in the economy.[5] This negative duration gap means that a decrease in the interest rate 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 if left unmanaged. The same duration gap also means that persistently low interest rates depress life insurers' net investment spread on new business and forces them to reinvest the proceeds from maturing bonds into bonds paying lower coupon rates, which further depresses their overall net investment spread and, in turn, adversely affects their financial condition. In addition, explicit and implicit options on both assets and liabilities contributes to life insurers' interest rate risk. Because the prospect of insolvency is incompatible with the sale of long-term life and longevity insurance, state insurance regulations, or both, life insurers must credibly manage interest rate risk.

We find that life insurer stock prices are largely *uncorrelated* with long-term (10-year) Treasury interest rates. This suggests that life insurers' interest rate risk management is effective most of the time. This finding is comforting given some of the largest life insurers in the U.S. have been managing interest rate risk for over a century. However, in some states of

[5] For example, the duration of a typical life annuity is ten years, while the median corporate bond duration is around 5 years.

the world, realized gamma is statistically significant, revealing that after managing their interest rate risk—with liability driven investment, capital structure, and derivatives—life insurers remain exposed to changes in long-term interest rates in some states of the world.

We contrast our analysis of life insurance companies with an analysis of publicly-listed property and casualty (P&C) insurance companies. P&C insurers provide an ideal alternative to life insurers because the structure of their business means that they are relatively less exposed to interest rate risk. For example, the vast majority of P&C premiums are renewable every year and, therefore, P&C insurers do not need to actively manage a duration gap between their assets and insurance liabilities. Consistent with this difference in business model, we find that life insurers are more sensitive to changes in long-term interest rates than P&C insurers.

We then show that a measure of the term premium—the compensation for the risk associated with holding longer-term bonds—helps to explain the difference between the estimated sensitivities of life insurers and P&C insurers. We use the estimate of the term premium from the term structure model of Adrian, Crump and Moench (2013). We control for the funding cost of life insurers and a measure of the corporate credit return on life insurers' assets. Our finding likely reflects the outsized importance of longer-term debt in life insurers' investment portfolios. We use these results to illustrate how our measure provides information about the impact that rapidly changing interest rates may have on insurers.

Lastly, we show that our finding that life insurers' interest rate risk management is generally effective is *not* due to low long-term interest rate volatility. We provide two alternative approaches to address the potential endogeneity between realized gamma and long-term interest rate volatility. Both approaches are based on the exogenous increase in interest rate volatility that occurs on scheduled Federal Open Market Committee (FOMC) meeting days.

1.1 Related literature

Our paper connects to three distinct strands of literature. First, our method contributes to the high-frequency financial econometrics literature. Conceptually, our method is an extension of the single-factor realized beta model of Andersen, Bollerslev, Diebold and Wu (2006) and Hansen, Lunde and Voev (2014). We include a second right-hand side variable, that is Treasury security returns, in the estimated regression specification. To the best of our knowledge we are the first to introduce a second right-hand side variable. Our computation of asymptotically valid standard errors using the subsampling approach is unusual in the high frequency financial econometrics literature because the approach is conservative and computationally intensive. Our realized gamma estimates do not suffer from bias due to non-synchronous trading—see, for example, Christensen, Kinnebrock and Podolskij (2010) and Barndorff-Nielsen, Hansen, Lunde and Shephard (2011)—since we use index data aggregated at the one-minute frequency. Our choice of five-minute sampling frequency and averaging should also immunize our estimates from market microstructure noise biases.

Second, our method relates to—but is distinct from—studies of interest rate risk that measure the effects of realized changes in interest rates. These studies differ from other interest rate risk assessments that use balance sheet information to describe scenarios of potential effects associated with *hypothetical* changes in interest rates, e.g., Möhlmann (2021). Other papers that study *actual* changes in interest rates tend to focus on banks. Flannery and James (1984) studies the correlation between bank stock prices and interest rates using a similar regression model and weekly data. English, Van den Heuvel and Zakrajšek (2018) identify the response of bank stock prices to FOMC interest rate shocks. Paul (2022) revisits the findings of English et al. (2018) by decomposing the effect of monetary policy surprises into changes in future expected short-term rates and changes in term premium. Hoffmann, Langfield, Pierobon and Vuillemey (2018) use supervisory bank balance sheet data to estimate interest rate risk and study its determinants in the cross section. Vuillemey (2019) and Begenau, Piazzesi and Schneider (2015) show that banks increase their exposure to interest rate risk using derivatives. Most of these papers study low-frequency data and, in some cases, attempt to identify interest rate shocks. By contrast, we exploit the information in high-frequency data and we use changes in interest rates rather than identified shocks.

Third, our paper adds to the extensive literature on risk management of financial institutions—e.g., Froot, Scharfstein and Stein (1993); Froot and Stein (1998). Our method is applicable to any financial intermediary. We chose to focus on the interest rate risk of life insurers, as they have received much less attention than, for example, banks. The theoretical foundation for our application to life insurers comes from recent work studying interest rate risk management at insurance companies (Foley-Fisher, Narajabad and Verani, 2016; Verani and Yu, 2021). In these papers, limited liability insurers manage the ex-ante risk of insolvency due to future movement in the interest rate by choosing an optimal insurance price, asset portfolio, and capital structure. Our method is an ex-post statistical test of the performance of insurers' ex-ante interest rate risk management strategy. As such, our analysis is closely related to empirical work that measures the residual interest rate risk exposure of insurers using a two-variable regression model of stock prices and low-frequency data (Brewer III, Mondschean and Strahan, 1993; Berends, McMenamin, Plestis and Rosen, 2013; Hartley, Paulson and Rosen, 2016; Ozdagli and Wang, 2019; Sen, 2021; Koijen and Yogo, 2022; Huber, 2022).[6] We show in

[6]Some of these papers use stock prices only as a motivation for subsequent analysis of insurer balance sheet measures of interest rate risk.

Appendix A that estimates obtained through low-frequency rolling window ordinary least squares (OLS) regressions are severely biased, inconsistent, and potentially misleading. For example, in contrast to our findings, inference on the OLS estimates suggests that life insurers are sensitive to any movement in long-term interest rates at almost all times in the post-crisis period.

The rest of our paper is structured as follows: Section 2 sets out the empirical framework for our estimation and explains how we construct our standard errors using subsampling. Section 3 describes our application to US life insurers, including institutional background and details on the data. We summarize our main findings in section 3.3 and offer some concluding remarks in section 4.

2 Methodology

2.1 A two-variable regression model

In this section, we introduce our new method to measure the residual interest rate risk exposure of financial intermediaries. Let r_{ijt} be the stock return of financial intermediary i indexed to minute j within day t. Let r_{mjt} be the return on aggregate market m and r_{yjt} be the return on Treasury security y.

Our framework is a regression model with two right-hand side variables using minute-by-minute financial market returns:

$$r_{ijt} = \alpha_t + \beta_t r_{mjt} + \gamma_t r_{yjt} + \epsilon_{ijt} \tag{1}$$

where $\{\alpha_t,\ \beta_t,\ \gamma_t\}$ are day-specific coefficients estimated using within-day returns. Our regression with the restriction $\gamma_t = 0$ is well established in the finance literature and is referred to as the one-factor capital asset pricing model (CAPM). In the CAPM regression, the coefficient β_t is interpreted as a dynamic measure of the comovement of individual stock returns with aggregate market or systematic returns.[7] We extend the one-variable CAPM regression to include a second right-hand side variable, that is Treasury security returns.[8]

The time-varying γ_t coefficient estimates the sensitivity of an individual firm's stock price returns to high-frequency realized changes in Treasury security returns. As Treasury security returns are inversely dependent on changes in interest rates, γ_t provides an estimate of that firm's interest rate sensitivity.

We label our γ_t estimates *realized gamma* because our method

[7]A full discussion of the extensive literature studying time-varying β_t and its determinants is beyond the scope of this paper. See Fama and French (2004) for an overview.

[8]To be sure, we are not assuming that the right-hand side variables in our regression model are orthogonal. We are estimating the general equilibrium relationship between the three variables in our regression, which is fully consistent with the standard one-factor CAPM and yields an unbiased estimate of γ_t. Other papers that adopt a similar approach include Fama and Schwert (1977) and Flannery and James (1984). We explore the effect of an exogenous increase in long-term Treasury rate volatility in section 3.4.

can also be cast in the nonparametric framework of realized variances and covariances (Meddahi, 2002; Barndorff-Nielsen and Shephard, 2004; Andersen, Bollerslev and Meddahi, 2004). Our estimates of daily gammas are based on realized daily variances and covariances after conditioning on the aggregate market returns. We first project out aggregate stock market returns from stock returns and Treasury security returns by running two auxiliary regressions for each day t:

$$r_{ijt} = \hat{\alpha}_t^1 + \hat{\beta}_t^1 r_{mjt} + \hat{\epsilon}_{ijt}, \tag{2}$$

$$r_{yjt} = \hat{\alpha}_t^2 + \hat{\beta}_t^2 r_{mjt} + \hat{\epsilon}_{yjt}. \tag{3}$$

The residuals from these auxiliary regressions $\{\hat{\epsilon}_{ijt}, \hat{\epsilon}_{yjt}\}$ are, respectively, the within-day conditional stock returns and Treasury security returns. The daily realized covariance of each financial intermediary's conditional stock returns and Treasury security conditional returns is given by:

$$\hat{v}_{i,y,t} = \sum_j \hat{\epsilon}_{ijt} \cdot \hat{\epsilon}_{yjt}.$$

And the daily realized variance of conditional Treasury security returns is given by:

$$\hat{v}_{y,t} = \sum_j \hat{\epsilon}_{yjt}^2.$$

So we can define realized gamma as the ratio of the conditional covariance to the daily realized conditional variance of Treasury security returns:

$$\gamma_t = \frac{\hat{\nu}_{i,y,t}}{\hat{\nu}_{y,t}}. \tag{4}$$

The γ_t estimates by equation 1 and equation 4 are identical by what is commonly-known as the Frisch-Waugh-Lovell Theorem (Davidson and MacKinnon, 1993, Section 1.4). However, care must be taken with interpretation. The simple regression shown in equation 1 yields a consistent estimate of the ex-post realized gamma coefficient. That said, obtaining asymptotically valid standard errors is not a simple process, as we will describe in Section 2.2.

2.1.1 Addressing market microstructure noise

Controlling for market microstructure noise that is prevalent in high frequency financial market data is an important issue (Aït-Sahalia and Yu, 2009). Microstructure noise naturally arises from a variety of features built in to financial market trading, including prices bouncing from bids to asks, variation in the size of trades, adjustment to new information contained in prices, order flow dynamics, and inventory management. Following Aït-Sahalia and Mykland (2009), we address the presence of market microstructure noise without discarding observations from our samples.

We employ two well-established techniques to mitigate concerns that market microstructure noise is clouding our ability to construct estimators and draw inference from high-frequency data. First, we calculate returns at five-minute intervals as their use as a benchmark for estimators generally outperforms all alternatives (Liu, Patton and Sheppard, 2015). We use every possible five-minute grid point in a trading day to exploit all available high-frequency information given the data structure as described in Zhang et al. (2005).[9] Second, we filter all of our returns time series through AR(1) processes estimated separately for each day. That is, we take the raw returns r_{kjt} for $k \in \{i, y, m\}$ and estimate $r_{k,j,t} = \rho + \phi r_{k,j-1,t} + \varepsilon_{k,j,t}$ for each day t. We then use the residuals $\varepsilon_{k,j,t}$ as our returns time series that has filtered out market microstructure noise.

2.2 Statistical inference

A key principle for our new methodology is to impose minimal assumptions about the data generating process. This principle underpins our use of high-frequency data to estimate nonparametrically the time-varying correlation between interest rates and financial intermediaries' stock prices. Similarly, we follow this principle when we consider what standard errors are appropriate for valid inference. We derive asymptotically valid standard

[9] Our approach is identical to the method commonly referred to as "subsampling" in high-frequency financial econometrics. We avoid using the term here to prevent confusion with the concept of "subsampling" that we use to construct asymptotically valid standard errors.

errors without imposing undue structure on the time series processes. Our choice of standard errors is a crucial part of our approach to estimate interest rate risk, as the data generating process underpinning our realized gamma estimates is nonstandard. For example, we use rolling five-minute windows to construct our time series of returns.

We adopt the subsampling methodology as it is a valid technique in extremely general cases (Politis et al., 1999).[10] The basic idea of subsampling in a time series context is to approximate the sampling distribution using all possible subsets of the time series. Theorem 4.3.1 from Politis et al. (1999), which we reproduce in Appendix B for completeness, shows that we can derive asymptotically valid confidence intervals for the daily estimator γ_t. In addition, we can draw asymptotically valid inference about the true γ_t by exploiting the familiar duality between the construction of confidence intervals for γ_t and the construction of hypothesis tests about γ_t. We can test the null hypothesis that our estimate of the daily γ_t is statistically different from 0. That is, under the null, financial intermediaries are hedged against interest rate risk as their stock prices are not sensitive to movements in interest rates.

Our algorithm for hypothesis testing uses within-day observations to construct subsamples. We follow Politis et al. (1999) and evaluate statistics

[10] Alternative methodologies based on the bootstrap technique could be devised, but they typically require additional assumptions, such as a finite fourth moment of the model residuals (Paparoditis and Politis, 2009).

on an exhaustive set of subsamples of size $b < n$ that are created from the original daily sample of size n. We estimate the distribution of this statistic after a suitable normalization for each day in our sample. Note that our limiting concept is that the number of observations in a day approaches infinity. To be clear, each subsample contains consecutive observations from the original time series sample. Therefore, each subsample of size b is drawn without replacement from the true data generating process. We calculate a confidence interval for each of the daily γ_t using subsampling following Politis et al. (1999) under the assumption that the errors are asymptotically stationary. Asymptotic stationarity is a weak condition that means, for example, the errors could follow an $AR(1)$ process with autocorrelation parameter strictly less than 1 and heteroskedastic innovations. The essence of the subsampling method is to approximate the sampling distribution of the (normalized) γ_t estimate with the empirical distribution generated by its subsample counterpart.

As we have a large daily sample size (roughly speaking, $n = 390$ observations per day), the choice of subsample size (b) should not have a large effect on the empirical distribution of our statistic. Nevertheless, we need to choose the size of our subsamples. We follow the algorithm proposed by Politis et al. (1999) in section 9.3.3. Let b^t be the subsample size for day t, which yields a confidence interval $\{I_{b^t,low}, I_{b^t,high}\}$. We construct a discrete grid of possible values $b^t_s \in \{b^t_{small}, \ldots, b^t_{large}\}$. For each subsample size b_s^t we consider a perturbation of small integer k around the subsample size and calculate a measure of variation in the confidence interval:

$$VI_{b_s^t} \equiv \text{var}\left(I_{b_s^t-k,low}, ..., I_{b_s^t+k,low}\right) + \text{var}\left(I_{b_s^t-k,high}, ..., I_{b_s^t+k,high}\right).$$

Finally, we pick the value of b that delivers stable confidence intervals for the most number of days in the entire sample:

$$b = \text{argmax}_b \sum_{t=0}^{T} \mathbb{1}(b^{t*} = b) \text{ where } b^{t*} = \text{argmin}_{b_s^t} VI_{b_s^t}.$$

Having determined the 'optimal' subsample size, we construct the empirical distribution of the normalized γ_t estimate for each day t. We use empirical distributions to obtain confidence intervals, which allow us to make inference about the statistical significance of each γ_t. With our new methodology in hand, we can turn to a specific application and data.

3 Application to U.S. life insurers

3.1 Institutional background

Life insurers play a major role in the financial system, holding \$6 trillion in total assets in their general accounts, of which roughly \$3 trillion are in corporate and foreign fixed income securities (Federal Reserve release Z.1 table L.116.g). Their overall business model consists of earning a spread between the yield they owe on their insurance liabilities and the yield they earn on the assets backing those liabilities. Life insurers write liabilities that are traditionally long-term, illiquid, and make fixed payments, such as fixed annuities. Life insurers tend to invest their premiums primarily in fixed rate corporate debt, in an effort to match their asset and liability cash flows and illiquidity profile and to offer a competitive return to policy holders.

Like other financial intermediaries, life insurers have multiple sources of exposure to interest rate risk. A key underlying reason for their exposure is that the duration of insurance liabilities is typically much longer than the duration of assets available in the economy. In the U.S., the typical duration of life insurance liabilities is 15–20 years (Huber, 2022). By contrast, in most countries, long-term fixed coupon bonds with more than two-year maturity do not exist (Gajek and Ostaszewski, 2004). Even in the U.S., which has the largest corporate bond market in the world, the supply of long-duration corporate bonds paying fixed interest rates is considerably smaller than the size of the life insurance industry (Verani and Yu, 2021). This means that, in practice, it is difficult for life insurers to hedge interest rate risk by investing in assets that have the same duration and greater cash flow variability than their insurance liabilities i.e., they cannot directly implement the classical immunization strategy of Redington (1952).

Convexity—the effect of changing interest rates on the duration—of life insurer assets and/or liabilities also contributes to interest rate risk. One well-known source of convexity stems from options on financial contracts. For life insurers, the option for corporate bond issuers to call their bonds creates convexity on the asset side of their balance sheet. Likewise, policyholders may have the option of surrendering their life insurance products–perhaps for some cost–that creates convexity on the liability side of the balance sheet. The combination of these options creates a short straddle position for investors in the life insurer, which means they suffer when volatility is high (Babbel and Stricker, 1987).[11]

A natural way for life insurers to manage their interest rate risk consists of choosing a price for their insurance liabilities, an asset portfolio to back their insurance liabilities, and a capital structure to prevent insolvency along different paths for interest rates (Verani and Yu, 2021). For example, life insurers can hedge interest rate risk by charging a markup on the actuarially fair cost of their insurance products. The present value of the markup adds to the insurer's 'net worth'. Net worth allows the insurer to close its duration gap by financing bonds whose present value is greater than the present value of its insurance liabilities. Or, put differently, net worth acts as precautionary savings and helps cushion the effect of interest

[11]Briys and de Varenne (1997) provide an alternative formulation for the investor straddle position in which insurance liabilities are more convex than assets.

rate changes that disproportionately affect the value of the insurance liabilities.[12]

Large and sophisticated life insurers also manage interest rate risk by adding net-positive duration to their balance sheets synthetically using derivatives (Sen, 2021; Verani and Yu, 2021) and nontraditional lines of business (Foley-Fisher et al., 2016; Foley-Fisher, Narajabad and Verani, 2020), which amounts to using leverage instead of net worth to close the duration gap. For example, life insurers can add positive duration to their balance sheet by entering into a long-term *fixed-for-float* interest rate swaps or by financing long-term fixed interest rate assets with nontraditional liabilities such as overnight securities lending cash collateral (Gissler, Foley-Fisher and Verani, 2019; Foley-Fisher et al., 2016) and funding agreement-backed short-term funding (Foley-Fisher et al., 2020). All these interest rate hedging strategies amount to closing the insurer's natural negative duration gap by either directly or synthetically financing fixed-maturity assets with short-term floating rate debt.

Nevertheless, insurers typically carry residual interest rate risk after they have implemented their hedging strategies. Investors in the insurers—either the policyholders in the case of mutual insurers or shareholders in the case of publicly-listed insurers—provide additional risk-bearing capital and receive compensation for bearing the insurer's residual interest rate

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

risk (Allen, 1993). When investment takes place through traded equity, the market price for the equity reflects the interest rate risk compensation.

One real-world example when the residual interest rate risk carried by life insurers was realized occurred in the early 1980s. At that time, the Federal Reserve sharply increased short-term interest rates amid persistently high inflation. Life insurers' financial condition deteriorated as policyholders surrendered their claims or took out policy loans in search of higher interest rates on alternative saving vehicles (Briys and de Varenne, 2001). Life insurers responded by rewriting existing business at a loss and selling new products that offered higher-than-current long-term rates (negative spreads) (NAIC, 2013). While locking in huge losses—eroding their net worth—they avoided even greater losses they would have incurred had they sold their fixed income assets at far-below costs given the rise in current rates. The surge in short-term interest rates occurred after a relatively long period of low interest rate volatility, making these sharp rises largely unexpected. The significance of the episode is underscored by subsequent efforts to develop new tools for managing interest rate risk (Doffou, 2005).

Adverse scenarios such as the early 1980s create a need for researchers and policymakers to monitor and assess the effects of rising interest rates on life insurers. However, they do not have access to the complete set of balance sheet information needed to precisely identify the effectiveness of life insurers' interest risk management and their residual interest rate risk. For example, information about the interest rate sensitivity of life insurance liabilities is difficult to gauge, although it is easier in some non-U.S. jurisdictions (Huber, 2022; Möhlmann, 2021; Kirti, 2017; Domanski, Shin and Sushko, 2017). Furthermore, it is hard to incorporate balance sheet information about the interest rate sensitivity of derivative positions, off-balance sheet liabilities (such as those in offshore captive reinsurers), and nontraditional liabilities.

To overcome this problem, researchers turned to analyzing the sensitivity of life insurer stock returns to changes in long-term interest rates. An insurer's equity valuations reflect the market price for its residual interest rate risk, after it has implemented its hedging strategies. That is, the ex-post effectiveness of life insurers' management of ex-ante interest rate risk.[13] To the best of our knowledge, this approach was first adopted by Brewer III et al. (1993). To assess the *dynamics* of interest rate risk exposure, some papers run OLS on rolling windows of stock returns e.g., Hartley et al. (2016). Although conceptually valid, the OLS implementation can lead to biased estimates in the presence of heteroskedasticity.[14] In Appendix A, we show the bias is extremely large

[13] Here, again, the term 'effectiveness' should not be taken to imply that investors think insurers should target any particular level of interest rate risk. Rather, it's investors' assessment of the effect that actual interest rate changes had on the net worth of the insurer.

[14] Brewer III, Carson, Elyasiani, Mansur and Scott (2007) recognised this concern and allowed for time-varying volatility in a GARCH-M process.

by imposing some structure on the data generating processes. We will now apply our preferred methodology described in Section 2 to obtain consistent estimates without imposing such structure.

3.2 Data

All the price data for our empirical application to life insurers come from Refinitiv. The underlying data are timed to the microsecond and recorded from data feeds covering both over-the-counter and exchange traded instruments on more than 500 trading venues and third parties. We use a preprocessed version of the underlying data aggregated by Refinitiv to a minutely frequency using the last trade during each minute. We construct the data so as to follow the *previous tick method*, that is, if there are no transactions during a specific minute, the last transaction is used.

The dataset identifier for each dataseries typically combines a ticker with a code indicating the primary trading market. For example, MetLife's identifier is `MET.N` as it trades on the New York Stock Exchange. The list of the individual insurer identifiers and their mapping to the life and P&C insurers used in our analysis is provided in Table 1. Column 2 of Table 1 shows the insurers included in each index. Our list of publicly-listed life insurers almost completely overlaps with the list of "publicly traded U.S. variable annuity insurers" used by Koijen and Yogo (2022). This is not surprising because virtually all large listed life insurers offer variable annuities contracts at some point in the sample period.[15] In addition, our analysis uses Standard and Poor's S&P500 index as our measure of the aggregate market. The identifier for the index is `.SPX`. We also use Refinitiv evaluated prices for 10-year Treasury securities. The identifier for the series is `US10YT=RRPS`. Evaluated prices contain information from actual trades, quotes, and other sources within a model-based methodology.

We use minutely data for each trading day beginning at 9:30am through 4pm.[16] Except for 9:30am, we use closing prices recorded for each minute. For 9:30am, we use the opening price of 9:31am to avoid concerns about jumps following overnight information and trading. We calculate five-minute log returns of all time series using every possible five-minute grid point in a trading day. That is, we calculate the returns $ln(p_{i,j,t}) - ln(p_{i,j-5,t})$ for each day t, data series i, and all $j \in \{9.35am,\ 9.36am,\ \cdots,\ 3.59pm,\ 4.00pm\}$.

We construct high-frequency price indexes separately for life insurers and P&C insurers, weighting each individual insurer's intraday market price by its end-of-day market capitalization. We obtain daily data on market capitalization from the Center for Research in Security Prices hosted by Wharton Research Data Services. Figure 1 shows the life insurer index as a red solid line and the P&C insurer index as a dotted blue

[15]As we will discuss in Section 3.4, this means that it is not possible to attribute the residual interest rate risk exposure to variable annuities.

[16]We exclude holidays, weekends, emergency closures, and partial trading days.

Table 1: Mapping insurance groups to identifiers. This table shows the insurance groups that we use in our empirical application, with their respective NAIC Group codes, and identifiers.

Name	Code	Life/P&C	Identifier	Ticker	Notes
Alleghany Group	501	P&C	Y.N	Y	
American Financial Group	84	P&C	AFG.N	AFG	
American Intl Group, Inc.	12	Life/P&C	AIG.N	AIG	
Assurant, Inc.	19	Life/P&C	AIZ.N	AIZ	
The Allstate Corporation	8	Life/P&C	ALL.N/BK.N	ALL/BX	Identifier change for Life in 2021
Ameriprise Financial, Inc.	4	Life	AMP.N	AMP	
American National Financial Group	408	Life	ANAT.OQ	ANAT	
Apollo Global Management, Inc.	4734	Life	ATH.N/APO.N	ATH/APO	Identifier change in 2022
Brighthouse Financial, Inc.	4932	Life	BHF.OQ	BHF	
Berkshire Hathaway Inc.	31	P&C	BRKb.N	BRK.B	
Chubb Ltd.	626	P&C	ACE.N/CB.N	ACE/CB	Identifier change in 2016
Cigna Health Group	901	Life	CI.N	CI	
Cincinnati Financial Corporation	244	P&C	CINF.OQ	CINF	
CNA Financial Corporation	218	P&C	CNA.N	CNA	
CNO Financial Group	233	Life	CNO.N	CNO	
Erie Insurance Group	213	P&C	ERIE.OQ		
Equitable Holdings, Inc.	4965	Life	EQH.N	EQH	
FBL Financial Group Inc.	513	Life	FFG.N	FFG	Ceased trading in 2021
Fidelity and Guaranty Life	4731	Life	FGL.N	FGL	Ceased trading in 2017
Fidelity National Financial, Inc.	670	Life	FNF.N	FNF	
Genworth Financial, Inc.	4011	Life	GNW.N	GNW	
Hanover Insurance Group, Inc.	88	P&C	THG.N	THG	
The Hartford Fin. Svcs Group, Inc.	91	Life/P&C	HIG.N	HIG	Remove identifier from Life in 2018
Horace Mann Group	300	Life	HMN.N	HMN	
Kansas City Life Insurance Group	588	Life	KCLI.OQ	KCLI	Delisted in 2015
Kemper Corporation Group	215	P&C	KMPR.N	KMPR	
Lincoln National Corporation	20	Life	LNC.N	LNC	
Mercury General Group	660	P&C	MCY.N	MCY	
Markel Corporation Group	785	P&C	MKL.N	MKL	
MetLife, Inc.	241	Life	MET.N	MET	
Manulife Financial Corporation	904	Life	MFC.TO	MFC	
Nationwide Corporation Group	140	Life	NFS.N	NFS	Ceased trading in 2008
The Phoenix Companies, Inc.	403	Life	PNX.N	PNX	Ceased trading in 2016
Primerica Group	4750	Life	PRI.N	PRI	
Principal Financial Group, Inc.	332	Life	PFG.OQ	PFG	
Protective Life Corporation	458	Life	PL.N	PL	Ceased trading in 2015
The Progressive Corporation	155	P&C	PGR.N	PGR	
Prudential Financial, Inc.	304	Life	PRU.N	PRU	
Selective Insurance Group	88	P&C	THG.N	THG	
Symetra Financial Corp.	4855	Life	SYA.N	SYA	Ceased trading in 2016
The Travelers Companies, Inc. Group	3548	P&C	TRV.N	TRV	
Voya Financial, Inc.	4832	Life	VOYA.N	VOYA	
W. R. Berkley Corporation	98	P&C	BER.N/WRB.N	BER/WRB	Identifier change in 2008

line. The dotted blue line lies above the red solid line as P&C insurers have generally outperformed life insurers in the post-crisis low interest rate environment.

Figure 1: Insurer price indexes. Each line is a weighted average high-frequency price for large publicly-traded insurers listed in Table 1. The weights for each series are the daily market capitalization of insurers. Source: Authors' calculations based on data from Refinitiv and the Center for Research in Security Prices.

Table 2 shows summary statistics for the high-frequency data used in our analysis. Column 1 shows that the first day that data are available is different for each of our variables. The S&P500 Index is earliest available, while our high frequency data on long-term Treasury bond prices ('US10YT') begin only in 2007. Our indexes of large life and P&C insurers stock prices also begin in 2007. By construction, our sample ends on October 31, 2022. In addition to the first date available, we report the number of days, and the total number of minutely five-minute returns in our data. We also report that there are no zero returns in our data, alleviating concerns about downward bias in our estimates due to zero returns (Bandi, Kolokolov, Pirino and Renò, 2020; Kolokolov and Renò, 2023). Across these returns, we report the mean, median, standard deviation, percentiles, and higher-order moments for each time series. Life insurers' returns have a higher standard deviation than P&C insurers, but the kurtosis of life insurers' returns is far lower.

Table 2: Summary statistics. For each returns series in our sample, the table shows the first observation date, the number of days, the number of five-minute returns, the number of returns equal to zero, as well as the mean, median, standard deviation, percentiles, skewness, and kurtosis. The statistics reported in columns 6 through 10 are multiplied by $1e+4$ for legibility. Source: Authors' calculations based on data from Refinitiv and the Center for Research in Security Prices.

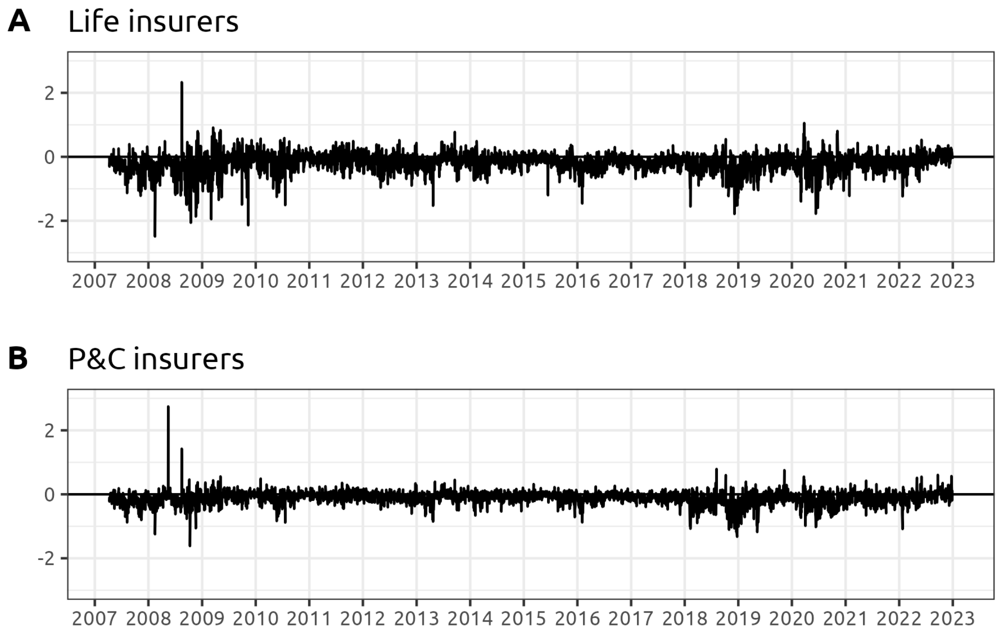
Series	First date	No. days	No. obs.	#Zeroes	Mean	Median	Std. dev.	p25	p75	Skew.	Kurt.
S&P500	2000-01-03	5,768	2,218,825	0	-0.01	0.05	10.74	-4.06	4.11	-0.04	38.74
US10YT	2007-04-10	3,961	1,524,682	0	0.01	0.00	3.91	-1.67	1.70	0.99	80.66
Life	2007-01-03	3,996	1,533,761	0	-0.01	0.06	16.55	-5.76	5.79	0.44	35.25
P&C	2007-01-03	3,996	1,533,761	0	0.02	0.03	10.44	-4.00	4.02	2.06	147.25

3.3 Results

In this section, we apply the methodology laid out in Section 2 to the data described in the previous section. Panel A of Figure 2 shows the daily point estimate of realized gamma for life insurers, and Panel B of the same figure shows the daily point estimate of realized gamma for P&C insurers.

Both panels exhibit volatility, which is a well-known feature of time-varying coefficients estimated using realized variances and covariances (Hansen et al., 2014). Nevertheless, life insurers' realized gamma evidently has a higher level of volatility than P&C insurers'.

Figure 2: Daily realized gammas. The panels show daily realized gammas for life insurers and P&C insurers from 2007 through to the end of 2022. Source: Authors' calculations based on data from Refinitiv and the Center for Research in Security Prices.



We obtain confidence intervals from the empirical distributions, which are estimated for each day. Table 3 shows the results from applying the algorithm described in subsection 2.2 to determine the block size. While any block size satisfying the conditions of Theorem B.1 is valid, the ideal block size is the one that produces the most stable i.e., least variable,

Table 3: Optimizing block size. Each of columns 2-4 shows the number of days on which the block size (column 1) produces the most stable, i.e. least variable, confidence intervals. The measures of variation used in columns 2 and 4 is the standard deviation, while columns 3 and 5 use the difference between the minimum and the maximum values. The row with the highest count of days reveals the ideal block size for life insurers (columns 2-3) and P&C insurers (columns 4-5). Source: Authors' calculations based on data from Refinitiv and the Center for Research in Security Prices.

Block	Life		P&C	
size (%)	Std. dev.	Min-max	Std. dev.	Min-max
15	481	475	519	513
20	419	418	381	379
25	992	1000	988	995
30	676	673	702	700
35	597	594	618	615
40	472	480	452	454
45	418	415	395	399

confidence intervals given a small perturbation in the size of the block.[17] Each of columns 2-4 shows the number of days on which the block size (column 1) produces the most stable confidence intervals. The measure of variation used in columns 2 and 4 is the standard deviation, while columns 3 and 5 use the difference between the minimum and the maximum values. The row with the highest count of days reveals the ideal block size for life insurers (columns 2-3) and P&C insurers (columns 4-5). For both life insurers and P&C insurers, the optimal block size is 25 percent of the daily observations, corresponding to about 100 consecutive observations in each

[17] Note that there is no reason to expect variation across grid points to follow a monotonic function or have a global optimum (Politis et al., 1999).

block and about 300 points in the empirical distribution. These relatively large values alleviate concerns about the power of the test.

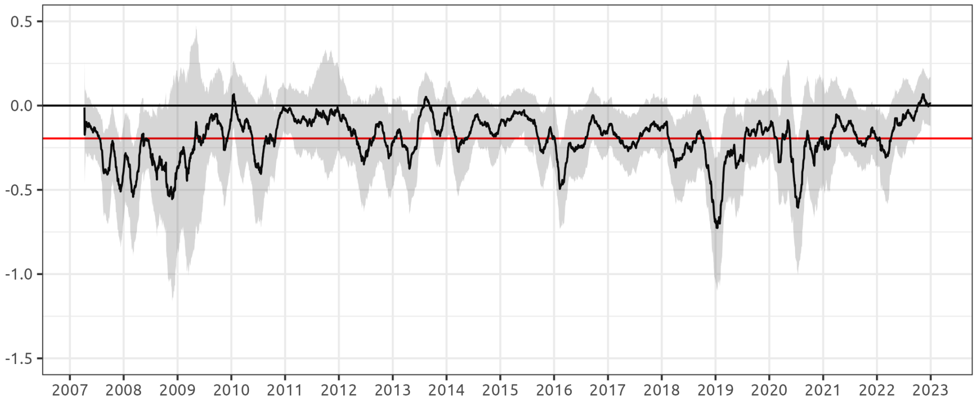
To smooth out the volatility in both the point estimates of realized gammas and the confidence intervals, we calculate time-series averages. The smoothed series are simply easier to read. We construct averages using a rolling window of two months. Panels A and B in Figure 3 show the smoothed time series. The red horizontal lines represent the sample means of the respective series.

The smoothed time series reveal that realized gamma for life insurers is statistically significant on only 1,261 days, equivalent to roughly 32 percent of the sample. Realized gamma is always negative whenever it is statistically significant, which means that life insurers would benefit from higher long-term interest rates. For the majority of our sample, life insurer stock prices are uncorrelated with long-term interest rates. This suggests that life insurers' interest rate risk management is effective most of the time. These results should not be interpreted as a normative assessment of life insurers' interest rate risk management, neither by us nor by equity market participants. The measure is a reflection of how actual changes in interest rates affected—or did not affect—equity investors in life insurers, who expect compensation for bearing interest rate risk.

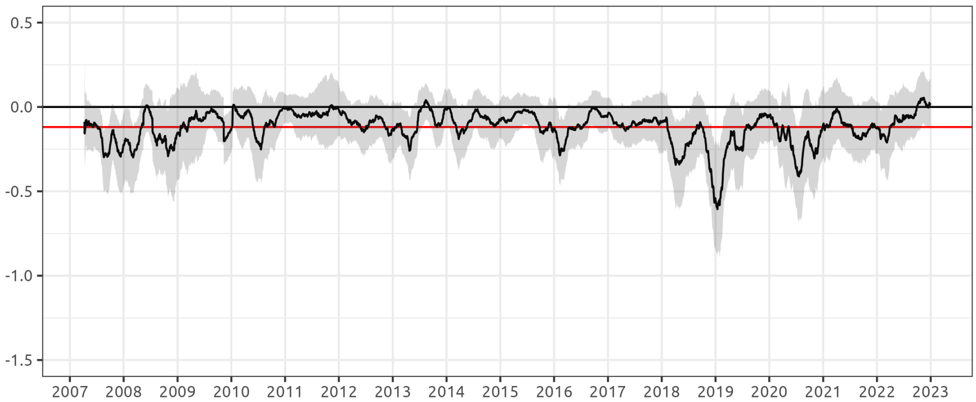
The time series also reveal that life insurers are more sensitive to interest rate changes than P&C insurers. In contrast to life insurers, realized

Figure 3: Smoothed daily realized gammas. The panels show daily realized gammas averaged using a rolling window of two months for life insurers and P&C insurers from 2007 through to the end of 2022. The shaded region in both panels represents the 90 percent confidence intervals for each daily estimate. The underlying data are shown in Figure 2. The red horizontal lines represent the sample means of the respective series. A negative realized gamma means that insurers would benefit from higher long-term interest rates. Source: Authors' calculations based on data from Refinitiv and the Center for Research in Security Prices.

A Life insurers



B PC insurers



gamma for P&C insurers is significant on only 705 days (about 18 percent of the sample). Like life insurers, realized gamma for P&C insurers is always negative whenever it is statistically significant. This finding could be interpreted as evidence that P&C insurers carry less residual interest rate risk, or as evidence that life insurers are exposed to different kinds of interest rate risk. In the next section, we offer some support for the latter interpretation by analyzing individual components of long-term interest rates.

3.4 Analysis

3.4.1 When is life insurer hedging not effective?

In this section, we study macroeconomic variables during periods when realized gamma is statistically significant. Our findings help to explain why life insurers' realized gamma is more often statistically significant that P&C insurers' realized gamma. While we offer an interpretation of our findings, we do not claim causal identification, as we recognize that long-term yields are a general equilibrium outcome of supply and demand (Schneider, 2022). Life insurers' interest rate sensitivity is potentially endogenous to their demand for compensation to hold longer-term debt and, as we noted earlier, life insurers are important investors in the long-term debt market. All the analysis in this section is conducted at a daily frequency out of necessity.

In an ideal empirical experiment, we would use intraday data to analyze the force(s) behind the results described in the previous section. However, we do not know of any high-frequency measures of the variables described below.

We focus on three key variables based on Verani and Yu (2021), who showed that the relative cost of hedging interest rate risk is determined by the long-term investment grade bond spread relative to life insurers' cost of funding. As measures of the return on life insurers' long-term assets, we use the term premium and Moody's Baa-Aaa seasoned corporate spread.[18] We use the term structure model of Adrian et al. (2013) to decompose long-term yields and obtain an estimate of the term premium. While the term premium contributes to the slope of the yield curve, it is more specifically the component that compensates investors for holding longer-term debt instead of rolling over short-term debt. In addition to these measures of asset returns, we use the ICE BoA Single-A U.S. corporate index option-adjusted spread as a proxy for life insurers' average cost of funding because life insurers are rated around A. Summary statistics for all the variables used in this analysis are provided in Appendix C.

We construct a binary variable that takes the value 1 if the estimated

[18]The corporate bonds used to construct this spread all have at least 20 years of maturity. The yield on Aaa-rated corporate bonds with at least 20 years of maturity is a quasi-risk free benchmark. Under state insurance regulation, corporate bonds rated by Moody's to be Baa or higher are designated as NAIC 1 and uniformly attract the lowest statutory risk-based capital charge.

realized gamma (γ_{it}) is statistically significant on day t for insurer type $i \in$ {Life, P&C}, and takes the value 0 otherwise. We then estimate a linear probability model using as independent variables the term premium (TP_t) estimates from Adrian et al. (2013), the Moody's Baa-Aaa seasoned corporate spread ($Baa - Aaa_t$), and a measure of the funding cost of life insurance companies (FC_t) that is the ICE BoA Single-A U.S. corporate index option-adjusted spread. In more technical terms, we estimate:

$$\widehat{P}(\mathbb{1}(\gamma_{it} < 0)|TP_t, Baa - Aaa_t, FC_t) = \alpha_i + \beta_i^1 TP_t + \beta_i^2 Baa - Aaa_t + \beta_i^3 FC_t$$

where $\widehat{P}(\mathbb{1}(\gamma_{it} < 0)|TP_t, Baa - Aaa_t, FC_t)$ is the predicted probability that $\gamma_{it} < 0$ given TP_t, $Baa - Aaa_t$, FC_t, and a linear functional form.

The results are shown in Table 4 where we report the coefficient estimates and standard errors in parentheses. Column 1 shows the bivariate relationship between the term premium and the statistical significance of realized gamma for life insurers. Columns 2-4 provide the main result under a range of standard error estimates, as indicated at the bottom of the table. HC are heteroskedasticity consistent standard errors, HAC are heteroscedasticity and autocorrelation consistent standard errors, and NW are Newey-West standard errors. The dependent variable in column 5 is a binary variable for statistical significance of P&C insurers' realized gamma. This column acts as a placebo test of the main result for life insurers: The key variables we focus on for life insurers are not statistically important for P&C insurers, consistent with our prior expectations.

Noting again that the results are not causal, the estimates nevertheless suggest that there is a strong economic relationship between the variables, in addition to the statistical significance indicated in the table. We use as a benchmark for the economic effects the 32 percent unconditional probability that realized gamma for life insurers is statistically significant (see Table 6 in Appendix C). A one standard deviation increase in the term premium, which compensates investors for holding longer-term debt, reduces the probability that realized gamma for life insurers is statistically significant by about 17 percentage points—equivalent to about half of the unconditional probability that realized gamma for life insurers is statistically significant. A one standard deviation increase in Moody's Baa-Aaa seasoned corporate spread reduces the probability that realized gamma for life insurers is statistically significant by about 24 percentage points. And a one standard deviation increase in the ICE BoA Single-A U.S. corporate index option-adjusted spread raises the probability that realized gamma for life insurers is statistically significant by about 28 percentage points.

Our analysis provides support for the view that a flattening yield curve can drive realized gamma below zero. This can be seen, for example, around September 2019 when short-term interest rates rose and the 10-

Table 4: When are insurers not hedged? The table shows the results from estimating a linear probability model with a binary dependent variable that takes the value 1 if the estimated realized gamma (γ_{it}) is statistically significant on day t for $i \in$ {Life, P&C} and 0 otherwise. Other variables are the term premium (TP_t) estimates from Adrian et al. (2013), the Moody's Baa-Aaa seasoned corporate spread ($Baa - Aaa_t$), a measure of the funding cost of life insurance companies (FC_t) that is the ICE BoA Single-A U.S. corporate index option-adjusted spread, and the daily realized volatility (σ_t^{10yt}) of the ten-year Treasury interest rate. Columns 2-4 show the main result for three different standard error specifications. HC are heteroskedasticity consistent, HAC are heteroscedasticity and autocorrelation consistent, and NW are Newey-West standard errors. Columns 7–8 show a two-stage least squares estimation where the daily realized volatility (σ_t^{10yt}) is instrumented using a binary variable ($FOMC_t$) that takes the value 1 if there was a scheduled FOMC meeting on that day and 0 otherwise. Source: Authors' calculations based on data from Refinitiv, the Center for Research in Security Prices, FRED, and Adrian et al. (2013). *** p<0.01, ** p<0.05, * p<0.1

							2SLS: 1	*2SLS: 2*
Dep. var.:	Life				PC	Life	σ_t^{10yt}	Life
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TP_t	-0.09^{***}	-0.15^{***}	-0.15^{***}	-0.15^{***}	-0.04	-0.15^{***}	-0.03^{*}	-0.14^{***}
	(0.01)	(0.01)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)
$BAA - AAA_t$		-0.52^{***}	-0.52^{**}	-0.52^{***}	-0.10	-0.53^{***}	-0.25^{***}	-0.51^{***}
		(0.05)	(0.25)	(0.18)	(0.12)	(0.19)	(0.07)	(0.19)
FC_t		0.28^{***}	0.28^{**}	0.28^{***}	0.06	0.29^{***}	0.25^{***}	0.27^{***}
		(0.02)	(0.12)	(0.09)	(0.06)	(0.09)	(0.04)	(0.09)
σ_t^{10yt}						-0.06		0.01
						(0.05)		(0.09)
$FOMC_t$							0.32^{***}	
							(0.09)	
α	0.37^{***}	0.55^{***}	0.55^{***}	0.55^{***}	0.22^{***}	0.56^{***}	0.13^{***}	0.55^{***}
	(0.01)	(0.02)	(0.13)	(0.09)	(0.06)	(0.10)	(0.03)	(0.09)
Std. Err.	Robust	Robust	HAC	NW	NW	NW	NW	NW
Observations	3,896	3,892	3,892	3,892	3,892	3,892	3,892	3,892
Adjusted R^2	0.04	0.07	0.07	0.07	0.01	0.08	0.22	0.07
F Statistic	169.28^{***}	106.05^{***}	106.05^{***}	106.05^{***}	8.40^{***}	81.17^{***}	277.42^{***}	

year Treasury yield fell. Similarly, our findings chime with the broad consensus that an upward shift of the entire yield curve is generally good for life insurers. For example, realized gamma remained statistically close to zero during the rapid rise in short-term interest rates that occurred as the Federal Reserve tightened monetary policy in 2022. Our measure suggests that market participants focused on the positive effect on life insurers' profitability from rising long-term interest rates and widening spreads on long-term investment grade bonds. In summary, our realized gamma measure of stock price sensitivity to long-term interest rates serves as a useful barometer for market sentiment about the effectiveness of insurers' interest rate risk management.

3.4.2 Is realized gamma low due to interest rate volatility?

Column 6 of Table 4 shows that the daily realized volatility of 10-year Treasury security returns is not correlated with the statistical significance of realized gamma. This finding should be intuitive, as we are estimating realized gamma conditional on intraday 10-year Treasury security returns, but is important to emphasize: It means life insurers' interest rate risk management is generally effective *not* as a consequence of generally low interest rate volatility. In this section, we provide further evidence for this key result.

We provide additional tests as we recognize the potential endogeneity of long-term interest rate volatility and realized gamma. Life insurers are important investors in the long-term debt market, as we noted above. Their willingness to lend at long terms may simultaneously affect their own sensitivity to long-term interest rates and long-term interest rate volatility. We address the potential endogeneity with a source of plausibly exogenous variation in long-term interest rate volatility.

Scheduled Federal Open Market Committee (FOMC) meeting days are a well-known source of volatility in interest rates, that is sometimes used as a exogenous source of variation (Rigobon and Sack, 2004; Foley-Fisher and Guimaraes, 2013).[19] FOMC meeting days are exogenous to the supply-side variables that give rise to endogeneity concern in our setting. We exploit this source of exogenous variation in two ways. First, we use scheduled FOMC meeting days as an instrumental variable (IV) to obtain exogenous variation in long-term interest rate volatility. Second, we test the difference in means between realized gamma on scheduled FOMC meeting days and on other days when long-term interest rate volatility is lower.

Our first test using an IV is reported in columns 7 and 8 of Table 4, where we show the results from estimating a two-stage least squares regression specification. The IV for the endogenous interest rate volatility variable (σ_t^{10yt}) is a dummy variable ($\boldsymbol{FOMC_t}$) that takes the value 1 on

[19]Note that monetary policy shocks are the root cause of the exogenous increase in interest rate volatility, but we do not need to identify the size of those shocks to implement our tests.

days when the FOMC holds a scheduled meeting and the value 0 otherwise. The first stage, reported in column 7, shows that the FOMC variable is a strong instrument for σ_t^{10yt}. The coefficient estimate is highly statistically significant and positive, consistent with rising interest rate volatility on days with scheduled FOMC meetings. The F-statistic for the first stage regression is 277.4, indicating a strong IV. The second stage, which includes the fitted values from the first stage as a right-hand side variable to replace σ_t^{10yt}, is reported in column 8. The coefficient on long-term interest rate volatility remains statistically insignificant.

Our second test addresses two limitations of our IV approach: (1) other right-hand side variables in our specification may be invalid as instruments in the first stage, and (2) our left-hand side variable is a dummy variable for the statistical significance of realized gamma. We focus on the statistical property of the average difference in realized gammas between high-volatility days when the FOMC has its scheduled meetings and low-volatility days just before the FOMC meetings. Specifically, in our data sample we have 125 FOMC meeting days from May 2007 to December 2022. We pair these days with two alternative low-volatility samples: (1) the days that are one day before the scheduled FOMC meetings, and (2) the days that are one week before the scheduled FOMC meetings. Our null hypothesis (H_0) is that the average difference between the paired high-volatility realized gammas and low volatility realized gammas is zero.[20] We implement this test using the sub-sampling approach, which does not require making strong assumptions about the unknown distribution of realized gammas or estimating the sample mean variances. All that is required to obtain an asymptotically valid test is that the sampling distribution of the difference in paired realized gammas converges to some unknown distribution and that each *pair* of realized gammas is independent and identically distributed. The former is an extremely weak condition and the latter is natural as we estimate realized gamma using the ratio of daily realized covariances.

The results are reported in Table 5. Columns 1 and 2 show the 99-percent confidence intervals obtained by sub-sampling for the average difference in paired realized gammas for life insurers and P&C insurers, respectively. The test rejects $\boldsymbol{H}_0$ when the confidence intervals do not contain zero. The first row of the table shows the results when the FOMC meeting days are paired with one-day earlier days. The second row of the table shows the results from pairing FOMC meeting days with one-week earlier days. In both rows, the confidence intervals contain zero and we cannot reject the null hypothesis that the paired realized gammas are the same. For comparison, column 3 reports the confidence intervals from testing the difference in 10-year Treasury security realized volatility

[20] In addition to calculating the paired difference, we also tested the difference between the average realized gamma on scheduled FOMC meeting days and non-meeting days.

between paired days. Column 3 shows that there was a statistically significant increase in σ_t^{10yt}, as should be expected.

Table 5: Comparing realized gammas on days with high and low interest rate volatility. We test the statistical significance of the average difference in realized gammas between high-volatility days when the FOMC has its scheduled meetings and low-volatility days just before the FOMC meetings. Column 1 reports the test of life insurer gammas. Column 2 reports the test of P&C insurer gammas. Column 3 reports the test of daily realized volatility of 10-year Treasury security returns, multiplied by $1e+4$ for legibility. The first row pairs the scheduled FOMC meeting days with one-day earlier days. The second row pairs the scheduled FOMC meeting days with one-week earlier days. The sub-sampled confidence intervals are calculated using 10,000 combinations of 15 paired dates. Source: Authors' calculations based on data from Refinitiv, the Center for Research in Security Prices and the St Louis Fed's FRASER database.

	99% confidence interval		
	Life insurers	P&C insurers	10yr Treasury
FOMC days vs. 1 day before	[-0.072, 0.102]	[-0.027, 0.086]	[0.23, 0.606]
FOMC days vs. 7 days before	[-0.063, 0.096]	[-0.038, 0.074]	[0.271, 0.629]

In summary, the additional tests we implemented to address the potential endogeneity of realized gamma and σ_t^{10yt} underscore that low interest rate volatility is not the reason for our finding that life insurers' interest rate risk hedging is generally effective i.e., that realized gamma is generally statistically insignificant.

4 Concluding remarks

In this paper, we introduced a new method to measure the time-varying *residual* interest rate risk of financial intermediaries after they have executed their risk management strategies. Our estimates are daily partial correlations obtained using a nonparametric approach on high-frequency financial market data. We then showed how to conduct statistical inference on our estimates by calculating confidence intervals that are asymptotically valid under extremely weak conditions. Our method can be adapted to include additional variables in the regression model that underpins our framework. Another potential future extension would be to allow for 'jumps' when estimating realized variances and covariances (Andersen, Bollerslev and Diebold, 2007).

Our measure can be used to evaluate the interest rate risk vulnerabilities of any financial intermediary with high-frequency stock prices almost in real time, which is a useful tool for market analysts, supervisors, and policymakers. We applied our method to life insurers, whose exposure to interest rate risk has received less attention than, for example, banks. In doing so, we offered an alternative to the biased and inconsistent low-frequency rolling window OLS regressions that are prevalent in the existing literature. We find that life insurers are generally well-hedged against long-term interest rate movements. That said, they are more sensitive to changes in long-term interest rates than P&C insurers. We then showed that a measure of the term premium helps to explain the difference in estimated sensitivities between the two types of insurer. Lastly, we provided evidence that our finding that insurers are generally well-hedged against interest rate risk is *not* because long-term interest rate volatility is low. Comparing these results with those of other financial intermediaries, such as banks, is another avenue for further research.

References

Adrian, Tobias, Richard Crump, and Emanuel Moench, "Pricing the term structure with linear regressions," *Journal of Financial Economics*, 2013, *110* (1), 110–138.

Aït-Sahalia, Yacine and Jialin Yu, "High frequency market microstructure noise estimates and liquidity measures," *The Annals of Applied Statistics*, 2009, *3* (1), 422–457.

_ **and Per Mykland**, "Estimating Volatility in the Presence of Market Microstructure Noise: A Review of the Theory and Practical Considerations," in Thomas Mikosch, Jens-Peter Kreiß, Richard A. Davis, and Torben Gustav Andersen, eds., *Handbook of Financial Time Series*, Springer Berlin Heidelberg, January 2009, chapter 25, pp. 577–598.

Allen, Franklin, "Estimating Divisional Cost of Capital for Insurance Companies," in J. David Cummins and Joan Lamm-Tennant, eds., *Financial Management of Life Insurance Companies*, Springer Netherlands, 1993, pp. 101–123.

Andersen, Torben, Tim Bollerslev, and Francis Diebold, "Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility," *The Review of Economics and Statistics*, 2007, *89* (4), 701–720.

_ , _ , and Nour Meddahi, "Analytical Evaluation of Volatility Forecasts," *International Economic Review*, 2004, *45* (4), 1079–1110.

_ , _ , Francis Diebold, and Ginger Wu, "Realized beta: Persistence and predictability," in "Econometric Analysis of Financial and Economic Time Series," Emerald Group Publishing Limited, 2006.

Babbel, David and Robert Stricker, "Asset/Liability Management for Insurers," *Insurance Perspectives*, May 1987.

Bandi, Federico M., Aleksey Kolokolov, Davide Pirino, and Roberto Renò, "Zeros," *Management Science*, 2020, *66* (8), 3466–3479.

Barndorff-Nielsen, Ole and Neil Shephard, "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 2004, *72* (3), 885–925.

_ , Peter Reinhard Hansen, Asger Lunde, and Neil Shephard, "Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading," *Journal of Econometrics*, 2011, *162* (2), 149–169.

Begenau, Juliane, Monika Piazzesi, and Martin Schneider, "Banks' Risk Exposures," *NBER Working Paper 21334*, 2015.

Berends, Kyal, Robert McMenamin, Thanases Plestis, and Richard J Rosen, "The sensitivity of life insurance firms to interest rate changes," *Economic Perspectives*, 2013, *37* (2).

Brainard, Lael, "Global Financial Stability Considerations for Monetary Policy in a High-Inflation Environment," *Speech at "Financial Stability Considerations for Monetary Policy," a research conference organized by the Board of Governors of the Federal Reserve System and the Federal Reserve Bank of New York*, 2022.

Brewer III, Elijah, James Carson, Elyas Elyasiani, Iqbal Mansur, and William Scott, “Interest Rate Risk and Equity Values of Life Insurance Companies: A GARCH-M Model,” *Journal of Risk and Insurance*, 2007, *74* (2), 401–423.

_ , Thomas Mondschean, and Philip Strahan, “Why the life insurance industry did not face an" S&L-type" crisis,” *Economic Perspectives*, 1993, *17* (5), 12–24.

Briys, Eric and François de Varenne, “On the Risk of Insurance Liabilities: Debunking Some Common Pitfalls,” *The Journal of Risk and Insurance*, 1997, *64* (4), 673–694.

_ and _ , “Insurance: from underwriting to derivatives,” *Springer US*, 2001.

Brunnermeier, Markus and Yuliy Sannikov, “A macroeconomic model with a financial sector,” *American Economic Review*, 2014, *104* (2), 379–421.

Christensen, Kim, Silja Kinnebrock, and Mark Podolskij, “Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data,” *Journal of Econometrics*, 2010, *159* (1), 116–133.

Davidson, Russell and James MacKinnon, *Estimation and inference in econometrics*, Oxford New York, 1993.

Doffou, Ako, “New Perspectives in Asset-Liability Management for Insurers,” *Journal of Business and Behavioral Sciences*, 2005, *12* (2).

Domanski, Dietrich, Hyun Song Shin, and Vladyslav Sushko, “The hunt for duration: not waving but drowning?,” *IMF Economic Review*, 2017, *65* (1), 113–153.

English, William, Skander Van den Heuvel, and Egon Zakrajšek, “Interest rate risk and bank equity valuations,” *Journal of Monetary Economics*, 2018, *98*, 80–97.

Fama, Eugene and Kenneth French, "The Capital Asset Pricing Model: Theory and Evidence," *Journal of Economic Perspectives*, September 2004, *18* (3), 25–46.

Fama, Eugene F. and G.William Schwert, "Asset returns and inflation," *Journal of Financial Economics*, 1977, *5* (2), 115–146.

Flannery, Mark and Christopher James, "The Effect of Interest Rate Changes on the Common Stock Returns of Financial Institutions," *The Journal of Finance*, 1984, *39* (4), 1141–1153.

Foley-Fisher, Nathan and Bernardo Guimaraes, "U.S. Real Interest Rates and Default Risk in Emerging Economies," *Journal of Money, Credit and Banking*, 2013, *45* (5), 967–975.

_ , Borghan Narajabad, and Stéphane Verani, "Securities lending as wholesale funding: Evidence from the us life insurance industry," Technical Report, National Bureau of Economic Research 2016.

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

Froot, Kenneth and Jeremy Stein, "Risk management, capital budgeting, and capital structure policy for financial institutions: an integrated approach," *Journal of Financial Economics*, 1998, *47* (1), 55–82.

_ , David Scharfstein, and Jeremy Stein, "Risk management: Coordinating corporate investment and financing policies," *Journal of Finance*, 1993, *48* (5), 1629–1658.

Gajek, Leslaw and Krzysztof Ostaszewski, *Financial risk management for pension plans*, Elsevier, 2004.

Gissler, Stefan, Nathan Foley-Fisher, and Stéphane Verani, "Over-the-Counter Market Liquidity and Securities Lending," *Review of Economic Dynamics*, 2019.

Hansen, Peter, Asger Lunde, and Valeri Voev, "Realized Beta GARCH: A Multivariate GARCH Model with Realized Measures of Volatility," *Journal of Applied Econometrics*, 2014, *29* (5), 774–799.

Hartley, Daniel, Anna Paulson, and Richard Rosen, "Measuring interest rate risk in the life insurance sector," *The economics, regulation, and systemic risk of insurance markets*, 2016, p. 124.

Hoffmann, Peter, Sam Langfield, Federico Pierobon, and Guillaume Vuillemey, "Who Bears Interest Rate Risk?," *The Review of Financial Studies*, 11 2018, *32* (8), 2921–2954.

Holmstrom, Bengt and Jean Tirole, "Financial intermediation, loanable funds, and the real sector," *the Quarterly Journal of economics*, 1997, *112* (3), 663–691.

Huber, Maximilian, "Regulation-Induced Interest Rate Risk Exposure," *Working Paper*, 2022.

Kirti, Divya, "When gambling for resurrection is too risky," *IMF Working Paper No. 2017/180*, 2017.

Koijen, Ralph and Motohiro Yogo, "The fragility of market risk insurance," *The Journal of Finance*, 2022, *77* (2), 815–862.

Kolokolov, Aleksey and Roberto Renò, "Jumps or staleness?," *Journal of Business & Economic Statistics*, 2023, *0*, 1–23.

Liu, Lily, Andrew Patton, and Kevin Sheppard, "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes," *Journal of Econometrics*, 2015, *187* (1), 293–311.

Meddahi, Nour, "A theoretical comparison between integrated and realized volatility," *Journal of Applied Econometrics*, 2002, *17* (5), 479–508.

Möhlmann, Axel, "Interest rate risk of life insurers: Evidence from accounting data," *Financial Management*, 2021, *50* (2), 587–612.

NAIC, "Historical Evolution of Life Insurance," in CIPR, ed., *State of the Life Insurance Industry: Implications of Industry Trends*, Center for Insurance Policy and Research, 2013.

Ozdagli, Ali and Zixuan Wang, "Interest rates and insurance company investment behavior," *Available at SSRN 3479663*, 2019.

Paparoditis, Efstathios and Dimitris Politis, "Resampling and Subsampling for Financial Time Series," in Thomas Mikosch, Jens-Peter Kreiß, Richard A. Davis, and Torben Gustav Andersen, eds., *Handbook of Financial Time Series*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 983–999.

Paul, Pascal, "Banks, maturity transformation, and monetary policy," *Journal of Financial Intermediation*, 2022, p. 101011.

Politis, Dimitris, Joseph Romano, and Michael Wolf, *Subsampling*, Springer-Verlag, 1999.

Redington, F. M., "Review of the Principles of Life-office Valuations," *Journal of the Institute of Actuaries*, 1952, *78*, 286–340.

Rigobon, Roberto and Brian Sack, "The impact of monetary policy on asset prices," *Journal of Monetary Economics*, 2004, *51* (8), 1553–1575.

Schneider, Andrés, "Risk-Sharing and the Term Structure of Interest Rates," *The Journal of Finance*, 2022, *77* (4), 2331–2374.

Sen, Ishita, "Regulatory Limits to Risk Management," *Review of Financial Studies*, 2021, *forthcoming.*

Verani, Stéphane and Pei Cheng Yu, "What's Wrong with Annuity Markets?," *mimeo*, 2021.

Vuillemey, Guillaume, "Bank Interest Rate Risk Management," *Management Science*, 2019, *65* (12), 5933–5956.

Zhang, Lan, Per Mykland, and Yacine Aït-Sahalia, "A tale of two time scales: Determining integrated volatility with noisy high-frequency data," *Journal of the American Statistical Association*, 2005, *100* (472), 1394–1411.

Appendix for online publication

A How large is the rolling window bias?

In this appendix, we demonstrate the size of the bias from estimating the two-variable regression model on a rolling window of daily data for insurance companies. We start from the specification:

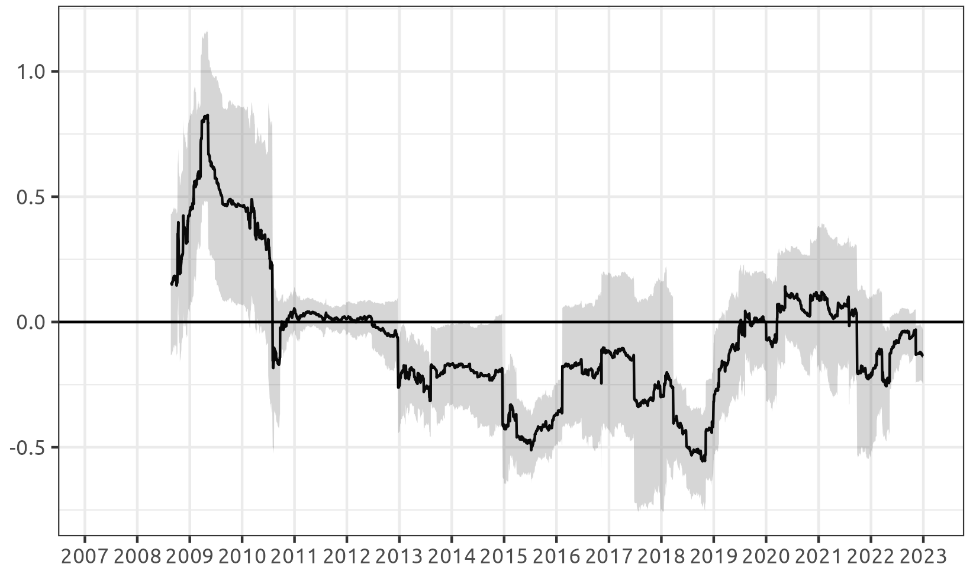
$$r_{i,t} = \alpha + \beta r_{m,t} + \gamma r_{y10,t} + \epsilon_{i,t}$$

where $r_{i,t}$ is the stock price return on the index of life insurers (described in section 3.2) on day t, $r_{m,t}$ is the return on the benchmark S&P500, and $r_{y10,t}$ is the return on the 10-year Treasury security. The γ coefficient in this specification is termed *rolling gamma* and is a low-frequency counterpart to the realized gamma described in section 2 of the main paper.

Selecting the size for the rolling window is typically framed as a tradeoff between (i) including more data to reduce standard errors and (ii) being forced to assume the parameter is stable within the window (Robertson, 2018). We follow the standard approach in the empirical literature estimating interest rate risk for life insurers, and assume a rolling window of two years (Sen, 2021; Huber, 2022).

Figure 4 shows the time series of rolling gammas. The shaded region indicates the heteroskedasticity-corrected 90 percent confidence interval for

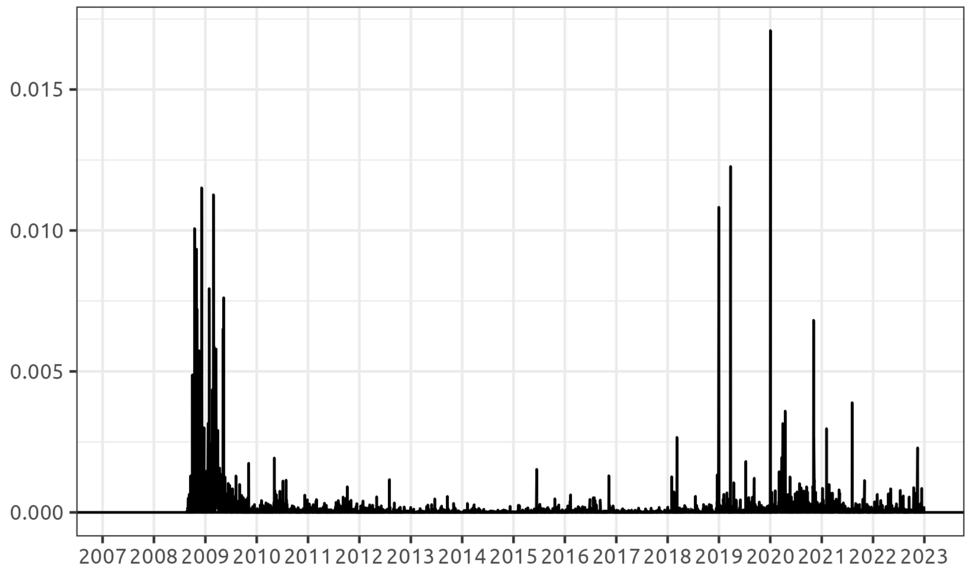
Figure 4: Rolling window regression results. The black line shows the rolling gamma estimates using end-of-day data and a two-year rolling window. The shaded region indicates the heteroskedasticity-corrected 90 percent confidence interval for the estimates. Source: Authors' calculations based on data from Refinitiv, the Center for Research in Security Prices.



the estimates. There are two main takeaways from the figure. First, the estimates are almost always statistically significant in the post-crisis period. This finding led researchers to conclude that life insurers' risk management became less effective in the aftermath of the GFC and spurred a research agenda to understand the cause of this regime switch—e.g., Sen (2021); Koijen and Yogo (2022); Huber (2022). Second, there are large "jumps" in the time series corresponding to periods of market volatility, such as the beginning of the financial crisis (2008) and the pandemic (2020).

Jumps in the time series hint at a problem of time-varying conditional volatility in the underlying data. Figure 5 shows the problem by plotting the square of the residuals $(r_{i,t} - \hat{r}_{i,t})^2$. Volatility clustering, which is clearly present in our data, is a long-known empirical feature of financial time series (Bollerslev, Chou and Kroner, 1992).

Figure 5: Squared residuals from rolling regression. Source: Authors' calculations based on data from Refinitiv, the Center for Research in Security Prices.



The potential effects of conditional heteroskedasticity for OLS regressions are well known. In some applications, such as when the primary concern is estimating the conditional mean, a common view is that inference can be made using the standard corrections proposed by White (1980) or Newey and West (1987). However, as Hamilton (2008) points out, misspecifying the errors will produce inefficient estimates and incorrect inference. The specific case of the rolling window OLS estimator was studied by Cai and Juhl (2021), who showed that a bias can exist even asymptotically with well-behaved errors. Intuitively, the rolling window OLS estimates are weighted averages of the time-varying parameter and the weights depend on the time-varying volatility. The asymptotic bias arises when the two time series (parameter and volatility) are correlated. In simulations, the rolling window OLS estimates are often unstable and the bias can be substantial (Robertson, 2018).

One solution to the problem is to assume some structure for the variance processes. By explicitly modeling the heteroscedasticity in the variance-covariance matrix, we address the bias in the time series of parameter estimates and gain efficiency. A typical approach in financial econometrics is to appeal to autoregressive conditional heteroskedasticity (ARCH) models (Bollerslev, Engle and Nelson, 1994). This class of flexible models and its wide range of extensions are straightforward to implement in off-the-shelf statistical packages. In practice, the generalized ARCH, or 'GARCH', model that allows for greater serial dependence in the error term is an extremely common choice. The conditional variance of the process for a GARCH(r, p) is given by:

$$
\begin{aligned}
Var(\epsilon_t|\Omega_{t-1}) = h_t = a_0 + a_1\epsilon_{t-1}^2 + a_2\epsilon_{t-2}^2 + \cdots + a_p\epsilon_{t-p}^2 + \\
b_1 h_{t-1} + b_2 h_{t-2} + \cdots + b_r h_{t-r}.
\end{aligned}
$$

As an exercise to gauge the size of the rolling window OLS bias in the estimates reported in Figure 4, we assume that our three time series of daily returns follow a multivariate GARCH(1,1) process. We specify the joint process:

$$r_{i,t} = \alpha_i + u_{i,t}$$

$$r_{m,t} = \alpha_m + u_{m,t}$$

$$r_{y10,t} = \alpha_{y10} + u_{y10,t}$$

so

$$E\left[\begin{pmatrix} r_{i,t} \\ r_{m,t} \\ r_{y10,t} \end{pmatrix} \mid \Omega_{t-1}\right] = \begin{pmatrix} \alpha_i \\ \alpha_m \\ \alpha_{y10} \end{pmatrix}$$

and

$$Var\left[\begin{pmatrix} r_{i,t} \\ r_{m,t} \\ r_{y10,t} \end{pmatrix} \mid \Omega_{t-1}\right] = Var\left[\begin{pmatrix} u_{i,t} \\ u_{m,t} \\ u_{y10,t} \end{pmatrix} \mid \Omega_{t-1}\right] = \begin{pmatrix} \sigma^2_{11,t} & \sigma_{12,t} & \sigma_{13,t} \\ \sigma_{21,t} & \sigma^2_{22,t} & \sigma_{23,t} \\ \sigma_{31,t} & \sigma_{32,t} & \sigma^2_{33,t} \end{pmatrix}.$$

The last matrix—known as the dynamic conditional variance-covariance matrix—can be used to form the dynamic conditional ratio:

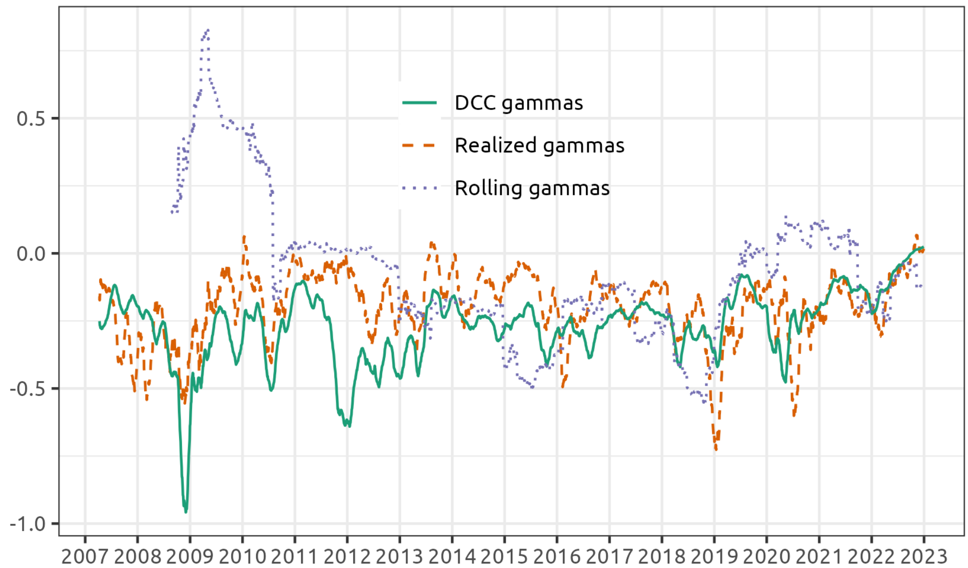
$$\gamma_t^{DC} = \frac{\sigma_{13,t}}{\sigma_{33,t}^2},$$

which is obtained after specifying GARCH(1,1) processes for each second moment. We call the ratio γ_t^{DC} the *dynamic conditional gamma*, following the literature that uses the same technique to estimate dynamic conditional betas (Engle, 2016).

Figure 6 compares the different estimates of gamma. The blue dotted line shows the rolling gamma estimates, while the green solid line shows the dynamic conditional gamma estimates. The difference between the two estimates is particularly striking during periods of high volatility, such as the financial crisis and the global pandemic, revealing that the rolling gamma is highly biased and misleading. For completeness, we include the realized gamma estimates as the brown dashed line in the figure. The relative proximity of the realized gamma estimates and the dynamic conditional gamma estimates during those periods of stress is a reassuring sign that both approaches are solving the underlying problem of conditional heteroscedasticity. Note that dynamic conditional gamma and realized gamma use completely different data and approaches to address the same underlying problem.

Although dynamic conditional gamma and realized gamma deliver

Figure 6: Comparing rolling gamma, dynamic conditional gamma, and realized gamma. The figure shows three different estimates of the sensitivity of life insurers' stock prices to changes in interest rates. Source: Authors' calculations based on data from Refinitiv, the Center for Research in Security Prices.



similar parameter estimates, they are not the same. In particular, the empirical approach that underpins the dynamic conditional gamma is known to suffer from substantial limitations (Caporin and McAleer, 2013). As a stated data representation—rather than derived model—the dynamic conditional gamma has no moments or desirable asymptotic properties. It serves our purposes as a diagnostic tool that reveals a huge bias in rolling gamma. But to avoid reliance on the imposed structure and—most importantly—to conduct valid inference, we strongly prefer the empirical approach that uses realized variances and covariances in our paper.

B Hypothesis testing using the subsampling method

In each day $t \in \{1, \ldots, T\}$, we estimate γ_t using the following linear regression model

$$\tilde{r}_{i,j,s} = \alpha_t + \beta_t \tilde{r}_{m,j,s} + \gamma_t \tilde{r}_{y10,j,s} + \epsilon_{i,j,s}$$

on a sample of $n = 388$ observations corresponding to each of the 388 trading minutes for day t between 9:31am and 3:59pm indexed by s. We calculate a confidence interval for each of the daily γ_t using subsampling following Politis et al. (1999) under the assumption that the errors are *asymptotically* *stationary*. Asymptotic stationarity means that, for example, the errors could follow an AR(1) process with autocorrelation parameter strictly less than one and heteroskedastic innovations.

To simplify the exposition of subsampling, we rewrite the linear regression model in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ and $\boldsymbol{\epsilon}$ are $n \times 1$ vectors, $\boldsymbol{\beta}$ is a $p \times 1$ vector which includes γ_t as an element and $\mathbf{X}$ is an $n \times p$ matrix of five-minute returns and a constant. The estimator of $\boldsymbol{\beta}$ based on $\mathbf{X}$ and $\mathbf{y}$ is given by $\hat{\boldsymbol{\beta}} \equiv (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

For any $b < n$ such that $b > p$, define the subvectors and submatrices

$$\mathbf{y}_{b,s} \equiv (y_s, \ldots, y_{s+b-1})' \, , \, \boldsymbol{\epsilon}_{b,s} \equiv (\epsilon_s, \ldots, \epsilon_{s+b-1})' \text{ and} \tag{5}$$

$$\mathbf{X}_{b,s} \equiv \begin{pmatrix} \boldsymbol{x}'_s \\ \vdots \\ \boldsymbol{x}'_{s+b-1} \end{pmatrix}, \text{ where } \mathbf{X} \equiv \begin{pmatrix} \boldsymbol{x}'_1 \\ \vdots \\ \boldsymbol{x}'_n \end{pmatrix} \tag{6}$$

The estimator of $\boldsymbol{\beta}$ based on $\mathbf{X}_{b,s}$ and $\mathbf{y}_{b,s}$ is given by

$$\hat{\boldsymbol{\beta}}_{n,b,s} \equiv (\mathbf{X}'_{b,s}\mathbf{X}_{b,s})^{-1}\mathbf{X}'_{b,s}\mathbf{y}_{b,s}.$$

Denote by $J_b(P)$ the sampling distribution of the normalized statistic

$\sqrt{b}(\hat{\boldsymbol{\beta}}_{n,b,s} - \boldsymbol{\beta})$, where P is the probability law governing the estimator $\hat{\boldsymbol{\beta}}_{n,b,s}$, which is unknown. For any Borel set $A \in \mathbb{R}^p$, let

$$J_b(A, P) = Prob_P\{\sqrt{b}(\hat{\boldsymbol{\beta}}_{n,b,s} - \boldsymbol{\beta}) \in A\}.$$

The approximation to $J_n(A, P)$ is defined by

$$L_{n,b}(A) = \frac{1}{n-b+1} \sum_{s=1}^{n-b+1} 1\{\sqrt{b}(\hat{\boldsymbol{\beta}}_{n,b,s} - \hat{\boldsymbol{\beta}}) \in A\}.$$

Therefore, subsampling consist of evaluating a statistics on an exhaustive set of subsamples of size $b < n$ that are created from the original sample of size n and estimating the distribution of this statistics normalized by $\sqrt{b}$. As should be clear, each subsample contains consecutive observations from the original time series sample. Therefore, each subsample is drawn from the true data generating process.

In what follows we summarize the main result from subsampling related to the estimation of a daily γ_t using intra-day time series observation. Note that our limiting concept is that the number of equally spaced intraday returns approaches infinity. We refer the readers to Politis et al. (1999) for details and proofs.

Assumption 1 *There exists a limiting law $J(P)$ such that*

1. *$J_n(P)$ converges weakly to $J(P)$ as $n \to \infty$. This means that for*

any Borel set A whose boundary has mass zero under $J(P)$, we have $J_n(A, P) \longrightarrow J(A, P)$ as $n \to \infty$.

2. *For every Borel set A whose boundary has mass zero under $J(P)$ and for any index sequence $\{s_b\}$, we have $J_{b,s_b}(A, P) \longrightarrow J(A)$ as $b \to \infty$.*

Theorem B.1 (Politis et al. (1999) Theorem 4.3.1) *Let $\{(\mathbf{x}_s, \epsilon_s)\}$ be a sequence of random vectors defined on a common probability space. Denote the mixing coefficients for the $\{(\mathbf{x}_s, \epsilon_s)\}$ sequence by $\alpha(\cdot)$. Define*

$$T_{k,s} \equiv \frac{1}{\sqrt{k}} \sum_{a=s}^{s+k-1} \mathbf{x}_a \epsilon_a \ , \ V_{k,s} \equiv Cov(T_{k,s}) \ , \ \text{and } M_{k,s} \equiv E(\mathbf{X}'_{k,s}\mathbf{X}_{k,s}/k).$$

Assume the following conditions hold. For some $\delta > 0$,

- $E(\mathbf{x}_s \epsilon_s) = 0$ *for all* s,
- $E|\mathbf{x}_{s,j}\epsilon_s|^{2+2\delta} \leq \Delta_1$ *for all* s *and all* $1 \leq j \leq p$,
- $E|\mathbf{x}_{s,j}|^{4+2\delta} \leq \Delta_2$ *for all* s *and all* $1 \leq j \leq p$,
- $V_{k,s} \longrightarrow V > 0$ *uniformly in* s *as* $k \to \infty$,
- $M_{k,s} \longrightarrow M > 0$ *uniformly in* s *as* $k \to \infty$,
- $C(4) \equiv \sum_{k=1}^{\infty}(k+1)^2 \alpha^{\frac{\delta}{4+\delta}}(k) \leq K$.

Furthermore, assume that $b/n \longrightarrow 0$ and $b \to \infty$ as $n \to \infty$. Letting $J(P) = N(0, M^{-1}VM^{-1})$. Then:

i. $L_{n,b}(A) \to J(A,P)$ *in probability for each Borel set A whose boundary has mass zero under* $J(P)$.

ii. Let Z *be a random vector with* $\mathcal{L}(Z) = J(Z)$. *For a norm* $||\cot||$ *on* $\mathbb{R}^k$, *define univariate distributions* $L_{n,||\cot||}$ *and* $J_{||\cot||}(P)$ *in the following way:*

$$L_{n,b,||\cot||}(x) = \frac{1}{n-b+1}\sum_{s=1}^{n-b+1} 1\{||\sqrt{b}(\hat{\boldsymbol{\beta}}_{n,b,s} - \hat{\boldsymbol{\beta}}||) \leq x\}$$

$$J_{||\cot||}(x,P) = Prob\{||Z|| \leq x\}.$$

For $\alpha \in (0,1)$, *let*

$$c_{n,b,||\cdot||}(1-\alpha) = \inf\{x : L_{n,b,||\cdot||}(x) \geq 1-\alpha\}.$$

Correspondingly, define

$$c_{||\cdot||}(1-\alpha,P) = \inf\{x : J_{||\cdot||}(x,P) \geq 1-\alpha\}.$$

If $J_{||\cdot||}(\cdot,P)$ *is continuous at* $c_{||\cdot||}(1-\alpha,P)$ *then*

$$Prob_P\{||\sqrt{b}(\hat{\boldsymbol{\beta}}_{n,b,s} - \hat{\boldsymbol{\beta}}|| \leq c_{n,b,||\cdot||}(1-\alpha)\} \to 1-\alpha \text{ as } n \to \infty.$$

Thus, the asymptotic coverage probability under $\boldsymbol{P}$ *of the region*

$$\{\boldsymbol{\beta} : ||\sqrt{b}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}|| \leq c_{n,b,||\cdot||}(1 - \alpha)\}$$

is the nominal level $1 - \alpha$*.*

Theorem B.1 shows that we can derive asymptotically valid confidence intervals for the daily estimator $\hat{\boldsymbol{\beta}}$ using $L_{n,b}(A)$ because it is a consistent estimator of $J(A, P)$. By exploiting the usual duality between the construction of a confidence interval for γ_t and the construction of a hypothesis test about γ_t, subsampling allows us to make asymptotically valid inference about the true γ_t. In our application, we wish to test the null hypothesis that the daily γ_t is statistically different from 0. Under the null, insurers are hedged against interest rate risk as their stock price is not sensitive to movement in the ten-year treasury rate. If the value of the estimated daily γ_t falls outside the daily confidence interval, we reject the null hypothesis on that day.

Subsampling is not as well known as the bootstrap method in economics and finance, which warrants a cursory comparison—see Politis et al. (1999) for textbook-length treatment. The most relevant bootstrap method for our time series application is the so-called Moving Blocks Bootstrap (MBB). As with subsampling, MBB breaks down the original time series to smaller blocks of consecutive observations, which preserves the serial correlation structure within each block. Practically, the main difference is that MBB draws samples with replacement from the blocks and connects the sampled blocks together to form a bootstrap sample of size n. Therefore, by construction, MBB imposes the assumption that blocks of an arbitrary size b are uncorrelated. This assumption about the unknown data generating process is rather strong and likely to be violated in our application. From a technical point view, the bootstrap method requires that the distribution of the statistic of interest be locally smooth as a function of the unknown model. Establishing this result, even if it is indeed true, would be non-trivial. With subsampling, we do not need to make these assumptions or verify the smoothness of the distribution under the true model to draw asymptotically valid inferences. All that is required is that our normalized statistic has a limit distribution under the true model.

C Summary statistics for Section 3.4

Table 6: Summary statistics. This table reports summary statistics for the variables used to analyze the determinants of the significance of realized gamma. $\gamma_{i,t}$ is realized gamma for insurer type $i \in$ {Life, P&C}. The binary variable $\mathbb{1}(\gamma_{i,t}<0)$ takes the value 1 when realized gamma for insurer type i is statistically significant and 0 otherwise. TP_t is the term premium estimate from Adrian et al. (2013), $Baa - Aaa_t$ is the Moody's Baa-Aaa seasoned corporate spread, and FC_t is the ICE BoA Single-A US corporate index option-adjusted spread. σ_t^{Life} is the realized volatility of the intraday returns of life insurers. σ_t^{10yt} is the realized volatility of the intraday returns on 10-year Treasury. The statistics for σ_t^{Life} and σ_t^{10yt} are multiplied by 1e+4 for legibility. Source: Authors' calculations based on data from Refinitiv, the Center for Research in Security Prices, FRED, and Adrian et al. (2013).

Variable	No. obs.	Mean	Median	Std. Dev.	p25	p75
$\gamma_{\text{Life},t}$	3,901	-0.19	-0.16	0.31	-0.34	-0.02
$\mathbb{1}(\gamma_{\text{Life},t}<0)$	3,923	0.32	0	0.47	0	1
$\gamma_{P\&C,t}$	3,901	-0.12	-0.09	0.22	-0.21	0.01
$\mathbb{1}(\gamma_{\text{P\&C},t}<0)$	3,923	0.18	0	0.38	0	0
TP_t	3,896	0.54	0.30	1.10	-0.33	1.51
$Baa - Aaa_t$	3,897	1.08	0.96	0.47	0.82	1.19
FC_t	3,921	1.49	1.17	1.01	0.94	1.64
σ_t^{10yt}	3,923	0.24	0.16	0.31	0.10	0.28

D Data citations

- Refinitiv
- Center for Research in Security Prices, CRSP 1925 US Indices Database, Wharton Research Data Services, `http://www.whartonwrds.com/datasets/crsp/`
- FRED API, accessed using third party R software package fredr, `https://fred.stlouisfed.org/docs/api/fred/`
  - 'BAA' — Moody's, Moody's Seasoned Baa Corporate Bond Yield [BAA], retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/BAA`
  - 'AAA' — Moody's, Moody's Seasoned Aaa Corporate Bond Yield [AAA], retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/AAA`

Online appendix references

Bollerslev, Tim, Ray Chou, and Kenneth Kroner, "ARCH modeling in finance: A review of the theory and empirical evidence," *Journal of Econometrics*, 1992, *52* (1), 5–59.

_ **, Robert Engle, and Daniel Nelson**, "ARCH Models," in "Handbook of Econometrics," Vol. 4, Elsevier, 1994, pp. 2959–3038.

Cai, Zongwu and Ted Juhl, "The Distribution Of Rolling Regression Estimators," *mimeo*, 2021.

Caporin, Massimiliano and Michael McAleer, "Ten things you should know about the dynamic conditional correlation representation," *Econometrics*, 2013, *1* (1), 115–126.

Engle, Robert, "Dynamic conditional beta," *Journal of Financial Econometrics*, 2016, *14* (4), 643–667.

Hamilton, James, "Macroeconomics and ARCH," *NBER Working Paper 14151*, 2008.

Huber, Maximilian, "Regulation-Induced Interest Rate Risk Exposure," *Working Paper*, 2022.

Koijen, Ralph and Motohiro Yogo, "The fragility of market risk insurance," *The Journal of Finance*, 2022, *77* (2), 815–862.

Newey, Whitney and Kenneth West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 1987, *55* (3), 703–708.

Politis, Dimitris, Joseph Romano, and Michael Wolf, *Subsampling*, Springer-Verlag, 1999.

Robertson, Donald, "Estimating β," *mimeo*, 2018.

Sen, Ishita, "Regulatory Limits to Risk Management," *Review of Financial Studies*, 2021, *forthcoming.*

White, Halbert, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 1980, *48* (4), 817–838.