Journal of Financial Economics 123 (2017) 225–250

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

The price of variance risk☆

Ian Dew-Becker[a,e], Stefano Giglio[b,f,*], Anh Le[c], Marius Rodriguez[d]

[a] Kellogg School of Management, Northwestern University, USA
[b] Booth School of Business, University of Chicago, USA
[c] Smeal College of Business, Pennsylvania State University, USA
[d] Board of Governors of the Federal Reserve System, USA
[e] National Bureau of Economic Research
[f] National Bureau of Economic Research and the Centre for Economic Policy Research

ARTICLE INFO

Article history:
Received 7 August 2015
Revised 9 March 2016
Accepted 11 April 2016
Available online 23 November 2016

JEL classifications:
G12
G13

Keywords:
Variance risk premium
Recursive preferences
Disasters

ABSTRACT

Between 1996 and 2014, it was costless on average to hedge news about future variance at horizons ranging from 1 quarter to 14 years. Only unexpected, transitory realized variance was significantly priced. These results present a challenge to many structural models of the variance risk premium, such as the intertemporal CAPM and recent models with Epstein–Zin preferences and long-run risks. The results are also difficult to reconcile with macro models in which volatility affects investment decisions. At the same time, the data allows us to distinguish between different disaster models; a model in which the stock market has a time-varying exposure to disasters and investors have power utility fits the major features of the variance term structure.

1. Introduction

The recent explosion of research on the effects of volatility in macroeconomics and finance shows that economists care about uncertainty shocks. It appears that 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 not been required to pay for insurance against volatility news. Many economic theories—both in macroeconomics and in finance—have the opposite prediction. The recent consumption-based asset pricing literature is heavily influenced by Epstein and Zin (1991) preferences, which in standard calibrations, with a preference for early resolution of uncertainty, imply that investors have a strong desire to hedge news about future uncertainty, and hence should be willing to pay large premia for insurance against volatility shocks. Furthermore, in recent macroeconomic models, shocks to uncertainty about the future can induce large fluctuations in the economy.[1] But if increases

☆ 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, Steve Davis, Itamar Drechsler, Greg Duffee, Darrell Duffie, Cam Harvey, Jing-zhi Huang, Dana Kiku, Loriano Mancini, Ian Martin, Tim McQuade, Ivan Shaliastovich, Fabio Trojani, Harald Uhlig, Joel Vanden, Andrea Vedolin, Jessica Wachter, Liuren Wu, 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, MIT Sloan, Ohio State, Southern Denmark University, Copenhagen Business School, Penn State Smeal, SFS Cavalcade, SITE, ITAM, WFA, FIRS, SoFie, CEPR Gerzensee and AFA. We are particularly grateful to Gerald Lim and Markit Group for providing variance swap data.

* Corresponding author.
E-mail address: Stefano.giglio@chicagobooth.edu (S. Giglio).

http://dx.doi.org/10.1016/j.jfineco.2016.04.003
0304-405X/

[1] See, e.g., Bloom (2009), Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014), Christiano, Motto, and Rostagno (2014), Fernandez-Villaverde, Guerron, Rubio-Ramirez, and Uribe (2011), and Gourio (2012) Gourio (2013).

in economic uncertainty can drive the economy into a recession, we would expect that investors would want to hedge those shocks.[2] The fact that shocks to expected volatility have not earned a risk premium thus presents a challenge to a wide range of recent research.

As a concrete example, consider the legislative battles over the borrowing limit of the US in the summers of 2010 and 2011. Those periods were associated with increases in both financial measures of uncertainty, e.g., the Chicago Board Options Exchange's Volatility Index (VIX), and also the measure of policy uncertainty from Baker, Bloom, and Davis (2014). Between July and October, 2011, the 1-month variance swap rate—a measure of investor expectations for Standard & Poor's (S&P) 500 volatility over the next month—rose every month, from 16.26 to 42.32% (annualized, computed at the beginning of the month). But those shocks also had small effects on *realized* volatility in financial markets; for example, realized volatility actually decreased between August and September of 2011. The debt ceiling debate caused uncertainty about the future to be high during the whole period, but did not correspond to high contemporaneous volatility during the same period. It is precisely this imperfect correlation between realized volatility and expectations of future volatility that allows us to disentangle the pricing of their shocks. In this paper, we directly measure how much people pay to hedge shocks to expectations of future volatility. We find that news shocks have been unpriced: any investor could have bought 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 one month and ten years. The data cover 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. Importantly, though, we show that our results hold in a range of other markets, including index options, which are both more liquid and traded on exchanges.

The analysis of the pricing of variance swaps yields two simple but important results. First, news about future volatility is unpriced in our sample—exposure to volatility news did not earn a risk premium. Second, exposure to *realized variance* is strongly priced in our data, with an annualized Sharpe ratio of −1.3—four times larger than the Sharpe ratio on equities. 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) and Segal, Shaliastovich, and Yaron (2015). We conclude that over our sample, investors paid 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. 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.[3]

Consumption-based models with Epstein and Zin 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.[4] 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.[5]

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. While that 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: it predicts that shocks to future expected volatility should be strongly priced, counter to what we observe empirically.

We obtain similar results in a version of Wachter's (2013) model of time-varying disaster risk with Epstein–Zin preferences. The combination of fluctuations in the 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. Du's (2011) model of disaster risk and habit formation also fails to match the data.[6]

[2] See Berger, Dew-Becker, and Giglio (2016) for an analysis of the effects of volatility shocks on the real economy, finding that news about future volatility is not contractionary.

[3] Recently, Campbell, Giglio, Polk, and Turley (2013) and Bansal, Kiku, Shaliastovich, and Yaron (2013) estimate an ICAPM model with stochastic 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, Giglio, Polk, and Turley (2013) 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.

[4] This 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.

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

[6] Similar 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 a version of Gabaix's (2012) model of rare disasters, which builds on the work of Rietz (1988), Barro (2006), and many others, can match the stylized fact that Sharpe ratios on variance claims are large at the very short end of the term structure and fall to zero rapidly with maturity. Intuitively, when investors have power utility, they invest myopically in that they do not price shocks that only affect expectations about the future. Disaster risk and high risk aversion help the model generate the large risk premia that we observe on short-term claims. That said, our calibration of Gabaix's (2012) model is not a complete quantitative description of financial markets, as it does not perfectly match all the patterns in the data; we simply view it as giving a set of sufficient conditions that allow a model to match the economically relevant features of the variance swap term structure.[7]

An 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 to one asset class.

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, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014), Christiano, Motto, and Rostagno (2014), Fernandez-Villaverde, Guerron, Rubio-Ramirez, and Uribe (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 that macro models should focus on shocks to *realized* rather than *expected* volatility.

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), and Bansal, Kiku, Shaliastovich, and Yaron (2013).[8]

Finally, there is a large 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 swaps with maturities from two to 24 months, including Egloff, Leippold, and Wu (2010) and Ait-Sahalia, Karaman, and Mancini (2014), who study no-arbitrage term structure models. The pricing models we estimate are less technically sophisticated than that of Ait-Sahalia, Karaman, and Mancini (2014), but we complement and advance their work in two 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.

The remainder of the paper is organized as follows. Section 2 describes the novel data sets we obtain for variance swap prices. Section 3 reports unconditional moments 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 a standard asset pricing framework. In Section 5, we discuss what structural models can fit the data. We calibrate four 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

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 asset occurring until maturity. In this paper, the underlying is the S&P 500 index unless otherwise specified. The payment at expiration

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).

[7] In a paper that is contemporaneous to this one, Eraker and Wu (2016) propose a simple consumption-based model with volatility shocks that matches some of the features of the volatility market. We leave a comparison of that model to our data to future work.

[8] Andries, Eisenbach, and Schmalz (2015) analyze a model consumption-based model that matches broad features of the variance market, while van Binsbergen and Koijen (2015) discuss other recent work on related topics.

[9] A number of papers study the pricing of volatility in options markets, e.g., Jackwerth and Rubinstein (1996), Coval and Shumway (2001), Bakshi and Kapadia (2003), Broadie, Chernov, and Johannes (2009), Christoffersen, Jacobs, Ornthanalai, 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. Johnson (2016) studies the predictability of returns on option portfolios. 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.

of a variance swap initiated at time τ and with maturity m is

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

where periods here denote 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 data sets of quoted prices for S&P 500 variance swaps. Both data sets were obtained from industry sources. Data set 1 is obtained from a hedge fund. Data set 2 is obtained from Markit Totem, and reports means of quotes (11, on average) obtained from dealers in the variance swap market. Data set 1 contains monthly variance swap prices for contracts expiring in one, two, three, six, 12, and 24 months, and includes data from December, 1995, to October, 2013. Data set 2 contains data on variance swaps with expirations that are fixed in calendar time, instead of fixed maturities. Common maturities are clustered around one, three, and six months, and one, two, three, five, ten, and 14 years. Data set 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 data set to obtain the prices of variance swaps with standardized maturities covering all months between one month and 12 months for Data set 1 and between one month and 120 months for Data set 2 (though in estimating the no-arbitrage model in the Appendix we use the original price data without interpolation).

Both variance swap data sets are novel to the literature. Variance swap data with maturities up to 24 months as in Data set 1 have been used before (Ait-Sahalia, Karaman, and Mancini, 2014; Amengual and Xiu, 2014; Egloff, Leippold, and Wu, 2010; Filipovic, Gourier, and Mancini, 2013), but the shortest maturity previous studies observe is two months. The one-month variance swap is special in this market because it is the exclusive claim to next month's realized variance, which is by far the most strongly priced risk in this market.

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, 2013), having claims with very long maturities is important for effectively testing the central predictions of Epstein–Zin preferences.

The variance swap market is sizeable: the notional value of outstanding variance swaps at the end of 2013 was $4 billion of notional vega, which means that an increase in annualized realized volatility of 1% induces total payments of $4 billion.[10] This market is thus small relative to the aggregate stock market, but it is non-trivial economically.

We obtained information about average bid-ask spreads by maturity from a large market participant. Typical bid-ask spreads are reported to be 1–2% for maturities up to one year, 2–3% between one and two years, and 3–4% for maturities up to ten years. The bid-ask spreads are thus non-trivial, but also not so large as to prohibit trading. Moreover, they are small relative 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).

Table 1 shows the total volume in notional vega terms for all transactions between March 2013 and June 2014, obtained from the Depository Trust & Clearing Corporation (DTCC; see Appendix Section A.1). In little more than a year, the variance swap market saw $7.2 billion of notional vega traded. Only 11% of the volume was traded in short maturity contracts (one to three months); the bulk of the transactions occurred for maturities between six months and five years, and the median maturity was 12 months.

A recent paper by Mixon and Onur (2015) studies the liquidity of the variance swap market and the VIX futures market using proprietary data from the Commodity Futures Trading Commission (CFTC). They document that trading in these (essentially interchangeable) products occurs mostly in the VIX futures markets for maturities below one year, and in the variance swap market for higher maturities. We show below that these two markets are tightly integrated—prices for maturities present in both markets are virtually identical; we will show below that our results will hold in each of these two markets taken separately.

Since these data sets are new to the literature, we devote Appendix Section A.1 to a battery of tests to ensure the quality of the data. In particular, we verify that: neither data set contains stale prices (at the monthly frequency, which is the one we observe); the two data sets contain essentially the same information when they overlap (correlation above 0.997); quotes from the two data sets correspond closely to the prices for actual trades we observe since 2013; and prices in the variance swap

Table 1
Volume of variance swaps across maturities.

Maturity (months)	Volume (million vega)	Volume (%)
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

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

[10] See 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.

market are extremely highly correlated with other related markets (synthetic variance swaps constructed from options as described below, and VIX futures).

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

It is well known that variance swaps can be synthesized as a portfolio of all available out-of-the money options (Carr and Wu, 2009; Jiang and Tian, 2005). The synthetic variance swap portfolio is used to construct the CBOE's VIX index. Options thus give an alternative source of information about the pricing of variance risk.

The VIX is usually reported for a 30-day maturity, but the formulas are valid at any horizon (see Appendix Section A.2 for details on construction). 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.[11] Since options are exchange-traded, they involve minimal counterparty risk, so we can use them to check 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.

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 outstanding notional vega during 2015 averaging $332 million.[12] Bid-ask spreads are smaller than what we observe in the variance swap market, at roughly 0.1%, 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. As of 2014, these funds had an aggregate notional exposure to the VIX of roughly $5 billion, making them comparable in size to the variance swap market.

2.4. Liquidity across markets

An obvious concern with any study of derivatives prices, especially derivatives traded over the counter, is that the prices are affected by liquidity. Liquidity effects can cause the prices to be stale and can generate risk premia. As noted above, the variance swap market is large in terms of notional values, but the bid/ask spreads are also larger than in other markets.

To check the accuracy of the variance swap prices, we compare their behavior to that of VIX futures for the dates and maturities where they overlap. We show in appendix section A.1.4.2 that the VIX futures and variance swap prices are extremely highly correlated: in levels, the correlation is on average 0.993, while for monthly changes it is 0.98. So even though variance swaps are less liquid than other assets, their prices are nearly identical to those for VIX futures.

Mixon and Onur (2015) show that for maturities of six months or less, there is far more volume in the VIX futures market than in variance swaps. And Fig. A.2 in the Appendix shows that the 75th percentile of maturity for CBOE-traded S&P 500 options weighted by open interest is six months. On the other hand, Mixon and Onur (2015) find that the mean maturity (weighted by open interest) for S&P 500 variance swaps is four years. So VIX futures and options appear to be used to trade volatility at relatively short maturities, while variance swaps are used at longer maturities. In the analysis below, we therefore emphasize that our primary results are clear even in the options and futures markets, and even at maturities shorter than six months.

So it is ultimately important that we examine a wide range of data sources, since they have trade concentrated at different maturities. We will show below, though, that our results and conclusions are the same across the various variance claims. And as noted above, where the different data sources overlap, the prices we measure are nearly identical.

Finally, we show in Appendix Section A.4 that our main results on the difference between short-term and long-term variance claim returns are robust to accounting for the bid-ask spread and can be obtained in trading strategies that minimize the amount of trading needed (in particular, using holding periods of six months). This robustness test helps mitigate the concern that liquidity issues across the term structures may be the driver of our empirical results.

3. The term structure of variance claims

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

[11] Even in 1990, Vijh (1990) noted that the CBOE was highly liquid and displayed little evidence of price impact for large trades. George and Longstaff (1993) study options on the S&P 100 in 1989, and document that at-the-money calls and puts had bid-ask spreads of 2–3% at all maturities they analyze. Volume and liquidity in the options market have grown over time, but these earlier studies are important because they document that even earlier than our main sample begins, options markets were developed and liquid.

[12] According to the CBOE futures exchange market statistics. See: http://cfe.cboe.com/Data/HistoricalData.aspx.

log returns—during month t. Subscripts from here forward index 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

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

where RV_{t+m} is the sum of daily squared returns in month $t+m$ and E_t^Q denotes the mathematical expectation under the risk-neutral measure conditional on information available at the end of month t. So VS_t^n is the expected sum of daily squared returns between months $t+1$ and $t+n$.

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 forward claims on realized variance. We define an n-month variance forward as an asset with a payoff equal to realized variance in month $t+n$. The absence of arbitrage implies

$$F_t^n \equiv E_t^Q[RV_{t+n}] \tag{3}$$

$$= VS_t^n - VS_t^{n-1}. \tag{4}$$

F_t^n represents the market's risk-neutral expectation of realized variance n months in the future (at the end of month t). We use the natural convention that

$$F_t^0 = RV_t \tag{5}$$

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

Fig. 1 plots the time series of variance forward prices for maturities between one month and ten years. The figure shows all series in annualized percentage volatility units, rather than variance units: $100 \times \sqrt{12 \times F_t^n}$ instead of F_t^n. The top panel plots variance forward prices for maturities below one year and maturities longer than one year are in the bottom panel.

The term structure of variance forward 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.

Fig. 2 reports the average term structure of variance forwards 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, 1996–2014, is in the bottom panel. The first point on the graph (maturity 0) corresponds to the average realized volatility, whereas all points from 1 on are forward claims of different maturity.

Fig. 2 shows that the term structure of variance forwards has been upward sloping on average, but also concave, flattening out very quickly as the maturity increases. For example, the top panel shows that the three-month forward was 30% more expensive than realized volatility on average, but from the three-month forward to the 120-month forward, the price rose only by another 20%. The bottom panel shows that the 12-month forward was only 5% more expensive than the three-month forward over the longer sample.

The average variance term structures in Fig. 2 provide the first indication that the compensation for bearing risk associated with news about future volatility has been small in this market. Since the return on holding a variance forward for a single month is $\frac{F_{t+1}^{n-1} - F_t^n}{F_t^n}$, it is clearly closely related to the slope of the forward variance term structure.

So if the average term structure is upward sloping between maturities $n-1$ and n, forward claims of maturity n will tend to 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.

To see whether the shape of the curve is well measured statistically, Fig. 3 plots the average slope ($F_t^n - F_t^{n-1}$) and curvature ($(F^{n+1} - F_t^n) - (F_t^n - F_t^{n-1})$) at each maturity along with confidence intervals calculated using the Newey and West (1987) method with six lags. The top panel of Fig. 3 shows that the slopes are well identified—the slope falls from 3.7 annualized percentage points at the one-month maturity to an insignificant 0.3 percentage points at three months. The slope is also uniformly declining with maturity. The bottom panel of Fig. 3 plots the average curvature of the term structure. The term structure is concave on average at every maturity (statistically significantly at seven of 11 maturities). Fig. 3 thus confirms that the basic intuition from Fig. 2, that the term structure was steep at short maturities and nearly flat on average at longer maturities, is well measured statistically.

The top and bottom panel of Fig. 2 differ in both the time period and the maturities displayed. To check the robustness of our conclusions about the average shape of the term structure of variance forwards over the period used to construct it, Fig. 4 examines the average term structure in different subsamples, focusing on the maturities up to 12 months to make the comparison easier. The figure shows that after 2008 the curve became slightly steeper for maturities above one month. However, even after 2008 the curve is still much flatter at maturities above three months than it is at the very short end, displaying the same pattern as in the full sample. The results are similar when we eliminate the financial crisis. Finally, we also use data from the CME to construct the VIX for maturities up to six months going back to 1983. Before 1996, the overall level of the curve was lower, but the shape of the curve was the same.

Of course, the economic significance of the "flatness" of the curve must be understood within the context of a model. In Section 5, we show formally that the curve of forward variance swaps is too flat in both subperiods relative to the implications of workhorse asset pricing models.

3.2. Returns on variance forwards

The return on an n-month variance forward corresponds to a strategy that buys the n-month forward and

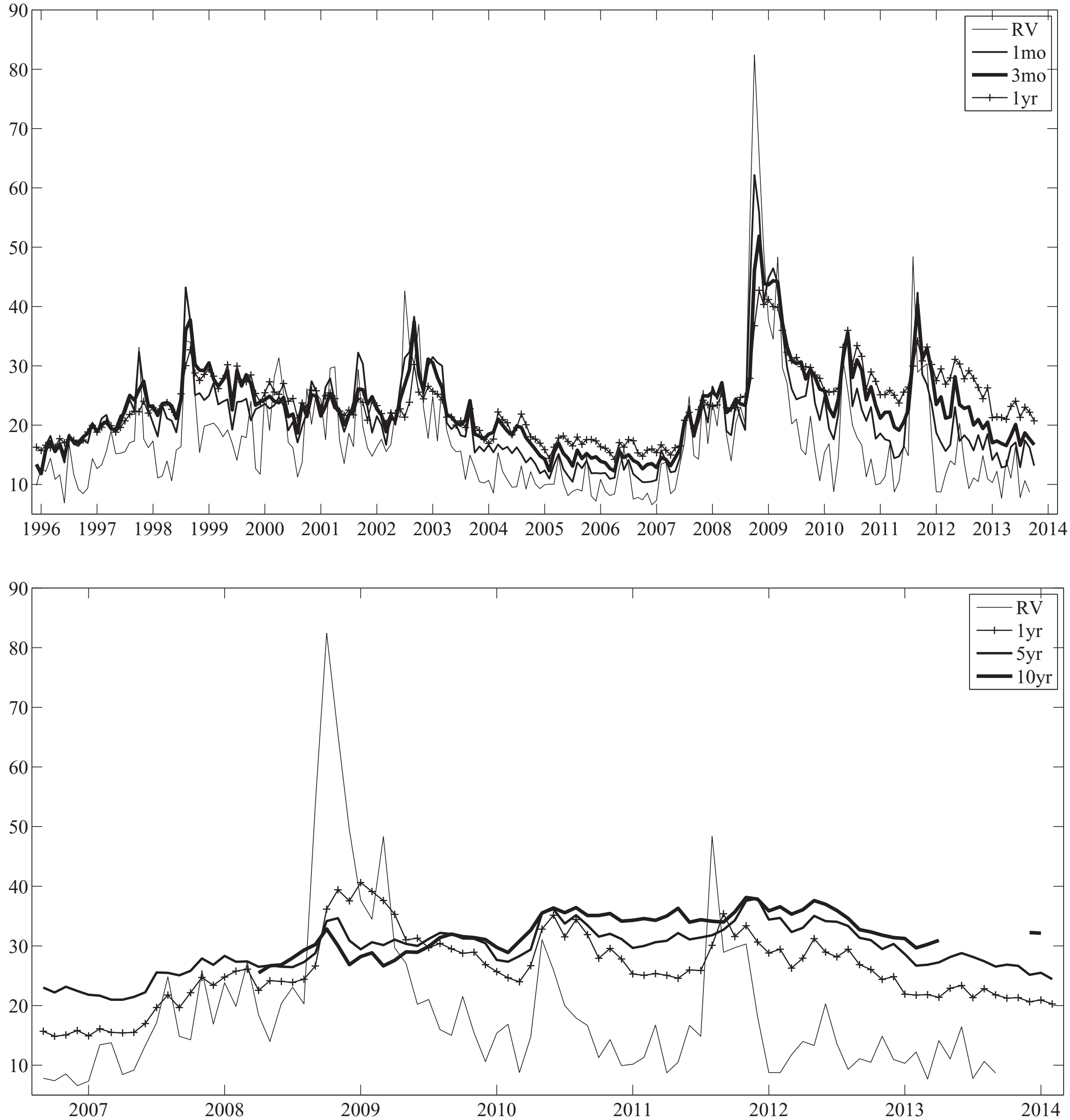


Fig. 1. Time series of forward variance claim prices. The figure shows the time series of forward variance claim prices of different maturities. For readability, each line plots the prices in annualized volatility terms, $100 \times \sqrt{12 \times F_t^n}$, for a different n. The top panel plots forward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

sells it one month later as an $(n-1)$-month forward, reinvesting then again in a new n-month forward. We define the excess return of an n-period variance forward following Gorton, Hayashi, and Rouwenhorst (2013).[13]

$$R_{t+1}^n = \frac{F_{t+1}^{n-1} - F_t^n}{F_t^n}. \tag{6}$$

Given the definition that $F_t^0 = RV_t$, the return on a one-month forward, R_{t+1}^1 is simply the percentage return on a

[13] Note that $F_{t+1}^{n-1} - F_t^n$ 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, Jonathan E. Ingersoll, and Ross (1985). We have reproduced our analysis using the unscaled excess return $F_{t+1}^{n-1} - F_t^n$ as well and confirmed that the results hold in that case; we report the Sharpe ratios in Appendix Fig. A.3.

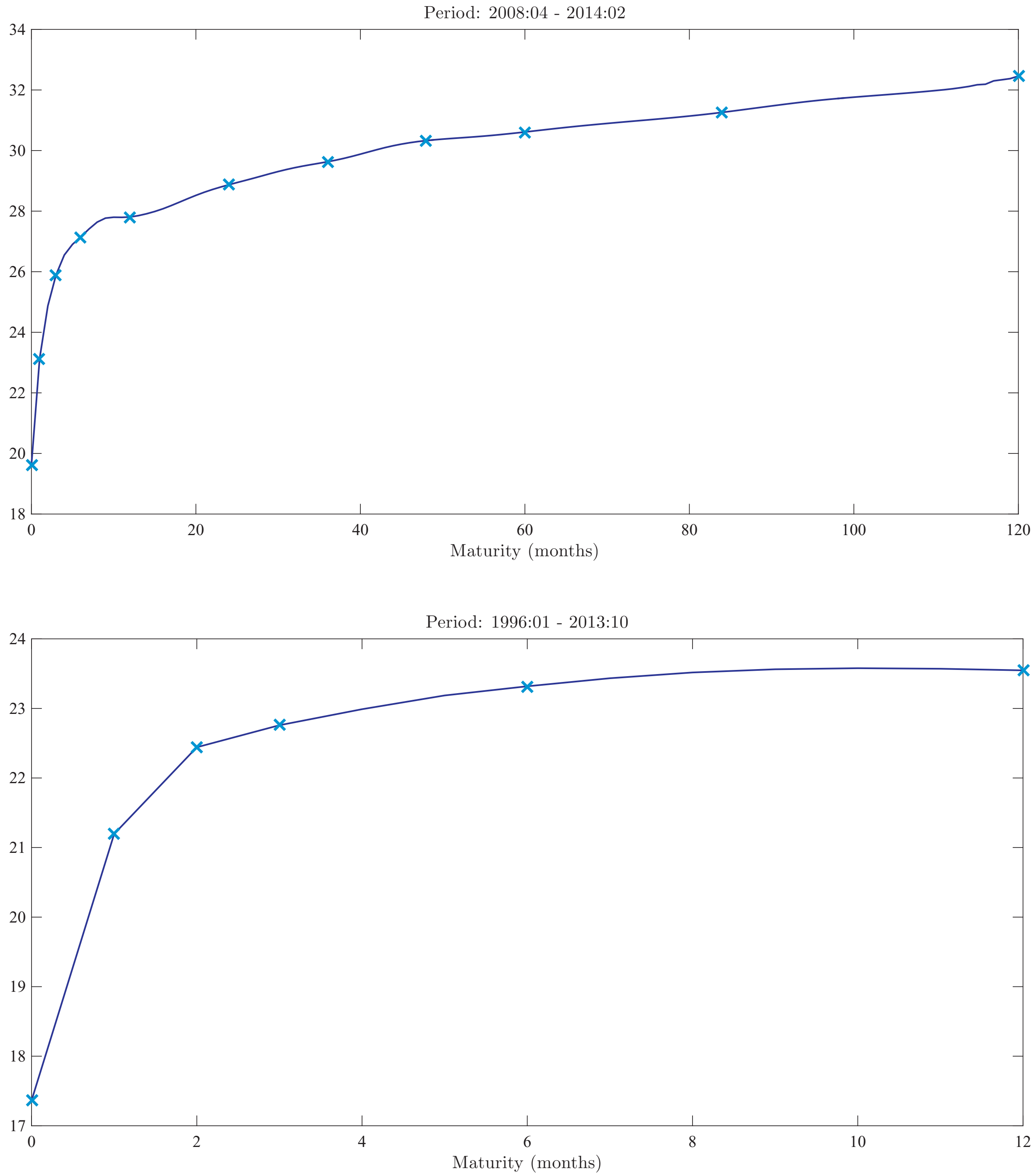


Fig. 2. Average forward variance claim prices. The figure shows the average prices of forward variance claims of different maturity, across different periods. The top panel shows average prices between 2008 and 2013, when we observe maturities up to 10 years (Dataset 2). The bottom panel shows averages between 1996 and 2013, for claims of up to 1 year maturity (Dataset 1). 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, $100 \times \sqrt{12 \times F_t^n}$. Maturity zero corresponds to average realized volatility, $100 \times \sqrt{12 \times F_t^0}$.

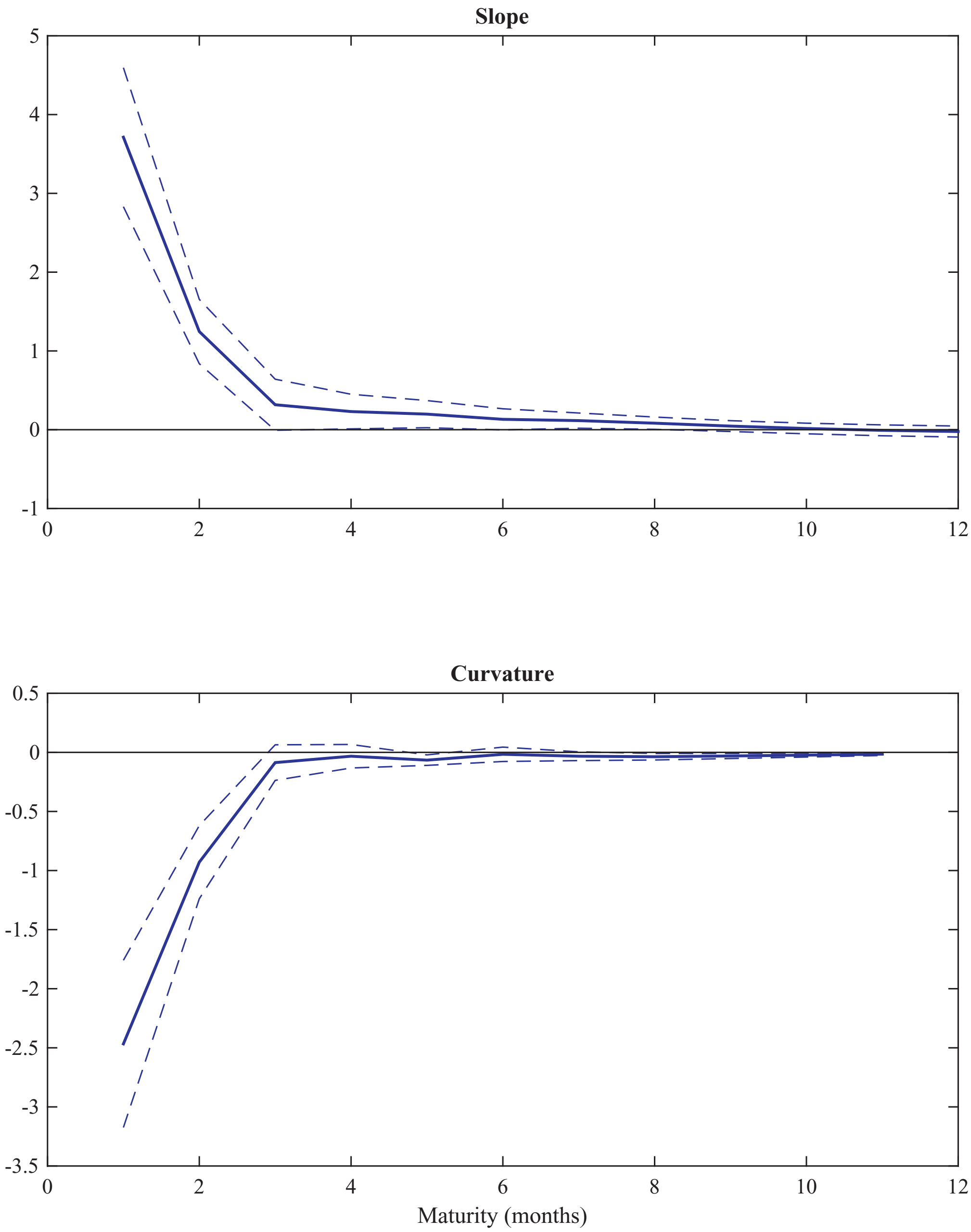


Fig. 3. Slope and curvature of the term structure of forward variance claims. The top panel plots the slope of the term structure of variance swaps (Fig. 2) at each maturity. The bottom panel plots the curvature of the same curve at each maturity. Dotted lines are 95% confidence intervals constructed using Newey-West with 6 lags.

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 one- and two-month maturities are negative, while all the others are weakly positive. Return volatilities are also much higher at short maturities, though the long end still displays significant variability—returns on the 12-month forward have an annual standard deviation of 17%, which indicates that expectations of 12-month volatility fluctuate 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

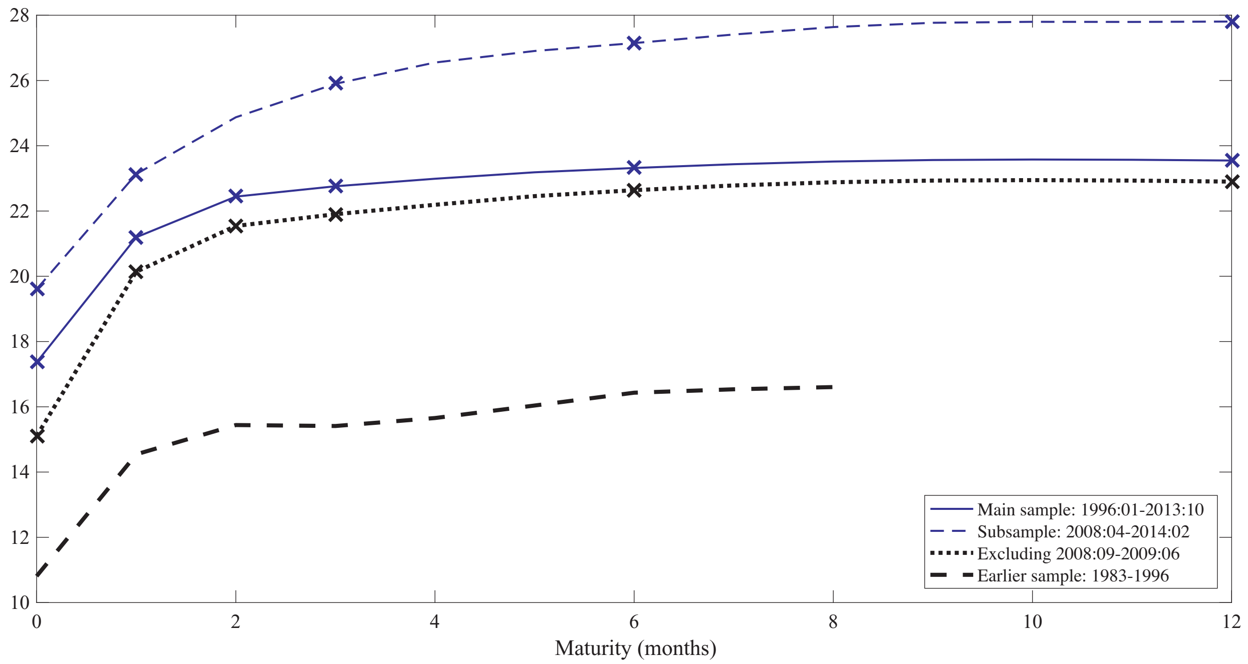


Fig. 4. Subsample analysis of forward variance claims. The figure compares the average prices of forward variance claims for maturities up to 1 year, for the two subsamples of the top and bottom panel of Fig. 2, as well as for the period that excludes the financial crisis and for the period 1983–1996 (the latter uses options data from the CME to construct the term structure of the VIX as opposed to variance swaps).

Table 2
Characteristics of returns.

Maturity (months)	Mean	Std.	Min.	25th p.	Median	75th p.	Max.	Skew	Exc.Kurt.
1	−25.7	67.9	−85.5	−58.4	−40.2	−16.0	686.4	6.2	56.5
2	−5.8	47.7	−59.3	−32.9	−18.4	9.3	376.0	3.9	23.4
3	0.7	33.9	−46.1	−21.4	−5.3	14.5	249.4	2.7	14.1
4	0.6	27.4	−42.2	−17.3	−5.6	11.2	170.4	2.0	7.6
5	0.1	22.5	−37.3	−14.0	−3.7	9.8	126.7	1.6	5.2
6	0.5	19.6	−31.0	−12.2	−3.8	12.9	100.6	1.3	3.3
7	0.6	18.6	−31.4	−12.4	−2.5	11.0	90.7	1.1	2.5
8	0.7	17.4	−29.8	−11.4	−2.9	11.6	81.6	1.0	2.0
9	0.9	16.2	−27.7	−10.2	−1.9	9.2	74.6	0.9	1.7
10	1.1	15.6	−30.0	−9.6	−2.0	9.8	70.8	0.9	1.5
11	1.4	16.0	−32.6	−9.9	−1.9	11.2	69.7	0.9	1.3
12	1.8	17.4	−35.0	−10.3	−2.4	12.1	70.4	1.0	1.4

The table reports descriptive statistics of the monthly returns for forward variance claims (in percentage points). For each maturity n, returns are computed each month as $R^n_{t+1} = \frac{F^{n-1}_{t+1} - F^n_t}{F^n_t}$. Given the definition that $F^0_t = RV_t$, the return on a one-month claim, R^1_{t+1} is the percentage return on a one-month variance swap.

end of the term structure would bias our estimate towards *not* finding large negative expected returns. On the other hand, returns at longer maturities have much lower skewness and kurtosis, which indicates that counterparty risk is substantially less relevant. Finally, we note that we obtain the same results below using options, which are exchange traded and have far less counterparty risk.

Given the different volatilities of the returns at different ends of the term structure, it is more informative to examine Sharpe ratios, which measure compensation earned per unit of risk. Fig. 5 shows the annual Sharpe ratios of the 12 forwards. The Sharpe ratios are negative for the one- and two-month maturities (at around −1.3 and −0.4, respectively), but all other Sharpe ratios are insignificantly different from zero, and in fact slightly positive. More importantly, the lower bounds of the confidence intervals are economically small. We can statistically reject the hypothesis that the Sharpe ratios are meaningfully negative at all maturities above 3; for example, we can reject at the 95% level that the annual Sharpe ratio on a 12-month claim is below −0.11.[14]

[14] One may also worry that some of our results depend on the interpolation between observed maturities. To make sure this does not affect our results, we have constructed six-month holding period returns of a claim to variance six to 12 months forward (which we refer to as the 6/12 portfolio), which does not depend on interpolated data. Of course, the return of the one-month claim (Sharpe ratio of −1.3 as reported in the figure) also does not depend on interpolated data. See Appendix Section A.4 for the empirical results. Appendix Section A.4 also shows that our results hold when the bid-ask spread is explicitly taken into account in computing the returns of these trading strategies.

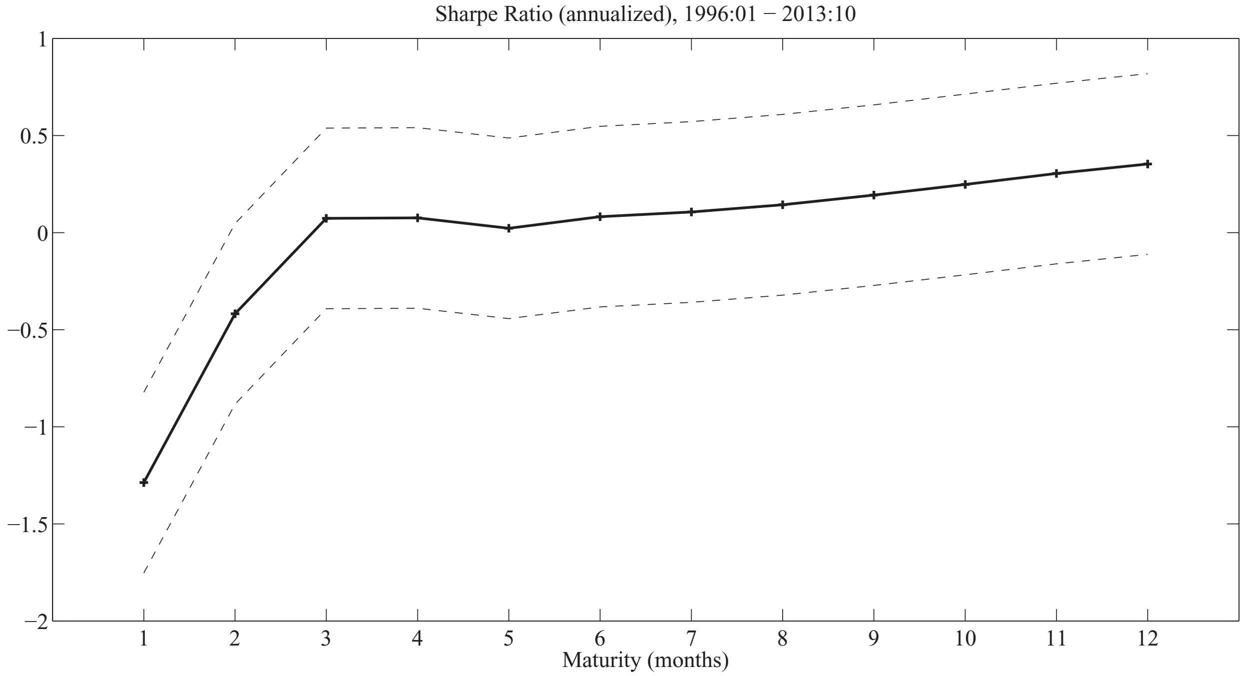


Fig. 5. Annualized Sharpe ratios for forward variance claims. The figure shows the annualized Sharpe ratio for the forward variance claims. The returns are calculated assuming that the investment in an n-month variance claim is rolled over each month. Dotted lines represent 95% confidence intervals. All tests for the difference in Sharpe ratio between the one-month variance swap and any other maturity confirm that they are statistically different with a p-value of 0.03 (for the second month) and $<$ 0.01 (for all other maturities). The sample used is 1996–2013.

Despite the relatively short sample, there are also strongly statistically significant differences between the Sharpe ratios at the very short end of the curve and everywhere else. The annual Sharpe ratio of the one-month variance claim is more negative by 0.9 than the two-month claim (the p-value for the difference is 0.03), and at least 1.3 lower than the Sharpe ratio at all higher horizons (the p-values of the differences are all less than 0.01). These are enormous differences, considering for example that the annual Sharpe ratio of the aggregate stock market has historically been approximately 0.3.

Any 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. Specifically, in the absence of arbitrage, $F_t^n = E_t^Q RV_{t+n}$, and so R_{t+1}^n follows

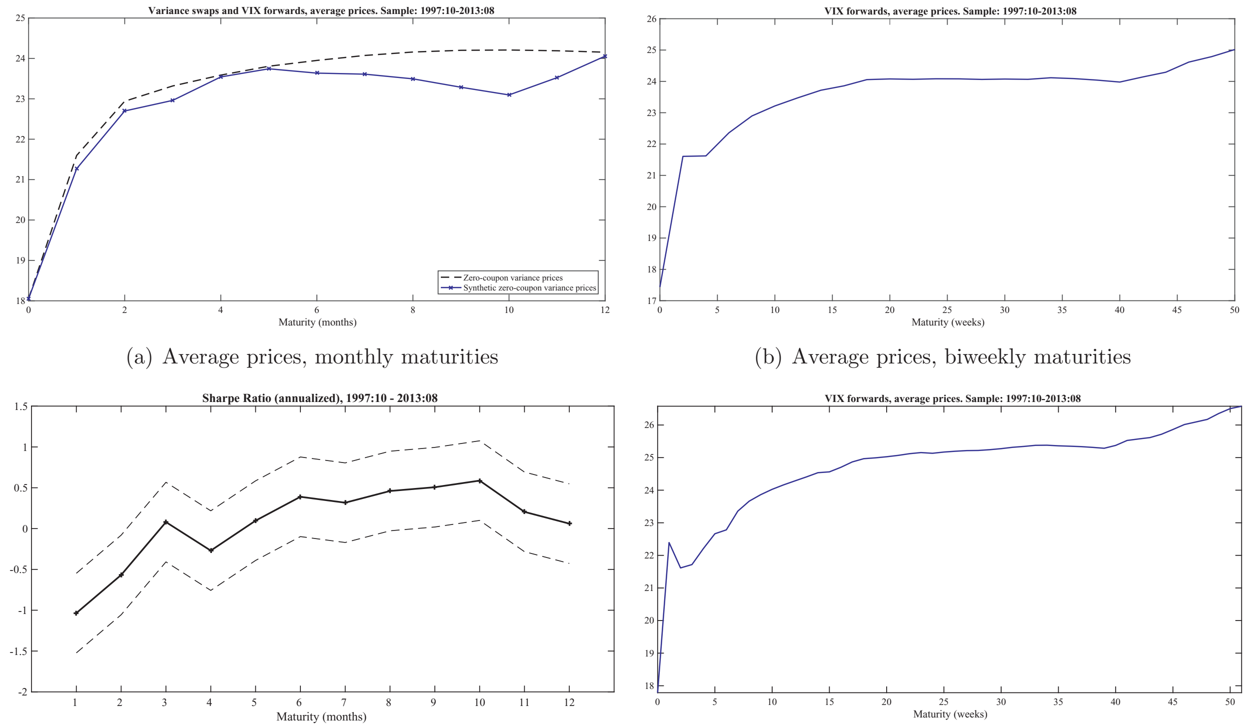
$$R_{t+1}^n = \frac{E_{t+1}^Q RV_{t+n} - E_t^Q RV_{t+n}}{E_t^Q RV_{t+n}} \tag{7}$$

and is determined by the change in expectations of volatility in month $t+n$ (for all $n > 1$). 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 commands a small to zero risk premium in our data.

The results at the short end of the curve indicate that investors were willing to pay a large premium to hedge realized volatility. What is new and surprising in this picture is the fact that investors were willing to pay much less to hedge any innovations in *expected* volatility. The estimated Sharpe ratio is actually positive at *every* point on the curve above maturity three months. Moreover, these declining Sharpe ratios are consistent with the findings of van Binsbergen and Koijen (2015), who find that Sharpe ratios in a range of markets decline with maturity.[15] Like them, we show below that our results are difficult to reconcile with standard theories, thus further extending the puzzle originally set forth by van Binsbergen, Brandt, and Koijen (2012).

Finally, looking down the columns of Table 1, one can also notice a clear correlation between skewness and average returns. Specifically, the correlation between returns and skewness across the various maturities is −0.92, while the correlation between Sharpe ratios and skewness is −0.96. This suggests that one possible explanation of our results is that investors have preferences over skewness, as suggested by Kraus and Litzenberger (1976), Harvey and Siddique (2000), and, more recently, by Schneider, Wagner, and Zechner (2015). We investigate the connection between downside risk and variance swap prices in the context of theoretical models below.

[15] The 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).



(a) Average prices, monthly maturities

(b) Average prices, biweekly maturities

(c) Sharpe ratios

(d) Average prices, weekly maturities

Fig. 6. Synthetic forward variance claims (VIX). The solid line in panel A plots average prices of forward 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 forward variance claims constructed from variance swap prices. Both curves are constructed at monthly maturities as in Fig. 2. Panel C plots annualized Sharpe ratios for forward variance claims returns with prices calculated using the VIX formula and CBOE option data at monthly maturities. Dotted lines in panel C represent 95% confidence intervals. Panels B and D construct the VIX term structure at biweekly and weekly maturities, respectively. The sample covers the period 1997–2013.

3.3. Evidence from other markets

Fig. 6 shows the term structure of prices and Sharpe ratios of variance forwards 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.[16] Appendix Fig. A.6 shows that we obtain similar results with VIX futures.[17]

We focus on monthly maturities because the past literature has mostly studied those maturities and they appear to have high liquidity. But options data have the advantage that it includes maturities shorter than one month, which allows us to measure the point at which the variance forward curve flattens out more precisely. The two right-hand panels of Fig. 6 plot average prices for variance forwards at maturities that are multiples of one and two weeks. These panels provide evidence that the steep slope of the forward curve inside one month is mostly due to the one- to two-week maturities. The curve flattens out noticeably even at two weeks forward. This result remains consistent with, and in fact strengthens, the intuition above, that it is primarily realized variance that is priced, while news about future variance, even at very short horizons, are not.

Our results also extend to international markets. Fig. 7 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. 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.[18]

[16] Given 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. Appendix Section A.1 shows that our synthetic variance swaps are highly correlated with those constructed by the CBOE for the VXV and VXMT indexes. Appendix Fig. A.4 repeats the exercise constructing monthly VIX forward returns from daily data (i.e., with overlapping monthly windows); the results are qualitatively and quantitatively consistent with those in Fig. 6.

[17] VIX 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.

[18] In the Appendix (Fig. A.5), 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.

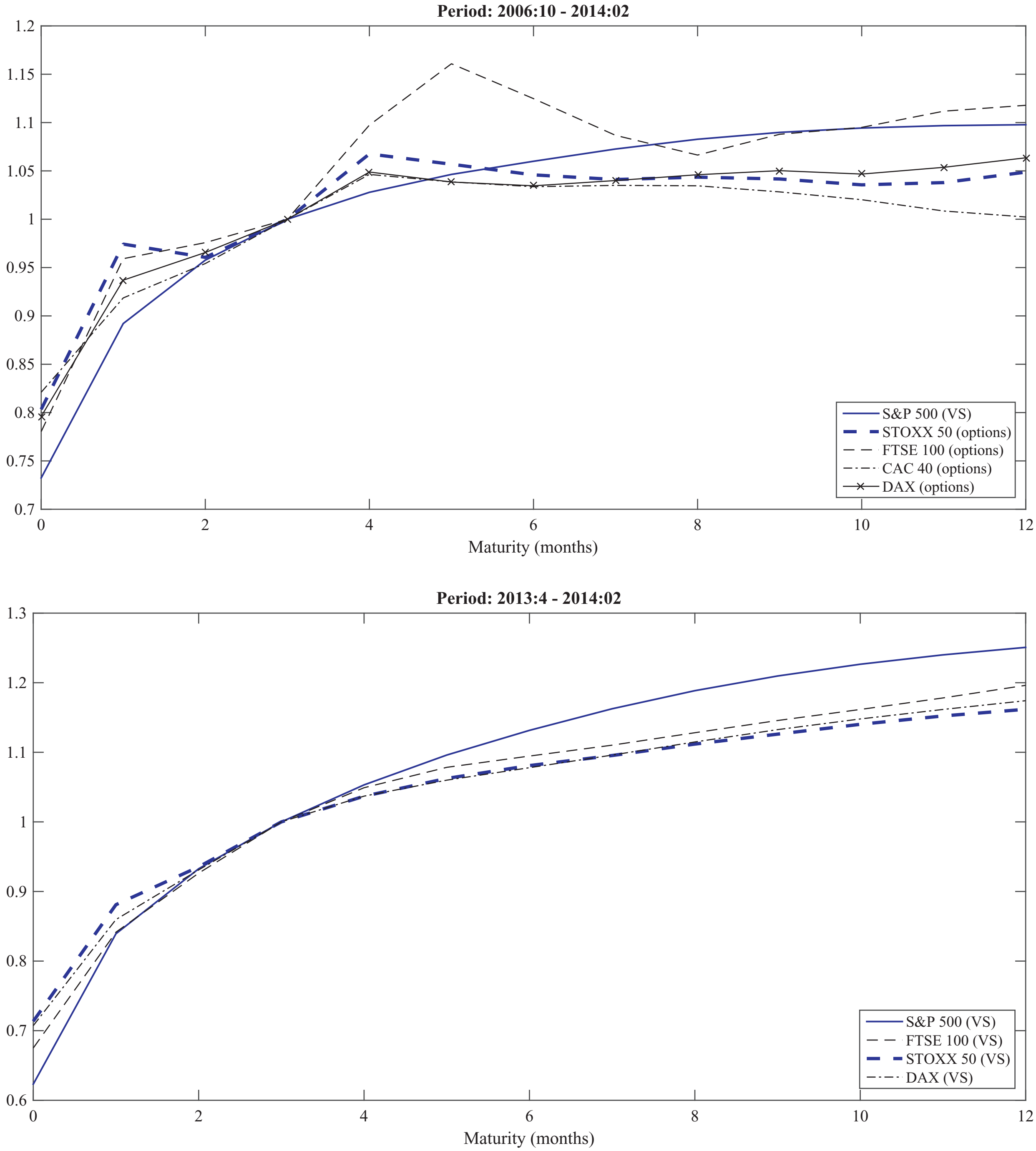


Fig. 7. Average forward variance claim prices for international markets. The figure plots the average prices of forward variance claims as in Fig. 2 for different international indices. The series for the S&P 500 (both in the top and bottom panel) is obtained from variance swaps (as in Fig. 2). The top panel shows international curves obtained using option prices, using the same methodology used to construct the VIX for the S&P 500 (as in Fig. 6). 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 three-month forward variance price.

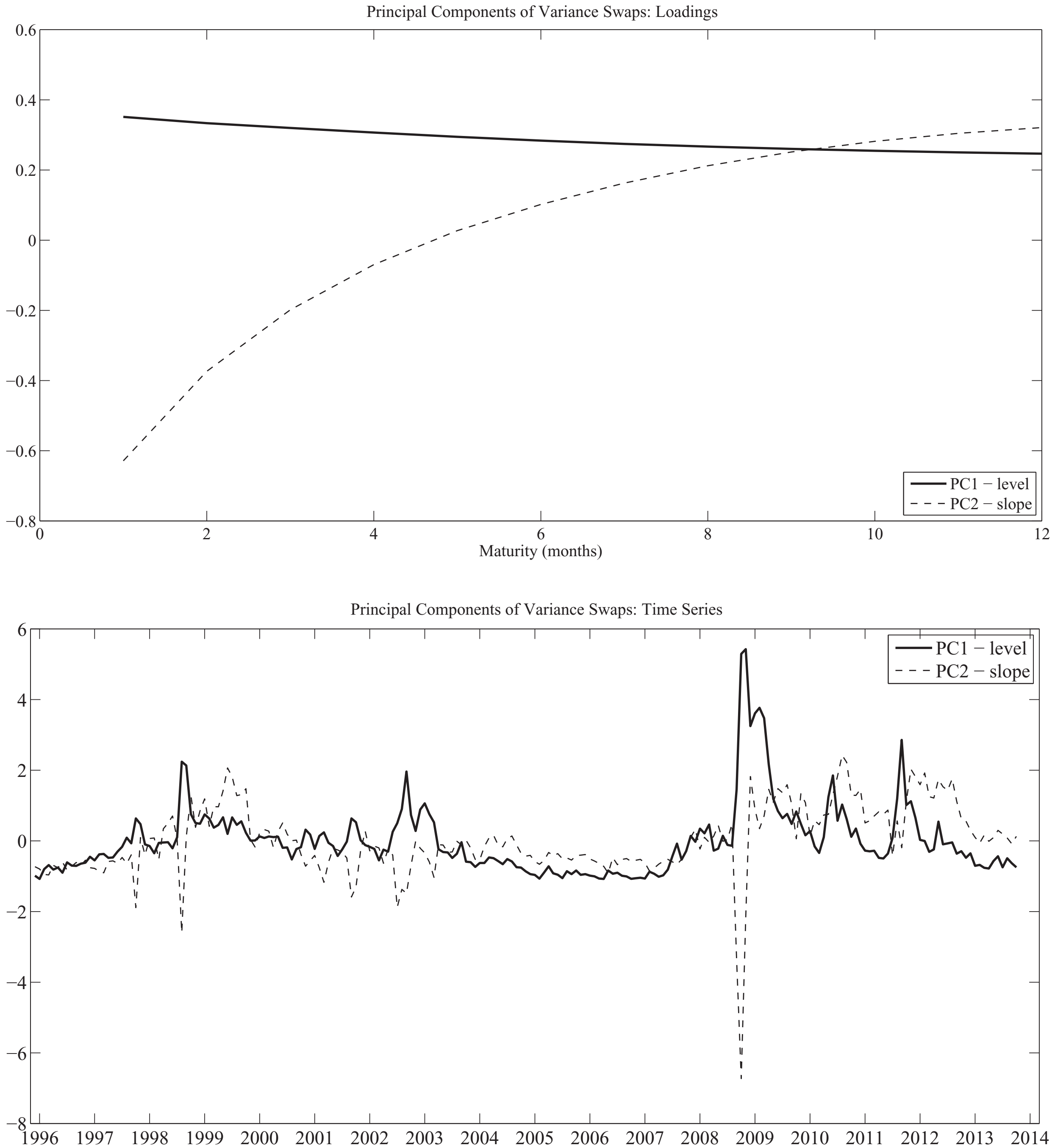


Fig. 8. Principal components of variance swap prices. The top panel plots the loadings of the variance swap prices on the level and slope factors (first two principal components). The bottom panel plots the time series of the level and slope factors. Both are normalized to have zero mean and unit standard deviation and are uncorrelated in the sample. The sample covers the period 1996–2013.

4. Asset pricing

4.1. Reduced-form estimates

We now formally estimate the pricing of volatility risk. As usual in the term structure literature, we begin by extracting principal components (PCs) from the term structure of variance swaps. Throughout this section, we examine models specified both in terms of levels and logs of variance swap prices. The top panel of Fig. 8 plots the loadings of the variance swaps on the PCs in levels—the results in logs are highly similar. The first factor explains 97.1% of the variation in the term structure and the second explains an additional 2.7%. The time series of the factors are shown in the bottom panel of Fig. 8. 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.

We model innovations in the pricing kernel as a linear function of the innovations in realized volatility and the principal components (since we are examining excess returns, the expectation of the SDF is irrelevant),

$$\Delta E_{t+1}M_{t+1} = -b_{RV}\frac{\Delta E_{t+1}RV_{t+1}}{std(\Delta E_{t+1}RV_{t+1})} - b_{PC1}\frac{\Delta E_{t+1}PC1_{t+1}}{std(\Delta E_{t+1}PC1_{t+1})} - b_{PC2}\frac{\Delta E_{t+1}PC2_{t+1}}{std(\Delta E_{t+1}PC2_{t+1})}, \quad (8)$$

where, for any variable X, $\Delta E_{t+1}X_{t+1} \equiv X_{t+1} - E_t X_{t+1}$ is the change in expectations, and $PC1$ and $PC2$ are the first two principal components (PCs) of the term structure, which capture innovations in expectation of future volatility. The innovations are all standardized to have unit variance to aid the interpretation of the coefficients in terms of price of risk per unit of volatility in each factor. In the specification of the model in logs, RV_{t+1} above is replaced with $\log RV_{t+1}$, and the PCs are calculated from the log variance swap prices.

To extract shocks to RV_{t+1} and the two PCs, we estimate a first-order vector autoregression (VAR) with the two principal components and realized variance (RV). The risk prices, b_{RV}, b_{PC1} and b_{PC2}, represent partial derivatives of the pricing kernel with respect to each innovation in the VAR. Note that, as is standard in asset pricing models, we allow for the pricing factors to be correlated with each other.

We estimate risk prices for the three shocks using GMM. Panels A and B of Table 3 report the estimation results using one-step and two-step efficient GMM. The results are consistent across the two panels, and indicate that in the cross section of variance swaps the only factor with a significant price of risk is the realized variance shock. Shocks to expectations of future variance are not priced in this cross-section; they are statistically insignificant and their magnitude is several times smaller than the price of risk of realized variance shocks.

The third row in each panel reports the p-value from a test of whether either of the coefficients on the PCs is the same as the coefficient on RV. That hypothesis is strongly rejected in almost all of the cases (in panel B, b_{PC2} is not significantly different from b_{RV}, but that is simply due to a very large standard error on b_{PC2}).[19]

Panels C and D report the results for the version of the model specified in logs. The results are similar to those in panels A and B: the shock to RV is the only one that is significantly priced, while those to the two PCs are not. The coefficients on RV remain economically and statistically significant, though they are slightly smaller.

[19] Despite the good fit of the model in terms of R^2, the GMM and the Gibbons-Ross-Shanken 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.

Table 3
Reduced-form pricing estimates.

Panel A: level specification, 2-step GMM	RV	$PC1$	$PC2$
Risk prices	-1.32^{***}	0.41	-0.42
Standard error	0.27	0.26	0.36
Difference from RV (p-value)		<.001	<.001
Cross-sectional R^2		.70	
Panel B: level specification, 1-step GMM	RV	$PC1$	$PC2$
Risk prices	-1.23^{**}	0.34	-0.59
Standard error	0.55	0.33	0.63
Difference from RV (p-value)		.003	.17
Cross-sectional R^2		.99	
Panel C: log specification, 2-step GMM	RV	$PC1$	$PC2$
Risk prices	-0.84^{***}	0.38	-0.07
Standard error	0.15	0.24	0.19
Difference from RV (p-value)		<.001	<.001
Cross-sectional R^2		.99	
Panel D: log specification, 1-step GMM	RV	$PC1$	$PC2$
Risk prices	-0.81^{***}	0.42	-0.01
Standard error	0.25	0.33	0.34
Difference from RV (p-value)		.007	.067
Cross-sectional R^2		.99	

Results of GMM estimation of the risk prices for the shocks to RV and to the first two principal components of the term structure of variance swap prices, using Newey-West GMM standard errors with six monthly lags. The three priced innovations are the reduced-form innovations from a VAR with the RV and two PCs. The table also reports the p-values of a test for the differences between the risk prices for $PC1$ and $PC2$ and the risk price for RV. Panels A and C use two-step efficient GMM; panels B and D use one-step GMM with the identity matrix as weighting matrix. Panels A and B use level RV as a first factor, and level prices to construct $PC1$ and $PC2$. Panels C and D use log RV as a first factor, and log prices to construct $PC1$ and $PC2$. *** denotes significance at the 1-percent level.

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 (we also add the market return as a test asset to impose discipline on its risk premium.) Table 4 shows that the return on the market portfolio does not help price the variance claims—the risk price on the market is insignificant across the four specifications, while b_{RV} remains highly significant.

In addition to these reduced-form estimates, we have explored different variations of an affine term structure model with three factors—realized variance and two latent factors that govern its dynamics. The results of the affine model—reported in Appendix Section A.5— confirm those of the reduced-form estimates in this section.

4.1.1. Upside and downside volatility

A natural question is whether investors desire to hedge all volatility shocks, or whether they primarily desire to hedge volatility during downturns. Segal, Shaliastovich, and Yaron (2015), for example, discuss such a model. 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 and lower RV in each month as

Table 4
Controlling for the market return.

Panel A: level specification, 2-step GMM	RV	*PC* 1	*PC*2	*Rm*
Risk prices	−1.35***	0.11	−0.40*	−0.55
Standard error	0.27	0.42	0.24	0.39
Cross-sectional R^2			.88	
Panel B: level specification, 1-step GMM	RV	*PC* 1	*PC*2	*Rm*
Risk prices	−1.25**	−0.02	−0.44	−0.68
Standard error	0.58	0.67	0.43	0.72
Cross-sectional R^2		.99		
Panel C: log specification, 2-step GMM	RV	*PC* 1	*PC* 2	*Rm*
Risk prices	−0.75***	−0.05	−0.01	−0.56
Standard error	0.17	0.39	0.12	0.51
Cross-sectional R^2		.97		
Panel D: log specification, 1-step GMM	RV	*PC* 1	*PC* 2	*Rm*
Risk prices	−0.89**	−0.43	0.12	−1.25
Standard error	0.43	0.64	0.17	0.89
Cross-sectional R^2		.99		

Same as Table 3, but adding the market portfolio as a test asset and as a pricing factor. * Denotes significance at the 10-percent level, ** denotes significance at the 5-percent level, and *** denotes significance at the 1-percent level.

$$RV_t^U = \sum_{j\in t}(r_j)^2 1\left\{P_j > P_0\right\} \tag{9}$$

$$RV_t^D = \sum_{j\in t}(r_j)^2 1\left\{P_j \le P_0\right\}, \tag{10}$$

where $j \in t$ indicates days j in month t, and $1\{\cdot\}$ is the indicator function. RV^U is realized variance calculated only on days when the market is above its level at the beginning of the month, and RV^D is realized variance on the remaining days.

Andersen and Bondarenko (2007) discuss two useful properties of these realized semivariances: the two components sum to RV_t, and their prices sum to the squared VIX,

$$RV_t = RV_t^U + RV_t^D \tag{11}$$

$$VIX_t^2 = (VIX_t^U)^2 + (VIX_t^D)^2, \tag{12}$$

where VIX_t^U and VIX_t^D are the prices of claims to RV_{t+1}^U and RV_{t+1}^D, respectively. 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.

Fig. 9 plots the term structure of the variance forwards 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.[20] This confirms the intuition that the reason investors hedge realized volatility is due to its downside component (which Bollerslev and Todorov, 2011 show is dominated by downward jumps), and is consistent with investors displaying aversion to skewness.

[20] Note 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.

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.

First, there is strong evidence in the literature that volatility is predictable three months ahead. See, for example, Andersen, Bollerslev, Diebold, and Labys (2003). Indeed, the volatility literature has demonstrated predictability at horizons much longer than three months.[21] But the sample mean Sharpe ratios on variance forwards are insignificantly different from zero even for maturities as short as three months, and the point estimates for some are even positive. This suggests that even at shorter horizons where the evidence for volatility predictability is strongest, volatility news has not been priced.

To quantify the magnitude of the predictability of volatility at different horizons Table 5 reports R^2s from predictive regressions for realized volatility at different frequencies and horizons. Specifically, we run the regression

$$RV_t = b_0 + b_{RV}RV_{t-j} + b_{PE}PE_{t-j} + b_{DEF}DEF_{t-j} + \varepsilon_t, \tag{13}$$

where PE_t is the aggregate market's price/earnings ratio and DEF_t is the default premium—the spread between the yields on Aaa and Baa bond yields reported by Moody's.

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^2s for monthly volatility range from 45% at the one-month horizon to 20% at the 12-month horizon. In predicting annual volatility, R^2s range between 56 and 21% for horizons of one to ten years.

It is important to note here that the monthly or annual specification of the regression (13) cannot distinguish between whether diffusive or jump risk is predictable. What we show is simply that total realized volatility is predictable. This is sufficient for our purposes for two reasons. First, RV_t is what actually determines the payoffs of variance swaps. Second, in the theoretical models based on Epstein–Zin preferences that we examine below, fluctuations in both jump risk and diffusive volatility should be priced. So it is important just to know whether any source of realized variance is predictable.

In order to gauge the economic magnitude of the predictability of realized variance, the third pair of columns in Table 5 reports the results of forecasts of dividend growth (i.e., replacing RV_t with dividend growth in Eq. (13)). R^2s

[21] Andersen, Bollerslev, Diebold, and Labys (2003), Ait-Sahalia and Mancini (2008), Federico Bandi 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, Giglio, Polk, and Turley (2013) focus on longer horizons (up to ten years) and show that both the aggregate price-earnings ratio and the Baa-Aaa default spread are useful predictors of long-run volatility.

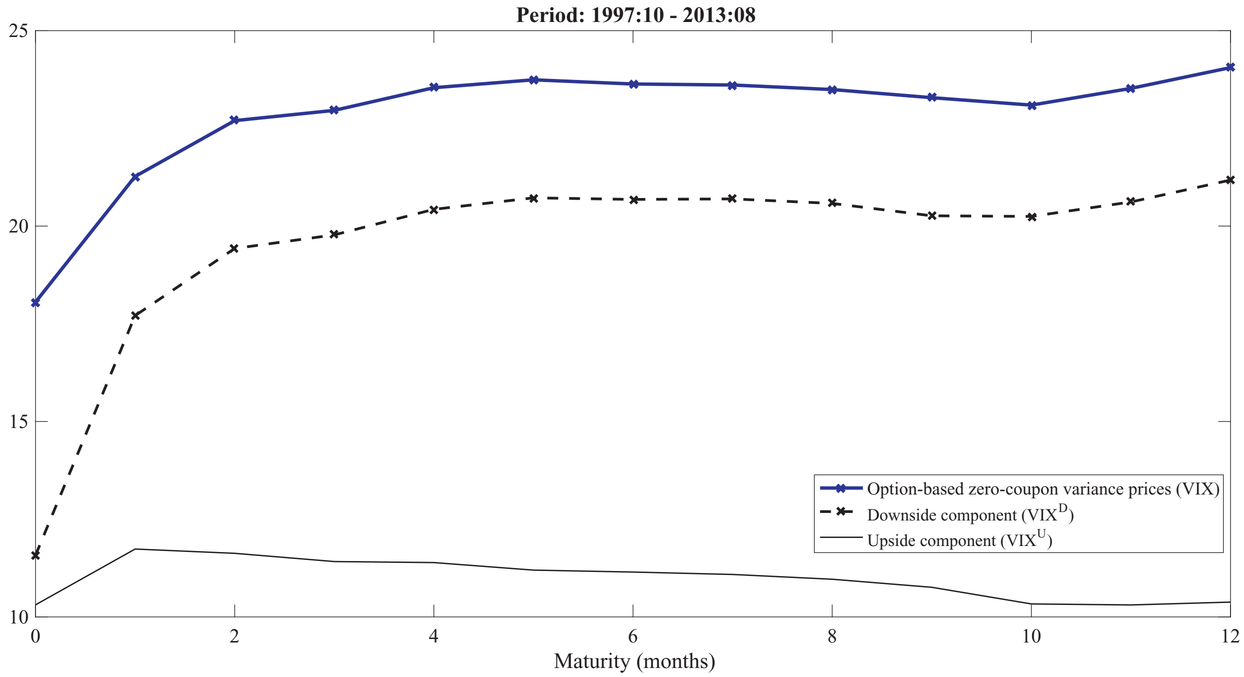


Fig. 9. Decomposing the upward and downward volatility components. The solid thick line plots average prices of forward variance claims calculated using the formula for the VIX index. The dashed line plots the forward prices of the downside component of the VIX, VIX^D. The thin solid line plots the forward prices of the upside component of the VIX, VIX^U. All series are constructed using option data from CBOE. The sample covers the period 1997–2013.

Table 5
Forecasting volatility at different horizons: *R*2.

Predictor: with PE_t, DEF_t Months	Monthly RV_{t+n} RV_t	$\surd$	Years	Yearly RV_{t+n} RV_t	$\surd$	Years	Yearly Δd_{t+n} Δd_t	$\surd$
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

The first column of the table reports R^2 of predictive regressions of monthly volatility n months ahead at the monthly frequency. The second column reports R^2 of predictive regressions of yearly volatility n years ahead at the yearly frequency. The second column reports R^2 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.

for dividend growth are never higher than 9%. So in the context of financial markets, there is an economically large amount of predictability of volatility. The Appendix (section A.3) takes an extra step beyond Table 5 and provides evidence, 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 three 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 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; similarly the macro literature showing that volatility news can drive the business cycle would seem irrelevant if there is no volatility news.

5. Economic interpretation

This section examines simulations of four major structural asset pricing models to understand how our data on the variance term structure can help test and distinguish among models. Among the models we consider, only the long-run risk model of Drechsler and Yaron (2011) was originally calibrated to match the one-month variance risk premium. We therefore calibrate that model as in the original specification. The other models we study did not originally target any moments of the variance risk premia, and tend to predict too low a level of the risk premium for one-month variance swaps. In order to focus on the predictions of the models for the slope of the term structure in these models (rather than just the level of risk premia), whenever possible we raise risk aversion to help the models match the one-month variance risk premium, and then study the implications for the higher maturities.

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 one-month 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

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

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

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

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

where Δc_t is log consumption growth, the shocks ε are mean-zero and normally distributed, and the shocks J are Poisson-distributed 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

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

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.

5.1.2. Time-varying disaster risk and Epstein–Zin preferences

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

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

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

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

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.[22] We calibrate the model similarly to Wachter (2013) and Barro (2006), with the main exception being the increase in risk aversion motivated above. 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% as in Wachter (2013). σ_F is set to 0.0056 (ε_F is a standard normal), and $\rho_F = 0.92^{1/12}$, 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, $J_t \sim N(-0.15, 0.1^2)$. It is important to note that this distribution is not identical to what is used in Wachter (2013), which is an actual empirical distribution of disaster sizes (we use the normal distribution for analytic tractability).

Finally, dividends are a claim to aggregate consumption with a leverage ratio of 2.6. Note that the occurrence of a disaster shock implies that equity values decline 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 −10%. This allows for a slightly delayed diffusion of information and also potentially realistic factors such as exchange circuit-breakers.[23] Our results are qualitatively unchanged as long as the jump in stock prices in a disaster is as large as 35% per day.[24] The Appendix provides more details on how the results depend on this choice.

We follow Wachter (2013) in assuming the elasticity of intertemporal substitution is 1 and raise risk aversion to 4.9 to give the model the best chance of generating Sharpe ratios and a slope for the term structure as large as we see in the data (it cannot be raised further because of equilibrium existence constraints, as discussed in the Appendix).

5.1.3. Disaster risk and habit formation

Du (2011) and Christoffersen, Du, and Elkamhi (2015) study an extension of Campbell and Cochrane's (1999) model of habit formation, adding rare disasters (where Campbell and Cochrane assumed consumption growth was normally distributed). The model is specified in continuous time. The representative agent maximizes

$$E_0\left[\int_0^\infty \exp(-\rho t)\log(C_t - H_t)dt\right], \tag{21}$$

[22] Seo and Wachter (2015) consider an extension of the model here that allows multiple factors to drive the probability of a disaster. In unreported results, we find that the two-factor model predicts behavior for volatility claims that is highly similar to that for the single-factor specification.

[23] For example, a jump of 20% would occur over 2 consecutive days, with a 10% decline per day. Note that this choice has only a minimal effect on the Sharpe ratios predicted by the model; for example, Sharpe ratios with a maximum daily loss of −20% are essentially the same (the p-values discussed below are unchanged, for example). The small shocks $\varepsilon_{\Delta c,t}$ are treated as though they occur diffusively over the month, as in Drechsler and Yaron (2011).

[24] The largest negative single-day return in the Center for Research in Security Prices value-weighted index is −17.4% on October 19, 1987. The second and third largest negative returns are −11.29% and −12.01% on October 28, 1929 and October 29, 1929, respectively. The largest negative daily returns in 2008 were all smaller than 10%. A single-day decline of 35% would thus be nearly twice as large as any return experienced in the US.

where ρ is the pure rate of time preference, C_t is consumption, and H_t is the level of habit. The implied coefficient of relative risk aversion at time t, γ_t, is

$$\gamma_t = \frac{C_t}{C_t - H_t}. \tag{22}$$

So when consumption is higher above the habit, risk aversion is lower. Following Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), γ_t is specified to follow a continuous-time AR(1) process with time-varying sensitivity to the consumption growth innovation,

$$d\gamma_t = k(\bar{\gamma} - \gamma_t)dt - \alpha(\gamma_t - \beta)(dc_t - E_t[dc_t]), \tag{23}$$

which then implies a process for H_t. dc_t is consumption growth, which is independent and identically distributed over time with both a small diffusive component and a large jump component,

$$dc_t \equiv d\log C_t = \mu dt + \sigma dB_t + bdN_t, \tag{24}$$

where B_t is a standard Brownian motion and N_t is a Poisson process with a constant jump intensity λ. b is the size of the jump in consumption on the impact of a disaster shock.

Note that the volatility of γ_t depends on γ_t itself. Following negative shocks, when γ_t is high, the future volatility of innovations to γ_t (and hence also of stock returns) is also high. This is thus a model with endogenously time-varying volatility, which is why there is no need to include time variation in disaster risk. Following a disaster, consumption is low, γ_t is high, and future volatility is high. So not only is realized volatility high in disasters, but so is expected future volatility. We would thus expect claims to both realized and expected future volatility to earn large negative returns since they both have high payoffs following disasters.

Given this setup, Du (2011) derives the price and return of a consumption claim in closed form. We then calculate prices of variance claims numerically.

We calibrate the model exactly as in Du (2011). The steady-state level of risk aversion, $\bar{\gamma}$, is 34, and the size of a disaster, $-b$ is 17.2%.[25] As we will see later, the model cannot match the Sharpe ratio on the one-month variance swap. However, it turns out that raising risk aversion in this model does not actually increase risk premia, since in the model this also affects the dynamics of the habit precisely in a way that offsets the increase in risk aversion. We therefore use the original specification of the model.[26]

5.1.4. Time-varying recovery rates

The final model we study is based on Gabaix's (2012) model of disasters with time-varying recovery rates. Because the probability of a disaster is constant, power utility and Epstein-Zin preferences are equivalent in terms of their implications for risk premia. We model the consumption process identically to Eq. (19) above, but with the probability of a disaster, F_t, fixed at 1% per year (Gabaix's calibration). Following Gabaix, dividend growth is

$$\Delta d_t = \mu_{\Delta d} + \lambda\varepsilon_{\Delta c,t} - L_t \times 1\{J_{\Delta c,t} \neq 0\}. \tag{25}$$

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. L_t represents the recovery rate of stocks in a disaster and is assumed to follow the process

$$L_t = (1 - \rho_L)\bar{L} + \rho_L L_{t-1} + \varepsilon_{L,t}. \tag{26}$$

We calibrate $\bar{L} = 0.5$ and $\rho_L = 0.87^{1/12}$, and $\varepsilon_{L,t} \sim N(0, 0.04^2)$, which means that the standard deviation of L is 0.25.[27] We set the coefficient of relative risk aversion to 7 to match the Sharpe ratio on one-month variance swaps (as we did for the other disaster models; as pointed out before, for the long-run risk model, there was no need to adjust the calibration since the paper already targeted the behavior of the one-month variance swap). Other than the change in risk aversion, our calibration of the model is nearly identical to Gabaix's (2012). 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 four calibrated models for variance forwards. Fig. 10 plots population moments from the models against the values observed empirically. The top panel reports annualized Sharpe ratios for forward variance claims with maturities from one month to 12 months. Our calibration of the model with time-varying recovery rates with power utility matches the main stylized facts well: it generates a Sharpe ratio for the one-month claim of −1.3, while all the forward claims earn Sharpe ratios of zero, economically similarly to what we observe in the data.

On the other hand, the other three models significantly underprice variance risk at the short end relative to the longer end of the variance curve. In these models, the Sharpe ratio on the one month forward is far smaller—at approximately 0.3—than in the data (approximately −1.3). By contrast, the models generate Sharpe ratios for claims on variance more than three months ahead that are counterfactually large, almost as large as the one-month forward. In the data, instead, they are zero or even positive at all horizons above three months.

The underpricing of risks at the short end is caused by the fact that these models do not generate pricing kernels sufficiently volatile to give *any* asset a Sharpe ratio of 1.3. However, simply increasing the volatility of the pricing kernel by increasing risk aversion will not solve the problem,

[25] As in the time-varying disaster model, 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 −10%.

[26] In particular, it is easy to see from Eq. (3) in Du (2011) that scaling up the entire risk aversion process requires scaling proportionally $\bar{\gamma}$ as well as β. But this scaling has no effect on risk prices, as evident from Eq. (7) in Du (2011).

[27] As for the time-varying disaster model, 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 −10%.

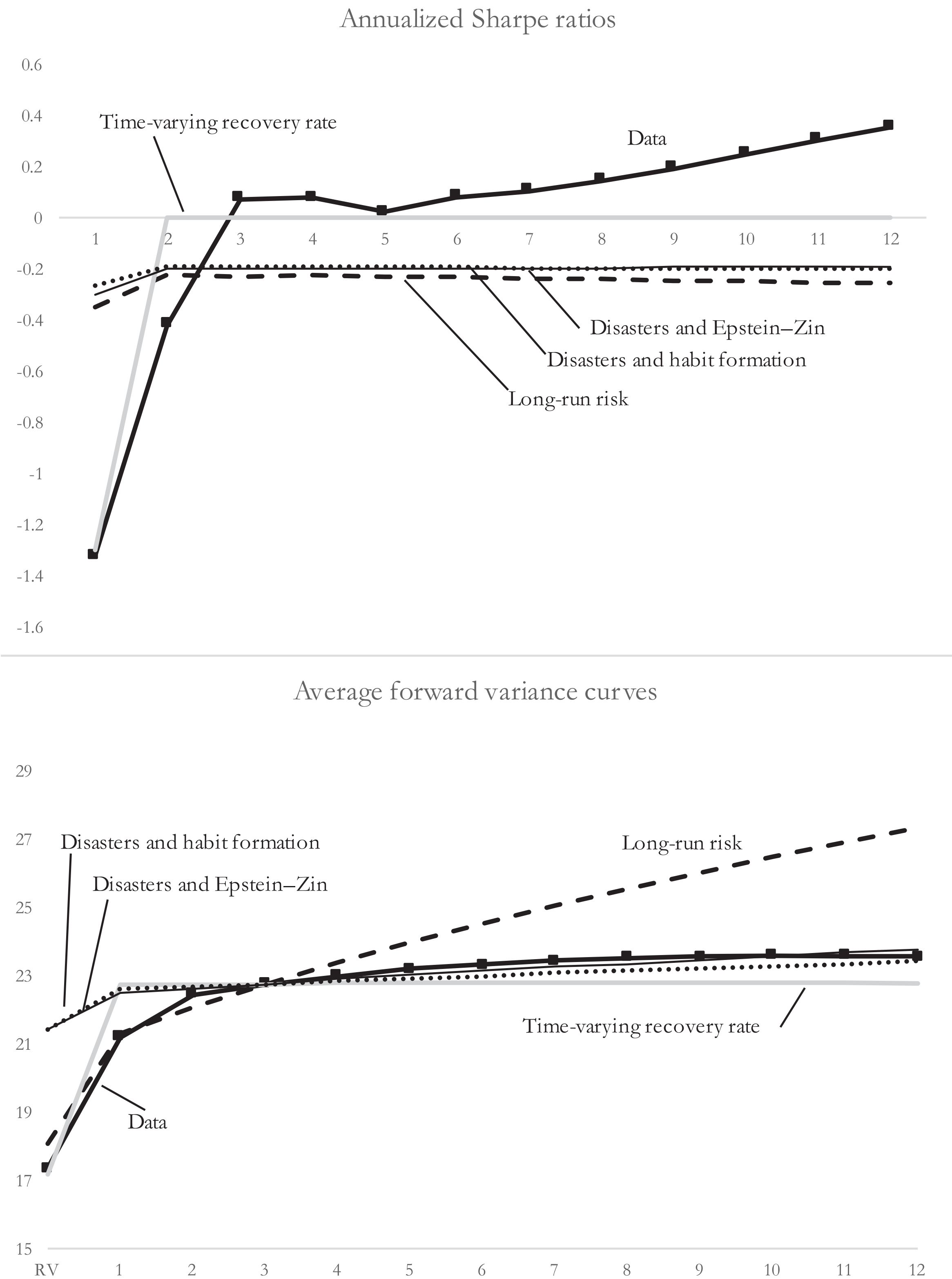


Fig. 10. Sharpe ratios and average term structure in different models. The top panel gives the population Sharpe ratios from the four models and the sample values from the data. The bottom panel plots population means of the prices of forward claims. All the curves are normalized to have the same value for the three-month forward claim.

as it will simply increase the Sharpe ratio at all maturities and exacerbate the mispricing at horizons longer than three months.

The economic intuition for the result with Epstein–Zin preferences is straightforward. If investors are risk-averse, then periods of high expected future consumption volatility are periods of low lifetime 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 consumption volatility, and in these models forward variance claims allow them to do so. Moreover, in these models volatility in all future periods (discounted at a rate close to the rate of pure time preference, and therefore close to zero in standard calibrations) affects lifetime utility, which is why investors in these models pay nearly the same amount to hedge volatility at any horizon.

The expected returns on the variance forwards are closely related to the average slope of the variance term structure. The bottom panel of Fig. 10 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.[28] The model with habit formation is similar to Epstein–Zin with time-varying disaster risk. 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. Of course, as is clear from the figure, according to the model the term structure should be always exactly flat at maturities higher than one—the model is therefore technically speaking unable to generate the small but positive slope observed in the data. Economically speaking, however, the model matches well the flatness of the term structure that we have documented empirically for the period 1996–2014.

The comparison between the calibrated models and the data reported in Fig. 10 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 (similar to the analysis in van Binsbergen and Koijen (2015)). In particular, we focus on the ability of the models to match the high slope at the short end and the flatness at the long end of the curve.

Table 6 reports results from those simulations. We examine 215-month simulations to compare to our full sample since 1996, and 70-month simulations to compare to the shorter sample in which we have ten-year swaps available. For each simulation, we calculate the averages of the simulated values of $(F_t^3 - F_t^0)$, $(F_t^{12} - F_t^3)$, and, in the long sample, $(F_t^{120} - F_t^3)$. Table 6 reports the fraction of simulated samples in which the sample mean of $(F_t^3 - F_t^0)$ is at least as *large* as we see in the data, the sample mean of $(F_t^{12} - F_t^3)$ is *smaller* than in the data, or the sample mean of $(F_t^{120} - F_t^3)$ is *smaller* than in the data. These fractions are one-sided p-values: they measure the probability that the model would have generated slopes as extreme as we observe in the data. Furthermore, the bottom rows report the fraction of samples in which the models simultaneously generate slopes as high as we observe below three months and as flat as we observe above three months. They are thus p-values for tests of whether the models can match the observed concavity of the term structure.

The long-run risk model does a relatively good job of generating a large slope at the short end—20% of the long samples and 38% of the short samples are at least as steep as in our data. However, the slopes after the three-month maturity struggle to match the data—the sample mean of $(F_t^{12} - F_t^3)$ is as small as observed empirically in the long sample less than 0.1% of the time. When we ask how many samples generate both the steep slope below three months and the flat slope after three months, the p-value is less than 0.005. In other words, the long-run risk model generates a large short-maturity slope, but significantly fails to match the flatness of the term structure after three months.

To explore this result in greater detail, the top panel of Fig. 11 reports a four-year moving average estimate of the slope of the forward curve between three months and one year (solid line). The width of the window was chosen so that the very last points cover exclusively the period after the financial crisis. The solid line reports the average slope with 95 confidence intervals, while the dashed line reports the average slope according to the long-run risk model.

The figure shows several interesting patterns. The 12- versus three-month slope has been quite stable in the last 20 years; it has never—in any four-year period in the last 20 years—taken the value that should be the overall unconditional *average* according to the long-run risks model. After the financial crisis the slope has in fact increased above its historical mean, but even then, not high enough to reach the average value implied by the model.

The bottom panel of Fig. 11 plots the six- versus three-month slope using S&P 500 options data from the CME. This sample is slightly different from our other sources and extends back to 1983. At these shorter maturities, we see that over a 30-year period, the four-year moving average

[28] We note that increasing the maximum daily return possible in the disaster model with Epstein-Zin preferences increases the short-term slope of the term structure. Unless the maximum daily return is −40%, though (more than twice as large as any return observed in the US since 1926), the p-values do not change. The appendix reports results for different calibrations of the maximum daily returns. Changing the maximum daily return does not affect the term structure of Sharpe ratios, which also are not consistent with the data. One way to get the Sharpe ratios for the variance swaps to change is to assume that the data we observe is a period without disasters (the financial crisis notwithstanding). The Appendix also reports results for that case. Then it is possible to generate large negative Sharpe ratios for variance swaps, but they are too large by an order of magnitude – −18 or more. Furthermore, without disasters, realized volatility is not sufficiently volatile – the one-month variance swap return in the model has a standard deviation less than half as large as that of the six-month forward, whereas in the data the six-month forward return is only one-third as volatile as the one-month variance swap.

Table 6
Model tests using the variance swap data.

	70-month simulations, up to 12mo maturity			
	Long-run risks p-value	Disasters and Epstein-Zin p-value	Disasters and habit formation p-value	Time-varying recovery p-value
Simulated 3mo/RV slope ≥ empirical slope	0.20	<0.01	<0.01	0.69
Simulated slope 12mo/3mo ≤ empirical slope	<0.01	0.49	0.16	1.00
Simulated slope 120mo/3mo ≤ empirical slope	–	–	–	
Joint test: 3mo/RV≥ data and 12mo/3mo≤ data	<0.01	<0.01	<0.01	0.69
Joint test: 3mo/RV≥ data and 120mo/3mo≤ data	–	–	–	–
	215-month simulations, up to 120mo maturity			
	Long-run risks p-value	Disasters and Epstein-Zin p-value	Disasters and habit formation p-value	Time-varying recovery p-value
Model 3mo/RV slope ≥ empirical slope	0.38	<0.01	<0.01	0.82
Model slope 12mo/3mo ≤ empirical slope	0.05	0.90	0.38	1.00
Model slope 120mo/3mo ≤ empirical slope	0.02	0.20	0.42	1.00
Joint test: 3mo/RV≥ data and 12mo/3mo≤ data	<0.01	<0.01	<0.01	0.82
Joint test: 3mo/RV≥ data and 120mo/3mo≤ data	<0.01	<0.01	<0.01	0.82

We simulate 10,000 70- and 215-month samples from the four models (respectively, in the top and bottom panels). In each simulation, we calculate 3–0 (RV), 12-3, and 120–3 month slopes of the variance forward term structure. The numbers in the first row of each panel are the fraction of samples in which the models generate a slope at the short end of the curve at least as large as observed empirically. The second and third rows are the fraction of samples in which the models generate slopes at the long end of the curve at least as flat as observed empirically. The bottom rows are the fraction of samples in which both conditions are satisfied.

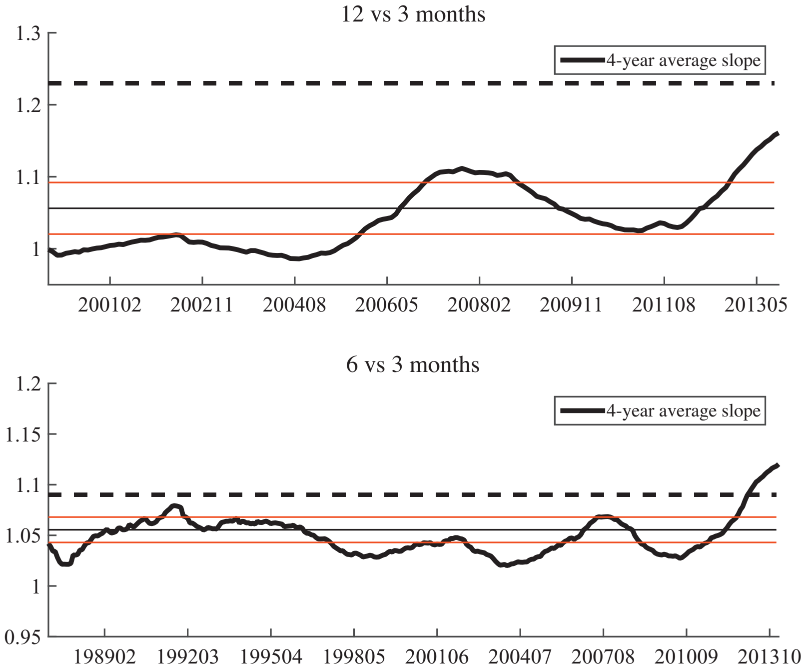


Fig. 11. Four-year moving averages of the slope vs. long-run risk model. The top panel reports a four-year moving average of the slope of the variance forward term structure between three and 12 months of maturity (thick solid line). The thin solid lines report the full-sample average slope with 95% confidence bands. The dashed line reports the average slope implied by the Drechsler and Yaron (2011) model. The sample period covers 1996–2013. The bottom panel uses CME data on options up to six months maturity to construct a four-year moving average of the VIX forward term structure, between maturities three and six months. The sample period covers 1984–2013.

never approaches the unconditional mean of the DY model except for a short period at the very end of the sample.[29]

[29] The better performance of the long-run risks model at the very end of the sample is an interesting fact, visible also at higher maturities, which is difficult to interpret given the short period in which the slope has increased. It could be a temporary deviation from the much flatter historical slope we have observed since the 80s, or it could be a permanent change that will persist in the future. Whether in the future the data will behave

Table 7
Higher moments of variance forward returns in data and models.

	Maturity (months)	Data	Disasters with Epstein-Zin	Disasters with habit	Long-run risks and Epstein-Zin	Time-varying recovery
Mean	1	−308.98	−110.85	−131.65	−176.00	−476.66
	3	8.58	−7.07	−20.28	−33.11	0.07
	6	5.58	−6.90	−17.79	−26.41	0.07
	12	21.29	−6.60	−16.74	−20.25	0.06
SD	1	235.34	416.87	603.37	547.73	401.75
	3	117.51	36.90	93.82	148.35	5.99
	6	67.99	35.40	82.29	115.78	5.78
	12	60.33	32.92	77.49	79.55	5.38
Skewness	1	6.17	49.17	20.23	26.24	32.54
	3	2.68	0.02	5.85	9.77	0.01
	6	1.27	0.02	5.22	9.48	0.01
	12	0.99	0.01	4.94	8.67	0.01
Kurtosis	1	59.49	3580.10	461.96	1078.00	1091.20
	3	17.13	3.80	119.05	157.42	3.18
	6	6.33	3.68	103.72	154.07	3.17
	12	4.37	3.52	97.41	139.69	3.14

The data moments are estimated on the full sample. The model-implied moments are the average values across simulations with the same length as our empirical sample.

The model with time-varying disaster risk and Epstein–Zin preferences and Du's model with disasters and habit formation have the opposite problem from the long-run risk model: they generate relatively flat term structures at maturities longer than three months, but they both fail to match the steep slope observed below three months. The p-values are similar to those for the long-run risk model—the probability that the time-varying disaster model generates the steep slope below three months is less than 0.1%, while the probability that it generates a slope as flat as we see between three and 12 months is 49%. In none of our simulations do the disaster models with Epstein–Zin or habit formation preferences simultaneously match the slopes both below and above three months.

Finally, Table 6 shows that the model with time-varying recovery can in fact match well both the slopes below and above three months. It has a slope as steep as we observe empirically between zero and three months in 69% of the short samples and 82% of the long samples. It also has a slope after three months as flat as we observe empirically in 100% of the samples. It therefore matches the slopes both below and above three months in 69% and 80% of the short and long samples, respectively.

To step beyond the means and standard deviations summarized by Fig. 10 and Tables 6 and7 reports also the skewness and kurtosis of the returns of variance claims at horizons of one, three, six, and 12 months. For the one-month variance claim, all four models we examine overstate the standard deviation, skewness, and kurtosis of returns. The models with very large disasters—Wachter (2013) and Gabaix (2012)—generate the greatest skewness and kurtosis. Du (2011) and Drechsler and Yaron (2011) have much lower levels of skewness, since they have much smaller jumps in stock prices. To the extent that the sample that we observe empirically does not feature major wars or natural disasters, it is not surprising that the models all predict higher skewness than what we have observed: our sample simply does not contain a major disaster (though it certainly has what might be called a minor disaster in the financial crisis).

At higher maturities, the models all replicate the empirical observation that the volatility, skewness, and kurtosis of returns fall rapidly after the first month.

To summarize, based on the ability to generate a term structure steep enough at the short end and flat enough at higher maturities, we can reject the long-run risk, time-varying disaster, and disasters plus habits models with p-values of less than 1%, while the time-varying recovery model is not rejected. We thus take the results in Fig. 10 and Table 6 as providing further support for Gabaix's model of time-varying recovery rates.

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 where agents have preferences for early resolution of uncertainty, investors will pay to hedge shocks to expected future consumption volatility, especially at long horizons. If the equity market is modeled as being related to a consumption claim, then long-term forward variance claims should have large negative returns because they hedge volatility news. But in the data, we observe shocks to future expected volatility and find that their price has been close to zero.

While it is true that there exist parameterizations of Epstein–Zin preferences for which agents are *not* averse to bad news about future expected volatility, or even *enjoy* news about high future volatility, these are degenerate or nonstandard cases. The very motivation behind using Epstein–Zin preferences in asset pricing models is to model investors who are averse to bad news about the future, i.e., agents that have an intertemporal hedging motive. It is that force, generated by standard calibration of Epstein–Zin preferences, with preference for early resolution of uncertainty, that is at odds with the term structure of variance swaps.

closer to the model's prediction is an interesting question that only time can answer, and we leave for future research.

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

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, Eichengreen, Klingebiel, and Martinez-Peria (2001).

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. The presence of jumps in returns (due to the occurrence of a disaster) induces skewness in returns for the one-month variance swap and is the main reason that investors pay a high risk premium for short-term variance claims in this model.

5.3. The historical 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.[30]

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 (including, for the US, the Great Depression). 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 (and from CRSP since 1926 for the US). We compute realized volatility in each month for each country. To identify disasters, 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, Eichengreen, Klingebiel, and Martinez-Peria (2001) as financial crises.[31] 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% during disasters and 15% 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 Fig. 12 we perform an event study around the peak of volatility during a disaster. For

[30] In particular, see Man Group's TailProtect product (Inc., 2014), Deutsche Bank's ELVIS product (Deutsche Bank, 2010) and the JP Morgan Macro Hedge index.

[31] See Giglio, Maggiori, and Stroebel (2015) for a more detailed description of the data sources.

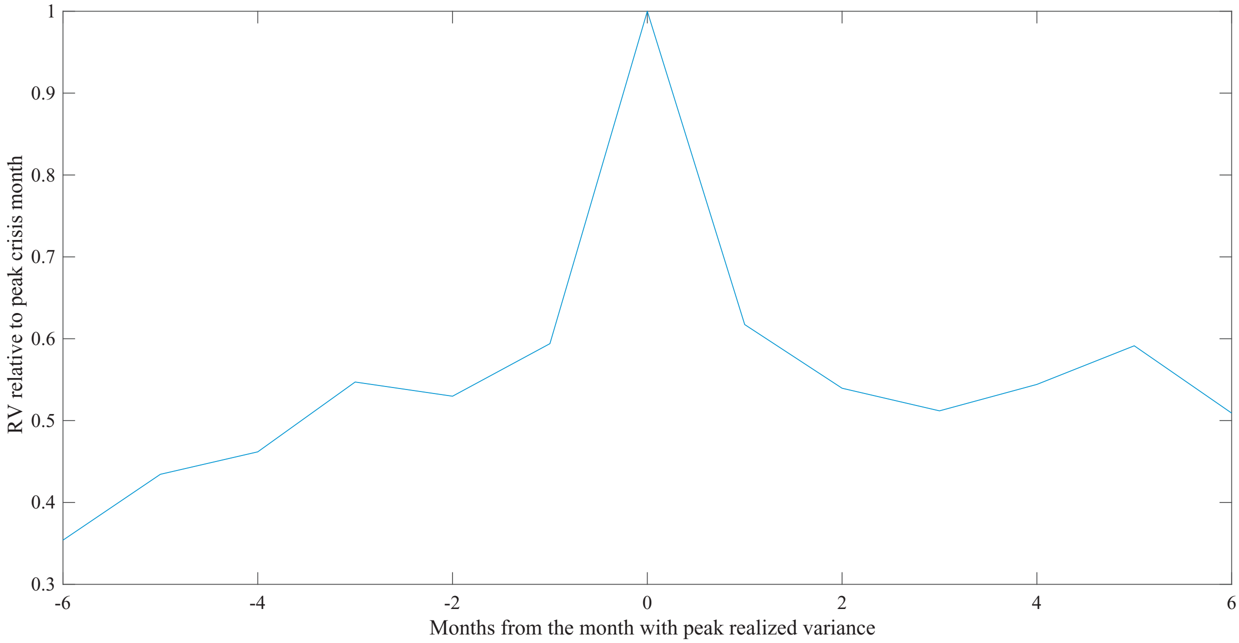


Fig. 12. Average behavior of RV during consumption disasters and financial crises. We calculate realized variance in each month of a crisis and scale it by the maximum realized variance in each crisis. The figure plots the average of that scaled series for each country and crisis in terms of months relative to the one with the highest realized variance.

each country and for each disaster episode, we identify the month of the volatility peak during that crisis (month 0) and the six 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 one 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% lower than its peak, and it is lower by half or more both six months before and after the worst month.

6. Conclusion

This paper shows that it is only the transitory part of realized variance that was priced in the period 1996–2014. That fact is inconsistent with a broad range of structural asset pricing models. It is qualitatively consistent with a model in which investors desire to hedge rare disasters, but not news about the future probability of disaster. 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 in 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 in volatility markets has been 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

Ait-Sahalia, Y., Karaman, M. Mancini, L., 2014. The term structure of variance swaps, risk premia and the expectations hypothesis. Working paper.

Ait-Sahalia, Y., Mancini, L., 2008. Out of sample forecasts of quadratic variation. Journal of Econometrics 147, 17–33.

Amengual, D. Xiu, D., 2014. Resolution of policy uncertainty and sudden declines in volatility. Working paper.

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Andersen, T. G. Bondarenko, O., 2007. Construction and interpretation of model-free implied volatility. Working paper.

Andries, M., Eisenbach, T. Schmalz, M., 2015. Asset pricing with horizon dependent risk aversion. Working paper.

Backus, D., Chernov, M., Martin, I., 2011. Disasters implied by equity index options. Journal of Finance 66 (6), 1969–2012.

Baker, S. R., Bloom, N. Davis, S. J., 2014. Measuring economic policy uncertainty. Working Paper.

Bakshi, G., Kapadia, N., 2003. Delta-hedge gains and the negative market volatility risk premium. The Review of Financial Studies 16 (2), 527–566.

Bakshi, G., Panayotov, G., Skoulakis, G., 2011. Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios. Journal of Financial Economics 100 (3), 475–495.
Bansal, R., Kiku, D., Shaliastovich, I., Yaron, A., 2013. Volatility, the macroeconomy and asset prices. Journal of Finance 69 (6), 2471–2511.
Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59 (4), 1481–1509.
Bao, J., Pan, J., Wang, J., 2011. The illiquidity of corporate bonds. The Journal of Finance 66 (3), 911–946.
Barras, L. Malkhozov, A., 2014. Variance risk premium dynamics in equity and option markets. Working paper.
Barro, R.J., 2006. Rare disasters and asset markets in the twentieth century. Quarterly Journal of Economics 121 (3), 823–866.
Berger, D., Dew-Becker, I. Giglio, S., 2016. Contractionary volatility or volatile contractions?Working paper.
van Binsbergen, J.H., Brandt, M.W., Koijen, R., 2012. On the timing and pricing of dividends. American Economic Review 102 (4), 1596–1618.
van Binsbergen, J. H. Koijen, R. S., 2015. The term structure of returns: facts and theory. Working paper.
Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77 (3), 623–685.
Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I. Terry, S. J., 2014. Really uncertain business cycles. Working paper.
Bollerslev, T., Todorov, V., 2011. Tails, fears, and risk premia. Journal of Finance 66 (6), 2165–2211.
Bordo, M., Eichengreen, B., Klingebiel, D., Martinez-Peria, M.S., 2001. Is the crisis problem growing more severe? Economic policy 16 (32), 51–82.
Branger, N. Volkert, C., 2010. What is the equilibrium price of variance risk? a long-run risk model with two volatility factors. Working paper.
Broadie, M., Chernov, M., Johannes, M., 2009. Understanding index option returns. The Review of Financial Studies 22 (11), 4493–4529.
Brownlees, C., Engle, R., Kelly, B., 2011. A practical guide to volatility forecasting through calm and storm. Journal of Risk 14 (2), 1–20.
Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107 (2), 205–251.
Campbell, J. Y., Giglio, S., Polk, C. Turley, R., 2013. An intertemporal capm with stochastic volatility. Working paper.
Campbell, J.Y., Shiller, R.J., 1991. Yield spreads and interest rate movements: a bird's eye view. Review of Economic Studies 58 (3), 495–514.
Carr, P., Wu, L., 2009. Variance risk premiums. Review of Financial Studies 22 (3), 1311–1341.
Christiano, L.J., Motto, R., Rostagno, M., 2014. Risk shocks. American Economic Review 104 (1), 27–65.
Christoffersen, P., Du, D. Elkamhi, R., 2015. Rare disasters and credit market puzzles. Rotman School of Management. Working paper.
Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with long-run and short-run volatility components. Journal of Financial Economics 90 (3), 272–297.
Cieslak, A. Povala, P., 2014. Information in the term structure of yield curve volatility. Working paper.
Coval, J.D., Shumway, T., 2001. Expected option returns. The Journal of Finance 56 (3), 983–1009.
Cox, J.C., Jonathan E. Ingersoll, J., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.
Dew-Becker, I. Giglio, S., 2013. Asset pricing in the frequency domain: theory and empirics. Working paper.
Drechsler, I., Yaron, A., 2011. What's vol got to do with it? The Review of Financial Studies 24 (1), 1–45.
Du, D., 2011. General equilibrium pricing of options with habit formation and event risks. Journal of Financial Economics 99 (2), 400–426.
Duffee, G. R., 2011. Sharpe ratios in term structure models. Working paper.
Egloff, D., Leippold, M., Wu, L., 2010. The term structure of variance swap rates and optimal variance swap investments. Journal of Financial and Quantitative Analysis 45 (5), 1279–1310.
Epstein, L.G., Zin, S.E., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis. The Journal of Political Economy 99 (2), 263–286.
Eraker, B. Wu, Y., 2016. Explaining the negative returns to volatility claims: an equilibrium approach. Working paper.
Fama, E.F., Bliss, R.R., 1987. The information in long-maturity forward rates. The American Economic Review 77 (4), 680–692.
Federico Bandi, J.R., Yang, C., 2008. Realized volatility forecasting and option pricing. Journal of Econometrics 147 (1), 34–46.
Fernandez-Villaverde, J., Guerron, P., Rubio-Ramirez, J.F., Uribe, M., 2011. Risk matters: the real effects of volatility shocks. American Economic Review 101 (6), 2530–2561.
Filipovic, D., Gourier, E., Mancini, L., 2013. Quadratic variance swap models. Swiss Finance Institute. Research Paper.
Gabaix, X., 2012. Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. Quarterly Journal of Economics 127 (2), 645–700.
Gabaix, X., Krishnamurthy, A., Vigneron, O., 2007. Limits of arbitrage: theory and evidence from the mortgage-backed securities market. Journal of Finance 62 (2), 557–595.
George, T.J., Longstaff, F.A., 1993. Bid-ask spreads and trading activity in the s&p 100 index options market. Journal of Financial and Quantitative Analysis 28 (3), 381–397.
Giglio, S., Maggiori, M., Stroebel, J., 2015. Very long-run discount rates. Quarterly Journal of Economics 130 (1), 1–53.
Gorton, G.B., Hayashi, F., Rouwenhorst, K.G., 2013. The fundamentals of commodity futures returns. Review of Finance 17 (1), 35–105.
Gourio, F., 2012. Disaster risk and business cycles. American Economic Review 102 (6), 2734–2766.
Gourio, F., 2013. Credit risk and disaster risk. American Economic Journal: Macroeconomics 5 (3), 1–34.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. The Journal of Finance 55 (3), 1263–1295.
Inc., M. I., 2014. The case for tailprotect strategy. https://www.man.com/US/the-case-for-tailprotect-strategy.
Jackwerth, J.C., Rubinstein, M., 1996. Recovering probability distributions from option prices. The Journal of Finance 51 (5), 1611–1631.
Jiang, G.J., Tian, Y.S., 2005. The model-free implied volatility and its information content. Review of Financial Studies 18 (4), 1305–1342.
Johnson, T. L., 2016. Risk premia and the vix term structure. Working paper.
Kelly, B., Pastor, L. Veronesi, P., 2014. The price of political uncertainty: theory and evidence from the option market. Working paper.
Kraus, A., Litzenberger, R.H., 1976. Skewness preference and the valuation of risk assets. The Journal of Finance 31 (4), 1085–1100.
Lu, Z., Zhu, Y., 2010. Volatility components: the term structure dynamics of vix futures. The Journal of Futures Markets 30 (3), 230–256.
Martin, I., 2014. What is the expected return on the market? Working paper.
Martin, I., 2015. Simple variance swaps. Working paper.
Mencia, J., Sentana, E., 2013. Valuation of vix deivatives. Journal of Financial Economics 108 (2), 367–391.
Menzly, L., Santos, T., Veronesi, P., 2004. Understanding predictability. Journal of Political Economy 112 (1), 1–47.
Merton, R.C., 1973. An intertemporal capital asset pricing model. Econometrica 41 (5), 867–887.
Mixon, S., Onur, E., 2015. Volatility Derivatives in Practice: Activity and Impact. Working paper.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Reinhart, C.M., Rogoff, K.S., 2009. This Time is Different: Eight Centuries of Financial Folly. Princeton University Press, Princeton, NJ.
Rietz, T.A., 1988. The equity risk premium: a solution. Journal of Monetary Economics 22 (1), 117–131.
Schneider, P., Wagner, C., Zechner, J., 2015. Low Risk Anomalies? Working paper.
Schularick, M., Taylor, A.M., 2012. Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870–2008. American Economic Review 102 (2), 1029–1061.
Segal, G., Shaliastovich, I., Yaron, A., 2015. Good and bad uncertainty: macroeconomic and financial market implications. Journal of Financial Economics 117 (2), 369–397.
Seo, S.B., Wachter, J.A., 2015. Option prices in a model with stochastic disaster risk. Tech. rep.. National Bureau of Economic Research.
Deutsche Bank Pension Strategies & Solutions, 2010. Tail Risk Hedging: A Roadmap for Asset Owners. Deutsche Bank Pension Strategies & Solutions.
Vijh, A.M., 1990. Liquidity of the CBOE equity options. Journal of Finance 45 (4), 1157–1179.
Wachter, J.A., 2013. Can time-varying risk of rare disasters explain aggregate stock market volatility? Journal of Finance 68 (3), 987–1035.
Zhou, G., Zhu, Y., 2012. Volatility trading: what is the role of the long-run volatility component? Journal of FInancial and Quantitative Analysis 47 (2), 273–307.