

The price of variance risk

Ian Dew-Becker, Stefano Giglio, Anh Le, and Marius Rodriguez*

April 29, 2015

Abstract

In the period 1996–2014, the average investor in the variance swap market was indifferent to news about future variance at horizons ranging from 1 month to 14 years. It is only purely transitory and unexpected realized variance that were priced. These results present a challenge to most structural models of the variance risk premium, such as the intertemporal CAPM, recent models with Epstein–Zin preferences and long-run risks, and models where institutional investors have value-at-risk constraints. The results also have strong implications for macro models where volatility affects investment decisions, suggesting that investors are not willing to pay to hedge shocks in expected economic uncertainty.

*Northwestern Kellogg, Chicago Booth, UNC Kenan–Flagler, and the Federal Reserve Board. The views in this paper are solely the those of the authors and do not reflect the views of the Board of Governors of the Federal Reserve System. We appreciate helpful comments from and conversations with Dante Amengual, Jean-Francois Bacmann, Brandon Bates, Nick Bloom, John Campbell, John Cochrane, Greg Duffee, Darrell Duffie, Cam Harvey, Ian Martin, Ivan Shaliastovich and Harald Uhlig, Andrea Vedolin, Jessica Wachter, and seminar participants at the Booth Econometrics Workshop, University of New South Wales, University of Sydney, University of Technology, Sydney, University of Adelaide, Northwestern Kellogg, the 2014 Macro Finance Society Meeting, the Office for Financial Research, the Midwest Finance Association, ASU Sonoran Winter Finance Conference, Oxford, LSE, HEC, CSEF, Bocconi, and MIT Sloan. We are particularly grateful to Gerald Lim and Markit Group for providing variance swap data.

1 Introduction

The recent explosion of research on the effects of volatility in macroeconomics and finance shows that economists care about uncertainty shocks. Investors, on the other hand, do not. In the period since 1996, it has been costless on average to hedge news about future volatility in aggregate stock returns; in other words investors have been unwilling to pay anything for insurance against volatility news. Many economic theories – both in macroeconomics and in finance – have the opposite prediction. In recent macroeconomic models, for example, shocks to uncertainty about the future can induce large fluctuations in the economy.¹ But if increases in economic uncertainty can drive the economy into a recession, we would expect that investors would want to hedge those shocks. The fact that volatility shocks have been unpriced thus presents a challenge to the recent macro literature on the effects of volatility shocks.

As a concrete example, consider the legislative battles over the borrowing limit of the United States in the summers of 2010 and 2011. Those periods were associated with increases in both financial measures of uncertainty, e.g. the VIX, and also the measure of policy uncertainty from Baker, Bloom, and Davis (2014). Between June and July, 2011, the 1-month variance swap rate – a measure of investor expectations for S&P 500 volatility over the next month – rose from 16.26 to 25.96 percent (annualized, computed at the end of each month). However, those shocks also had small effects on *realized* volatility in financial markets: annualized realized volatility during June and July, 2011, was 14.59 and 15.23 percent, respectively. The debt ceiling debate caused uncertainty about the future to be high, but did not correspond to high contemporaneous volatility.²

Those facts make the debt-ceiling shocks the exact type of shock that is studied in the recent literature. It is precisely changes in expectations of *future* uncertainty that can have strong macroeconomic effects, because they affect all forward-looking decisions. In

¹See, e.g., Bloom (2009), Bloom et al. (2014), Christiano, Motto, and Rostagno (2014), Fernandez-Villaverde et al. (2011), and Gourio (2012, 2013)

²The table below reports realized volatility and the 1-month variance swap rate (nearly identical to the VIX) for June to October of 2011. Realized volatility is computed as the sum of squared daily log returns during a month, and the variance swap rate is the price traded on the last day of the month.

	June	July	Aug.	Sept.	Oct.
1-month variance swap	16.26	25.96	31.68	42.32	28.53
Realized volatility	14.59	15.23	47.18	28.80	29.80

Except for August, when both the variance swap rate and realized volatility rose in tandem, for all other months the changes in the two series are essentially unrelated.

Some of the volatility in the fall of 2011 was also due to uncertainty about the state of the European economy.

this paper, we directly measure how much people are willing to pay to hedge shocks to expectations of future volatility. We find that those news shocks are unpriced: any investor can buy insurance against volatility shocks for free, and therefore any investor could have freely hedged the increases in uncertainty during the debt ceiling debate.

We measure the price of variance risk using novel data on a wide range of volatility-linked assets both in the US and around the world, focusing primarily on variance swaps with maturities between 1 month and 14 years. The data covers the period 1996–2014. Variance swaps are assets that pay to their owner the sum of daily squared stock market returns from their inception to maturity. They thus give direct exposure to future stock market volatility and are the most natural and direct hedge for the risks associated with increases in aggregate economic uncertainty.

The analysis of the pricing of variance swaps yields two simple but important results. First, news about future volatility is unpriced – exposure to volatility news has not earned a risk premium. Second, exposure to *realized variance* is strongly priced, with an annualized Sharpe ratio of -1.7 – five times larger than the Sharpe ratio on equities.³ We find that it is the downside component of realized volatility that investors are specifically trying to hedge, consistent with the results of Bollerslev and Todorov (2011) showing that realized variance is priced due to its correlation with large negative jumps. We conclude that over our sample, investors were willing to pay a large amount of money for protection from extreme negative shocks to the economy (which mechanically generate spikes in realized volatility), but they did not pay to hedge news that uncertainty or the probability of a disaster has changed.

The results present a challenge to a wide range of models. In macroeconomics, there is now a large literature following Bloom (2009) (who also studies the variance of aggregate stock returns) arguing that shocks to uncertainty can have important effects on the aggregate economy. If increases in future uncertainty have sufficiently important effects on the economy that they affect investor utility, though, we would expect them to carry a risk premium. The fact that they do not implies that volatility shocks are not a major driver of welfare.

From a finance perspective, Merton’s (1973) intertemporal capital asset pricing model says that assets that have high returns in periods with good news about future investment opportunities are viewed as hedges and thus earn low average returns. Since expected future volatility is a natural state variable for the investment opportunity set, the covariance of an asset’s returns with shocks to future volatility should affect its expected return, but it does not.⁴

³The Sharpe ratio is equal to the average excess return divided by the standard deviation. It thus gives a scale-free measure of the risk premium on an asset.

⁴Recently, Campbell et al. (2014) and Bansal et al. (2013) estimate an ICAPM model with stochastic

Consumption-based models with Epstein–Zin (1991) preferences have similar predictions. Under Epstein–Zin preferences, marginal utility depends on lifetime utility, so that assets that covary positively with innovations to lifetime utility earn high average returns.⁵ If high expected volatility is bad for lifetime utility (either because volatility affects the path of consumption or because volatility reduces utility simply due to risk aversion), then volatility news should be priced.⁶

As a specific parameterized example with Epstein–Zin preferences, we study variance swap prices in Drechsler and Yaron’s (2011) calibrated long-run risk model. Drechsler and Yaron (2011) is a key benchmark because it is a quantitative model that can match a wide range of features of the dynamics of consumption growth, stock returns, and volatility. While the model represents a major innovation in being able to both generate a large variance risk premium (the average gap between the 1-month variance swap rate and realized variance) and match results about the predictability of market returns, we find that its implications for the term structure of variance swap prices and returns are distinctly at odds with the data: as one would expect, it predicts that shocks to future expected volatility should be strongly priced, counter to what we observe empirically.

We obtain similar results in Wachter’s (2013) model of time-varying disaster risk with Epstein–Zin preferences. The combination of predictability in the long-run probability of disaster and Epstein–Zin preferences results in a counterfactually high price for insurance against shocks to expected future volatility relative to current volatility. In both Wachter (2013) and Drechsler and Yaron (2011), Sharpe ratios earned by claims on future variance from 3 months to 14 years ahead are similar to those earned by claims to realized variance over the next month, whereas in the data the Sharpe ratios are all near zero (or positive) for claims to variance more than two months in the future. So both models fail to match our key stylized fact that only very short-term variance claims earn large negative Sharpe ratios.⁷

volatility and find that shocks to expected volatility (and especially long-run volatility) are priced in the cross-section of returns of equities and other asset classes. Although the focus on their paper is not the variance swap market, Campbell et al. (2014) test their specification of the ICAPM model also on straddle returns and synthetic volatility claims, and find that the model manages to explain only part of the returns on these securities. This suggests that the model is missing some high-frequency features of the volatility market.

⁵This is true in the most common calibrations with a preference for early resolution of uncertainty. When investors prefer a late resolution of uncertainty the risk prices are reversed.

⁶Also see Branger and Völkert (2010) and Zhou and Zhu (2012) for discussions. Barras and Malkhozov (2014) study the determinants of changes in the variance risk premium over time.

⁷Similar problems with matching term structures of Sharpe ratios in structural models have been studied in the context of claims to aggregate market dividends by van Binsbergen, Brandt, and Koijen (2012). Our results thus support and complement theirs in a novel context. See also van Binsbergen and Koijen (2015) for a recent review of the broad range of evidence on downward sloping term structures. Our paper also

More positively, we show that Gabaix’s (2012) model of rare disasters, which builds on the work of Rietz (1988), Barro (2006), and many others, can qualitatively (though not precisely quantitatively) match the stylized fact that Sharpe ratios on variance claims fall to zero rapidly with maturity. In Gabaix’s model, the probability of a disaster is constant, but the exposure of the stock market – its expected decline if a disaster occurs – varies over time. The realization of a disaster is inevitably a state with high realized volatility (since if returns are highly negative, squared daily returns will mechanically be high), so variance swaps provide a direct hedge against the occurrence of a disaster, meaning they earn a large negative Sharpe ratio. But since changes in the exposure of the stock market to consumption disasters, which drive expected future return variance, are uncorrelated with the *current level* of consumption, they are not priced shocks. The model is thus able to simultaneously generate a large negative premium on realized stock return variance and zero premium on news about future variance, just like in the data.⁸ That said, Gabaix’s (2012) model is not a complete quantitative description of financial markets; we simply view it as giving a set of sufficient conditions a model must satisfy to match the behavior of the variance swaps.

Our work is related to three main strands of the literature. First, there is the recent work in macroeconomics on the consequences of shocks to volatility, such as Bloom (2009), Bloom et al. (2014), Christiano, Motto, and Rostagno (2014), Fernandez-Villaverde et al. (2011), and Gourio (2012, 2013). We argue that if shocks to volatility are important to the macroeconomy, then investors should be willing to pay to hedge them. The lack of a risk premium on volatility news thus argues against theories in which aggregate volatility news is a major driver of business cycles.

Second, we build on the consumption-based asset pricing literature that has recently focused on the pricing of volatility, including Bansal and Yaron (2004), Drechsler and Yaron (2011), Wachter (2013), Campbell et al. (2014), and Bansal et al. (2013). We argue that consumption-based models with Epstein–Zin preferences are unlikely to explain the pricing of volatility claims.

Finally, there is a large extant literature studying the pricing of volatility in financial markets.⁹ Most closely related to us is a small number of recent papers with data on variance

relates to a large literature that looks at derivative markets to learn about general equilibrium asset pricing models, for example Backus, Chernov and Martin (2011) and Martin (2014, 2015).

⁸An alternative possibility is that the variance market is segmented from other markets, as in, e.g., Gabaix, Krishnamurthy, and Vigneron (2007). In that case, the pricing of risks might not be integrated between the variance market and other markets. We show, however, that our results hold not only with variance swaps, but also in VIX futures and in the options market, which is large, liquid, and integrated with equity markets, making it less likely that our results are idiosyncratic.

⁹A number of papers study the pricing of volatility in options markets, e.g. Jackwerth and Rubinstein

swaps with maturities from two to 24 months, including Egloff, Leippold, and Wu (2010) and Aït-Sahalia, Karaman, and Mancini (2014), who study no-arbitrage term structure models. The pricing models we estimate are less technically sophisticated than that of Aït-Sahalia, Karaman, and Mancini (2014), but we complement and advance their work in three ways. First, we examine a vast and novel range of data sources. For S&P 500 variance swaps, our panel includes data at both shorter and longer maturities than in previous studies – from one month to 14 years. The one-month maturity is important for giving a claim to shorter-term realized variance, which is what we find is actually priced. Having data at very long horizons is important for testing models, like Epstein–Zin preferences, in which expectations at very long horizons are the main drivers of asset prices. In addition, we are the first to examine the term structure of variance swaps for major international indexes, as well as for the term structure of the VIX obtained from options on those indexes. We are thus able to confirm that our results hold across a far wider range of markets, maturities, and time periods than previously studied.

Our second contribution to the previous term structure literature is that rather than working exclusively within the context of a particular no-arbitrage pricing model for the term structure of variance claims, we derive from the data more general and model-independent pricing facts. Our results can be directly compared against the implications of different structural economic models, which would be more difficult if they were only derived within a specific no-arbitrage framework. Our key finding, that purely transitory realized variance is priced while innovations to expectations are not, can be obtained from a simple reduced-form analysis and in data both for the United States and other countries. Nevertheless, we also confirm our results in a more formal no-arbitrage setting, whose main advantage is to yield much more precise estimates of risk prices.

Our third and most important contribution is to explore the implications of variance swaps for testing structural economic models. Our theoretical analysis leads us to the conclusion that the empirical facts in the variance swap market are most consistent with a model in which variance swaps are used to hedge the realization of market crashes and in which variation in expected future stock market volatility is not priced by investors, counter to the predictions of recent asset pricing and macroeconomic models.

The remainder of the paper is organized as follows. Section 2 describes the novel datasets

(1996), Coval and Shumway (2001), Bakshi and Kapadia (2003), Broadie, Chernov and Johannes (2009), Christofferson, Jacobs, Ornathanalai, and Wang (2008), and Kelly, Pastor, and Veronesi (2014). Lu and Zhu (2010) and Mencia and Sentana (2013) study VIX futures markets, while Bakshi, Panayotov, and Skoulakis (2011) show how to construct forward claims on variance with portfolios of options. In the Treasury bond market, Cieslak and Povala (2014) find, similar to us, that short-run volatility is more strongly priced than long-run volatility. See also Amengual and Xiu (2014) for an important recent study of jumps in volatility.

we obtain for variance swap prices. Section 3 reports unconditional means for variance swap prices and returns, which demonstrate our results in their simplest form. Section 4 analyzes the cross-sectional and time-series behavior of variance swap prices and returns more formally in standard asset pricing frameworks. In section 5, we discuss what structural general-equilibrium models can fit the data. We calibrate three leading models from the literature, comparing them to our data, showing that only one matches the key stylized facts. Section 6 concludes.

2 The data

This section discusses various ways that an investor can obtain exposure to volatility. They are all obviously closely related. We have data on all of the major markets, both in the US and internationally.

2.1 Variance swaps

We focus primarily on variance swaps. Variance swaps are contracts in which one party pays a fixed amount at maturity, which we refer to as price of the variance swap, in exchange for a payment equal to the sum of squared daily log returns of the underlying occurring until maturity. In this paper, the underlying is the S&P 500 index unless otherwise specified. The payment at expiration of a variance swap initiated at time τ and with maturity m is

$$Payoff_{\tau}^m = \sum_{j=\tau+1}^{\tau+m} r_j^2 - VS_{\tau}^m \quad (1)$$

where time here is indicated in days, r_j is the log return on the underlying on date j , and VS_{τ}^m is the price on date τ of an m -day variance swap. We focus on variance swaps because they give pure exposure to variance, their payoffs are transparent and easy to understand, they have a relatively long time-series, and they are relatively liquid.

Our main analysis focuses on two proprietary datasets of quoted prices for S&P 500 variance swaps.¹⁰ Dataset 1 contains monthly variance swap prices for contracts expiring in 1, 2, 3, 6, 12, and 24 months, and includes data from December, 1995, to October, 2013. Dataset 2 contains data on variance swaps with expirations that are fixed in calendar time,

¹⁰Both datasets were obtained from industry sources. Dataset 2 is obtained from Markit Totem, and reports averages of quotes obtained from dealers in the variance swap market. Since the prices we observe are a composite of quotes from many different dealers (on average 11), the quality of this dataset is very high, and comparable to that of the widely used CDS dataset from Markit.

instead of fixed maturities. Common maturities are clustered around 1, 3, and 6 months, and 1, 2, 3, 5, 10, and 14 years. Dataset 2 contains prices of contracts with maturities up to five years starting in September, 2006, and up to 14 years starting in August, 2007, and runs up to February, 2014. We apply spline interpolation to each dataset to obtain the prices of variance swaps with standardized maturities covering all months between 1 and 12 months for Dataset 1 and between 1 and 120 months for Dataset 2 (though in estimating the no-arbitrage model below we use the original price data without interpolation).¹¹

Both variance swap datasets are novel to the literature. Variance swap data with maturities up to 24 months as in Dataset 1 has been used before (Egloff, Leippold, and Wu, 2010, Ait-Sahalia, Karaman, and Mancini, 2014, and Amengual and Xiu, 2014), but the shortest maturity previous studies observed was two months. We show that the one-month variance swap is special in this market because it is the exclusive claim to next month’s realized variance, which is the *only* risk priced in this market. Observing the one-month variance swap is critical for precisely measuring the price of realized-variance risk.

This is also the first paper to observe and use variance swap data with maturity longer than two years. Since Epstein–Zin preferences imply that it is the very low-frequency components of volatility that should be priced (Branger and Volkert, 2010; Dew-Becker and Giglio, 2014), having claims with very long maturities is important for effectively testing the central predictions of Epstein–Zin preferences.

The variance swap market is large: the notional value of outstanding variance swaps at the end of 2013 was \$4 billion of notional vega.¹² A notional of \$4 billion vega means that an increase in annualized realized volatility of one percentage point induces total payments of \$4 billion. This market is thus small relative to the aggregate stock market, but it is non-trivial economically.

From a large market participant we obtained information about average bid-ask spreads by maturity. Typical bid-ask spreads are 1 to 2 percent for maturities up to 1 year, 2 to 3 percent between 1 and 2 years, and 3 to 4 percent for maturities up to 10 years. The bid-ask spreads are thus non-trivial, but also not so large as to prohibit trading. Moreover, they are small relatively to the volatility of the prices of these contracts. At the short end, the spreads are comparable to those found for corporate bonds by Bao, Pan, and Wang (2011) (which is particularly notable given that the volatility of variance swap returns is far larger

¹¹For the times and maturities for which we have both datasets, the prices are effectively identical: the correlations between the two datasets are never below 0.996. We will also show below that the prices are well explained by only two principal components, suggesting that interpolation should accurately recover prices.

¹²See the Commodity Futures Trading Commission’s (CFTC) weekly swap report. The values reported by the CFTC are consistent with data obtained from the Depository Trust & Clearing Corporation that we discuss below.

than that of corporate bonds).

Table 1 shows the total volume (in notional vega) for all transactions between March 2013 and June 2014. In little more than a year, the variance swap market saw \$7.2 billion of notional vega traded. Only 11 percent of the volume was traded in short maturity contracts (1-3 months); the bulk of the transactions occurred for maturities between 6 months and 5 years, and the median maturity was 12 months.

To check the accuracy of the quoted prices that we obtained, we compare them to those reported for actual trades by Depository Trust & Clearing Corporation (DTCC), which has collected data on all trades of variance swaps in the US since 2013.¹³ Appendix Figure A.1 shows the distribution of the percentage difference between our quotes and the transaction prices for different maturity baskets. Quotes and transaction prices are in most cases very close, with the median absolute percentage difference across all maturities approximately 1 percent.¹⁴

In addition to the prices of S&P 500 variance swaps, we also obtained prices for variance swaps in 2013 and 2014 for the FTSE 100 (UK), Euro Stoxx 50 (Europe), and DAX (Germany) indexes. This is the first paper to examine volatility claims in international markets and we show that our main results are consistent globally.

2.2 Options

Next, we examine claims to variance constructed using option prices. We focus on the VIX index, which is constructed using all available out-of-the money options and measures the risk-neutral expectation of squared variation (e.g. variance if the asset follows a diffusion, variance plus squared jumps if it follows a jump diffusion) integrated over the life of the options (Jiang and Tian (2005); Carr and Wu (2009)). The VIX is usually reported for a 30-day maturity, but the formulas are valid at any horizon. While other portfolios of options can be constructed that are also exposed to volatility in some way (for example straddles),

¹³DTCC was the only swap data repository registered under the Dodd–Frank act to collect data on variance swaps in 2013. The Dodd–Frank act requires that all swaps be reported to a registered data repository.

¹⁴Since variance swaps are traded over the counter, it is possible that counterparty risk could influence their prices. Given that variance swaps are standardized contracts covered by the ISDA Master Agreement, their margining follows standard procedures: an initial margin is posted by both parties, and variational margin is exchanged regularly depending on the value of the position. The residual counterparty risk in these contracts depends on the possibility of jumps in the value of the contracts between exchanges of collateral, and is therefore only a material issue when returns have high skewness and kurtosis at short horizons. As we will discuss later, only short-term variance swaps have payoffs that are far from Gaussian, and are therefore exposed to counterparty risk, and we argue below that for these contracts counterparty risk would push *against* the results we observe. We conclude that if counterparty risk was indeed priced by market participants, accounting for it would in fact make our results stronger.

the VIX is unique in that it represents a pure claim to the variance of the chosen horizon (or the squared variation when there are jumps), without any direct exposure to price movement of the underlying, independently of the horizon.¹⁵

The VIX is calculated based on an extraordinarily deep market. Options are traded in numerous venues, have notional values outstanding of trillions of dollars, and have been thoroughly studied.¹⁶ Since options are exchange-traded, they involve no counterparty risk, so we can use them to whether our results for variance swaps are affected by counterparty risk.

We construct VIX-type portfolios for the S&P 500, FTSE 100, Euro Stoxx 50, DAX, and CAC 40 indexes using data from Optionmetrics. We confirm our main results by showing that term structures and returns obtained from investments in options are similar to those obtained from variance swaps.¹⁷

2.3 VIX futures

Futures have been traded on the VIX since 2004. The VIX futures market is significantly smaller than the variance swap market, with current outstanding notional vega of approximately \$500 million.¹⁸ Bid/ask spreads are smaller than what we observe in the variance swap market, at roughly 0.1 percent, but as the market is smaller, we would expect price impact to be larger (and market participants claim that it is). We collected data on VIX futures prices from Bloomberg since their inception and show below that they yield nearly identical results to variance swaps.

More recently, a market has developed in exchange-traded notes and funds available to retail investors that are linked to VIX futures prices. These funds currently have an aggregate notional exposure to the VIX of roughly \$5 billion, making them comparable in size to the

¹⁵An alternative strategy to obtain variance exposure are straddles. Relative to the VIX, straddles are easier to construct, but they are claims on the absolute value of the return, not its variance, which makes the term structure more difficult to interpret given that expected absolute values are not linearly summable across periods. In addition, obtaining exposure to realized variance through straddles requires constant rebalancing, while variance swaps require no rebalancing throughout their lives (and the VIX is just the price of a synthetic variance swap).

¹⁶Even in 1990, Vijn (1990) noted that the CBOE was highly liquid and displayed little evidence of price impact for large trades.

¹⁷Recently, Boguth et al. (2012a,b) argue that returns measured on options portfolios can be substantially biased by noise, one potential source of which is the bid/ask spread. The majority of our results pertain directly to prices of volatility claims, as opposed to their returns, meaning that the issues noted by Boguth et al. are unlikely to affect our analysis. Furthermore, when we analyze returns, the portfolios are not levered to the degree that Boguth et al. argue causes biases in results.

¹⁸According to the CBOE futures exchange market statistics. See: <http://cfe.cboe.com/Data/HistoricalData.aspx>

variance swap market.

3 The term structure of variance claims

In this section we study average prices and returns of variance swaps. The key result that emerges is that only very short-duration variance claims earn a risk premium in our data.

3.1 Variance Swap Prices

The shortest maturity variance swap we consistently observe has a maturity of one month, so we treat a month as the fundamental period of observation. We define RV_t to be realized variance – the sum of squared daily log returns – during month t . The subscript from here forward always indexes months, rather than days.

Given a risk-neutral (pricing) measure Q , the price of an n -month variance swap at the end of month t , VS_t^n , is

$$VS_t^n = E_t^Q \left[\sum_{j=1}^n RV_{t+j} \right] \quad (2)$$

where E_t^Q denotes the mathematical expectation under the risk-neutral measure conditional on information available at the end of month t .¹⁹

Since an n -month variance swap is a claim to the sum of realized variance over months $t + 1$ to $t + n$, it is straightforward to compute prices of zero-coupon claims on realized variance. Specifically, we define an n -month zero-coupon variance claim as an asset with a payoff equal to realized volatility in month $t + n$. The absence of arbitrage implies

$$Z_t^n \equiv E_t^Q [RV_{t+n}] \quad (3)$$

$$= VS_t^n - VS_t^{n-1} \quad (4)$$

Z_t^n represents the market's risk-neutral expectation of realized variance n months in the future. We use the natural convention that

$$Z_t^0 = RV_t \quad (5)$$

¹⁹In the absence of arbitrage, there exists a probability measure Q such that the price of an asset with payoff X_{t+1} is $\frac{1}{R_{f,t+1}} E_t^Q [X_{t+1}]$, where $R_{f,t+1}$ is the risk-free interest rate. Under power utility, for example, we have $E_t^Q [X_{t+1}] = E_t^P \left[\frac{(C_{t+1}/C_t)^{-\rho}}{E_t^P [(C_{t+1}/C_t)^{-\rho}]} X_{t+1} \right]$, where ρ is the coefficient of relative risk aversion, C is consumption, and P is the physical probability measure. The price of a variance swap does not involve the interest rate because money only changes hands at the maturity of the contract.

so that Z_t^0 is the variance realized *during* the current month t . A one-month zero-coupon variance claim is exactly equivalent to a one-month variance swap, $Z_t^1 = VS_t^1$.

Figure 1 plots the time series of zero-coupon variance claim prices for maturities between one month and ten years. The figure shows all series in annualized percentage volatility units, rather than variance units: $100 \times \sqrt{12 \times Z_t^n}$ instead of Z_t^n . It also plots annualized realized volatility, $100 \times \sqrt{12 \times Z_t^0}$, in each panel. The top panel plots zero-coupon variance claim prices for maturities below one year, while the bottom panel focuses on maturities longer than one year.

The term structure of variance claim prices is usually weakly upward sloping. In times of distress, though, such as during the financial crisis of 2008, the short end of the curve spikes, temporarily inverting the term structure. Volatility obviously was not going to continue at crisis levels, so markets priced variance swaps with the expectation that it would fall in the future.

Figure 2 reports the average term structure of zero-coupon variance claims for two different subperiods – 2008–2014, a relatively short sample for which we have data for longer maturities, is in the top panel, while the full sample, 1996–2013, is in the bottom panel. The term structure of zero-coupon variance claim prices is upward sloping on average, but the figure also shows that it is concave, flattening out very quickly as the maturity increases. In particular, the curve is much steeper at the very short end than everywhere else.

The top and bottom panel of Figure 2 differ in both the time period and the maturities displayed. To better compare the two time periods, Figure 3 shows the average prices in both periods for maturities up to 12 months. The figure shows that after 2008 the level of the whole curve increased, and the curve became slightly steeper for maturities above 1 month. However, even after 2008 the curve is still much flatter at maturities above 3 months than it is at the very short end, displaying the same pattern as in the full sample. Of course, the economic significance of the "flatness" of the curve is easiest to understand within the context of a model. In section 5 we show formally that the curve of zero-coupon variance swaps is "too flat" in both subperiods relative to the implications of workhorse asset pricing models.

The average zero-coupon term structures in Figures 2 and 3 provide the first indication that the compensation for bearing risk associated with news about future volatility is small in this market. The return on holding a zero-coupon variance claim for a single month is $\frac{Z_{t+1}^{n-1} - Z_t^n}{Z_t^n}$. The average return is therefore closely related to the slope of the variance term structure. If the variance term structure is upward sloping between maturities $n - 1$ and n , zero-coupon claims of maturity n will have negative average returns, implying that it is

costly to buy insurance against increases in future expected volatility $n - 1$ months ahead. The fact that the curve is very steep at short horizons and flat at long horizons is a simple way to see that it is only the claims to variance in the very near future that earn significant negative returns.

3.2 Returns on zero-coupon variance claims

We now study the monthly returns on zero-coupon variance claims. The return on an n -month zero-coupon claim corresponds to a strategy that buys the n -month claim and sells it one month later as an $(n - 1)$ -month claim, reinvesting then again in new n -month zero-coupon variance claims. We define the excess return of an n -period variance claim following Gorton, Hayashi, and Rouwenhorst (2013)²⁰

$$R_{t+1}^n = \frac{Z_{t+1}^{n-1} - Z_t^n}{Z_t^n} \quad (6)$$

Given the definition that $Z_t^0 = RV_t$, the return on a one-month claim, R_{t+1}^1 is simply the percentage return on a one-month variance swap. We focus here on the returns for maturities of one to 12 months, for which we have data since 1995. All the results extend to higher maturities in the shorter sample.

Table 2 reports descriptive statistics for our panel of monthly returns. Only the average returns for the first and the second zero-coupon claims are negative, while all the others are zero or slightly positive. Return volatilities are also much higher at short maturities, though the long end still displays significant variability – returns on the 12-month zero-coupon claim have an annual standard deviation of 17 percent, which indicates that markets’ expectations of 12-month volatility fluctuates significantly over time.

Finally, note that only very short-term returns have high skewness and kurtosis. A buyer of short-term variance swaps is therefore potentially exposed to counterparty risk if realized variance spikes and the counterparty defaults. This should induce her to pay less for the insurance, i.e. we should expect the average return to be *less negative*. Therefore, the presence of counterparty risk on the short end of the term structure would bias our estimate towards not finding the large negative expected returns that we instead find. On the other hand, returns on longer-maturity zero-coupon claims have much lower skewness and kurtosis,

²⁰Note that $Z_{t+1}^{n-1} - Z_t^n$ is also an excess return on a portfolio since no money changes hands at the inception of a variance swap contract. Following Gorton, Hayashi, and Rouwenhorst (2013), we scale the return by the price of the variance claim bought. This is the natural scaling if the amount of risk scales proportionally with the price, as in Cox, Ingersoll, and Ross (1985). We have reproduced all of our analysis using the unscaled excess return $Z_{t+1}^{n-1} - Z_t^n$ as well and confirmed that all the results hold in that case.

which indicates that counterparty risk is substantially less relevant for longer maturities. Finally, we note that we obtain the same results using options, which are exchange traded and not affected by counterparty risk.

Given the different volatilities of the returns at different ends of the term structure, it is perhaps more informative to examine Sharpe ratios (average excess returns scaled by standard deviations), which measure compensation earned per unit of risk. Figure 4 shows the annual Sharpe ratios of the 12 zero-coupon claims. The Sharpe ratios are negative for the one- and two-month claims (at -1.4 and -0.5, respectively), but all other Sharpe ratios are insignificantly different from zero. The results at the short end of the curve indicate that investors are willing to pay a large premium to hedge realized volatility. What is new and surprising in this picture is the fact that investors are *not* willing to pay to hedge any innovations in *expected* volatility, even two or three months ahead. A claim to volatility at a horizon beyond one month is purely exposed to news about future volatility: its return corresponds exactly to the change in expectations about volatility at its maturity. Pure news about future expected volatility will therefore affect its return, whereas purely transitory shocks to volatility that disappear before its maturity will not affect it at all. Our results therefore show that news about future volatility has commanded a small to zero risk premium.²¹

3.3 Evidence from other markets

The results for variance swaps can also be confirmed in the options market. We exploit the well-known fact that if realized variance is calculated using sufficiently small time periods, a variance swap can be replicated by a portfolio of options with the same maturity (Jiang and Tian (2005)).²² The term structure of synthetic zero-coupon variance claim prices constructed from options should then align well with the term structure of actual variance swap prices. The appendix reports details of the construction of the synthetic variance swap prices.

²¹The declining term structure of Sharpe ratios on short positions in volatility is consistent with the finding of van Binsbergen, Brandt, and Koijen (2012) that Sharpe ratios on claims to dividends decline with maturity, and that of Duffee (2011) that Sharpe ratios on Treasury bonds decline with maturity. For a review, see van Binsbergen and Koijen (2015).

²²The most famous use of that result is in the construction of the VIX index, which uses 1-month options, and corresponds to the price of a 1-month variance swap. Due to the discretization inherent in the practical calculation of the payoff of a variance swap (using daily returns, rather than instantaneous returns), the VIX is not identical to the price of a variance swap. While the difference in prices of variance swaps and option-based synthetic contracts is economically informative, its magnitude is far smaller than the differences in prices of zero-coupon volatility claims across maturities (Bollerslev and Todorov (2011) and Ait-Sahalia et al. (2014)).

Figure 5 shows the term structure of prices and Sharpe ratios of zero-coupon variance claims obtained from the variance swap data compared to the synthetic claims for maturities up to 1 year. While the curves obtained using options data seem noisier, the curves deliver the same message: the volatility term structure is extremely steep at the very short end but quickly flattens out for maturities above two months, and Sharpe ratios rapidly approach zero as the maturity passes two months.²³ Appendix Figure A.3 shows that we obtain similar results with VIX futures.^{24,25}

Figure 6 shows that our results also extend to international markets. Figure 6 plots average term structures obtained from both variance swaps and synthetic option-based variance claims for the Euro Stoxx 50, FTSE 100, CAC 40 and DAX indexes.²⁶ To ease the comparison across markets, in this figure we plot the term structures relative to the prices of the respective 2-month claims, so that all the curves are equal to 1 at the two-month maturity.

Both panels of the figure show that the international term structures have an average shape that closely resembles the one observed for the US (the solid line in both panels), demonstrating that our results using US variance swaps extend to the international markets.²⁷

4 Asset pricing

We now formally examine the pricing of risks in the variance market.

4.1 Reduced-form estimates

We begin by exhibiting our main pricing result in a simple reduced-form setting: investors pay to hedge the immediate realized volatility but not shocks to expected volatility. To test that claim, we need to disentangle shocks to realized variance from shocks to expectations

²³Given the high liquidity of the options market, we might have expected option-based portfolios to be less noisy. However, the synthetic variance portfolios load heavily on options very far out of the money where liquidity is relatively low. This demonstrates another advantage of studying variance swaps instead of options.

²⁴VIX futures are not exactly comparable to variance swaps because they are claims on the VIX , not on VIX^2 . A convexity effect makes the prices of claims on variance and volatility different, but the Figure shows that it is quantitatively small.

²⁵We have also compared our data to the CBOE's 3-month and mid-term volatility indexes (VXV and VXMT) and find that our data is nearly identical to those two series in the period when the CBOE calculates them.

²⁶We do not plot Sharpe ratios for these markets because the data is of relatively poor quality, and the series of returns are very noisy.

²⁷In the appendix (Figure A.2) we also confirm that for the indexes for which we have both variance swap prices and synthetic prices obtained from options, the two curves align well.

of future volatility. This subsection focuses just on the returns of the variance claims with maturity of 12 months or less since they require less interpolation; all the economically interesting results are clearly visible in that maturity range.

4.1.1 Extracting innovations

As usual in the term structure literature, we begin by extracting principal components from the term structure of zero-coupon variance claims. The first factor explains 97.1 percent of the variation in the term structure and the second explains an additional 2.7 percent. The loadings of the variance swaps on the factors are plotted in the top panel of Figure 7, while the time series of the factors are shown in the bottom panel of the figure. The first factor captures the level of the term structure, while the second measures the slope. As we would expect, during times of crisis, the slope turns negative. The level factor captures the longer-term trend in volatility and clearly reverts to its mean more slowly.

The two factors explain 99.9 percent of the variation in variance swap prices and thus encode essentially all the information contained in variance swap prices. So if we find that the shocks to both factors are unpriced, then that means that no forward-looking information in the term structure, whether it is driven by expectations for volatility or risk premia, is priced.

To extract shocks to variance and expectations, we estimate a first-order vector autoregression (VAR) with the two principal components and realized variance (RV). Including RV in the VAR allows us to separately identify shocks to the term structure of variance swaps and transitory shocks to realized volatility. The three estimated innovations are positively correlated: the correlation between RV and level shocks is 0.7, and that between RV and slope shocks is 0.6.

We rotate the three shocks using a Cholesky factorization where the first shock affects all three variables, the second affects only the slope and RV , and the third shock affects only RV . We will therefore refer to the third shock as the pure RV shock. The pure RV shock allows us to measure the price of risk for a shock that has only a transitory effect on realized variance and no effect on the term structure of variance swap prices, while the other two rotated shocks affect both current realized variance and also expectations of future variance. Impulse response functions are reported in Appendix Figure A.4.

4.1.2 Risk prices

We estimate risk prices for the three shocks using the Fama–MacBeth (1973) procedure on 1- to 12-month zero-coupon variance claims.²⁸ The top panel of Table 3 reports the loadings of each variance swap return on the three orthogonalized shocks. Short-maturity variance swaps are exposed to all three shocks with the expected signs. The higher maturities are mostly exposed to the level and slope shocks, with essentially no exposure to the pure *RV* shock.

The bottom panel of Table 3 reports the estimated annualized risk prices. Of the three shocks, only the pure *RV* shock has a statistically significant risk price. The risk price is also economically highly significant: it implies that an asset that was exposed only to the pure *RV* shock would earn an annualized Sharpe ratio of -2.72. Since the three shocks all have the same standard deviation, the magnitudes of the risk prices are directly comparable. Those for shocks 1 and 2 are five to eight times smaller than that for the pure *RV* shock, and thus economically far less important.²⁹

Table 3 thus shows that investors do not price shocks to the level and slope, but they accept large negative returns to hedge transitory *RV* shocks. No forward-looking information about volatility is priced.

4.1.3 Controlling for the market return

One possible explanation for why realized variance is priced is that it provides a good hedge for aggregate market shocks. To test that possibility, we add the market return as an additional factor in the estimation.³⁰ The first column of Table 4 shows that indeed the zero-coupon volatility claims are heavily exposed to the market return. But when the pure *RV* shock is included, the market return is no longer significantly priced. The R^2 of the model for the cross-section of average returns also rises from 37.7 to 99.7 percent when the pure *RV* shock is included.

²⁸The results are robust to estimating the risk prices using one- and two-step GMM.

²⁹Despite the good fit of the model in terms of R^2 , the GMM and the GRS test reject the null that all the average pricing errors are zero. This is because the pricing errors, while being small relative to the overall average returns of these contracts, are still statistically different from zero. The same applies to the tests in the remaining sections of the paper.

³⁰We add the market return as a test asset to impose discipline on its risk premium. For readability and to ensure that the risk premium on the market is matched relatively closely, we increase the weight on the market return by a factor of 12 as a test asset in our cross-sectional tests. That way, the market return carries as much weight in the pricing tests as do all the variance claims combined. The market factor, though, is still the monthly market return, as are all our zero-coupon variance returns.

4.1.4 Upside and downside volatility

As discussed above, realized variance is a combination of the diffusive component of returns and the realization of jumps. Our pricing results show that investors want to hedge the transitory component of this combination, but not news about its future expected values. So far, we have not disentangled the two components: the most striking of our empirical results is that news about future expected RV is unpriced, which means that investor must not care about *both* shocks to expected diffusive volatility and shocks to the expected jump probability. We do not need to disentangle the two to conclude that investors care about neither.³¹

However, it is interesting to understand why investors are willing to pay so much to hedge transitory realized variance. Existing literature on the variance risk premium, focusing on the short-term variance risk premium (particularly, the 30-day VIX), has addressed this question by explicitly modeling the jump process. Bollerslev and Todorov (2011), for example, show that investors are mostly worried about unexpected price jumps, and that is the component of transitory volatility they are paying to hedge.

Here we follow a different approach to show that agents mostly desire protection against realized downside component of realized variance.³² Following Andersen and Bondarenko (2007), we decompose the realized variance in a month, RV_t , into an upper and a lower semivariance: the integrated realized variances computed only when prices are above or below a threshold. In particular, following Andersen and Bondarenko (2007) we construct the upper RV in each month as

$$RV_t^U = \sum_{j \in t} (r_j)^2 I_j(F_j > F_0)$$

where $j \in t$ indicates days j in month t , $I_j(F_j > F_0)$ is an indicator that the futures price F_j of the underlying in day j is above the starting point F_0 at the beginning of the month. Similarly, we construct

$$RV_t^D = \sum_{j \in t} (r_j)^2 I_j(F_j \leq F_0)$$

Andersen and Bondarenko (2007) discuss two useful properties of these realized barrier variances (or semivariances), which can be interpreted as the volatility of the upward and

³¹One alternative possibility would be that for some reason news about diffusive volatility and news about jump risk have risk prices with opposite signs. We leave investigation of that theory for future work and simply note for now that the extant asset pricing models we examine below do not consider it. In the models we are aware of, investors are averse to both diffusive and jump-related volatility.

³²See also Kilic and Shaliastovic (2015).

downward price movements. First, the two components sum to RV_t ,

$$RV_t = RV_t^U + RV_t^D$$

Second, the price of claims to RV_t^U and RV_t^D can be obtained from option prices in a manner similar to how the VIX is computed. We refer to these two prices as VIX_t^U and VIX_t^D . $(VIX_t^U)^2$ is the no-arbitrage price of a contract whose payoff is RV_{t+1}^U , and $(VIX_t^D)^2$ is the no-arbitrage price of a contract whose payoff is RV_{t+1}^D . Andersen and Bondarenko (2007) also derive a relation between the three prices,

$$VIX_t^2 = (VIX_t^U)^2 + (VIX_t^D)^2$$

Finally, just as in the case of the VIX , we can compute the prices of the two claims for different maturities and study the term structure.

Figure 8 plots the term structure of the zero-coupons obtained from VIX , as well as those for VIX^U and VIX^D . As before, maturity zero corresponds to the average RV_t , RV_t^U and RV_t^D , respectively. The slopes between the zero- and one-month maturities then represent precisely the returns on the 30-day VIX , VIX^U , and VIX^D . We can see that most of the negative average return that investors are willing to accept to hold the VIX comes from the extremely negative monthly return of the VIX^D (about -30% per month), while VIX^U commands a return much closer to zero.³³ This confirms the intuition that the reason investors dislike realized volatility is due to its downside component (which Bollerslev and Todorov (2011) show is dominated by downward jumps).

4.2 The predictability of volatility

Since the key result of the paper concerns the pricing of volatility shocks at different horizons, a natural question is how much news there actually is about future volatility. The total risk premium for assets that hedge volatility news – which we showed at the beginning of Section 3 to be zero – is the product of the quantity of risk (how much news about future volatility investors receive), and the price of this risk (how averse investors are to such news). Perhaps the reason that investors are willing to pay little to hedge volatility news is that there is not a lot of news in the first place. If volatility is not predictable, there is no news for investors to actually hedge, and the quantity of risk is low. We argue that it is the *price* of risk, not the quantity of risk, that explains why risk premia for variance news are low. Investors do

³³Note that contrary to the case of the VIX , for VIX^U and VIX^D the slope between maturities above one month cannot be interpreted exactly in terms of returns since the barrier is moving over time.

receive news about future volatility, but they seem indifferent to that news.

We begin by noting that the results in the previous section *already* explicitly focus on the *price* of risk of volatility, rather than the *quantity*. The reported risk prices measure compensation per unit of risk, so they are unaffected by how much news there is about future volatility. If volatility were not very predictable, the quantity of news risk would be low, but the price per unit of risk would still be estimated correctly from our cross-sectional regressions. *So our previous analysis already shows that the low risk premia are due to a low price of variance news risk.*³⁴

And indeed, a large empirical literature shows clearly that volatility is predictable. Andersen et al. (2003), Ait-Sahalia and Mancini (2008), Bandi, Russell, and Yang (2008), and Brownlees, Engle, and Kelly (2011) show that volatility is predictable based on lagged returns of the underlying and past volatility. Campbell et al. (2014) focus on longer horizons (up to 10 years) and show that both the aggregate price-earnings ratio and the Baa-Aaa default spread are useful predictors of long-run volatility.

Moreover, the literature has found that multiple factors are needed to explain the dynamics of volatility expectations. Standard stochastic volatility models (e.g. Adrian and Rosenberg (2008)) as well as the reduced-form VAR model of Campbell et al. (2014) all point to the existence of different factors driving the volatility process, with the factors differing by their level of persistence. The presence of different factors driving the current and future evolution of RV is what allows the separate estimation of the price of realized variance shocks and the prices of shocks to future expected variance. Our pricing results from the previous sections exploit the multivariate nature of the volatility process to price separately shocks to RV and shocks to future expectations of RV, and the standard errors on the risk prices reflect our ability to empirically distinguish between the different shocks.

To give a sense of the magnitude of the predictability of volatility at different horizons Table 5 reports R^2 s from predictive regressions for realized volatility at different frequencies and horizons. One of the advantages of working with predictors other than variance swap prices is that we can use a much longer sample to establish the predictability of volatility, since we can compute RV and the predictors all the way back to 1926 (we also look at the predictability by variance swap prices in the appendix, despite the much shorter sample available).

³⁴Of course, the lower the amount of risk, the harder it is to estimate the price of risk; this effect is fully captured by the standard errors on the estimates of the price of risk. In the limit, if there were no news about future volatility at all, our analysis would not be able to provide an estimate of the price of risk of volatility news. Instead, in the data there is sufficient news about future expected volatility to enable us to pin down the price of these shocks (more precisely so when imposing the restrictions of the term structure model, explored below).

The first pair of columns focuses on forecasts of monthly realized variance, while the second pair repeats the exercise at the annual frequency. The R^2 s for monthly volatility range from 45 percent at the 1-month horizon to 20 percent at the 12-month horizon. In predicting annual volatility, R^2 s range between 56 and 21 percent for horizons of 1 to 10 years.

The third pair of columns in Table 5 reports, as a comparison, the results of forecasts of dividend growth.³⁵ R^2 s for dividend growth are never higher than 9 percent. So in the context of financial markets, there is an economically large amount of predictability of volatility. The appendix takes an extra step beyond Table 5 and shows, using Fama and Bliss (1987) and Campbell and Shiller (1991) regressions, that nearly all the variation in variance swap prices is actually due to variations in expected volatility, rather than risk premia.

We conclude by noting that while there is ample evidence of the predictability of volatility at the horizons relevant for this analysis (from 3 months upwards), the result that the risk premium for volatility news is close to zero would have strong implications for macroeconomic and financial models even if it was driven by low quantity of expected volatility risk. If there is not much volatility news, then the macro literature showing that volatility news can drive the business cycle would seem irrelevant. Similarly, asset pricing models in which news about future volatility plays an important role (like the ICAPM or several versions of the long-run-risks model) would lose this source of priced risk. Nonetheless, in this paper we build on the wealth of evidence supporting the predictability of volatility, and show that the high *price* of expected volatility risk implied by many models is strongly at odds with what we observe in the data.

4.3 A no-arbitrage model

In this section, we extend the pricing results reported above by considering a more formal estimation. We analyze a standard no-arbitrage term structure model for variance swaps. The model delivers implications strongly supportive of our reduced-form results. Because the no-arbitrage model uses the prices of the variance swaps, rather than just their returns, and because it uses a full no-arbitrage structure, it is able to obtain much more precise estimates of risk prices. We show that not only are the risk prices on the level and slope factors statistically insignificant, but they are also economically small.

The no-arbitrage model has three additional advantages over the reduced-form analysis: it explicitly allows for time-variation in the volatility of shocks to the economy and risk prices,

³⁵We compare predictability of volatility to that of dividends since realized variance in each month is the stochastic payment of the variance swap contract in that month.

the standard errors for the risk prices take into account uncertainty about the dynamics of the economy (through the VAR), and it links us more directly to the previous literature. Furthermore, because the inputs to the estimation of the no-arbitrage model are the observed variance swap prices rather than monthly returns, the results in this section do not rely on any interpolation and we can simultaneously use the full time series from 1996 to 2013 and every maturity from one month to 14 years.

4.3.1 Risk-neutral dynamics

As above, we assume that the term structure of variance swaps is governed by a bivariate state vector $(s_t^2, l_t^2)'$. Rather than state the factors as a level and slope, we now treat them as a short- and a long-term component, which will aid in the estimation process. s_t^2 is the one-month variance swap price: $s_t^2 = E_t^Q [RV_{t+1}]$. The other state variable, l_t^2 , governs the central tendency of s_t^2 .

We begin by specifying the conditional risk-neutral mean of the states,

$$E_t^Q \left[\begin{pmatrix} s_{t+1}^2 \\ l_{t+1}^2 \\ RV_{t+1} \end{pmatrix} \right] = \begin{pmatrix} \rho_s^Q & 1 - \rho_s^Q & 0 \\ 0 & \rho_l^Q & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} s_t^2 \\ l_t^2 \\ RV_t \end{pmatrix} + \begin{pmatrix} 0 \\ v_l^Q \\ 0 \end{pmatrix} \quad (7)$$

where v_l^Q is a constant to be estimated which captures the unconditional mean of realized variance. l_t^2 can be viewed as the risk-neutral trend of s_t^2 . The first two rows of (7) are the discrete-time counterpart to the standard continuous-time setup in the literature, e.g. Egloff, Leippold, and Wu (2010) and Ait-Sahalia, Karaman, and Mancini (2014).³⁶ We diverge from Egloff, Leippold, and Wu (2010) and Ait-Sahalia, Karaman, and Mancini (2014) in explicitly specifying a separate process for realized variance, noting that it is not spanned by the other shocks. The specification of a separate shock to RV_{t+1} allows us to ask how shocks to both realized variance and the term structure factors are priced.³⁷

Given the assumption that $s_t^2 = E_t^Q [RV_{t+1}]$, the price of an n -period variance swap VS_t^n

³⁶For admissibility, we require that $0 < \rho_s^Q < 1$, $\rho_l^Q > 0$, and $v_l^Q > 0$. These restrictions ensure that risk-neutral forecasts of s_t^2 and l_t^2 , hence variance swap prices at various maturities, are strictly positive.

³⁷From a continuous-time perspective, it is not completely obvious how to think about a "shock" to realized variance that is completely transitory. There are two standard interpretations. One is that the innovation in RV_{t+1} represents the occurrence of jumps in the S&P 500 price. Alternatively, there could be a component of the volatility of the diffusive component of the index that has shocks that last less than one month. At some point, the practical difference between a jump and an extremely short-lived change in diffusive volatility is not obvious. The key feature of the specification is simply that there are shocks to the payout of variance swaps that are orthogonal to both past and future information contained in the term structure.

is

$$VS_t^n = E_t^Q \left[\sum_{i=1}^n RV_{t+i} \right] = E_t^Q \left[\sum_{i=1}^n s_{t+i-1}^2 \right] \quad (8)$$

which can be computed by applying (7) repeatedly, and which implies that VS_t^n is affine in s_t^2 and l_t^2 for any maturity.

4.3.2 Physical dynamics and risk prices

Define $X_t \equiv (s_t^2, l_t^2, RV_t)'$. We assume that X follows a VAR(1) under the physical measure:³⁸

$$\begin{pmatrix} s_{t+1}^2 \\ l_{t+1}^2 \\ RV_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ v_l^Q \\ 0 \end{pmatrix} + \begin{pmatrix} \rho_s & 1 - \rho_s^Q & 0 \\ 0 & \rho_l & 0 \\ \rho_{s,RV} & 0 & 0 \end{pmatrix} \begin{pmatrix} s_t^2 \\ l_t^2 \\ RV_t \end{pmatrix} + \varepsilon_{t+1} \quad (9)$$

$$\varepsilon_{t+1} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, V_t(X_{t+1}) \right) \quad (10)$$

In our main results, we follow Egloff, Leippold, and Wu (2010) and Ait-Sahalia, Karaman, and Mancini (2014) and assume that the market prices of risk are proportional to the states, so that the log SDF, m_{t+1} , is

$$m_{t+1} - E_t[m_{t+1}] = \Lambda_t' V_t(X_{t+1})^{-1/2} \varepsilon_{t+1} \quad (11)$$

$$\text{where } \Lambda_t = \begin{pmatrix} \lambda_s s_t \\ \lambda_l l_t \\ \lambda_{RV} s_t \end{pmatrix} \quad (12)$$

where the superscript $^{1/2}$ indicates a lower triangular Cholesky decomposition. The term $V_t(X_{t+1})^{-1/2}$ standardizes and orthogonalizes the shocks ε_{t+1} . Λ_t thus represents the price of exposure to a unit standard deviation shock to each component of X_{t+1} .

To maintain the affine structure of the model, we need the product $V_t(X_{t+1})^{1/2} \Lambda_t$ to be affine in X_t . The specification for Λ_t in (12) is therefore typically accompanied by a structure

³⁸Admissibility requires that v_l^Q and the feedback matrix in (9) be non-negative, which ensures that the forecasts of X_t , and hence future volatility, be strictly positive.

for the conditional variance similar to that of Cox, Ingersoll, and Ross (1985),

$$V_t(X_{t+1}) = \begin{pmatrix} \sigma_s^2 s_t^2 & 0 & \sigma_{s,RV} s_t^2 \\ 0 & \sigma_l^2 l_t^2 & 0 \\ \sigma_{s,RV} s_t^2 & 0 & \sigma_{RV}^2 s_t^2 \end{pmatrix} \quad (13)$$

which guarantees that $V_t(X_{t+1})^{1/2}\Lambda_t$ is affine in X_t .³⁹

4.3.3 Empirical results

The estimation uses standard likelihood-based methods. The appendix describes the details. We use both Dataset 1 and Dataset 2, meaning that the number of variance swap prices used in the estimation varies over time depending on availability.

Model fit Table 6 reports the means and standard deviations of the variance swap prices observed and fitted by our model together with the corresponding root mean squared errors (RMSE). The average RMSE across maturities up to 24 months is 0.73 annualized volatility points (i.e. the units in Figure 1).⁴⁰ For maturities longer than 24 months, since we do not have time series of variance swap prices with fixed maturities for the entire sample, we cannot report the sample and fitted moments for any fixed maturity. Instead, we stack all contracts with more than 24 months to maturity into one single series and compute the RMSE from the observed and fitted values of this series. The corresponding RMSE is reported in the last row of Table 6. At 0.87 percentage points, it compares favorably with the RMSE for the shorter maturities. Table 6 suggests that our models with two term structure factors plus RV are capable of pricing the cross-section of variance swap prices for an extended range of maturities. Even when maturities as long as 14 years are included in estimation, the data does not seem to call for extra pricing factors.

Risk prices The steady-state risk prices in the model are reported in Table 7 along with their standard errors. As in the previous analysis, we find clearly that it is the purely

³⁹It is important to note that the specifications of Λ_t in (12) and $V_t(X_{t+1})$ in (13) introduce tight restrictions on the difference $E_t(X_{t+1}) - E_t^Q(X_{t+1})$. In the appendix, we therefore consider two alternative specifications for the variance process $V_t(X_{t+1})$ and the risk prices Λ_t that are more flexible in certain dimensions. The results, both in terms of point estimates and standard errors, are essentially identical across the various specifications that we consider, so we report results for this specification here since it is most common in the literature.

⁴⁰When we exclude the financial crisis, using a sample similar to that of Egloff, Leippold, and Wu (2010), we obtain an RMSE of 0.33 percentage points, which is nearly identical to their reported value. The increase in fitting error in the full sample is, not surprisingly, brought about by the large volatility spikes that occurred during the crisis.

transitory shock to realized variance that is priced (*RV*-risk). The Sharpe ratio associated with an investment exposed purely to the transitory *RV* shock – analogous to the pure *RV* shock above – is -1.70.

In the VAR analysis in the previous section, the pure *RV* shock had no immediate impact on the level and slope factors, but it could potentially indirectly affect future expected variance through the VAR feedback. In the no-arbitrage model, that effect is shut off through the specification of the dynamics. That is, the *RV* shock here is completely transitory – it has no impact on expectations of volatility on any future date. The other two shocks are forced to account for all variation in expectations. The fact that the results are consistent between the no-arbitrage model and the reduced-form analysis in the previous section helps underscore the robustness of our findings to different modeling assumptions.

The short- and long-term factors earn risk premia of only -0.11 and -0.18, respectively, neither of which is significantly different from zero. The lack of statistical significance is not due to particularly large standard errors; the standard errors for the risk prices for the s_t^2 and l_t^2 shocks are in fact substantially smaller than that for the *RV* shock. Moreover, Sharpe ratios of -0.11 and -0.18 are also economically small. For comparison, the Sharpe ratio of the aggregate stock market in the 1996–2013 period is 0.43. So the risk premia on the short- and long-term components of volatility are between 25 and 42 percent of the magnitude of the Sharpe ratio on the aggregate stock market. On the other hand, the Sharpe ratio for the *RV* shock is nearly four times larger than that for the aggregate stock market and 10 to 15 times larger than the risk prices on the other two shocks. Our no-arbitrage model thus clearly confirms the results from the previous sections.

Time-series dynamics The estimated parameters determining the dynamics of the state variables under the physical measure are (equation 9):

$$\begin{pmatrix} s_{t+1}^2 \\ l_{t+1}^2 \\ RV_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0.99^{***} \\ 0 \end{pmatrix} + \begin{pmatrix} 0.82^{***} & 0.16^{**} & 0 \\ 0 & 0.98^{***} & 0 \\ 0.75^{***} & 0 & 0 \end{pmatrix} \begin{pmatrix} s_t^2 \\ l_t^2 \\ RV_t \end{pmatrix} + \varepsilon_{t+1}$$

The key parameter to focus on is the persistence of l_t^2 . The point estimate is 0.9814, with a standard error of 0.0013. At the point estimate, long-term shocks to variance have a half-life of 37 months. That level of persistence is actually *higher* than the persistence of consumption growth shocks in Bansal and Yaron’s (2004) long-run risk model, and only slightly smaller than the persistence they calibrate for volatility, 0.987. Our empirical model thus allows us to estimate risk prices on exactly the type of long-run shocks that have been considered in calibrations. As we discuss further below, the fact that we find that the

long-term shock to volatility is unpriced is strongly at odds with Epstein–Zin preferences.

5 Economic interpretation

The key message of our empirical analysis is that the average investor in the variance swap market is not willing to pay for protection against news about high future volatility. In other words, investors do not hedge volatility intertemporally. That fact immediately suggests that models based on Epstein–Zin (1991) preferences, where intertemporal hedging effects are central, will struggle to match the data. To confirm that intuition, we simulate two models with Epstein–Zin preferences. The first is the long-run risk model proposed by Drechsler and Yaron (2011), and the second is a discrete-time version of the model with time-varying disaster risk proposed by Wachter (2013). In both cases, we show that the models imply that the Sharpe ratios earned from rolling over long-term zero-coupon variance claims are almost as negative as those earned from holding just the one-month variance swap, counter to what we observe empirically in Figure 4.

The evidence we provide that there is no intertemporal hedging runs counter to many models beyond Epstein–Zin. Merton’s ICAPM, for example, implies that shocks to expected volatility should be priced since volatility affects the investment opportunity set.⁴¹ The variance swap market thus is not well explained by the ICAPM. Similarly, in models with value-at-risk or leverage constraints, the constraint on financial intermediaries depends on expected volatility, rather than realized volatility.⁴² In general, then, it is forward-looking volatility that is relevant in most asset pricing models.

The lack of intertemporal hedging in the variance swap market suggests a myopic model of investors. We therefore consider a simple model in which investors have power utility. While it is well known that the power utility model fails to match many asset pricing facts when consumption follows a process with low volatility, Rietz (1988), Barro (2006), Martin (2013), and others show that allowing for a small probability of a large decline in consumption can render the power utility model consistent with standard asset pricing moments. Gabaix (2012) extends the disaster model to allow for a time-varying exposure of the stock market to disasters. We find that Gabaix’s model is able to match both the qualitative and quantitative features of the variance swap market. This suggests that investors in the variance swap market are mostly worried about large negative shocks to the economy in which returns

⁴¹This is true even if volatility does not predict returns. If volatility rises but expected returns remain constant, then the investment opportunity set has deteriorated.

⁴²For example, financial intermediaries might be limited in the total amount of risk they may take. When expected volatility is higher, their demand for risky securities will fall.

collapse and variance spikes, and are purchasing variance swaps to hedge these, and only these, shocks.⁴³

5.1 Structural models of the variance premium

5.1.1 A long-run risk model

Drechsler and Yaron (2011), henceforth DY, extend Bansal and Yaron's (2004) long-run risk model to allow for jumps in both the consumption growth rate and volatility. DY show that the model can match the mean, volatility, skewness, and kurtosis of consumption growth and stock market returns, and generates a large variance risk premium that forecasts market returns, as in the data. DY is thus a key quantitative benchmark in the literature.

The structure of the endowment process is

$$\Delta c_t = \mu_{\Delta c} + x_{t-1} + \varepsilon_{c,t} \quad (14)$$

$$x_t = \mu_x + \rho_x x_{t-1} + \varepsilon_{x,t} + J_{x,t} \quad (15)$$

$$\bar{\sigma}_t^2 = \mu_{\bar{\sigma}} + \rho_{\bar{\sigma}} \bar{\sigma}_{t-1}^2 + \varepsilon_{\bar{\sigma},t} \quad (16)$$

$$\sigma_t^2 = \mu_{\sigma} + (1 - \rho_{\sigma}) \bar{\sigma}_{t-1}^2 + \rho_{\sigma} \sigma_{t-1}^2 + \varepsilon_{\sigma,t} + J_{\sigma,t} \quad (17)$$

where Δc_t is log consumption growth, the shocks ε are mean-zero and normally distributed, and the shocks J are jump shocks. σ_t^2 controls both the variance of the normally distributed shocks and also the intensity of the jump shocks. There are two persistent processes, x_t and $\bar{\sigma}_t^2$, which induce potentially long-lived shocks to consumption growth and volatility. We follow DY's calibration for the endowment process exactly.

Aggregate dividends are modeled as

$$\Delta d_t = \mu_d + \phi x_{t-1} + \varepsilon_{d,t} \quad (18)$$

Dividends are exposed to the persistent but not the transitory part of consumption growth. Equity is a claim on the dividend stream, and we treat variance claims as paying the realized variance of the return on equities.

DY combine that endowment process with Epstein–Zin preferences, and we follow their calibration. Because there are many parameters to calibrate, we refer the reader to DY for the full details. However, the parameters determining the volatility dynamics are obviously

⁴³We do not explicitly consider here habit-formation models as in Campbell and Cochrane (1999), since conditional on the habit the agent behaves as a standard power utility investor, with an effective risk aversion that depends on the current level of the habit.

critical to our analysis. Note that the structure of equations (16) and (17) is the same as the VAR in our no-arbitrage model in equation (9). The parameters governing volatility in DY's calibration and the corresponding values from our estimation are:

	DY	Estimates
ρ_σ	0.87	0.82
$\rho_{\bar{\sigma}}$	0.987	0.9814
$\text{stdev}(\varepsilon_{\bar{\sigma},t})$	0.10	0.05
$\text{stdev}(\varepsilon_{\sigma,t} + J_{\sigma,t})$	1.10	1.48

The two feedback coefficients, ρ_σ and $\rho_{\bar{\sigma}}$, are nearly identical to our estimated values. Their long-term component, $\bar{\sigma}^2$, has a persistence of 0.987, which compares favorably with our estimate of 0.9814. Similarly, their calibration of $\rho_\sigma = 0.87$ is comparable to our estimate of 0.82. The calibration deviates somewhat more in the standard deviations of the innovations.

Overall, though, DY's calibration implies volatility dynamics highly similar to what we observe empirically. The close match is not surprising as DY's model was calibrated to fit the behavior of the (one-month) VIX and realized variance. As a robustness check, though, in the appendix we also simulate DY setting the standard deviations of the innovations to match our empirical estimates and obtain implications for variance swap prices that are essentially unchanged.

Given the high quality of DY's calibration, if the long-run risk model fails to match the term structure of variance swap prices, it is not because it has an unreasonable description of the dynamics of volatility. Rather, we would conclude that the failure is due to the specification of the preferences, namely Epstein–Zin with a representative agent.

5.1.2 Time-varying disaster risk

The second model we study is a discrete-time version of Wachter's (2013) model of time-varying disaster risk. In this case, consumption growth follows the process,

$$\Delta c_t = \mu_{\Delta c} + \sigma_{\Delta c} \varepsilon_{\Delta c,t} + J_{\Delta c,t} \quad (19)$$

where $\varepsilon_{\Delta c,t}$ is a mean-zero normally distributed shock and J_t is a disaster shock. The probability of a disaster in any period is F_t , which follows the process

$$F_t = (1 - \rho_F) \mu_F + \rho_F F_{t-1} + \sigma_F \sqrt{F_{t-1}} \varepsilon_{F,t} \quad (20)$$

The CIR process ensures that the probability of a disaster is always positive in the continuous-time limit, though it can generate negative values in discrete time. We calibrate the model similarly to Wachter (2013) and Barro (2006). Details of the calibration are reported in the appendix. The model is calibrated at the monthly frequency. In the calibration, the steady-state annual disaster probability is 1.7 percent as in Wachter (2013). σ_F is set to 0.0075 (ε_F is a standard normal), and $\rho_F = 0.87^{1/12}$, which helps generate realistically volatile stock returns and a persistence for the price/dividend ratio that matches the data. If there is no disaster in period t , $J_t = 0$. Conditional on a disaster occurring, $J_t \sim N(-0.30, 0.15^2)$, as in Barro (2006). Finally, dividends are a claim to aggregate consumption with a leverage ratio of 2.8.⁴⁴

Wachter (2013) combines this specification of disasters with Epstein–Zin preferences. One of her key results is that a model with time-varying disaster risk and power utility has strongly counterfactual predictions for the behavior of interest rates and other asset prices. She thus argues that time-varying disaster risk should be studied in the context of Epstein–Zin preferences. We follow her in assuming the elasticity of intertemporal substitution is 1, and we set risk aversion to 3.6.⁴⁵

5.1.3 Time-varying recovery rates

The final model we study is a version of Gabaix’s (2012) model of disasters with time-varying recovery rates. Because the probability of a disaster is constant, power utility and Epstein–Zin are equivalent in terms of their implications for risk premia. We use power utility in our calibration, which eliminates the intertemporal hedging motives present in the two previous models. In this model, the expected value of firms following a disaster is variable. Specifically, we model the consumption process identically to equation (19) above, but with the probability of a disaster, F_t , fixed at 1 percent per year (Gabaix’s calibration). Following Gabaix, dividend growth is

$$\Delta d_t = \mu_{\Delta d} + \lambda \varepsilon_{\Delta c,t} - L_t \times 1 \{J_{\Delta c,t} \neq 0\} \quad (21)$$

⁴⁴The occurrence of a disaster shock implies that firm value declines instantaneously. To calculate realized variance for periods in which a disaster occurs, we assume that the shock occurs over several days with maximum daily return of -5 percent. For example, a jump of 20% would occur over 4 consecutive days, with a 5% decline per day. This allows for a slightly delayed diffusion of information and also potentially realistic factors such as exchange circuitbreakers. The small shocks $\varepsilon_{\Delta c,t}$ are treated as though they occur diffusively over the month, as in Drechsler and Yaron (2011).

⁴⁵Given the calibration of the endowment, if risk aversion is raised any higher the model does not have a solution. The upper bound on risk aversion is a common feature of models in which the riskiness of the economy varies over time.

λ here represents leverage. $1\{\cdot\}$ is the indicator function. Dividends are thus modeled as permanently declining by an amount L_t on the occurrence of a disaster. The value of L is allowed to change over time and follows the process

$$L_t = (1 - \rho_L) \bar{L} + \rho_L L_{t-1} + \varepsilon_{L,t} \quad (22)$$

We calibrate $\bar{L} = 0.5$ and $\rho_L = 0.87^{1/12}$ as in the previous model, and $\varepsilon_{L,t} \sim N(0, 0.16)$. We set the coefficient of relative risk aversion to 7 to match the Sharpe ratio on one-month variance swaps. Other than the change in risk aversion, our calibration of the model is nearly identical to Gabaix's (2012), which implies that we will retain the ability to explain the same ten puzzles that he examines. He did not examine the ability of his model to match the term structure of variance claims, so this paper provides a new test of the theory.

5.2 Results

We now examine the implications of the three models for the zero-coupon variance curve.⁴⁶ Figure 9 plots population moments from the models against the values observed empirically. The top panel reports annualized Sharpe ratios for zero-coupon variance claims with maturities from 1 to 12 months. Our calibration of Gabaix's model with time-varying recovery rates matches the data well: it generates a Sharpe ratio for the one-month claim of -1.3, while all the forward claims earn Sharpe ratios of zero, similarly to what we observe in the data.

The two models with Epstein–Zin preferences, on the other hand, both generate Sharpe ratios for claims on variance more than one month ahead that are counterfactually large, especially when compared to the Sharpe ratio of the 1-month variance swap. For both the long-run risk and the time-varying disaster model, the Sharpe ratio on the three-month variance claim is roughly three-fourths as large as that on the one-month claim, whereas the three-month claim actually earns a slightly positive return empirically.

The economic intuition for the result is straightforward. If investors are risk-averse, then periods of high volatility are periods of low utility. And under Epstein–Zin preferences, periods with low lifetime utility are periods with high marginal utility. Investors thus desire to hedge news about future volatility, and forward variance claims allow them to do so. Moreover, volatility in all future periods affects lifetime utility symmetrically (discounted by the rate of pure time preference), which is why investors in these models pay the same

⁴⁶We focus on variance swaps rather than the VIX or SVIX because we have a full term structure out to ten years, whereas the VIX and SVIX are limited by data on option prices. The conclusions are unchanged across the contracts.

amount to hedge volatility at any horizons.

The expected returns on the variance claims are closely related to the average slope of the term structure. The bottom panel of Figure 9 reports the average term structure in the data and in the models. The figure shows, as we would expect, that neither model with Epstein–Zin preferences generates a curve that is as concave as we observe in the data. Instead, the DY model generates a curve that is too steep everywhere (including on the very long end), while the time-varying disaster model generates a curve that is too flat everywhere.⁴⁷ On the other hand, the average term structure in the model with time-varying recovery rates qualitatively matches what we observe in the data – it is steep initially and then perfectly flat after the first month.

The comparison between the calibrated models and the data reported in Figure 9 does not take into account the statistical uncertainty due to the fact that we only observe variance swap prices in a specific sample. To directly test the models against the data, we simulate the calibrated models and verify how likely we would be to see a period in which the variance swap curve looks like it does in our data. In particular, the left-hand column of Figure 10 plots results from 10,000 215-month simulations of the models (the number of months in our full sample). In each simulation, we calculate the average term structure of the variance curve, and normalize the value at the third month to 1 so that we are sampling the shape of the term structure rather than its level. We then plot the median and 95-percent sampling interval of the term structure from the simulations.

Both the long-run risk model and time-varying disaster risk model have a hard time in matching the empirical shape of the variance term structure in the simulations, and particularly producing a steep slope at the short end of the curve and a small slope for higher maturities. Gabaix’s model with time-varying recovery rates performs *qualitatively* better: the 1-month variance swap is priced significantly higher than average realized volatility, but the slope is zero for all the rest of the curve. The long-run risk and time-varying disaster models are statistically rejected at the short end of the curve, while the long-run risk model is rejected at the long end of the curve. It is true that Gabaix’s model does not provide an accurate quantitative fit, but its success is nevertheless notable given how few free parameters it has.

The right-hand column of figure 10 simulates variance swap prices in the models out to maturities of 10 years. The sampling intervals are wider because our sample with 10-year maturities only runs for 70 months. The story is similar to that in the right-hand column, though: all three models fail quantitatively, but the time-varying recovery model is the one

⁴⁷The models have similar Sharpe ratios but different slopes of the term structure because the latter depends on the expected return, not the Sharpe ratio.

that best matches the qualitative features of the term structure.

DY’s long-run risk model is calibrated in such a way that it nearly exactly matches our estimated volatility dynamics and it still fails to match the basic features of the term structure of variance claims. We therefore conclude from Figures 9 and 10 that models with a major intertemporal hedging motive, such as Epstein–Zin preferences, do not match the main features of the variance swap market. On the other hand, a model in which investors have power utility, and hence make investment choices myopically, is able to better match our data.

The main features of the models that affect their ability to match our data can be summarized as follows. In models with Epstein–Zin preferences, investors will pay to hedge shocks to expected future volatility, especially at long horizons. Long-term zero-coupon variance claims should thus have large negative returns because they provide such a hedge. But in the data, we observe shocks to future expected volatility and find that their price is close to zero. Models with power utility, or where the variation in expected stock market volatility is independent of consumption volatility, solve that problem since investors are myopic and shocks to future expected volatility are not priced. However, the models also need to explain the high risk price associated with the realized volatility shock. In a power utility framework, this can be achieved if states of the world with high volatility are associated with large drops in consumption, as in a disaster model.

5.3 The behavior of volatility during disasters

In order for variance swaps to be useful hedges in disasters, realized volatility must be high during large market declines. A number of large institutional asset managers sell products meant to protect against tail risk that use variance swaps, which suggests that they or their investors believe that realized volatility will be high in future market declines.⁴⁸

In the spirit of Barro (2006), we now explore the behavior of realized volatility during consumption disasters and financial crises using a panel data of 17 countries, covering 28 events. We obtain two results. First, volatility is indeed significantly higher during disasters. Second, the increase in volatility is not uniform during the disaster period; rather, volatility spikes for one month only during the disaster and quickly reverts. It is those short-lived but extreme spikes in volatility that make variance swaps a good product to hedge tail risk.

We collect daily market return data from Datastream for a total of 37 countries since 1973. We compute realized volatility in each month for each country. To identify disasters,

⁴⁸In particular, see Man Group’s TailProtect product (Man Group (2014)), Deutsche Bank’s ELVIS product (Deutsche Bank, 2010) and the JP Morgan Macro Hedge index.

we use both the years marked by Barro (2006) as consumption disasters and the years marked by Schularick and Taylor (2012), Reinhart and Rogoff (2009) and Bordo et al. (2001) as financial crises.⁴⁹ Given the short history of realized volatility available, our final sample contains 17 countries for which we observe realized volatility and that experienced a disaster during the available sample. Table 8 shows for each country the first year of our RV sample and the years we identify as consumption or financial disasters.

The first three columns of Table 8 compare the monthly annualized realized volatility during disaster and non-disaster years. Column 1 shows the maximum volatility observed in any month of the year identified as a disaster averaged across all disasters for each country. Column 2 shows the average volatility during the disaster years, and column 3 shows the average volatility in all other years.

Comparing columns 2 and 3, we can see that in almost all cases realized volatility is indeed higher during disasters. For example, in the US the average annualized realized volatility is 25 percent during disasters and 15 percent otherwise. Column 1 reports the average across crises of the highest observed volatility. Within disaster years there is large variation in realized volatility: the maximum volatility is always much higher than the average volatility, even during a disaster. Disasters are associated with large spikes in realized volatility, rather than a generalized increase in volatility during the whole period.

To confirm this result, in Figure 11 we perform an event study around the peak of volatility during a disaster. For each country and for each disaster episode, we identify the month of the volatility peak during that crisis (month 0) and the three months preceding and following it. We then scale the volatility behavior by the value reached at the peak, so that the series for all events are normalized to 1 at the time of the event. We then average the rescaled series across our 28 events.

The figure shows that indeed, the movements in volatility that we observe during disasters are short-lived spikes, where volatility is high for essentially only a single month. In the single months immediately before and after the one with the highest volatility, volatility is 40 percent lower than its peak, and it is lower by half both three months before and after the worst month.

6 Conclusion

This paper shows that it is only the transitory part of realized variance that is priced. That fact is not consistent with a broad range of structural asset pricing models. It is

⁴⁹See Giglio et al. (2014) for a more detailed description of the data sources.

qualitatively consistent with a model in which investors desire to hedge rare disasters, but even that model does not match all the quantitative features of the data. Interestingly, the data is not consistent with all disaster models. The key feature that we argue models need in order to match our results is that variation in expected stock market volatility is not priced by investors, whereas the transitory component of volatility is strongly priced.

The idea that variance claims are used to hedge crashes is consistent with the fact that many large asset managers, such as Deutsche Bank, JP Morgan, and Man Group sell products meant to hedge against crashes that use variance swaps and VIX futures. These assets have the benefit of giving tail protection, essentially the form of a long put, but also being delta neutral (in an option-pricing sense). They thus require little dynamic hedging and yield powerful protection against large declines.

More broadly, shocks to expected volatility, such as that observed during the recent debt ceiling debate, are a major driving force in many current macroeconomic models. If aggregate volatility shocks are a major driver of the economy, we would expect investors to desire to hedge them. We find, though, that the average investor is indifferent to such shocks. The evidence from financial markets is thus difficult to reconcile with the view that volatility shocks are an important driver of business cycles or welfare.

References

- Adrian, Tobias and Joshua Rosenberg**, “Stock Returns and Volatility: Pricing the Short-Run and Long-run Components of Market Risk,” *The Journal of Finance*, 2008, *63*(6), 2997–3030.
- Ait-Sahalia, Yacine and Lorian Mancini**, “Out of Sample Forecasts of Quadratic Variation,” *Journal of Econometrics*, 2008, *147* (1), 17–33.
- , **Mustafa Karaman, and Lorian Mancini**, “The Term Structure of Variance Swaps, Risk Premia and the Expectations Hypothesis,” 2014. Working paper.
- Amengual, Dante and Dacheng Xiu**, “Resolution of policy uncertainty and sudden declines in volatility,” 2014. Working paper.
- Andersen, Torben G. and Oleg Bondarenko**, “Construction and Interpretation of Model-Free Implied Volatility,” 2007. Working paper.
- Andersen, Torben, Tim Bollerslev, Francis Diebold, and Paul Labys**, “Modeling and Forecasting Realized Volatility,” *Econometrica*, 2003, *71*(2), 579–625.
- Backus, David, Mikhail Chernov, and Ian Martin**, “Disasters Implied by Equity Index Options,” *Journal of Finance*, 2011, *66*(6), 1969–2012.
- Baker, Scott R., Nicholas Bloom, and Steven J. Davis**, “Measuring Economic Policy Uncertainty,” *Working Paper*, 2014.
- Bakshi, Gurdip and Nikunj Kapadia**, “Delta-Hedge Gains and the Negative Market Volatility Risk Premium,” *The Review of Financial Studies*, 2003, *16*(2), 527–566.
- , **George Panayotov, and Georgios Skoulakis**, “Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios,” *Journal of Financial Economics*, 2011, *100*(3), 475–495.
- Bandi, Jeffrey Russell Federico and Chen Yang**, “Realized Volatility Forecasting and Option Pricing,” *Journal of Econometrics*, 2008, *147*(1), 34–46.
- Bansal, Ravi and Amir Yaron**, “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles,” *Journal of Finance*, 2004, *59* (4), 1481–1509.
- , **Dana Kiku, Ivan Shaliastovich, and Amir Yaron**, “Volatility, the Macroeconomy and Asset Prices,” *Journal of Finance*, 2013.
- Bao, Jack, Jun Pan, and Jiang Wang**, “The Illiquidity of Corporate Bonds,” *The Journal of Finance*, 2011, *66* (3), 911–946.
- Barras, Laurent and Aytek Malkhozov**, “Variance Risk Premium Dynamics in Equity and Option Markets,” 2014. Working paper.

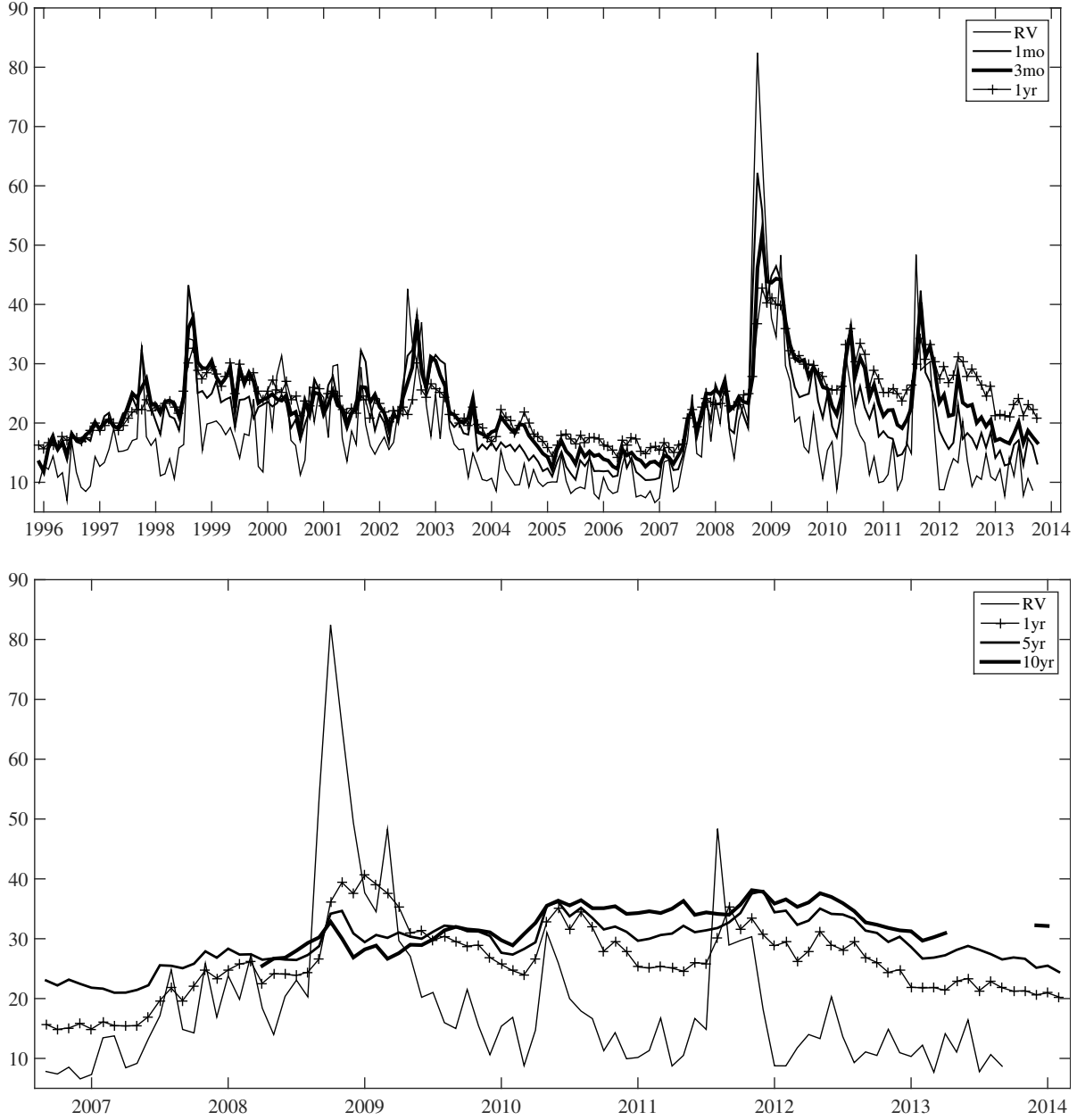
- Barro, Robert J.**, “Rare Disasters and Asset Markets in the Twentieth Century,” *Quarterly Journal of Economics*, 2006, *121*(3), 823–866.
- Bloom, Nicholas**, “The Impact of Uncertainty Shocks,” *Econometrica*, 2009, *77*(3), 623–685.
- , **Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry**, “Really Uncertain Business Cycles,” 2014. Working paper.
- Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin**, “Leverage and the Limits of Arbitrage Pricing: Implications for Dividend Strips and the Term Structure of Equity Risk Premia,” 2012. Working paper.
- , —, —, and —, “More on Levered Noise and Arbitrage Pricing,” 2012. Working paper.
- Bollerslev, Tim and Viktor Todorov**, “Tails, fears, and risk premia,” *Journal of Finance*, 2011, *66*(6), 2165–2211.
- Bordo, Michael, Barry Eichengreen, Daniela Klingebiel, and Maria Soledad Martinez-Peria**, “Is the crisis problem growing more severe?,” *Economic policy*, 2001, *16* (32), 51–82.
- Branger, Nicole and Clemens Volkert**, “What is the Equilibrium Price of Variance Risk? A Long-Run Risk Model with Two Volatility Factors,” 2010. Working paper.
- Broadie, Mark, Mikhail Chernov, and Michael Johannes**, “Understanding Index Option Returns,” *The Review of Financial Studies*, 2009, *22*(11), 4493–4529.
- Brownlees, Christian, Robert Engle, and Bryan Kelly**, “A Practical Guide to Volatility Forecasting Through Calm and Storm,” *Journal of Risk*, 2011, *14*(2), 1–20.
- Campbell, John Y. and John H. Cochrane**, “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, 1999, *107* (2), 205–251.
- , **Stefano Giglio, Christopher Polk, and Robert Turley**, “An Intertemporal CAPM with Stochastic Volatility,” 2013. Working paper.
- Carr, Peter and Liuren Wu**, “Variance Risk Premiums,” *Review of Financial Studies*, 2009, *22*(3), 1311–1341.
- Christiano, Lawrence J., Roberto Motto, and Massimio Rostagno**, “Risk Shocks,” *American Economic Review*, 2014, *104*(1), 27–65.
- Christoffersen, Peter, Kris Jacobs, Chayawat Ornathanalai, and Yintian Wang**, “Option Valuation with Long-run and Short-run Volatility Components,” *Journal of Financial Economics*, 2008, *90*(3), 272–297.

- Cieslak, Anna and Pavol Povala**, “Information in the Term Structure of Yield Curve Volatility,” 2014. Working paper.
- Coval, Joshua D. and Tyler Shumway**, “Expected Option Returns,” *The Journal of Finance*, 2001, *56*(3), 983–1009.
- Cox, John C., Jr. Jonathan E. Ingersoll, and Stephen A. Ross**, “A Theory of the Term Structure of Interest Rates,” *Econometrica*, 1985, *53*(2), 385–407.
- Deutsche Bank Pension Strategies & Solutions**, “Tail Risk Hedging: A Roadmap for Asset Owners,” May 2010.
- Dew-Becker, Ian and Stefano Giglio**, “Asset Pricing in the Frequency Domain: Theory and Empirics,” 2013. Working paper.
- Drechsler, Itamar and Amir Yaron**, “What’s Vol Got to Do with it?,” *The Review of Financial Studies*, 2011, *24*(1), 1–45.
- Duffee, Gregory R.**, “Sharpe ratios in term structure models,” 2011. Working paper.
- Egloff, Daniel, Madrkus Leippold, and Liuren Wu**, “The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments,” *Journal of Financial and Quantitative Analysis*, 2010, *45*(5), 1279–1310.
- Epstein, Larry G. and Stanley E. Zin**, “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis,” *The Journal of Political Economy*, 1991, *99*(2), 263–286.
- Exchange, Chicago Board Options**, “The CBOE Volatility Index - VIX,” Technical Report, Chicago Board Options Exchange 2009.
- Fama, Eugene F. and James D. MacBeth**, “Risk, Return, and Equilibrium: Empirical Tests,” *Journal of Political Economy*, 1973, *81*(3), 607–636.
- and **Robert R. Bliss**, “The Information in Long-Maturity Forward Rates,” *The American Economic Review*, 1987, *77*(4), 680–692.
- Fernandez-Villaverde, Jesus, Pablo Guerron, Juan F. Rubio-Ramirez, and Martin Uribe**, “Risk Matters: The Real Effects of Volatility Shocks,” *American Economic Review*, 2011.
- Gabaix, Xavier**, “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance,” *Quarterly Journal of Economics*, 2012, *127*(2), 645–700.
- , **Arvind Krishnamurthy, and Olivier Vigneron**, “Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market,” *Journal of Finance*, 2007, *62*(2), 557–595.

- Giglio, Stefano, Matteo Maggiori, and Johannes Stroebe**, “Very Long-Run Discount Rates,” *Unpublished Manuscript New York University, and University of Chicago*, 2014.
- Gorton, Gary B., Fumio Hayashi, and K. Geert Rouwenhorst**, “The Fundamentals of Commodity Futures Returns,” *Review of Finance*, 2013, *17*(1), 35–105.
- Gourio, Francois**, “Disaster Risk and Business Cycles,” *American Economic Review*, 2012, *102*(6), 2734–2766.
- , “Credit Risk and Disaster Risk,” *American Economic Journal: Macroeconomics*, 2013, *5*(3), 1–34. Working paper.
- Inc., Man Investments**, “The case for TailProtect Strategy,” 2014.
- Jackwerth, Jens Carsten and Mark Rubinstein**, “Recovering Probability Distributions from Option Prices,” *The Journal of Finance*, 1996, *51*(5), 1611–1631.
- Jiang, George J. and Yisong S. Tian**, “The Model-Free Implied Volatility and its Information Content,” *Review of Financial Studies*, 2005, *18* (4), 1305–1342.
- Kelly, Bryan, Lubos Pastor, and Pietro Veronesi**, “The Price of Political Uncertainty: Theory and Evidence from the Option Market,” 2014. Working paper.
- Kilic, Mete and Ivan Shaliastovich**, “Good and Bad Variance Premium & Expected Stock Returns,” 2015. Working Paper.
- Lu, Zhongjin and Yingzi Zhu**, “Volatility Components: The Term Structure Dynamics of VIX Futures,” *The Journal of Futures Markets*, 2010, *30*(3), 230–256.
- Martin, Ian**, “Consumption-Based Asset Pricing with Higher Cumulants,” *Review of Economic Studies*, 2013, *80*.
- , “What is the Expected Return on the Market?,” *Working Paper*, 2014.
- , “Simple Variance Swaps,” *Working Paper*, 2015.
- Mencia, Javier and Enrique Sentana**, “Valuation of VIX Deivatives,” *Journal of Financial Economics*, 2013, *108*(2), 367–391.
- Merton, Robert C.**, “An Intertemporal Capital Asset Pricing Model,” *Econometrica*, 1973, *41* (5), 867–887.
- Mixon, Scott**, “The implied volatility term structure of stock index options,” *Journal of Empirical Finance*, 2007, *14*, 333–354.
- Reinhart, Carmen M. and Kenneth S. Rogoff**, *This time is different: Eight centuries of financial folly*, Princeton University Press, 2009.
- Rietz, Thomas A.**, “The Equity Risk Premium: A Solution,” *Journal of Monetary Economics*, 1988, *22*(1), 117–131.

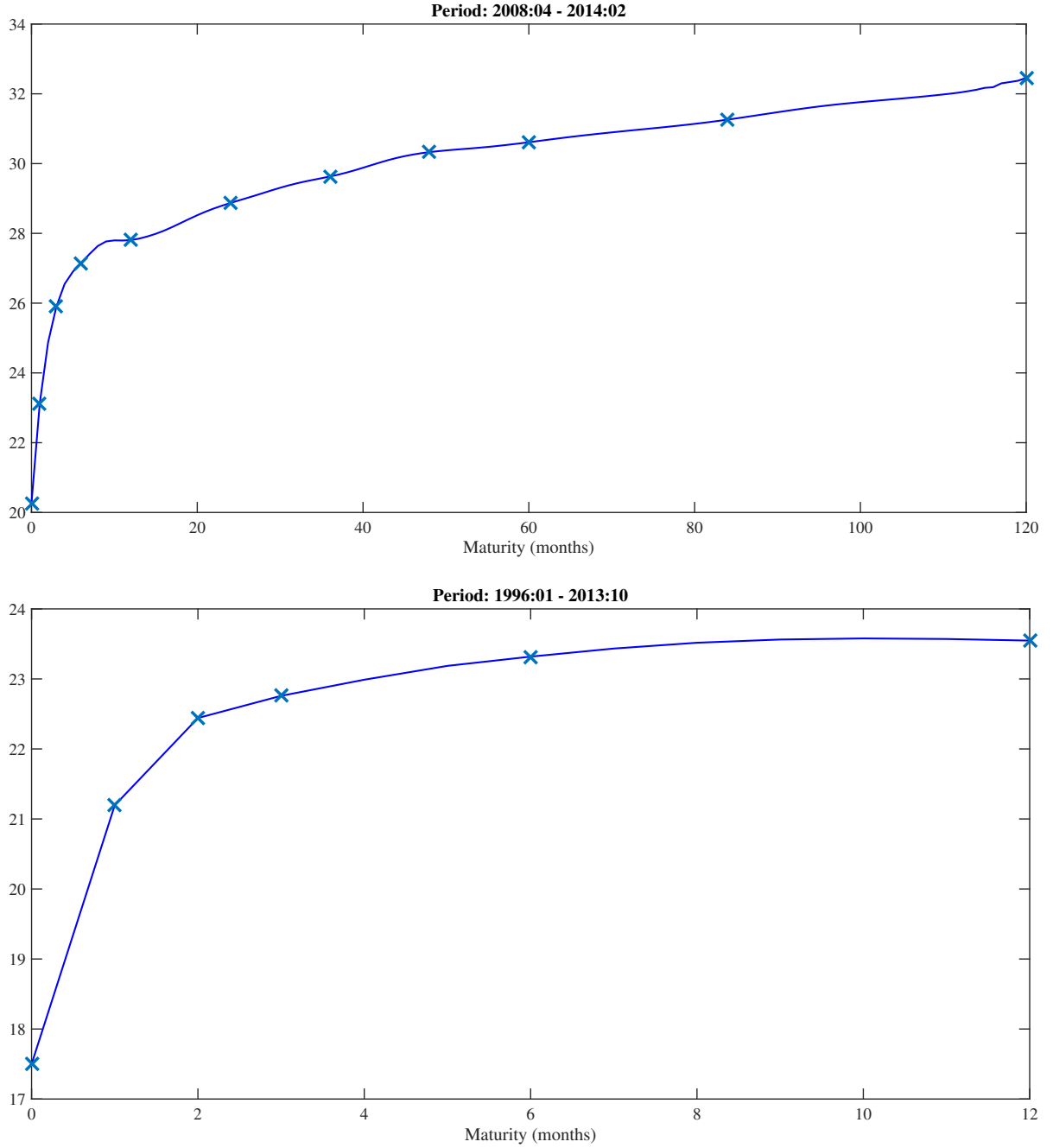
- Schularick, Moritz and Alan M. Taylor**, “Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008,” *American Economic Review*, 2012, *102* (2), 1029–1061.
- van Binsbergen, Jules H. and Ralph S.J. Koijen**, “The Term Structure of Returns: Facts and Theory,” 2015. Working paper.
- , **Michael W. Brandt, and Ralph Koijen**, “On the Timing and Pricing of Dividends,” *American Economic Review*, 2012, *102*, 1596–1618.
- Vijh, Anand M.**, “Liquidity of the CBOE Equity Options,” *Journal of Finance*, 1990, *45*(4), 1157–1179.
- Wachter, Jessica A.**, “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?,” *Journal of Finance*, 2013, *68*(3), 987–1035.
- Zhou, Guofu and Yingzi Zhu**, “Volatility Trading: What Is the Role of the Long-Run Volatility Component?,” *Journal of Financial and Quantitative Analysis*, 2012, *47*(2), 273–307.

Figure 1: Time series of zero-coupon variance claim prices



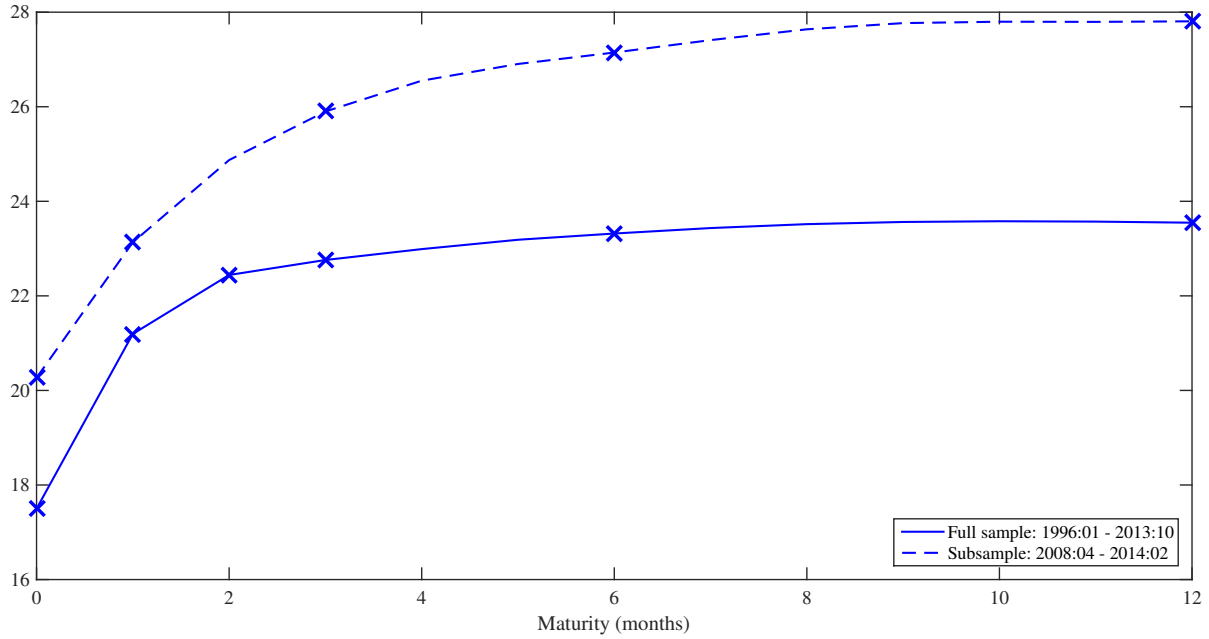
Note: The figure shows the time series of zero-coupon variance claim prices of different maturities. For readability, each line plots the prices in annualized volatility terms, $100 \times \sqrt{12 \times Z_t^n}$, for a different n . The top panel plots zero-coupon variance claim prices for maturities of 1 month, 3 months, and one year. The bottom panel plots zero-coupon variance claim prices for maturities of 1 year, 5 years and 10 years. Both panels also plot annualized realized volatility, $100 \times \sqrt{12 \times Z_t^0}$.

Figure 2: Average zero-coupon variance claim prices



Note: The figure shows the average prices of zero-coupon variance claims of different maturity, across different periods. The top panel shows average prices between 2008 and 2013, when we observe maturities up to 10 years. The bottom panel shows averages between 1996 and 2013, for claims of up to 1 year maturity. In each panel, the "x" mark prices of maturities we directly observe in the data (for which no interpolation is necessary). All prices are reported in annualized volatility terms, $100 \times \sqrt{12 \times Z_t^n}$. Maturity zero corresponds to average realized volatility, $100 \times \sqrt{12 \times Z_t^0}$.

Figure 3: Subsample analysis of zero-coupon variance claims



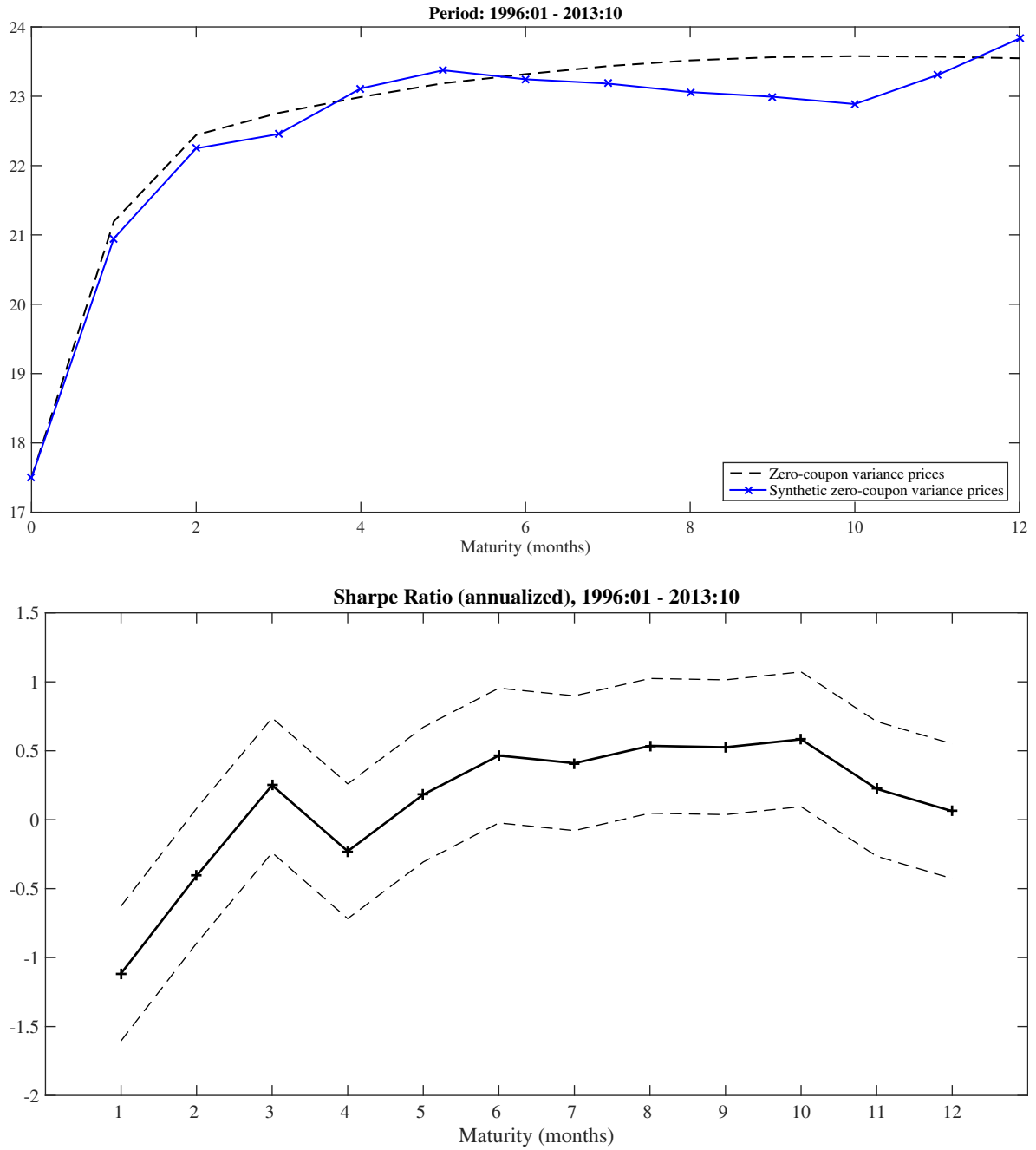
Note: The figure compares the average prices of zero-coupon variance claims for maturities up to 1 year, for the two subsamples of the top and bottom panel of figure 2.

Figure 4: Annualized Sharpe ratios for zero-coupon variance claims



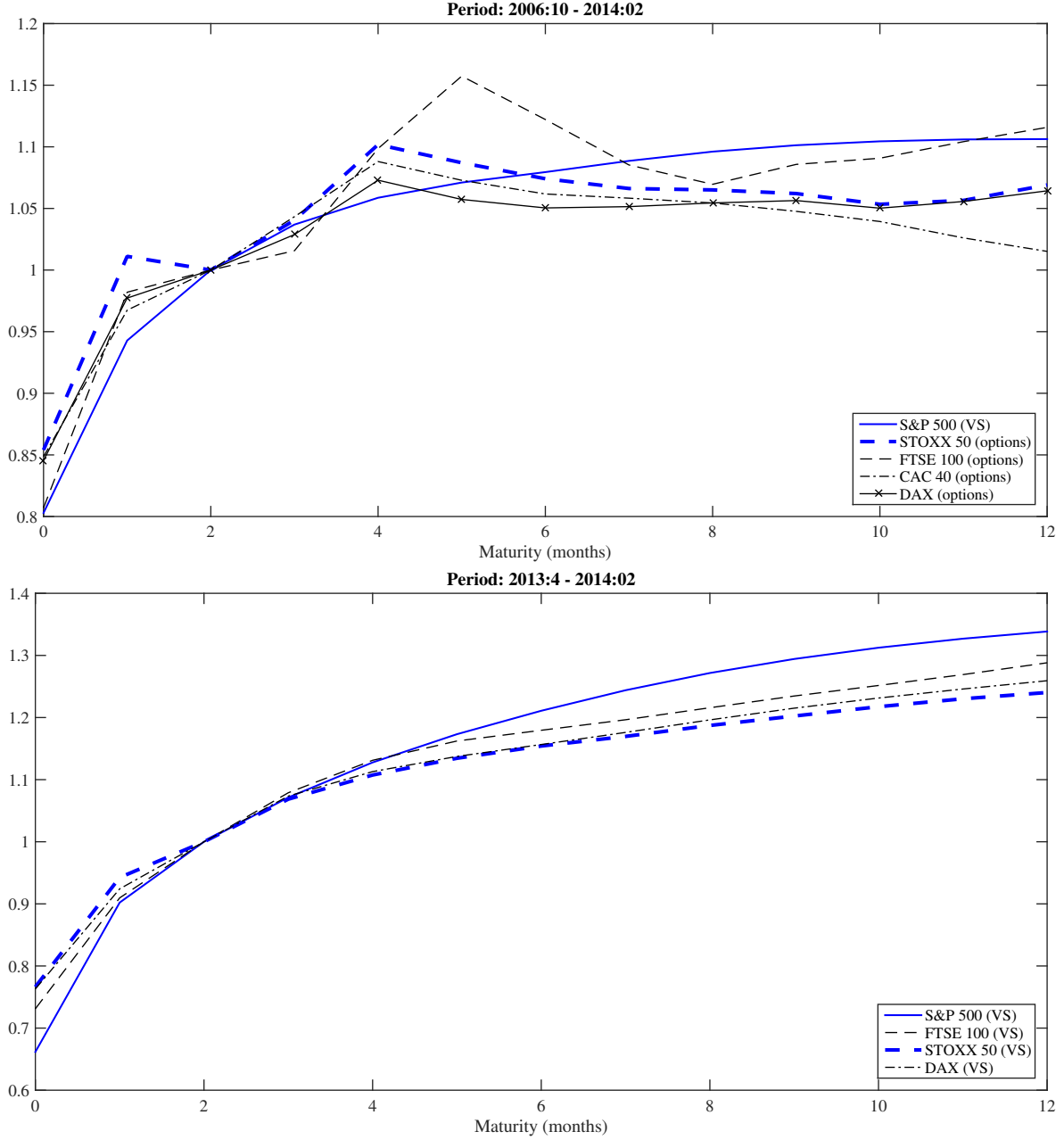
Note: The figure shows the annualized Sharpe ratio for the zero-coupon variance claims. The returns are calculated assuming that the investment in an n-month variance claim is rolled over each month. Dotted lines represent 95% confidence intervals. The sample used is 1996-2013.

Figure 5: Synthetic zero-coupon variance claims: prices and Sharpe ratios



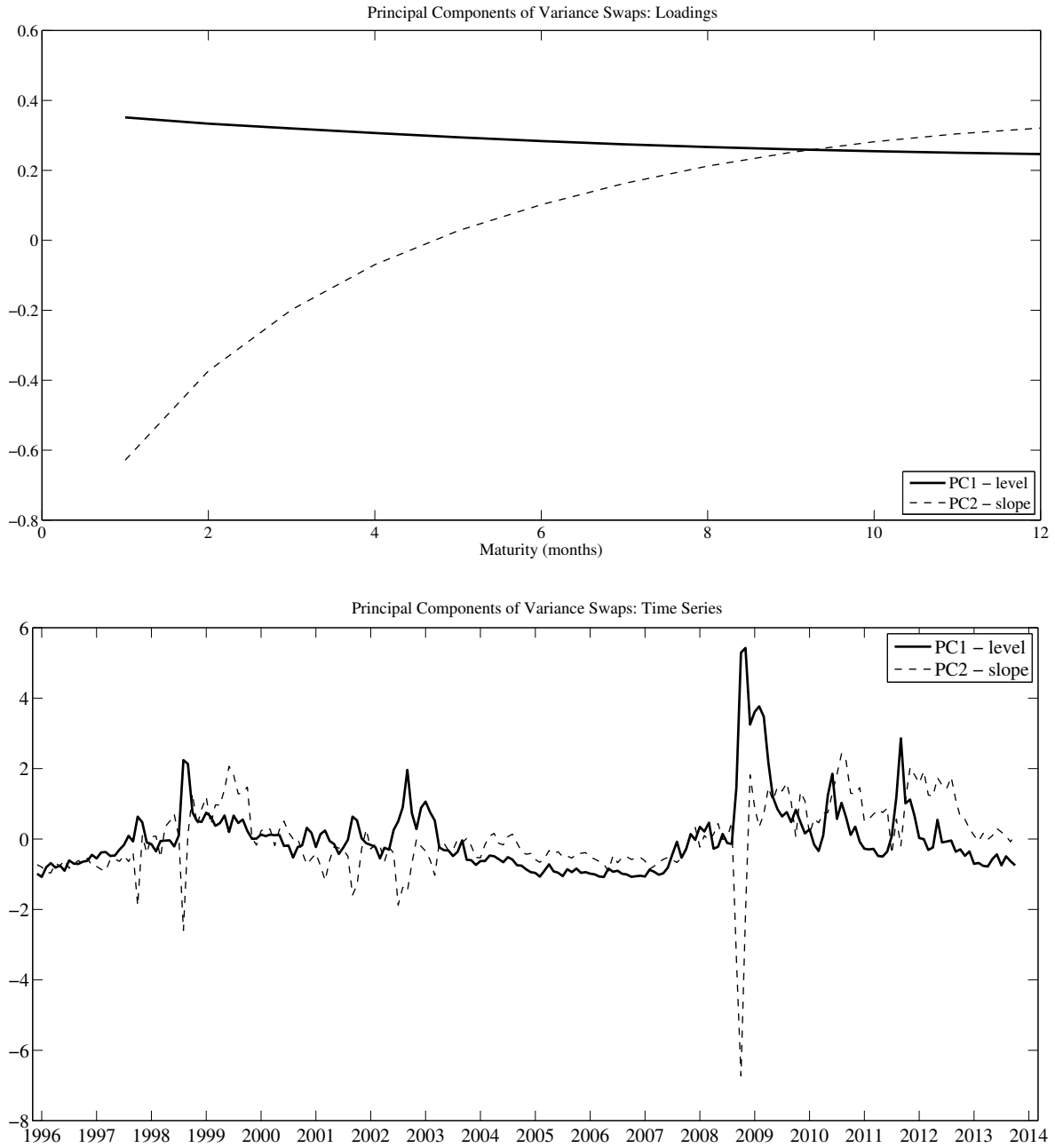
Note: See Figure 2. The solid line in the top panel plots average prices of zero-coupon variance claims calculated using the formula for the VIX index and data on option prices from the CBOE. The dotted line is the set of average prices of zero-coupon variance claims constructed from variance swap prices. The bottom panel plots annualized Sharpe ratios for zero-coupon variance claims returns with prices calculated using the VIX formula and CBOE option data. Dotted lines in the bottom panel represent 95% confidence intervals. The sample covers the period 1996-2013.

Figure 6: Average zero-coupon variance claim prices for international markets



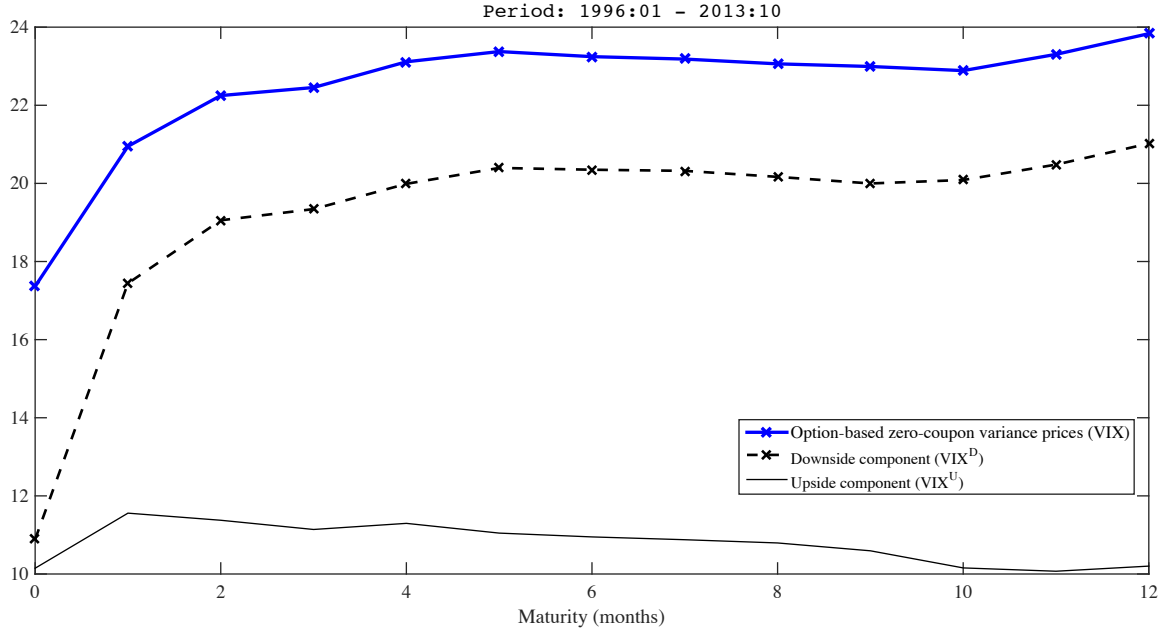
Note: The figure plots the average prices of zero-coupon variance claims as in Figure 2 for different international indices. The series for the S&P 500 (both in the top and bottom panel) is obtained from variance swaps (as in Figure 2). The top panel shows international curves obtained using option prices, using the same methodology used to construct the VIX for the S&P 500 (as in Figure 5). Options data is from OptionMetrics. The series cover FTSE 100, CAC 40, DAX, and STOXX 50, for the period 2006-2014. The bottom panel shows international curves obtained using variance swaps on the FTSE 100, DAX, and STOXX 50, for one year starting in April 2013. All series are rescaled relative to the price of the 2-month zero-coupon variance price (so they all cross 1 at maturity 2 months).

Figure 7: Principal components of variance swap prices



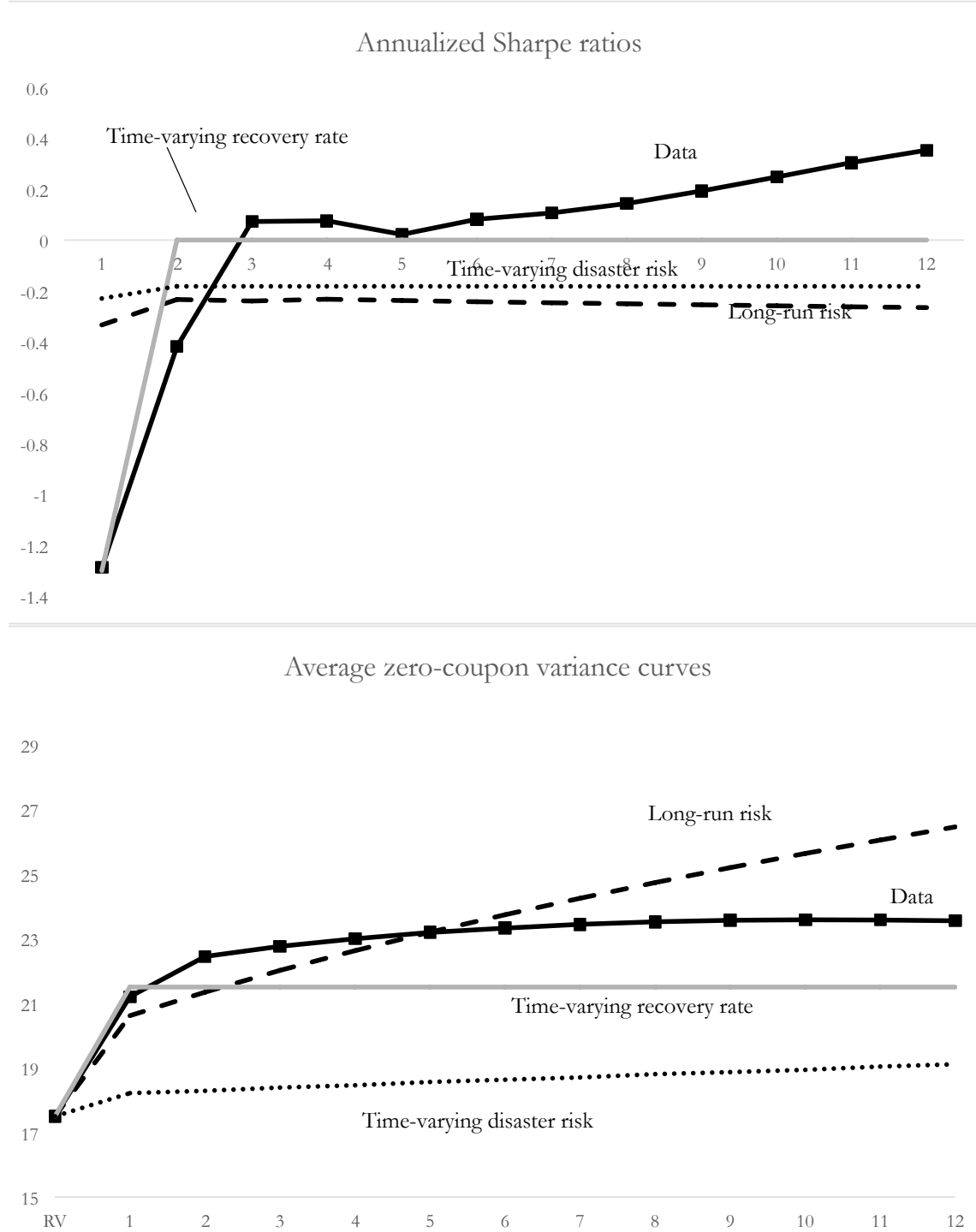
Note: The top panel plots the loadings of the variance swap prices on the level and slope factors (first two principal components). The bottom panel plots the time series of the level and slope factors. Both are normalized to have zero mean and unit standard deviation and are uncorrelated in the sample. The sample covers the period 1996-2013.

Figure 8: Decomposing the upward and downward volatility components



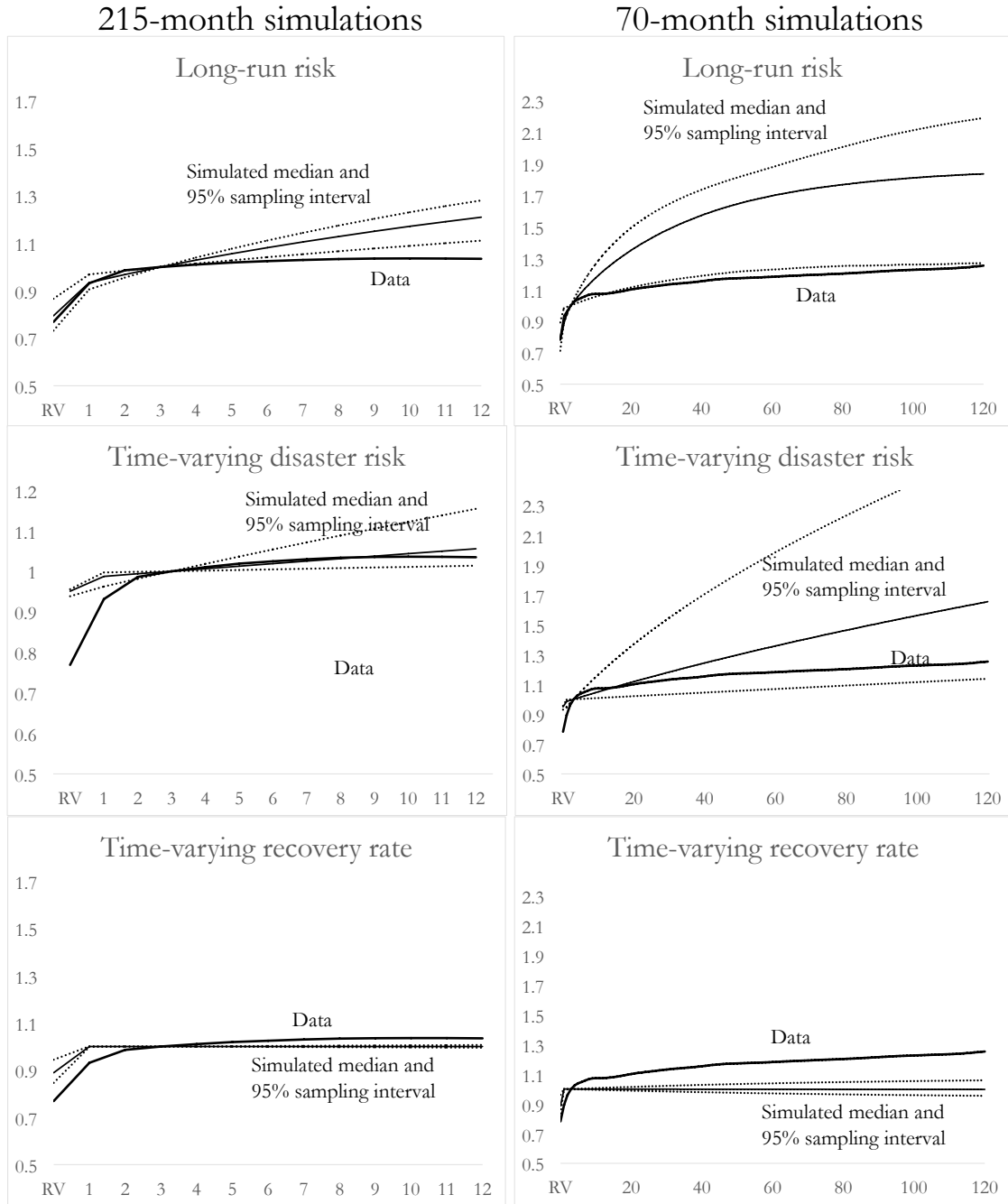
Note: The solid thick line plots average prices of zero-coupon variance claims calculated using the formula for the VIX index. The dashed line plots the zero-coupon prices of the downside component of the VIX, VIX^D . The thin solid line plots the zero-coupon prices of the upside component of the VIX, VIX^U . All series are constructed using option data from CBOE. The sample covers the period 1996-2013.

Figure 9: Sharpe ratios and average term structure in different models



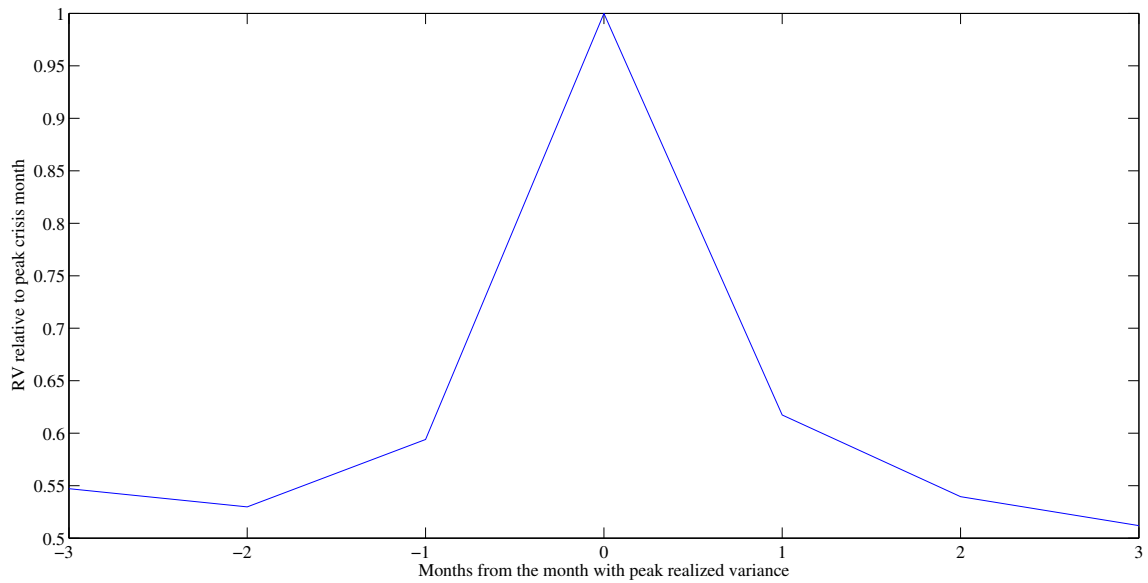
Note: Notes: The top panel gives the population Sharpe ratios from the three models and the sample values from the data. The bottom panel plots population means of the prices of zero-coupon claims. All the curves are normalized to have the same value for the realized variance.

Figure 10: Slope of the term structure in different models



Note: Simulated zero-coupon variance claim prices from structural models. All curves are normalized to equal 1 at three months. The dotted lines are 95-percent sampling intervals from the simulations. The left-hand column simulates the models for 215 months to match our sample of 1- to 12-month variance claims. The right-hand column simulates 70 months to match the sample with up to 10-year claims.

Figure 11: Average behavior of RV during consumption disasters and financial crises



Note: We calculate realized variance in each month of a crisis and scale it by the maximum realized variance in each crisis. The figure plots the average of that scaled series for each country and crisis in terms of months relative to the one with the highest realized variance.

Table 1: Volume of variance swaps across maturities

Maturity (months)	Volume (million vega)	Volume (percentage)
1	402	6%
2	403	6%
3	78	1%
4-6	1037	14%
7-12	1591	22%
13-24	2371	33%
25-60	1315	18%
60+	48	1%
Total	7245	100%

Note: Total volume of variance swap transactions occurred between March 2013 and June 2014 and collected by the DTCC.

Table 2: Characteristics of returns

Maturity (months)	Mean	Std	Min	25th p.	Median	75th p.	Max	Skew	Exc.Kurt.
1	-25.6	69.0	-86.2	-59.2	-40.7	-16.5	687.1	6.1	54.3
2	-5.6	47.8	-59.3	-32.5	-18.2	9.5	376.0	3.9	23.3
3	0.8	34.0	-46.1	-21.4	-4.9	15.0	249.4	2.7	14.1
4	0.7	27.4	-42.2	-17.1	-5.4	11.4	170.4	1.9	7.5
5	0.2	22.5	-37.3	-14.1	-3.6	9.9	126.7	1.6	5.2
6	0.5	19.7	-31.0	-12.3	-3.7	12.9	100.6	1.3	3.3
7	0.6	18.7	-31.4	-12.4	-2.4	11.0	90.7	1.1	2.5
8	0.8	17.4	-29.8	-11.4	-2.9	11.7	81.6	1.0	2.0
9	0.9	16.2	-27.7	-10.3	-1.8	9.6	74.6	0.9	1.7
10	1.2	15.6	-30.0	-9.7	-1.9	10.0	70.8	0.9	1.5
11	1.5	16.0	-32.6	-9.7	-1.9	11.3	69.7	0.9	1.3
12	1.8	17.4	-35.0	-10.2	-2.1	12.2	70.4	1.0	1.4

Note: The table reports descriptive statistics of the monthly returns for zero-coupon variance claims (in percentage points). For each maturity n , returns are computed each month as $R_{t+1}^n = \frac{Z_{t+1}^{n-1} - Z_t^n}{Z_t^n}$. Given the definition that $Z_t^0 = RV_t$, the return on a one-month claim, R_{t+1}^1 is the percentage return on a one-month variance swap.

Table 3: Reduced-form pricing estimates

<i>Panel A: Betas</i>		Shock 1	Shock 2	Pure RV
Maturities (months)	1	0.38	-0.19	0.32
	2	0.40	-0.11	0.05
	3	0.28	0.00	-0.01
	4	0.22	0.04	-0.01
	5	0.18	0.05	0.00
	6	0.15	0.06	0.00
	7	0.13	0.06	0.01
	8	0.12	0.06	0.01
	9	0.11	0.06	0.00
	10	0.11	0.06	0.00
	11	0.11	0.06	0.00
	12	0.11	0.06	0.00
<i>Panel B: Risk Prices</i>				
Risk prices		-0.51	-0.33	-2.72***
Standard error		0.44	0.64	0.47
R^2			0.998	

Note: Results of Fama–MacBeth regressions using the 12 zero-coupon claims as test assets and the three rotated VAR innovations as pricing factors. Shock 1 has effects on all three factors; shock 2 affects only the slope and RV, and pure RV only affects RV on impact. The top section reports betas on the three factors. The bottom section reports estimated risk prices and the Fama–MacBeth standard errors. *** denotes significance at the 1-percent level. Risk prices are annualized by multiplying by $\sqrt{12}$.

Table 4: Pricing results for the CAPM

		(1)	(2)	
		CAPM	CAPM + pure RV	
<i>Panel A: Betas</i>		Rm	Pure RV	Rm
VS, maturities	1	-7.11	0.36	-7.87
	2	-6.93	0.08	-7.11
	3	-5.04	0.01	-5.07
	4	-3.97	0.01	-3.98
	5	-3.14	0.01	-3.16
	6	-2.59	0.01	-2.62
	7	-2.32	0.02	-2.36
	8	-2.13	0.01	-2.17
	9	-2.03	0.01	-2.05
	10	-2.01	0.01	-2.03
	11	-2.06	0.01	-2.08
	12	-2.17	0.01	-2.18
Market		1.00	0	1.00
<i>Panel B: Risk Prices</i>				
Risk prices		0.0105***	-0.72***	0.0046
Standard error		0.0034	0.12	0.0032
R^2		0.377	0.997	

Note: See Table 3. Pricing results for the CAPM and for the CAPM with pure RV. The test assets are the 12 zero-coupon variance claims and the market portfolio. The market portfolio is given 12 times as much weight as the variance claims to ensure that it is priced correctly in the estimation.

Table 5: Forecasting volatility at different horizons: R2

Predictor: with PE_t, DEF_t	monthly RV_{t+n}		Years	yearly RV_{t+n}		Years	yearly Δd_{t+n}	
	RV_t	RV_t		RV_t	RV_t		Δd_t	Δd_t
	$\sqrt{\quad}$			$\sqrt{\quad}$			$\sqrt{\quad}$	
Months								
1	0.39	0.45	1	0.41	0.56	1	0.00	0.09
2	0.21	0.34	2	0.10	0.25	2	0.00	0.02
3	0.18	0.32	3	0.05	0.09	3	0.06	0.07
6	0.15	0.26	5	0.02	0.04	5	0.05	0.07
12	0.10	0.18	10	0.00	0.21	10	0.02	0.03

Note: The first column of the table reports R2 of predictive regressions of monthly volatility n months ahead at the monthly frequency. The second column reports R2 of predictive regressions of yearly volatility n years ahead at the yearly frequency. The second column reports R2 of predictive regressions of yearly log dividend growth n years ahead at the yearly frequency. The left side of each column reports univariate regressions using the lagged value of the target, while the right side of each column adds the market price-earnings ratio and the default spread as predictors. The sample is 1926-2014.

Table 6: Average prices and pricing errors for the no-arbitrage model

Maturity (months)	Sample		Fitted		RMSE
	Mean	Std	Mean	Std	
1	21.24	8.00	21.70	7.86	1.00
2	21.89	7.55	21.88	7.55	0.33
3	22.22	7.25	22.04	7.30	0.40
6	22.75	6.64	22.42	6.74	0.63
12	23.20	6.06	22.99	6.11	0.53
24	23.65	5.58	23.87	5.45	0.61
>24					0.87

Note: Prices are reported in annualized volatility terms. The RMSE is calculated using the deviation of the fitted price from the sample price in annualized volatility terms.

Table 7: Steady-state risk prices

Sources of risk	Estimates	Standard Error
s_t^2 -risk	-0.11	0.31
l_t^2 -risk	-0.18	0.12
RV-risk	-1.70***	0.48

Note: Estimates of steady-state risk prices from the no-arbitrage model. Risk prices are annualized by multiplying by $\sqrt{12}$. *** denotes significance at the 1-percent level.

Table 8: Realized volatility during disasters

Country	Peak Vol. during disaster	Mean Vol. during disaster	Mean Vol. outside disaster	Sample start year	Consumption disasters	Financial crises
US	47.5	25.2	14.9	1926	1933	1929, 1984, 2007
UK	24.6	16.4	15.1	1973		1974, 1984, 1991, 2007
France	72.1	31.4	16.6	1973		2008
Japan	40.9	21.4	15.1	1973		1992
Australia	33.7	13.8	15.1	1973		1989
Germany	83.1	28.1	14.3	1973		2008
Italy	55.1	23.0	19.2	1973		1990, 2008
Sweden	52.3	27.7	19.5	1982		1991, 2008
Switzerland	67.1	27.4	12.1	1973		2008
Belgium	66.1	32.0	12.4	1973		2008
Finland	29.3	18.9	25.0	1988	1993	1991
South Korea	80.0	43.6	24.6	1987	1998	1997
Netherlands	77.7	33.2	14.7	1973		2008
Spain	69.4	30.5	17.1	1987		2008
Denmark	37.2	14.7	14.4	1973		1987
Norway	44.2	20.2	20.7	1980		1988
South Africa	36.9	17.8	18.5	1973		1977, 1989

Note: Characteristics of annualized monthly realized volatility during and outside disasters across countries. Returns data used to construct realized volatility for the US is from CRSP, for all other countries from Datastream. Consumption disaster dates are from Barro (2006). Financial crisis dates are from Schularick and Taylor (2012), Reinhart and Rogoff (2009) and Bordo et al. (2001).

A.1 Synthetic variance swap prices

We construct option-based synthetic variance claims for maturity n , $VIX_{n,t}^2$ using the methods described by the CBOE (2009) in their construction of the VIX index, using data from Optionmetrics covering the period 1996 to 2012.

In particular, we construct $VIX_{n,t}^2$ for maturity n on date t as

$$VIX_{n,t}^2 \equiv \frac{2}{n} \sum_i \frac{\Delta K_i}{K_i^2} \exp(-nR_{n,t}) P(K_i) \quad (\text{A.1})$$

where i indexes options; K_i is the strike price of option i ; $\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$ unless i is the first or last option used, in which case ΔK_i is just the difference in strikes between K_i and its neighbor; $R_{n,t}$ is the n -day zero-coupon yield (from Fama and Bliss (1987)); and $P(K_i)$ is the midpoint of the bid-ask spread for the out-of-the-money option with strike K_i . The summation uses all options available with a maturity of n days. We deviate slightly here from the CBOE, which drops certain options with strikes very far from the current spot. For each t and n , we require the presence of at least 4 out-of-the-money calls and puts. We create $VIX_{n,t}^2$ for all monthly maturities by interpolating between available option maturities, using the same techniques as the CBOE.

Under the assumption that the price of the underlying follows a diffusion (i.e. does not jump), it is the case that

$$VIX_{n,t}^2 = E_t^Q \left[\int_t^{t+n} \sigma_j^2 dj \right] \quad (\text{A.2})$$

$$\approx E_t^Q \left[\sum_{j=1}^n RV_{t+j} \right] \quad (\text{A.3})$$

where σ_t^2 is the instantaneous volatility at time t .¹ The second line simply notes that the integrated volatility is approximately equal to the sum of squared daily returns (where the quality of the approximation improves as the sampling interval becomes shorter). In other words, when the underlying follows a diffusion, $VIX_{n,t}^2$ corresponds to the price of an idealized variance swap where the squared returns are calculated at arbitrarily high frequency.

Given $VIX_{n,t}^2$ prices at the monthly intervals, we also construct zero-coupon variance claims, $VIXZ_{n,t}$ as

$$VIXZ_{n,t} = VIX_{n,t}^2 - VIX_{n-1,t}^2 \quad (\text{A.4})$$

¹See Carr and Wu (2009).

The zero-coupon prices obtained from options are very close to those obtained from variance swaps. Figure 5 in the text compares the two curves for the S&P 500. Figure A.2 compares the curves for the STOXX 50, FTSE 100 and DAX, and shows that the curves are similar for international markets as well (though in that case there is more noise).

A.2 Decomposing the source of variation of variance swap prices

In this section we investigate whether the variation in variance swap prices is primarily driven by changes in expected future volatility or changes in risk premia.²

Using the definition of zero-coupon variance claims, the following identity holds, for each maturity n :

$$Z_t^n - Z_t^0 = [E_t RV_{t+n} - RV_t] + \left[-E_t \sum_{j=0}^{n-1} r_{t+1+j}^{n-j} \right] \quad (\text{A.5})$$

where r_t^n is the one-period return of the n -period zero-coupon claim. An increase in the n -period zero-coupon variance price must predict either an increase in future realized variance or lower future variance risk premia. Following Fama and Bliss (1987) and Campbell and Shiller (1991), we can then decompose the total variance of $Z_t^n - Z_t^0$ into the component that predicts future RV , and the component that captures movements in risk premia. Note that just as in Fama and Bliss (1987) and Campbell and Shiller (1991), we perform this decomposition in changes (predicting the change in volatility rather than the level of volatility), because the previous empirical literature and our term-structure estimates of Section 4.3 highlight the presence of a very persistent factor in the volatility process.

The right side of Table A.1 shows this decomposition for different maturities between 1 month and 1 year. We see that most of the variation in variance swap prices can be attributed to movements in the expectation of future realized variance, not risk premia. In particular, at horizons of 3 to 12 months, essentially all the variation in prices is due to variation in expected volatility rather than variation in risk premia.

At the same time, we know from Table 2 and Figure 1 that prices of both short-term and long-term claims vary substantially: this indicates that the expectation of future realized variance changes dramatically over time. For example, the standard deviation of innovations in the 12-month zero-coupon claim is 17% per year. Given the finding in this section that the variance of Z_t^{12} is driven entirely by changes in volatility expectations, we see that investors'

²A similar exercise was conducted by Mixon (2007), using S&P 500 options to predict future implied volatility.

expectations of future volatility in fact vary substantially over time.

A.3 Estimation of the no-arbitrage model

In this appendix, we provide more details on the estimation of the no-arbitrage model. We also report the estimated parameters and more detailed pricing results.

A.3.1 Estimation Strategy

For estimation purposes, a standard and convenient practice of the term structure literature is to assume that some fixed-weights “portfolios” of VS prices are priced perfectly. These portfolios in turn allow one to invert for the latent states which are needed in the computation of the likelihood scores of the data.

A challenge in implementing this practice in our context is that for parts of our sample the set of available maturities may change from one observation to the next. In the later part of our sample, we use VS prices with maturities up to 14 years, whereas the longest maturity for the earlier sample is only two years.

To tackle this issue, we maintain the assumption from the term structure literature that the current term structure of the VS prices perfectly reveals the current values of states. Nonetheless, we depart from the standard term structure practice by using some time-varying-weights “portfolios” of VS prices in identifying the states at each point in time. The portfolios weights are determined in a way to optimally accommodate different sets of maturities at each point in the sample.

To begin, let D_t denote the vector of observed data obtained by stacking up the vector of VS prices on top of the realized variance RV_t . Because VS prices are affine in states, we can write:

$$D_t = A + BX_t. \quad (\text{A.6})$$

All entries of the last column of B , except for the last row, are zeros because VS prices are only dependent on s_t^2 and l_t^2 . In addition, since the last entry of D_t corresponds to RV_t , the last row of A is 0 and the last row of B is (0,0,1). Keep in mind that the length of D_t can vary from time to time due to different maturity sets for the observed VS data.

We assume that D_t is observed with iid errors:

$$D_t^o = D_t + e_t \quad (\text{A.7})$$

where D_t^o denotes the observed counterpart of D_t . Adopting a standard practice in the

term structure literature, we assume that the observational errors for the VS prices are uncorrelated and have one common variance σ_e^2 . Because observed RV_t is used in practice to determine payoffs to VS contracts, it is natural to assume that RV_t is observed without errors. Combined, if a number of J VS prices are observed at time t , the covariance matrix of e_t , Σ_e , takes the following form:

$$\Sigma_e = \begin{pmatrix} I_J \sigma_e^2 & 0_{J \times 1} \\ 0_{1 \times J} & 0 \end{pmatrix}, \quad (\text{A.8})$$

where I_J denotes the identity matrix of size J .

We now explain how we can recursively identify the states. Assume that we already know X_t . Now imagine projecting X_{t+1} on D_{t+1}^o conditioning on all information up to time t . Our assumption (borrowed from the term structure literature) that the current term structure of the VS prices perfectly reveals the current values of the states implies that the fit of this regression is perfect. In other the words, the predicted component of this regression, upon observing D_{t+1}^o ,

$$E_t(X_{t+1}) + \text{cov}_t(X_{t+1}, D_{t+1}^o) \text{var}_t(D_{t+1}^o)^{-1} (D_{t+1}^o - E_t(D_{t+1}^o)), \quad (\text{A.9})$$

must give us the values for the states at time $t + 1$: X_{t+1} . All the quantities needed to implement (A.9) are known given X_t . Specifically, $E_t(X_{t+1}) = K_0 + K_1 X_t$ and $\text{var}_t(D_{t+1}^o) = V_t(X_{t+1}) + \Sigma_e$ where $V_t(X_{t+1})$ is computed according to each of our three specifications for the covariance. $E_t(D_{t+1}^o)$ is given by $A + B E_t(X_{t+1})$. And $\text{cov}_t(X_{t+1}, D_{t+1}^o)$ is given by $V_t(X_{t+1}) B'$. Clearly, the calculation in (A.9) can be carried out recursively to determine the values of X_t for the entire sample.

Our approach is very similar to a Kalman filtering procedure apart from the simplifying assumption that D_{t+1}^o fully reveals X_{t+1} . That is, $V_t(X_{t+1} | D_{t+1}^o) \equiv 0$. In a term structure context, Joslin, Le, and Singleton (2013) show that this assumption allows for convenient estimation, yet delivers typically highly accurate estimates.

We can view (A.9) as some “portfolios” of the observed data D_{t+1}^o with the weights given by: $\text{cov}_t(X_{t+1}, D_{t+1}^o) \text{var}_t(D_{t+1}^o)^{-1}$. As a comparison, whereas the term structure literature typically choose, prior to estimation, a fixed weight matrix corresponding to the lower principal components of the observed data, we do not have to specify *ex ante* any loading matrix. Our approach determines a loading matrix that optimally extracts information from the observed data and, furthermore, can accommodate data with varying lengths.

As a byproduct of the above calculations, we have available the conditional means and

variances of the observed data: $E_t(D_{t+1}^o)$ and $V_t(D_{t+1}^o)$. These quantities allow us to compute the log QML likelihood score of the observed data as (ignoring constants):

$$\mathcal{L} = \sum_t -\frac{1}{2} \|var_t(D_{t+1}^o)^{-1/2}(D_{t+1}^o - E_t(D_{t+1}^o))\|_2^2 - \frac{1}{2} \log |var_t(D_{t+1}^o)|. \quad (\text{A.10})$$

Estimates of parameters are obtained by maximizing \mathcal{L} . Once the estimates are obtained, we convert the above QML problem into a GMM setup and compute robust standard errors using a Newey West matrix of covariance.

A.3.2 Alternative variance specifications

Constant variance structure

In the first alternative specification, we let $V_t(X_{t+1})$ be a constant matrix Σ_0 . Since both $E_t(X_{t+1})$ and $E_t^Q(X_{t+1})$ are linear in X_t , Λ_t is also linear in X_t . We refer to this as the CV (for constant variance) specification.

Flexible structure

It is important to note that the specifications of Λ_t in (12) and $V_t(X_{t+1})$ in (13) introduce very tight restrictions on the difference: $E_t(X_{t+1}) - E_t^Q(X_{t+1})$. Simple algebra shows that the first entry of the product $V_t(X_{t+1})^{1/2}\Lambda_t$ is simply a scaled version of s_t^2 . This means that the dependence of $E_t(X_{t+1})$ and $E_t^Q(X_{t+1})$ on l_t^2 and RV_t and a constant must be exactly canceled out across measures. Similar arguments lead to the following restrictions on the condition mean equation:

$$E_t \begin{pmatrix} s_{t+1}^2 \\ l_{t+1}^2 \\ RV_{t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ v_l^Q \\ 0 \end{pmatrix}}_{K_0} + \underbrace{\begin{pmatrix} \rho_s & 1 - \rho_s^Q & 0 \\ 0 & \rho_l & 0 \\ \rho_{s,RV} & 0 & 0 \end{pmatrix}}_{K_1} \begin{pmatrix} s_t^2 \\ l_t^2 \\ RV_t \end{pmatrix}. \quad (\text{A.11})$$

So in the CIR specification, the conditional mean equation only requires three extra degrees of freedom in: ρ_s , ρ_l , and $\rho_{s,RV}$. The remaining entries to K_0 and K_1 are tied to their risk-neutral counter parts. By contrast, all entries of K_0 and K_1 are free parameters in the CV specification. Whereas CV offers more flexibility in matching the time series dynamics of X_t , the parsimony of CIR, if well specified, can potentially lead to stronger identification. However, this parsimony of the CIR specification, if mis-specified, can be

restrictive. For example, in the CIR specification, RV_t is not allowed to play any role in forecasting X .

In the flexible specification, we would like to combine the advantages of the CV specification in matching the time series dynamics of X and the advantages of the CIR specification in modeling time-varying volatilities.

For time-varying volatility, we adopt the following parsimonious structure:

$$V_t(X_{t+1}) = \Sigma_1 s_t^2, \quad (\text{A.12})$$

where Σ_1 is a fully flexible positive definite matrix. This choice allows for non-zero covariances among all elements of X .

The market prices of risks are given by:

$$\Lambda_t = V_t(X_{t+1})^{-1/2} (E_t(X_{t+1}) - E_t^Q(X_{t+1})), \quad (\text{A.13})$$

where the superscript $^{1/2}$ indicates a lower triangular Cholesky decomposition. Importantly, we do not require the market prices of risks to be linear, or of any particular form. This is in stark contrast to the CIR specification which requires the market prices of risks to be scaled versions of states. As a result of this relaxation, no restrictions on the conditional means dynamics are necessary. Aside from the non-negativity constraints, the parameters K_0, K_1 that govern $E_t(X_{t+1}) = K_0 + K_1 X_t$ are completely free, just as in the CV specification. In particular, RV_t is allowed to forecast X and thus can be important in determining risk premiums. We label this specification as the FLEX specification (for its flexible structure).

A.3.3 Additional estimation results

Table A.2 reports the risk neutral parameters of our no arbitrage models: ρ_s^Q , ρ_l^Q , and v_l^Q . As expected, these parameters are very strongly identified thanks to the rich cross-section of VS prices used in the estimation. Recall that our risk-neutral construction is identical for all three of our model specifications. As a result, estimates of risk neutral parameters are nearly invariant across different model specifications.

The effect of including the crisis is that the estimates for ρ_s^Q and v_l^Q are higher, whereas the estimate of ρ_l^Q is the same. A higher v_l^Q is necessary to fit a higher average VS curve. A higher estimate for ρ_s^Q implies that risk-neutral investors perceive the short-run factor s_t^2 as more (risk-neutrally) persistent. In other words, movements of the one-month VS price (s_t^2) will affect prices of VS contracts of much longer maturities.

We report the time series parameters – K_0, K_1 , and other parameters that govern the

conditional variance $V_t(X_{t+1})$ – for the CV, CIR, and FLEX specifications in Tables A.3, A.4, and A.5, respectively.

The values of annualized steady-state risk prices implied by each model specification are reported in A.6. Regardless of how the quantity of risk ($V_t(X_{t+1})$) is modeled, RV-risk is always very significantly negatively priced. The point estimates and standard errors are similar across the various specifications for the variance process (and hence the physical dynamics), emphasizing the results of our findings. The table also shows that the results are similar regardless of whether the financial crisis is included in the estimation sample. Since the financial crisis was a period when the returns on variance swaps were extraordinarily high, excluding it from the data causes to estimate risk prices that are even more negative than in the full sample.

A.4 Calibration and simulation of the models

This section gives the details of the three models analyzed in the main text.

A.4.1 Long-run Risk (Drechsler and Yaron)

Our calibration is identical to that of Drechsler and Yaron (DY; 2011), so we refer the reader to the paper for a full description of the model. We have confirmed that our simulation matches the moments reported in DY (tables 6, 7, and 8, in particular). Here we report an extra set of results that modifies DY’s calibration to match our empirical estimates. Specifically, we change the standard deviation of the innovations to $\bar{\sigma}^2$ from 0.10 to 0.05. We also multiply the volatility of both the diffusive and jump innovations in σ^2 by 1.35. Those two changes make the processes for σ^2 and $\bar{\sigma}^2$ match the one we estimate in the no-arbitrage model.

The following table reports results from three simulations of the model – the original DY calibration, a calibration with risk aversion raised to 16, and the calibration with the change in the volatility parameters:

Moment	Original	High RRA	Empirical volatility
E[var premium]	9.06	11.30	11.79
stdev[var premium]	12.28	18.78	22.42

The variance premium here measures the difference between the 1-month variance swap price and realized variance (it is the ”level difference” from DY). As we would expect, the

variance premium rises when we raise risk aversion. The “original” column corresponds to the calibration used in the main text. Shifting to the calibration that matches the estimated volatility process in the final column we can see reduces the mean and standard deviation of the variance premium, but only to a level similar to DY’s original calibration. The reason for this is that Epstein–Zin preferences emphasize long-run shocks, and the estimated volatility process puts more weight on the short-run than the long-run shock to volatility compared to DY’s original calibration.

Figures A.5 and A.6 replicate figures 9 and 10 but comparing now DY’s original calibration, our version with higher risk aversion, and the version with the calibration of volatility to the empirical estimates. In all three cases, the Sharpe ratios at all maturities are substantially negative, roughly the value of the Sharpe ratio of the aggregate stock market. The term structures are also far too steeply upward sloping.

A.4.2 Time-varying disasters (Wachter)

The key equations driving the economy are

$$\Delta c_t = \mu_{\Delta c} + \sigma_{\Delta c} \varepsilon_{\Delta c,t} + J_{\Delta c,t} \quad (\text{A.14})$$

$$F_t = (1 - \rho_F) \mu_F + \rho_F F_{t-1} + \sigma_F \sqrt{F_{t-1}} \varepsilon_{F,t} \quad (\text{A.15})$$

$$\Delta d_t = \lambda \Delta c_t \quad (\text{A.16})$$

where $\varepsilon_{\Delta c,t}$ and $\varepsilon_{F,t}$ are standard normal innovations. This is a discrete-time version of Wachter’s setup, and it converges to her model as the length of a time period approaches zero. The model is calibrated at the monthly frequency. Conditional on a disaster occurring, $J_{\Delta c,t} \sim N(-0.3, 0.15^2)$. The number of disasters that occurs in each period is a Poisson variable with intensity F_t . The other parameters are calibrated as:

Parameter	Value	Parameter	Value
$\mu_{\Delta c}$	0.02/12	$\sigma_{\Delta c}$	0.02/sqrt(12)
σ_F	0.0075	λ	2.8
ρ_F	$0.87^{1/12}$	μ_F	0.017/12

In the analytic solution to the model, the price/dividend ratio for a levered consumption claim takes the form $pd_t = z_0 + z_1 F_t$. The Campbell–Shiller approximation to the return

(which becomes arbitrarily accurate as the length of a time interval shrinks) is

$$r_{t+1} = \theta p d_{t+1} + \Delta d_{t+1} - p d_t \quad (\text{A.17})$$

$$= \theta p d_{t+1} + \lambda \Delta c_{t+1} - p d_t \quad (\text{A.18})$$

It is straightforward to show analytically (the derivation is available on request) that

$$p d_t = z_0 + z_1 F_t \quad (\text{A.19})$$

In the absence of a disaster, we treat the shocks to consumption and the disaster probability as though they come from a diffusion, so that the realized variance is $\theta^2 z_0^2 \sigma_F^2 F_{t-1} + \lambda \sigma_{\Delta c}^2$. When a disaster occurs, we assume that the largest daily decline in the value of the stock market is 5 percent. So, for example, a 30-percent decline would be spread over 6 days. The results are largely unaffected by the particular value assumed. The realized variance when a disaster occurs is then $\theta^2 z_0^2 \sigma_F^2 F_{t-1} + \lambda \sigma_{\Delta c}^2 - (0.05) J_{\Delta c, t}$ (assuming $J_{\Delta c, t} \leq 0$).

The model is solved analytically using methods similar to those in DY. Specifically, household utility, v_t , is

$$v_t = (1 - \beta) c_t + \frac{\beta}{1 - \alpha} \log E_t [\exp((1 - \alpha) v_{t+1})] \quad (\text{A.20})$$

α is set to 3.6 and $\beta = 0.98^{1/12}$. The recursion can be solved because the cumulant-generating function for a poisson mixture of normals (the distribution of $J_{\Delta c}$) is known analytically (again, see DY).

The pricing kernel is

$$M_{t+1} = \beta \exp(-\Delta c_{t+1}) \frac{\exp((1 - \alpha) v_{t+1})}{E_t [\exp((1 - \alpha) v_{t+1})]} \quad (\text{A.21})$$

Asset prices, including those for claims on realized variance in the future, then follow immediately from the solution of the lifetime utility function and cumulant-generating function for $J_{\Delta c}$. The full derivation and replication code is available upon request.

We attempted to keep the calibration as close as possible to Wachter's. The two differences are that we increase risk aversion somewhat in order to try to generate a larger variance risk premium and that we use a normal distribution for the disasters rather than the empirical distribution used by Wachter (to allow us to obtain analytic results). It is important to note that risk aversion cannot be increased past 3.7 because the model no longer has a solution.

As with the DY model, we checked that the moments generated by our solution to the model match those reported by Wachter. We confirm that our results are highly similar, in particular for her tables 2 and 3, though not identical since we use slightly different risk aversion and a different disaster distribution.

A.4.3 Disasters with time-varying recovery (Gabaix)

$$\Delta c_t = \mu_{\Delta c} + \varepsilon_{\Delta c,t} + J_{\Delta c,t} \quad (\text{A.22})$$

$$L_t = (1 - \rho_L) \bar{L} + \rho_L L_{t-1} + \varepsilon_{L,t} \quad (\text{A.23})$$

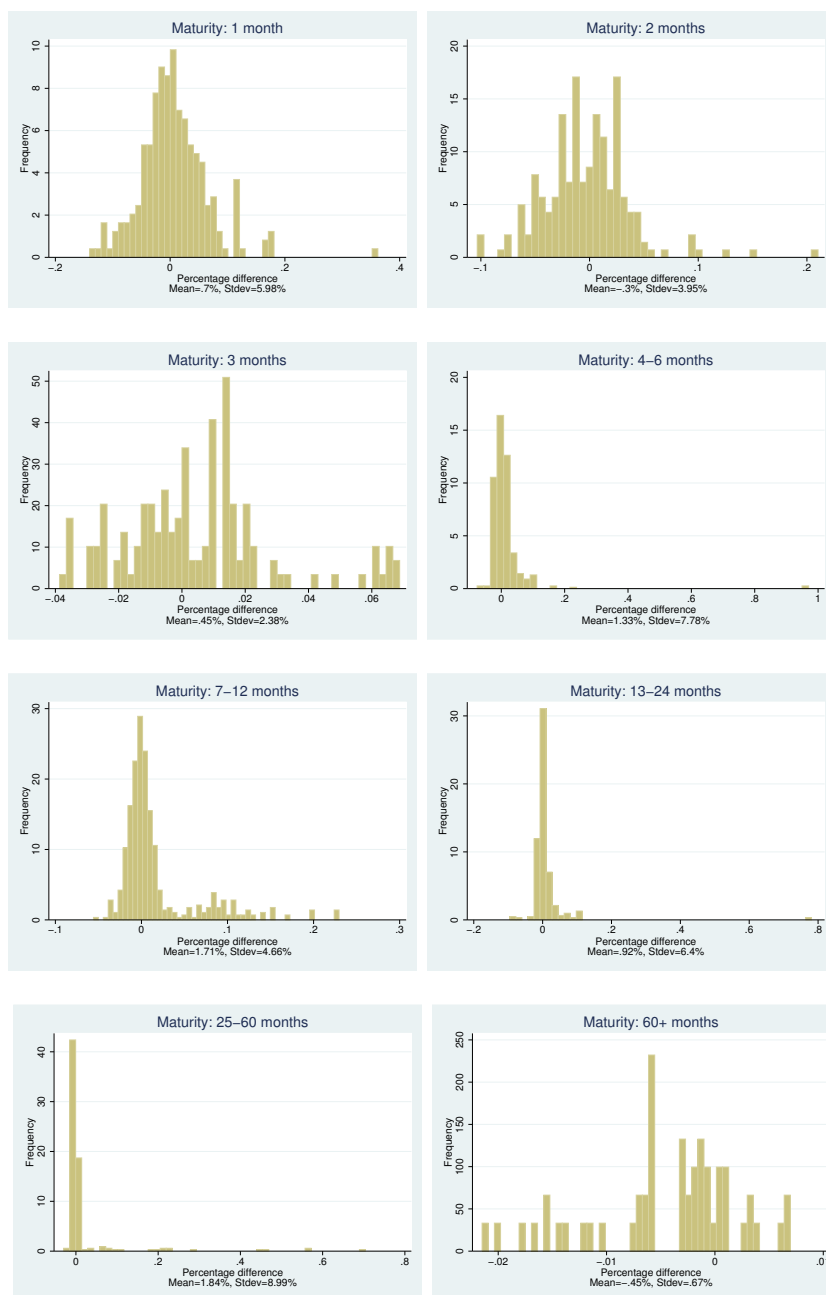
$$\Delta d_t = \lambda \varepsilon_{\Delta c,t} - L \times 1 \{J_{\Delta c,t} \neq 0\} \quad (\text{A.24})$$

The model is calibrated at the monthly frequency. Conditional on a disaster occurring, $J_{\Delta c,t} \sim N(-0.3, 0.15^2)$. The probability of a disaster in any period is 0.01/12. The other parameters are calibrated as:

Parameter	Value
$\mu_{\Delta c}$	0.1/12
$\text{stdev}(\varepsilon_{\Delta c})$	$0.02/\sqrt{12}$
ρ_L	$0.87^{1/12}$
\bar{L}	0.5
$\text{stdev}(\varepsilon_L)$	0.035
λ	5

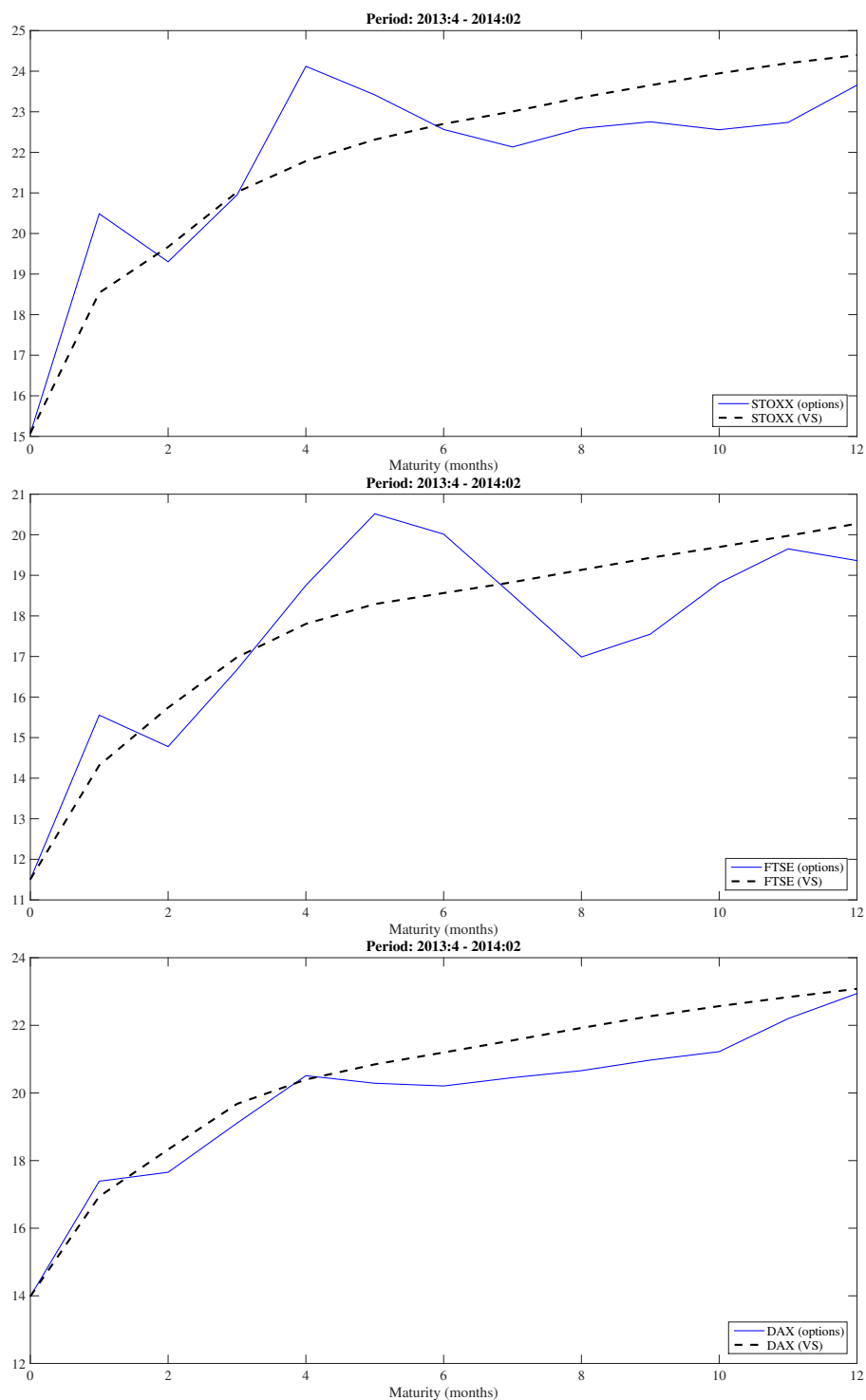
Agents have power utility with a coefficient of relative risk aversion of 7 and a time discount factor of $0.96^{1/12}$.

Figure A.1: Quotes vs. transaction prices



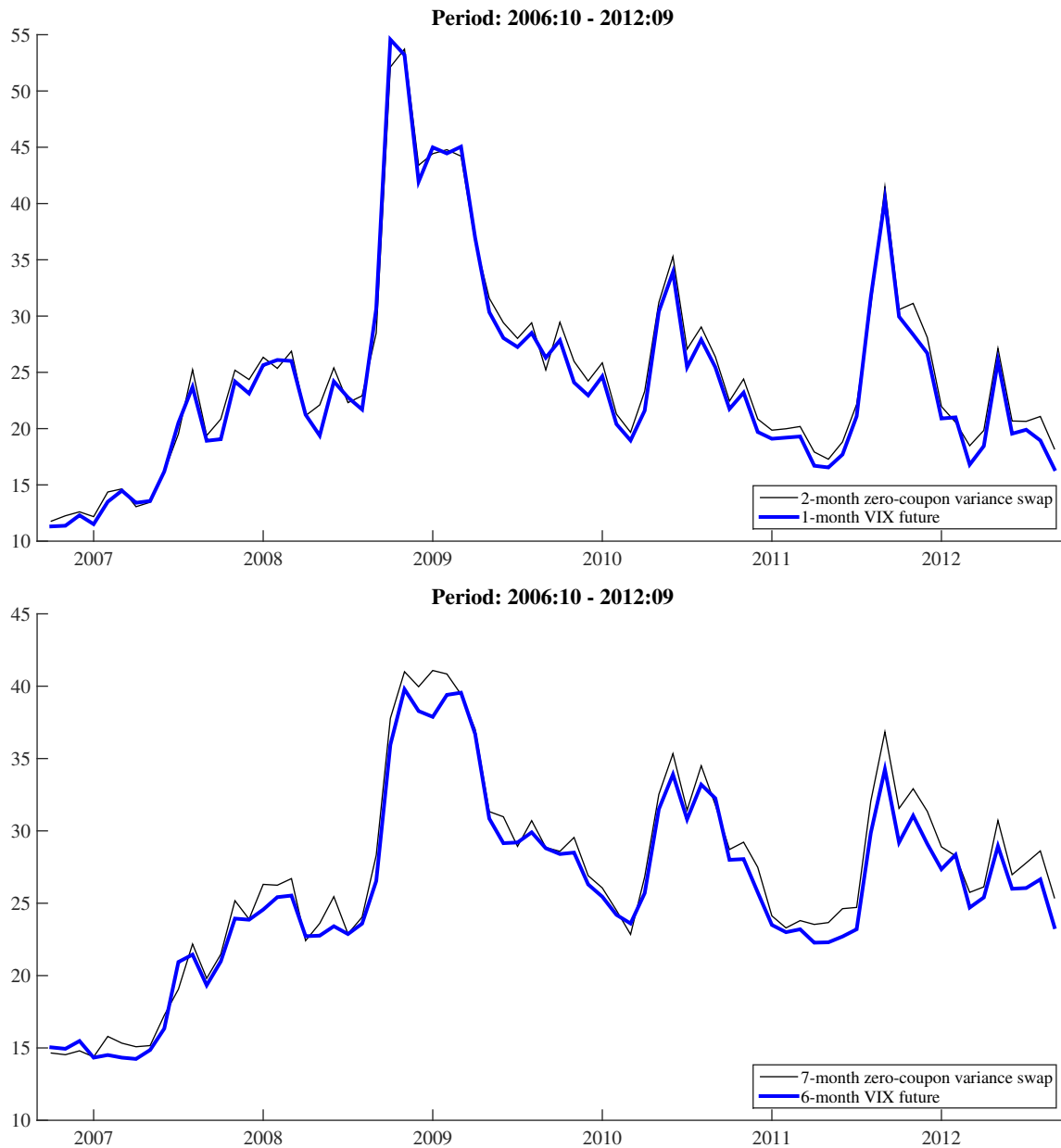
Note: The figure shows the distribution of percentage difference between variance swap price quotes and actual transaction prices, computed as $(\text{transaction price} - \text{quote}) / \text{quote}$. The quotes are our main sample, while transaction prices are obtained from the DTCC and begin in 2013. Each panel shows the histogram for a different bucket of maturity of the variance swap contracts.

Figure A.2: International zero-coupons variance claim prices from options and variance swaps



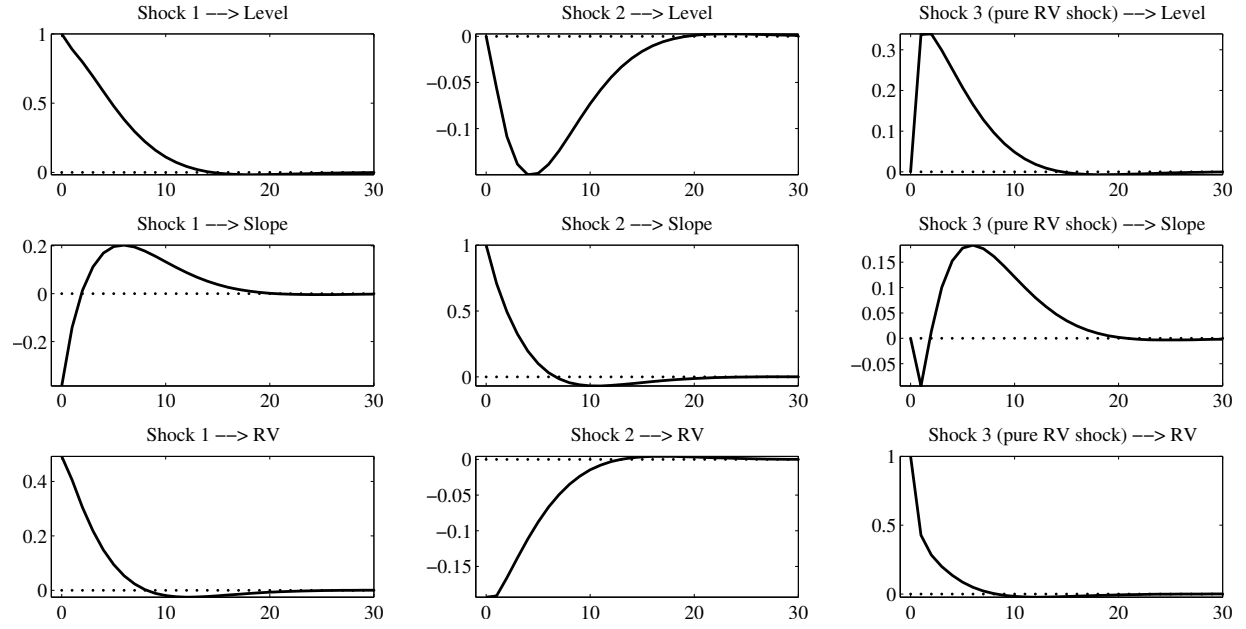
Note: In each panel (corresponding to STOXX 50, FTSE 100 and DAX), the solid line plots average prices of zero-coupon variance claims calculated using the formula for the VIX index and data on international option prices from Optionmetrics. The dotted line plots the average prices of the same claims computed from international variance swaps. The sample covers the period 2013:4-2014:2.

Figure A.3: VIX futures vs. zero-coupon variance swap prices



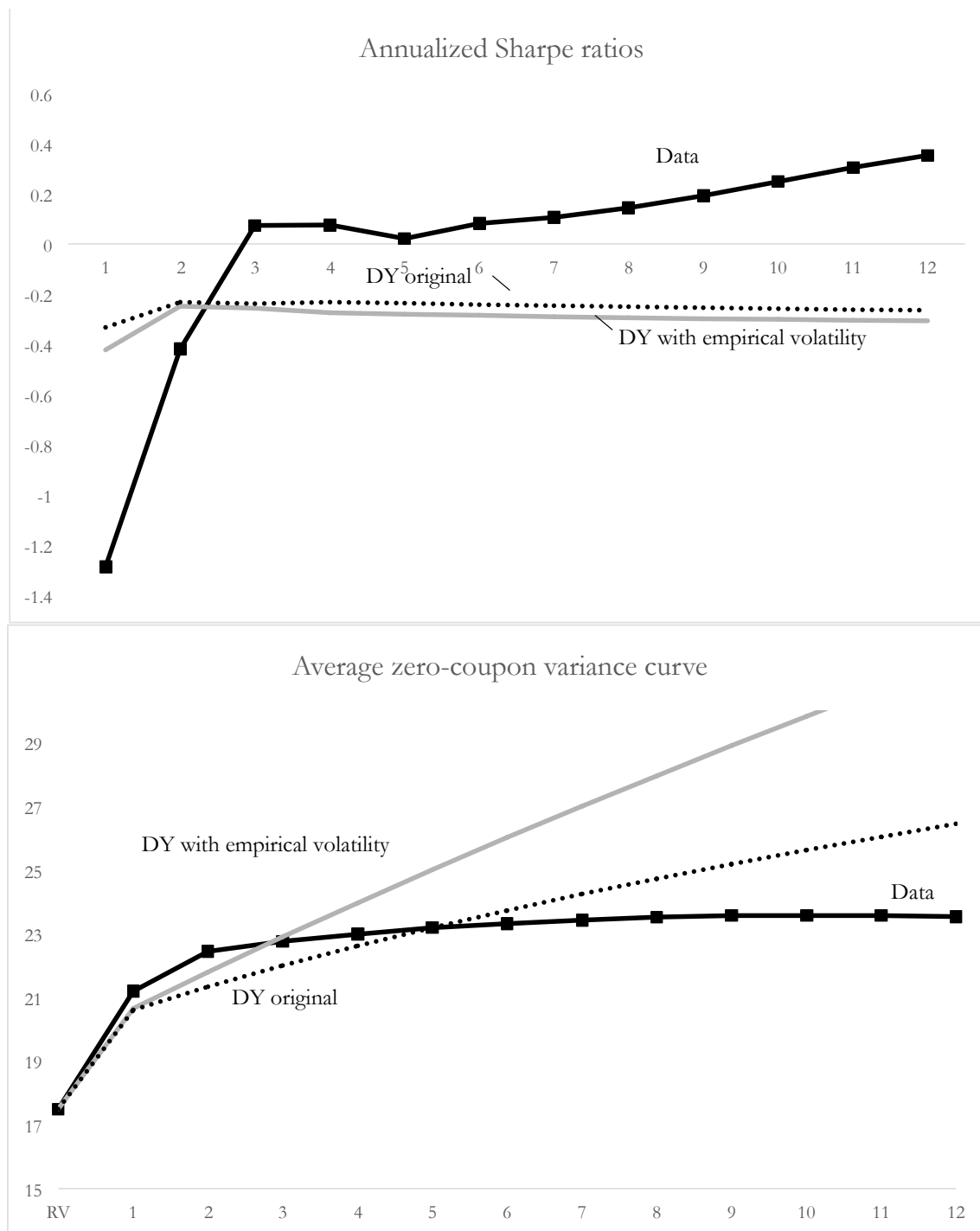
Note: the top panel plots the time series of the 2-month zero-coupon variance swap and the 1-month VIX future price from the CME, in annualized volatility terms. The bottom panel plots the time series of the 7-month zero-coupon variance swap and the 6-month VIX future price from the CME, in annualized volatility terms. The sample covers the period 2006:10-2012:9.

Figure A.4: Impulse response functions of level, slope and RV



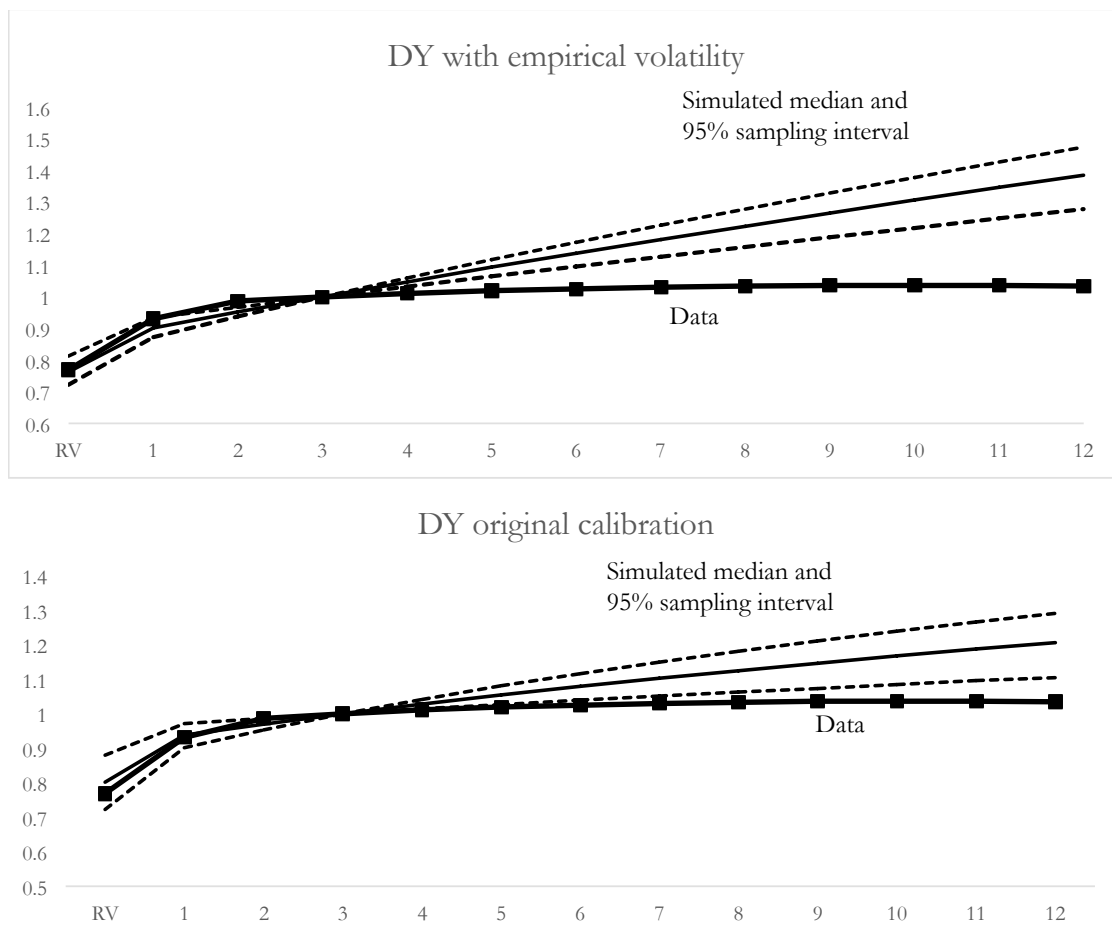
Note: Each panel plots the response of one of the variables in the VAR (level, slope, and RV) to one of the three orthogonalized shocks. The shocks are orthogonalized with a Cholesky factorization with the ordering level-slope-RV.

Figure A.5: Sharpe ratios and average term structure in the long-run risk model



Note: The top panel gives the population Sharpe ratios from the two models and the sample values from the data. The bottom panel plots population means of the prices of zero-coupon claims. All the curves are normalized to have the same value for the realized variance.

Figure A.6: Slope of the term structure in the long-run risk model



Note: Simulated zero-coupon variance claim prices from structural models. All curves are normalized to equal 1 at three months. The dotted lines are 95-percent sampling intervals from the simulations.

Table A.1: Predictive regressions

Horizon (months)	Predictor: Slope and Curvature	Predictor: $Z_t^n - Z_t^0$			
	Dep var: $RV_{t+n} - RV_t$ R^2	Dep var: $RV_{t+n} - RV_t$ R^2	coeff. R^2	Dep var: $-\sum_{j=0}^{n-1} r_{t+1+j}^{n-j}$ R^2	coeff. R^2
1	0.08	0.17	0.63*** (0.22)	0.07	0.37* (0.22)
2	0.22	0.18	0.71*** (0.17)	0.04	0.29* (0.17)
3	0.31	0.24	0.83*** (0.14)	0.01	0.17 (0.14)
6	0.36	0.34	0.95*** (0.09)	0.001	0.05 (0.09)
12	0.38	0.41	1.10*** (0.09)	0.001	-0.10 (0.09)

Note: Results of regressions forecasting changes in realized variance. The left side reports the R^2 of a regression of changes in realized volatility between month t and month $t + n$ on the level and the slope at time t . The right side reports the coefficients of univariate regressions of changes in realized volatility (left column) and returns to volatility claims from t to $t + n$ (right column) on the difference between the zero-coupon prices of maturity n (Z_t^n) and realized volatility (Z_t^0) at time t . * indicates significance at the 10-percent level, ** the 5-percent level, and *** the 1-percent level.

Table A.2: Risk neutral parameters

Specifications		1996:2007			1996:2013		
		ρ_s^Q	ρ_l^Q	v_l^Q	ρ_s^Q	ρ_l^Q	v_l^Q
CV	Est.	0.66***	0.99***	0.65***	0.83***	0.99***	1.01**
	Stderr.	0.06	0.00	0.14	0.07	0.01	0.48
CIR	Est.	0.68***	0.99***	0.65***	0.84***	0.99***	0.99**
	Stderr.	0.05	0.00	0.14	0.06	0.01	0.44
FLEX	Est.	0.66***	0.99***	0.64***	0.83***	0.99***	1.01**
	Stderr.	0.06	0.00	0.14	0.07	0.01	0.47

Note: The table reports the risk-neutral estimated dynamics of the term structure model, for the three specifications CV, CIR and FLEX, and separately for the full sample (1996-2013) and the pre-crisis sample (1996-2007).

Table A.3: Time series parameter estimates for the CV specification

Sample		K_0	K_1			Σ^*		
1996:2007	Est.	6.49***	0.65***	0.14**	0.03	15.42***	0	0
		3.36***	0.08	0.83***	0.02	3.87***	5.25***	0
		1.98	0.56***	0	0.12	10.57***	-2.62	14.08***
	Std. err.	2.42	0.08	0.07	0.12	3.23	-	-
		1.12	0.10	0.08	0.07	0.73	0.73	-
		2.11	0.07	-	0.10	2.36	2.63	2.07
1996:2013	Est.	8.99***	0.56***	0.08	0.19**	20.59***	0	0
		2.43***	0	0.91***	0.06***	3.26	7.04***	0
		5.67*	0.20	0	0.57***	30.00**	-2.14	25.49***
	Std. err.	2.48	0.13	0.05	0.09	5.07	-	-
		0.90	-	0.03	0.01	2.09	0.97	-
		2.92	0.13	-	0.15	13.38	2.94	4.23

Note: The table reports the time-series parameter estimates for the CV specification. Σ^* is the lower triangular Cholesky decomposition of Σ_0 . For admissibility, K_0 and K_1 are constrained to be non-negative. Those entries for which the non-negativity constraint is binding are set to zero and thus standard errors are not provided. The table reports the estimates separately for the full sample (1996-2013) and the pre-crisis sample (1996-2007).

Table A.4: Time series parameter estimates for the CIR specification

Sample		K_0	K_1			Σ^*		
1996:2007	Est.	0	0.66***	0.32***	0	5.16***	0	3.53***
		0.65***	0	0.99***	0	0	0.98***	0
		0	0.69***	0	0	3.53***	0	6.74***
	Std. err.	-	0.07	0.05	-	1.82	-	1.18
		0.14	-	0.00	-	-	0.23	-
		-	0.01	-	-	1.18	-	1.69
1996:2013	Est.	0	0.82***	0.16**	0	7.17**	0	9.42*
		0.99**	0	0.98***	0	0	1.42**	0
		0	0.75***	0	0	9.42*	0	23.56*
	Std. err.	-	0.04	0.06	-	2.84	-	5.42
		0.44	-	0.00	-	-	0.58	-
		-	0.09	-	-	5.42	-	12.88

Note: The table reports the time-series parameter estimates for the CIR specification. The diagonal elements of Σ^* correspond to the variance parameters σ_s^2 , σ_l^2 , and σ_{RV}^2 . The (3,1) entry of Σ^* correspond to the covariance parameter $\sigma_{s,RV}$. The table reports the estimates separately for the full sample (1996-2013) and the pre-crisis sample (1996-2007).

Table A.5: Time series parameter estimates for the FLEX specification

Sample		K_0	K_1			Σ^*		
1996:2007	Est.	2.31	0.74***	0.19**	0.01	2.26***	0	0
		2.18***	0.09*	0.87***	0	0.58***	0.78***	0
		0.11	0.60***	0	0.13	1.56***	-0.24	2.04***
	Std. err.	1.49	0.08	0.08	0.08	0.42	-	-
		0.81	0.05	0.06	-	0.08	0.08	-
		0.75	0.08	-	0.09	0.29	0.31	0.27
1996:2013	Est.	4.54***	0.68***	0.08**	0.19**	2.59***	0	0
		1.37**	0	0.93***	0.07***	0.44*	1.09***	0
		0.67***	0.36***	0	0.51**	3.32***	-0.34	3.20***
	Std. err.	1.30	0.08	0.04	0.08	0.48	-	-
		0.55	-	0.02	0.02	0.23	0.12	-
		0.05	0.11	-	0.20	1.24	0.43	0.48

Note: The table reports the time-series parameter estimates for the FLEX specification. Σ^* is the lower triangular Cholesky decomposition of Σ_0 . For admissibility, K_0 and K_1 are constrained to be non-negative. Those entries for which the non-negativity constraint is binding are set to zero and thus standard errors are not provided. The table reports the estimates separately for the full sample (1996-2013) and the pre-crisis sample (1996-2007).

Table A.6: Annualized steady state risk prices, all specifications

Specifications	Sources of risks	1996:2007		1996:2013	
		Estimates	Standard Error	Estimates	Standard Error
CV	s_t^2 -risk	-0.23	0.22	-0.08	0.14
	l_t^2 -risk	0.05	0.18	-0.21	0.14
	RV-risk	-2.78***	0.65	-1.44***	0.43
CIR	s_t^2 -risk	-0.18	0.36	-0.11	0.31
	l_t^2 -risk	0.04	0.14	-0.18	0.12
	RV-risk	-3.92***	0.91	-1.70***	0.48
FLEX	s_t^2 -risk	-0.23	0.30	-0.14	0.16
	l_t^2 -risk	0.06	0.22	-0.17	0.19
	RV-risk	-3.17***	0.78	-1.69**	0.73

Note: Estimates of steady-state risk prices from the no-arbitrage model. Risk prices are annualized by multiplying by $\sqrt{12}$. *** denotes significance at the 1-percent level. Results are reported for the full sample (1996-2013), and restricted to 1996-2007.