Variance Discount Rates: What Drives Preferences over Variance Risk?

Joren Koëter*

December 23, 2022

Abstract

I study time-variation in variance discount rates, defined as the expected returns for investing directly into stock market variance. Using a present value identity for S&P 500 variance swap prices, I document a strong term structure in the decomposition of variance swap price variation. Short-term variance swap prices mostly vary due to variation in expected stock market variance, whereas long-term variance swap prices are predominantly driven by variance discount rates. In contrast, prominent asset pricing models, which feature variance risk to capture time-variation in the equity premium, predict a flat term structure in the decomposition of variance swap price variation.

Keywords: Asset pricing, derivatives, variance pricing, variance discount rates.

JEL codes: G12, G13.

*Department of Finance at Rotterdam School of Management, Erasmus University. Email: koeter@rsm.nl.
I thank Lieven Baele, Martijn Boons, Fabio Braggion, Nicole Branger, Adrian Buss, Stefano Cassella, Julio Crego, Joost Driessen, Sebastian Ebert, Rik Frehen, Mohammad Ghaderi (discussant), Jasmin Gider, Stefano Giglio, Frank de Jong, Sicong (Allen) Li (discussant), Friedrich Lorenz, Shane Miller (discussant), Xuewu Wang (discussant), and Ole Wilms, and seminar/conference participants at MFA 2022, SFA 2021, Trans-Atlantic Doctoral Conference 2021, Derivative Markets Conference 2021, Paris-Dauphine University, Stockholm Business School and Rotterdam School of Management, Erasmus University for helpful comments and discussions on the topic.

1 Introduction

In this paper, I analyze time-series variation in expected returns for investing in stock market variance, which are called variance discount rates. More specifically, I analyze variation in variance discount rates in the S&P 500 variance swap market, and show that it drives a substantial fraction of the overall variation in prices in this market. I test whether prominent consumption-based asset pricing models are able to match the observed variation in variance discount rates, and therefore, I contribute to the recent literature that studies the predictions of these models with respect to derivative markets.[1]

The volume in the market for variance has increased significantly over the past decades; nevertheless, important properties of this market, like the time-variation in variance discount rates, have received little attention in the literature. One important reason to invest in stock market variance is that it can provide a hedge against stock market crashes. Furthermore, risk premia in this market are substantial,[2] which shows that investors worry about variance risk, and are willing to pay a premium to hedge these risks.

Variation in variance discount rates is either driven by variation in preferences over variance risk or by variation in the quantity of variance risk. Therefore, analyzing variance discount rates offers important insights into preferences of investors over variance risk and into how this risk varies over time. Many prominent asset pricing models feature variance risk as the main mechanism to have variation in the equity premium, and the variance market allows for a direct test of this mechanism. In the models considered in this paper,[3] variance risk is driven by consumption disasters or large and sudden movements in the agent's investment opportunity set. However, the literature has not settled on what the appropriate way is to incorporate variance risk in an asset pricing model. Analyzing the variation of variance discount rates allows me to contrast the empirical findings on variance discount rates with the predictions of these asset pricing models. Therefore, this analysis offers important insights into how variance risk should be incorporated in asset pricing models. In sum, my contribution to the literature is twofold.

[1] Many consumption-based asset pricing models (i.e. habit model by Campbell and Cochrane 1999, consumption disaster model by Rietz 1988, Barro 2006, Gabaix 2012 and Wachter 2013 and the long-run risk model by Bansal and Yaron 2004) are able to capture various moments of the equity premium. Therefore, many recent papers focus on derivative markets to discuss new predictions of these models, e.g. Bollerslev et al. (2009) discuss a long-run risk model that matches the variance premium in S&P 500 options, Dew-Becker et al. (2017) test which model is able to match the term structure of variance premia in S&P 500 variance swaps, Seo and Wachter (2018) and Seo and Wachter (2019) show that a time-varying disaster model matches the spreads on an index of credit default swaps and the implied volatility slope on S&P 500 options, respectively, and Bekaert and Engstrom (2017) and Bekaert et al. (2020) discuss habit models that are able to explain the relation between variance premium embedded in equity index options and consumption growth uncertainty.

[2] In the literature this phenomenon is referred to as the variance risk premium and documented in Bollerslev et al. (2009), Kozhan et al. (2013), Dew-Becker et al. (2017), and Aït-Sahalia et al. (2020).

[3] I study variance risk in the variable rare disaster model by Gabaix (2012), the time-varying disaster model by Wachter (2013), and the long-run risk model by Drechsler and Yaron (2011).

First, I show that variation in variance discount rates is an important determinant of variation in S&P 500 variance swap prices. Variance swaps are assets that pay their owner the sum of daily squared stock market returns and, thus, give a direct exposure toward stock market variance.[4] Variance discount rates are defined as the expected returns on variance swaps and equal the premium an investor pays to get exposure toward variance risk. I derive a log-linear pricing identity for variance swap prices in the spirit of Campbell and Shiller (1988a) for equity. The interpretation of this identity is as follows: Variation in variance swap prices is solely driven by variation in expected stock market variance and variation in variance discount rates. Using predictive regressions, I document a strong term structure in the decomposition of variance swap prices. Variation in short-term variance swaps prices (with maturity up to three months) is for at least 95.7% driven by variation in expected stock market variance, whereas variance discount rates account for at most 4.0% of the variation. In contrast, long-term variance swaps (with maturity up to 18 months) are for at most 72.8% driven by variance discount rates, and at least 24.5% of the remaining variation is attributed to variation in expected stock market variance. Therefore, I find strong evidence that not only expected stock market variance is time-varying, but also the premium associated with hedging variance risk. My results show that variation in short- and long-term variance discount rates is sizable, whereas the variation in expected short-term stock market variance is considerably larger than the variation in expected long-term stock market variance.

Second, I show that, in contrast to the data, prominent asset pricing models predict a more or less flat term structure in the decomposition of variance swap prices. In particular, the model by Gabaix (2012) predicts that all variation in variance swap prices is attributed to variation in variance discount rates, both for short-term and long-term variance swaps. The model by Wachter (2013) predicts that (almost) all variation in short-term and long-term variance swap prices is due to expected stock market variance. These results follow from the fact that the model by Wachter (2013) predicts a strong persistence in stock market variance, whereas the model by Gabaix (2012) predicts that, in absence of disasters, stock market variance is constant. Finally, the model by Drechsler and Yaron (2011) matches the empirical decomposition of long-term variance swap prices relatively well, and also, this model is able to match the empirical variation in variance discount rates. However, this model predicts that the variation in short-term variance swaps due to variance discount rates is substantially larger than documented in the data.

Besides an assessment of the models with respect to the decomposition of variance swap prices, I also analyze the predictions of the models with respect to the term structure of expected returns and return volatilities on variance swaps. I show that all of the analyzed

[4] The S&P 500 is a good, and often used, proxy for the U.S. stock market. Therefore, in the following I will refer to S&P 500 return variance as stock market variance.

models severely overestimate the one-month return volatility on variance swaps, an indication that the current calibrations are pretty extreme. However, despite this large degree of risk of the one-month contract, the models by Wachter (2013) and Drechsler and Yaron (2011) do not match the empirical risk premium on one-month variance swaps, and this indicates that the short-term risks which are hedged using these contracts are not sufficiently severe such that investors pay a premium similar to the data. All three models are able to match the expected returns on long-term variance swaps, and the models by Gabaix (2012) and Wachter (2013) understate the return volatility of long-term variance swaps whereas the model by Drechsler and Yaron (2011) overstates this volatility. However, this discrepancy between the model predictions and the data is much smaller compared to the discrepancy for the return volatility on short-term variance swaps. In sum, these models incorporate variance risk to capture empirical features of the equity premium; however, the models fail to directly match empirically documented features with respect to the pricing of variance risk.

I now explain the methodology to decompose variation in variance swap prices in more detail. In the spirit of the approach by Campbell and Shiller (1988a) for equity, I derive a present value identity for variance swap prices. More specifically, I show that the variance swap price with T months to maturity $vs_t^{(T)}$ approximately equals the difference between expected stock market variance $E_{\mathrm{rv},t}^{(T)}$ and variance discount rates $E_{\mathrm{vdr},t}^{(T)}$, as follows:

$$vs_t^{(T)} \approx E_{\mathrm{rv},t}^{(T)} - E_{\mathrm{vdr},t}^{(T)}, \tag{1}$$

where the full derivation and definitions are given in Section 2.2. Equation (1) offers a key insight into the pricing of variance risk, as it shows that the variance swap price is high, either due to high expected stock market variance or low variance discount rates (or both). Moreover, this identity for the variance swap price is similar to the pricing identity of the price-dividend ratio, which is presented in Campbell and Shiller (1988a), where variance expectations take the role of expected dividend growth and variance discount rates replace stock market discount rates. Similar to the analysis of Cochrane (2008) and Cochrane (2011) for equity, I decompose price variation in variance swap using predictive regressions in which the current variance swap price is used to predict stock market variance or variance swap returns. Variance swaps have the advantage of being finite cash flows, and therefore the future price does not play a role in these predictive regressions, whereas in the case of equity the future price is important. Moreover, I show that the results of the predictive regressions are robust to specifying a vector autoregression (VAR) to obtain expectations, and decompose variance swap prices in this way. Using the variance discount rates obtained in the VAR, I document the following about time-series variation in variance discount rates.

I show that the ex ante measure of variance discount rates obtained in the VAR predicts economically sizable variation in realized returns on variance swaps. Simple regressions show that a one-standard deviation increase in variance discount rates predicts an increase of 12.3% in average realized returns on a one-month variance swap, and this monotonically decreases to an increase 4.1% in average realized returns on an 18-month variance swap. These results indicate that there are periods during which investors are willing to pay a larger premium to hedge variance risk than during others. In particular, hedging variance risk was expensive in the period prior to the crash in 2000, and in the period post the crash of 2008, whereas in the period prior to the crash in 2008 it was relatively cheap to hedge variance risk. Moreover, the analysis of the term structure of variance discount rates shows that short- and long-term variance risk premia move in accordance, and this is supported by the result that the variation in short-term and long-term variance discount rates relate to several important macro-economic variables. Except for the one-month variance discount rate, which tends to go down during periods of rapidly increasing stock market variance, a result in line with Cheng (2019) who documents a similar finding for one-month VIX-futures.[5] I show that this negative correlation is a unique feature of the one-month variance risk premium, and this negative correlation disappears for long-term variance discount rates.

In the last part of this paper, I analyze the empirical drivers of the strong negative correlation between realized stock market and variance swap returns, also known as the leverage effect. In order to do so, I decompose returns on variance swaps into news about future variance and news about future variance discount rates and decompose stock market returns, following Campbell and Vuolteenaho (2004), into news about future dividends and news about future stock market discount rates. Because the correlation between variance swap and stock market returns is negative, it follows that the stock market beta of variance swaps is negative. First, I show that the stock market beta increases from -7.18 for short-term variance swaps to -2.21 for long-term variance swaps. Second, the low stock market beta is, both for short-term and long-term variance swaps, mostly explained by the positive correlation between news about stock market variance and news about stock market discount rates. This result indicates that during periods of increased stock market variance, stock market discount rates are revised upward to compensate for the increased risk of investing in the stock market. Third, I show that news about short-term variance discount rates is positively correlated to news about stock market discount rates, whereas news about long-term variance discount rates is negatively correlated to news about stock market discount rates.

Overall, I contribute to three strands of the literature. First, I contribute to the literature on the variance risk premium. Bollerslev et al. (2009), Kozhan et al. (2013), Dew-Becker et al.

[5] Lochstoer and Muir (2019) confirm this finding of Cheng (2019), and show that a model in which an agent has slow moving beliefs over stock market volatility can reconcile the evidence of the negative correlation of stock market variance and the risk premium. Slow moving volatility expectations lead the agent to initially underreact to volatility news followed by a delayed overreaction.

(2017), Eraker and Wu (2017), and Aït-Sahalia et al. (2020) show that this risk premium is, on average, sizable. Moreover, Bollerslev et al. (2009) find that the variance premium predicts future stock market returns, Bollerslev and Todorov (2011) show that a substantial fraction of the equity premium is compensation for variance risk, Martin (2017) shows that the risk-neutral variance is a lower bound for the equity premium, Johnson (2017) finds that there is predictability in the variance swap market, and Giglio and Kelly (2018) document variation in long-term variance forwards that exceeds the variation implied from short-term variance forwards. Second, I contribute to the literature on asset pricing models. Rietz (1988) and Barro (2006) propose that the equity premium is a compensation for consumption disasters. Gabaix (2012) and Wachter (2013) show how time-variation in consumption disaster risk is able to capture time-variation in the equity premium. On the other hand, Drechsler and Yaron (2011) generalize the long-run risk model by Bansal and Yaron (2004) to show that the equity premium and variance risk premium is explained by a long-run risk model that incorporates jumps in the investment opportunity set, rather than in the consumption growth process. Third, I contribute to the literature on discount rates in various financial markets of which the importance is stressed in Cochrane (2011). Shiller (1981), Campbell and Shiller (1988a), Campbell and Shiller (1988b), Cochrane (1992) and Cochrane (2008) present evidence in favor of variation in stock market discount rates. Variation in discount rates for bonds and corporate credits is analyzed in Shiller (1979), Fama and Bliss (1987), Campbell and Ammer (1993), and Nozawa (2017).

The remainder of this article is organized as follows. In Section 2, I derive the pricing identity for variance swap prices. Section 3 describes the data on variance swaps. The results of the decomposition are presented in Section 4, and afterward compared to the predictions of asset pricing models. I show in Section 5 that the results are robust to specifying a VAR to obtain expectations, and in Section 6 analyzes the time-series variation in variance discount rates. Section 7 concludes.

2 Methodology

In this section, I present the methodology to decompose S&P 500 variance swap prices into expected S&P 500 return variance and variance discount rates. A variance swap is a derivative security that pays the holder of the contract the realized variance of the underlying up to maturity. Variance swaps are used to manage market variance risk, and the variance discount rates correspond to expected returns for holding a variance swap.

In the next subsection, I formalize the cash flows of a variance swap and will afterward derive a log-linear pricing identity for the variance swap price in the spirit of the approach by Campbell and Shiller (1988a) for equity.

2.1 Variance swap contract

A variance swap pays its holder the realized variance of the underlying from the inception of the contract up to maturity. At maturity, the holder of the contract pays the variance swap price in exchange for the realized variance of the underlying from inception up to maturity. The payoff of a variance swap at maturity T-periods from origination time t is defined as follows:

$$\text{payoff}_{t+T} = \sum_{i=1}^{T} RV_{t+i} - VS_t^{(T)}, \tag{2}$$

where RV_{t+i} is the realized variance over period $t+i$ and $VS_t^{(T)}$ is the variance swap price at time t for a variance swap with T-periods to maturity. A variance swap can last for several periods, and, therefore, the total realized variance at the end of the contract equals the sum of the realized variance over each period. The holder of the variance swap receives the realized variance in exchange for a fixed rate at the end of the contract and, therefore, hedges variance risk until maturity of the contract. Given a risk-neutral pricing measure $\mathbb{Q}$, the variance swap price with T-periods to maturity at time t is given by the following:

$$VS_t^{(T)} = \sum_{i=1}^{T} \mathbb{E}_t^{\mathbb{Q}}\big(RV_{t+i}\big), \tag{3}$$

where $\mathbb{E}_t^{\mathbb{Q}}$ denotes the expectation under the risk-neutral measure conditional on information available at time t. Therefore, $VS_t^{(T)}$ corresponds to the risk-neutral expectation of the sum of realized variances from period $t+1$ until period $t+T$.

Next, I define the gross return over period t to $t+1$ on a variance swap with T-periods to maturity, as follows:

$$R_{t+1}^{(T)} = \frac{VS_{t+1}^{(T-1)} + RV_{t+1}}{VS_t^{(T)}}. \tag{4}$$

The gross return is as if the variance swap is bought for the current variance swap rate $VS_t^{(T)}$, and held for one period after which the one period realized variance RV_{t+1} and next period's variance swap price $VS_{t+1}^{(T-1)}$ are received. In the following period, the variance swap has $T-1$-periods remaining to maturity.[6] It follows from equation (4) that realized variance for returns on variance swaps is the equivalent to dividend payments for returns on the stock market; that is, realized variance is the cash flow component of a variance swap.

[6]The definition of the return in equation (4) corresponds to the return on a variance asset, which pays the realized variance at the end of each period rather than at the end of the contract. Under the assumption of no-arbitrage, the price of such an asset equals the variance swap price discounted with the T-period risk-free rate to time t, and the realized variance payment of such an asset is discounted in a similar way. It is possible to show that the logarithm of the return defined in (4) equals the log-return on this variance asset in excess of the risk-free rate.

2.2 Pricing identity for the variance swap price

In this section, I show how to derive the present value identity for variance swaps of equation (1). In this identity, the current variance swap price equals future expected stock market variance, and variance discount rates.

The first step in deriving the identity is to Taylor-expand the log-return on the variance swap:

$$r_{t+1}^{(T)} \approx k(T) + \rho(T) \cdot vs_{t+1}^{(T-1)} + \big[1 - \rho(T)\big] rv_{t+1} - vs_t^{(T)}, \tag{5}$$

where $r_{t+1}^{(T)} = \log\big(R_{t+1}^{(T)}\big)$, $vs_{t+1}^{(T-1)} = \log\big(VS_{t+1}^{(T-1)}\big)$, $rv_{t+1} = \log\big(RV_{t+1}\big)$, and $vs_t^{(T)} = \log\big(VS_t^{(T)}\big)$. In the following, unless stated otherwise, I refer to the logarithm of these variables as the return, variance swap price, and realized variance. $k(T)$ is the approximation constant and $\rho(T)$ governs the relative importance of the next period's price and the next period's realized variance in the calculation of the return of the variance swap. These constants depend on the maturity T of the variance swap, because the one-period return on a short-term variance swap is mostly driven by the next period's realized, whereas the one-period return on a long-term variance swap is mostly driven by the next period's price. Intuitively, $\rho(T)$ increases in the maturity of the variance swap, and thus, gets closer to one. In Section 4.1, I show that $\rho(T)$ can be estimated using a simple regression and, importantly, that equation (5) approximates the return on a variance swap really well.

The next step is to substitute the next period's variance swap price with the following approximation:

$$vs_{t+i}^{(T-i)} \approx k(T-i) + \rho(T-i) \cdot vs_{t+i+1}^{(T-i-1)} + \big[1 - \rho(T-i)\big] rv_{t+i+1} - r_{t+i+1}^{(T-i)},$$

where $k(T-i)$ and $\rho(T-i)$ are the log-linearization coefficients of a variance swap with $T-i$-periods to maturity. This equation allows me to substitute future variance swap prices up to the point that the current variance swap price depends on the one-month variance swap price $vs_{t+T-1}^{(1)}$. The following holds regarding the one-month variance swap price: $vs_{t+T-1}^{(1)} = rv_{t+T} - r_{t+T}^{(1)}$, i.e. $k(1) = \rho(1) = 0$. After this substitution, I obtain the following:

$$vs_t^{(T)} \approx K + \sum_{i=1}^{T} \big[1-\rho(T-i+1)\big] \Big(\prod_{j=1}^{i-1} \rho(T-j+1) \Big) \cdot rv_{t+i} - \sum_{i=1}^{T} \Big(\prod_{j=1}^{i-1} \rho(T-j+1) \Big) \cdot r_{t+i}^{(T-i+1)}, \tag{6}$$

where K is a constant and a function of the constants $k(T-i)$ and $\rho(T-i)$ from the individual log-linearizations. In the remainder of this paper, I will discard the constant K from the pricing identity because the focus of this paper is on time-series variation in variance swap prices.

The intuition of equation (6) is as follows: A high current variance swap price $vs_t^{(T)}$ is either due to high future realized stock market variance rv_{t+i} or low future returns on the variance swap $r_{t+i}^{(T-i+1)}$. Because equation (6) is an accounting identity, it also holds in expectation conditional on the information at time t, as follows:

$$vs_t^{(T)} \approx E_{\mathrm{rv},t}^{(T)} - E_{\mathrm{vdr},t}^{(T)}, \tag{7}$$

where

$$E_{\mathrm{rv},t}^{(T)} = \mathbb{E}_t \sum_{i=1}^{T} \left[1 - \rho(T-i+1)\right] \Big(\prod_{j=1}^{i-1} \rho(T-j+1) \Big) \cdot rv_{t+i} \quad \text{and} \tag{8}$$

$$E_{\mathrm{vdr},t}^{(T)} = \mathbb{E}_t \sum_{i=1}^{T} \Big(\prod_{j=1}^{i-1} \rho(T-j+1) \Big) \cdot r_{t+i}^{(T-i+1)}. \tag{9}$$

The intuition of equation (7) is similar to before: The current variance swap price $vs_t^{(T)}$ is high due to high expected stock market variance $E_{\mathrm{rv}}^{(T)}$ or low variance discount rates $E_{\mathrm{vdr}}^{(T)}$. Campbell and Shiller (1988a) show that the current price-dividend ratio increases in dividend growth expectations and decreases in stock market discount rates. Therefore, identity (7) for the variance swap price is similar to the pricing identity of the price-dividend ratio, where variance expectations take the role of expected cash flows and variance discount rates replace stock market discount rates. There are two main differences between the identity of equation (7) and the identity in Campbell and Shiller (1988a), and these are due to the fact that a variance swap is a finite cash flow, whereas equity is a perpetual cash flow. First, variance discount rates in equation (9) depend on the maturity of the variance swap, which is important because Dew-Becker et al. (2017) show that there is a strong term structure in risk premia on variance swaps. Dew-Becker et al. (2017) find that expected returns on short-term variance swaps are much lower than expected returns on long-term variance swaps. Second, in the derivation of the pricing identity for equity, Campbell and Shiller (1988a) assume the so-called no-bubble condition. This assumption is not needed in case of a variance swap, because it is a finite cash flow rather than a perpetual cash flow in the case of equity.

In Section 4, I take equation (7) to the data and decompose variance swap prices into variance expectations and variance discount rates for several maturities. This allows me to study the drivers of the variation in variance swap prices for various maturities.

3 Data

In this paper, I use data on S&P 500 options from January 1996 until June 2019 from OptionMetrics. Using the methodology described in Kozhan et al. (2013) and discussed more into details in Appendix A.1, I construct variance swaps with maturities ranging from one to 18 months. The maturity of 18 months to maturity is the longest for which I can calculate a variance swap price every month. I calculate variance swap prices at the end of each month in the sample and interpolate the variance swap prices linearly, such that the maturity equals T months. Note that interpolating variance swap prices linearly is equivalent to taking long positions in two variance swaps with maturity T_1 and T_2, such that the weighted average of the maturities equals T.

I use the methodology of Kozhan et al. (2013), because these variance swaps are most closely related to the variance swaps that are traded over-the-counter (OTC). In Appendix A.2, I show that my data on variance swaps is highly similar to data from the OTC market, which is analyzed in Dew-Becker et al. (2017). I show in Figure 6 of Appendix A.2 that the synthetic variance swap prices are highly similar to the variance swap prices in the OTC market, and simple regressions indicate R^2's above 97%. Table 6 in Appendix A.2 shows that also the returns on the synthetic variance swaps are highly similar to the returns on the OTC swaps as shown in Table 2 in Dew-Becker et al. (2017).

In addition to the pricing information on variance swaps, I obtain data that is used in the VAR. From the panel of variance swap prices, the first two principal components, $pc_t^{(1)}$ and $pc_t^{(2)}$, are calculated. $pc_t^{(1)}$ captures the level in the term structure of variance swap prices, and $pc_t^{(2)}$ captures the slope of the term structure of variance swap prices. Moreover, realized variance is defined as in Kozhan et al. (2013) and approximately equals the sum of daily squared returns within a month. Finally, the default spread is obtained from the Federal Reserve Bank of St. Louis and defined as the difference in yield on BAA and AAA credit-rated corporate bonds.

4 Results

In this section, I present the empirical term structure of the decomposition of variance swap prices, and relate the findings to the predictions of leading asset pricing models. First, I present the results of a variance decomposition on the basis of predictive regressions, and, these, serve as my benchmark results for the models. In section 5, I show that the results are robust to specifying a VAR, and decompose the variation in variance swap prices. Second, I consider the asset pricing models by: Gabaix (2012), Wachter (2013) and Drechsler and Yaron (2011), and repeat the empirical exercise in each of these models to analyze their predictions regarding the term structure of the decomposition.

4.1 Decomposition of variance swap prices in the data

In this section, I decompose variation in variance swap prices using predictive regressions. These regressions derive from the intuition of pricing identity (7): The current variance swap price is high either due to high future stock market variance or low future returns on the variance swap. Therefore, the current variance swap price should predict future stock market variance, future returns on the variance swap, or both. In order test this, I run predictive regressions of future stock market variance and future returns on the variance swap, with the current variance swap price as the predictor. This analysis is the equivalent of the analysis by Cochrane (2008) and Cochrane (2011) for the dividend-price ratio, where my analysis has the advantage that variance swaps have a finite maturity, and, thus, future prices do not play a role. By identity (7), the following relation between the regression coefficients should hold:

$$\sum_{i=1}^{T}\left[1-\rho(T-i+1)\right]\left(\prod_{j=1}^{i-1}\rho(T-j+1)\right)\cdot rv_{t+i} =: y_{\mathrm{rv},t+T} = a_{\mathrm{rv}} + b_{\mathrm{rv}}\cdot vs_t^{(T)} + \epsilon_{t+T}^{\mathrm{rv}}, \tag{10}$$

$$\sum_{i=1}^{T}\left(\prod_{j=1}^{i-1}\rho(T-j+1)\right)\cdot r_{t+i}^{(T-i+1)} =: y_{\mathrm{vdr},t+T} = a_{\mathrm{vdr}} + b_{\mathrm{vdr}}\cdot vs_t^{(T)} + \epsilon_{t+T}^{\mathrm{vdr}}, \tag{11}$$

$$1 \approx b_{\mathrm{rv}} - b_{\mathrm{vdr}}. \tag{12}$$

In equations (10) and (11), $y_{\mathrm{rv},t+T}$ and $y_{\mathrm{vdr},t+T}$ correspond to the realized variance and realized returns on the variance swap starting at time t up to maturity $t+T$. Equation (12) follows from the present value identity (7), and indicates whether variation in variance swap prices is driven by future realized variance or returns. The regression coefficients indicate whether a high current variance swap price predicts high future stock market variance or low future returns. Therefore, economic intuition suggests that $b_{\mathrm{rv}} > 0$ and $b_{\mathrm{vdr}} < 0$. Moreover, if $b_{\mathrm{rv}} \approx 1$, it indicates that variation in the variance swap price is exclusively driven by stock market variance expectations, whereas if $b_{\mathrm{vdr}} \approx -1$, it indicates that variation in the variance swap price is exclusively driven by variance discount rates. Due to the fact that a variance swap is a finite cash flow, these regressions provide a powerful test if the present value identity holds. This identity holds if the difference between the regression coefficients are, indeed, close to one. In the case of a perpetual cash flow like equity, the regressions (10) and 11 do not fully decompose the current price, because the future price of the asset is important.

In order to compute $y_{\mathrm{rv},t+T}$ and $y_{\mathrm{vdr},t+T}$, I need the log-linear approximation coefficients $\rho(T)$. I estimate $\rho(T)$ using a simple regression, the results of this exercise are presented in Table 1.

Table 1: This table shows the regression results of $r_{t+1}^{(T)} - rv_{t+1} + vs_t^{(T)} = k(T) + \rho(T) \cdot \left(vs_{t+1}^{(T-1)} - rv_{t+1}\right) + \epsilon_{t+1}^{(T)}$, for different maturities T. For each maturity, the log-linearization coefficient $\rho(T)$ is given as well as the R^2 of the regression.

Maturity	1	2	3	4	5	6	7	8	9
$\rho(T)$	0	0.634	0.777	0.841	0.876	0.898	0.914	0.925	0.933
R^2	100%	98.24%	99.28%	99.57%	99.70%	99.78%	99.82%	99.85%	99.88%
Maturity	10	11	12	13	14	15	16	17	18
$\rho(T)$	0.940	0.945	0.950	0.954	0.958	0.961	0.963	0.965	0.967
R^2	99.90%	99.91%	99.92%	99.94%	99.94%	99.95%	99.96%	99.96%	99.96%

As expected, the log-linearization coefficient $\rho(T)$ depends on the maturity of the variance swap. $\rho(T)$ increases in the maturity, which indicates that the next period's variance swap price is relatively more important than the one-period realized variance for long-term variance swaps. The second row in Table 1 shows that the log-linear approximation of the variance swap returns is in fact a very good approximation, as indicated by the large R^2's. The R^2's range from 98.24% to 99.96%, with an average of 99.7%.

The next step is to estimate the predictive regressions (10) and (11) which decompose the variance swap price into variance expectations and variance discount rates. Table 2 presents the results.

Table 2: This table shows the results of the predictive regressions of equations (10) and (11), in which the variance swap price is the independent variable. t-statistics are represented in brackets and are computed using Newey-West standard errors with number of lags equal to T.

Dependent variable:	$y_{\mathrm{rv},t+T}$		$y_{\mathrm{vdr},t+T}$	
Maturity	b_{rv} (t-stat.)	R^2	b_{vdr} (t-stat.)	R^2
18	0.245	0.030	-0.728	0.205
	(1.32)		(-3.87)	
12	0.558	0.142	-0.419	0.080
	(3.70)		(-2.65)	
6	0.833	0.315	-0.168	0.017
	(7.00)		(-1.38)	
3	0.957	0.432	-0.040	0.001
	(11.50)		(-0.46)	
1	1.101	0.577	0.101	0.013
	(19.52)		(1.79)	

Table 2 shows that there is a strong term structure in the decomposition of variance swap prices. Variation in short-term (maturity less than one year) variance swap prices is mostly attributed to variation in realized variance, as indicated by a coefficient b_{rv} between 0.833 and 1.101 which suggests that more than 83.3% of the variation is driven by variance expectations. The coefficient estimates of b_{vdr} for short-term variance swaps exceed -0.168, which indicates that at most 16.8% of the variation in short-term variance swap prices is driven by variance discount rates. Moreover, these short-term variance swap prices do not predict lower future variance discount rates with a coefficient significantly different from zero. However, in Section 5, where I show that the results are robust to specifying a VAR, I find that also short-term variance discount rates vary over time. The finding of a positive b_{vdr} to predict one-month variance swap returns indicates that a high one-month variance swap price predicts high future returns, a finding that is not in line with economic intuition. However, this result is in line with Cheng (2019), who shows that the variance risk premium tends to decline during market turmoil, and the result can be explained by a falling hedging demand for variance risk in those periods.

The predictive regressions for long-term (maturity of one year and beyond) variance swaps indicate that a much larger fraction of the total variation is due to variance discount rates. The coefficients b_{rv} for long-term variance swaps indicate that variance expectations account for, respectively, 55.8% and 24.5% of the total variation in variance swap prices with 12 or 18 months to maturity, whereas b_{vdr} indicate that variance discount rates account for 41.9% and 72.8% of the total variation. Long-term variance swap prices predict future returns on the variance swaps negatively and with a coefficient significantly different from zero, which indicates that high variance swap prices are mostly an indication of low variance discount rates. Furthermore, the 18-month variance swap price does not predict future realized variance over the lifetime of the contract with a coefficient significantly different from zero. In sum, these predictive regressions indicate a strong term structure in the decomposition of variance swap prices: Long-term variance swap prices mostly driven by variance discount rates, whereas short-term variance swap prices are mostly driven by stock market variance expectations.[7] Finally, in Appendix OA.2 I show that the results are robust to redoing the analysis at the quarterly rather than monthly frequency.

In Appendix A.4, I show that most of the variation in variance swap prices comes from variation in downside variance swap prices rather than upside variance swap prices.[8] The

[7] The result that $b_{\text{rv}} - b_{\text{vdr}}$ is, indeed, very close to one for each of the considered maturities indicates that the present value identity, which is an approximation, actually is a very good approximation for variance swap prices. The difference between the coefficients of the predictive regressions lies between 0.973 and 1.001.Therefore, the results show that variance swap prices are effectively decomposed into stock market variance expectations and variance discount rates.

[8] My definition of up- and downside variance swap prices closely follows the definitions of Andersen and Bondarenko (2009), Dew-Becker et al. (2017), Baele et al. (2019) and Kilic and Shaliastovich (2019) such that:

variance decomposition of Table 11 in Appendix A.4 shows that the variance swap prices are driven by downside variance swap prices for 63.4% to 74.2%, depending on the maturity of the contract. Also, the results suggest that long-term variance swaps are driven by a somewhat larger fraction by the downside variance swap than short-term variance swaps, however this term structure effect is relatively small.

4.2 Decomposition of variance swap prices in asset pricing models

In this subsection, I discuss the predictions of several prominent asset pricing models with respect to the pricing of variance risk. The ability of asset pricing models to match stylized facts of variance markets has been an important topic in the literature.[9] This is the case as variance risk is a central component of many asset pricing models, and variance markets allow researchers to study variation in variance risk in combination with preferences of investors over this risk. I discuss the implications of the following three models: the variable rare disaster model by Gabaix (2012), the time-varying rare disaster model by Wachter (2013), and the long-run risk model by Drechsler and Yaron (2011). These models are able to match some empirical results on the pricing of variance risk. Dew-Becker et al. (2017) show that the model by Gabaix (2012) matches the empirical results on the term structure of risk premia of variance risk, Seo and Wachter (2019) show that the model by Wachter (2013) matches the implied volatility slope on S&P 500 options, and the model by Drechsler and Yaron (2011) is designed to match the empirical results on the variance risk premium by Bollerslev et al. (2009).

In each of the considered models, I decompose the variation in variance swap prices using predictive regressions similar to the empirical exercise in Section 4.1. Afterward, I calculate the following moments in each model: the expected returns on variance swaps, the standard deviation of variance swap returns, and the variance of the variance swap prices. These results from the models are obtained from a simulation study.[10] Finally, in online appendices OA.4.1— OA.4.3 I discuss the models in more detail, and show that the analyzed moments are relatively stable across the simulation sets.

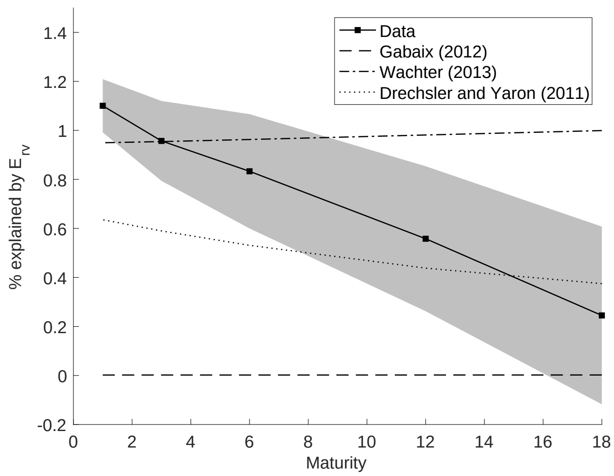
In the following, I show the results from the decomposition of variance swap prices in each of the considered models. Figure 1 plots the results.

$VS_t^{(T)} = VS_{u,t}^{(T)} + VS_{d,t}^{(T)}$, where $VS_{u,t}^{(T)}$ and $VS_{d,t}^{(T)}$ are the up- and downside variance swap, respectively. The formal definitions of these swaps are given in Appendix A.4.

[9] Bollerslev et al. (2009), Drechsler and Yaron (2011), Gabaix (2012), Bekaert and Engstrom (2017) and Lochstoer and Muir (2019) present consumption-based asset pricing models that match various moments of the variance premium. Seo and Wachter (2019) shows that an asset pricing models that features time-varying disaster risk matches the implied volatility slope on S&P 500 options. Dew-Becker et al. (2017) analyzes the ability of asset pricing models to match the term structure of variance risk premia.

[10] For each model, 1,000 independent simulation sets of a time-series with 1,000 data points are obtained. On the basis of these simulation sets, each of the considered statistics is calculated.

Figure 1: The left (right) figure plots how much of the variation in variance swap prices is driven by variance expectations (variance discount rates) in the data (solid line), the model by Gabaix (2012) (dashed line), the model by Wachter (2013) (dash-dotted line), and the model by Drechsler and Yaron (2011) (dotted line). The grey area corresponds to a 95% confidence interval. The results are plotted for variance swap prices with 1, 3, 6, 12, and 18 months to maturity. The y-axis corresponds to how much of the variation is attributed to variance expectations or variance discount rates in percentages.



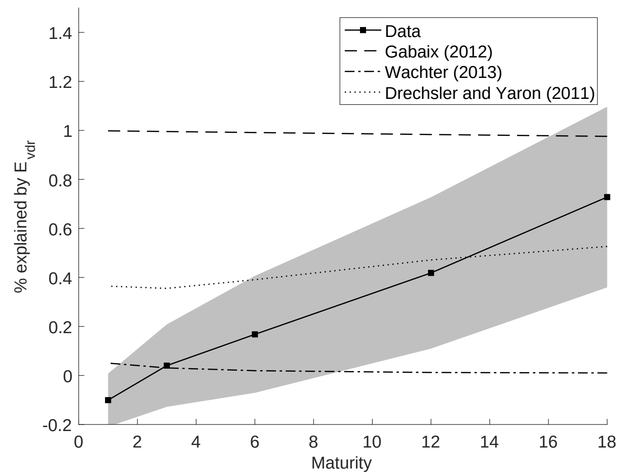


Figure 1 plots how much of the variation in variance swap prices is attributed to variance expectations (variance discount rates) in the left (right) graph. I show the results from the data (solid line) as well as the predictions of the models. Overall, the considered models struggle to produce the steep term structure in the decomposition of variance swap prices that is documented in the data. In the models by Wachter (2013) (dash-dotted line) and Drechsler and Yaron (2011) (dotted line) expected stock market variance drives a too large fraction of the variation in long-term variance swap prices. This is due to the strong persistence of state variables in the models, which makes stock market variance more persistent than in the data. A simple solution would be to decrease the persistence of the variables that drive stock market variance, i.e. the persistence of the disaster intensity in Wachter (2013) and the persistence in consumption volatility in Drechsler and Yaron (2011). However, the persistence of these state variables is the key mechanism of the models to match the equity and variance premium. Also the model by Gabaix (2012) (dashed line) is not able to match the observed term structure of the variance swap price decomposition. Instead, all the variation in variance swap prices, both short- and long-term, is driven by variance discount rates.

The main reason why the model by Gabaix (2012) is not in line with the data is that the time-variation in the disaster size only affects the realized variance, conditional on a disaster hitting the economy. Given that this probability is relatively low (1% p.a.), expected stock market variance only increases marginally when the disaster size increases. However, given that this consumption disaster is highly undesirable for the investor, variance discount rates adjust accordingly when the disaster size increases. As I will show in Figure 2, the model by Gabaix (2012) is not able to match the observed variation in variance swap prices, and, therefore, the overall variation in variance discount rates is still relatively small. In order to match the observed variation and empirical decomposition, the model has to incorporate an additional source of stochastic volatility.

The model by Wachter (2013) (dash-dotted line) predicts a large variation in stock market variance, because of the heteroskedastic nature of the disaster intensity process; that is high levels of the disaster intensity scale future variance of the disaster intensity upward. Therefore, stock market variance varies, even in the absence of disasters, over time. At the same time, high levels of the disaster intensity correspond to low variance discount rates; however, variation in variance discount rates drives a much smaller fraction in the total variation in variance swap prices, as seen in the right graph of Figure 1. Even if realizations in which the consumption disaster hits the economy are excluded, the variation in 18-month variance swap prices due to variance discount rates only increases to 6%. Therefore, even in the absence of consumption disasters, the model by Wachter (2013) is not able to match the data.

The model by Drechsler and Yaron (2011) (dotted line) matches the data most closely, as short-term variance swap prices are mostly driven by expected stock market variance, and long-term variance swap prices are mostly driven by variance discount rates. However, the term structure of variance price decomposition is considerably less steep compared to the data. For example, in the data the variation due to variance discount rates of the one-month variance swap price is close to zero (even negative), whereas in the model by Drechsler and Yaron (2011), variance discount rate variation accounts for 36% of the variation. The variation in short-term variance discount rates is considerably larger in the model by Drechsler and Yaron (2011), because variation in variance risk is sizable and it varies in a predictable way. At the same time, the model is not able to capture the empirical result that the variation in variance swap prices due to expected stock market variance strongly decreases in maturity. Again, this result is an indication that the state variables governing stock market variance are too persistent.

Figure 2: The figure plots the term structure of the variance of variance swap prices in the data (solid line), the model by Gabaix (2012) (dashed line), the model by Wachter (2013) (dash-dotted line), and the model by Drechsler and Yaron (2011) (dotted line). The grey area corresponds to a 95% confidence interval. The variance of variance swap prices with 1, 3, 6, 12, and 18 months to maturity. The y-axis corresponds to monthly volatility.

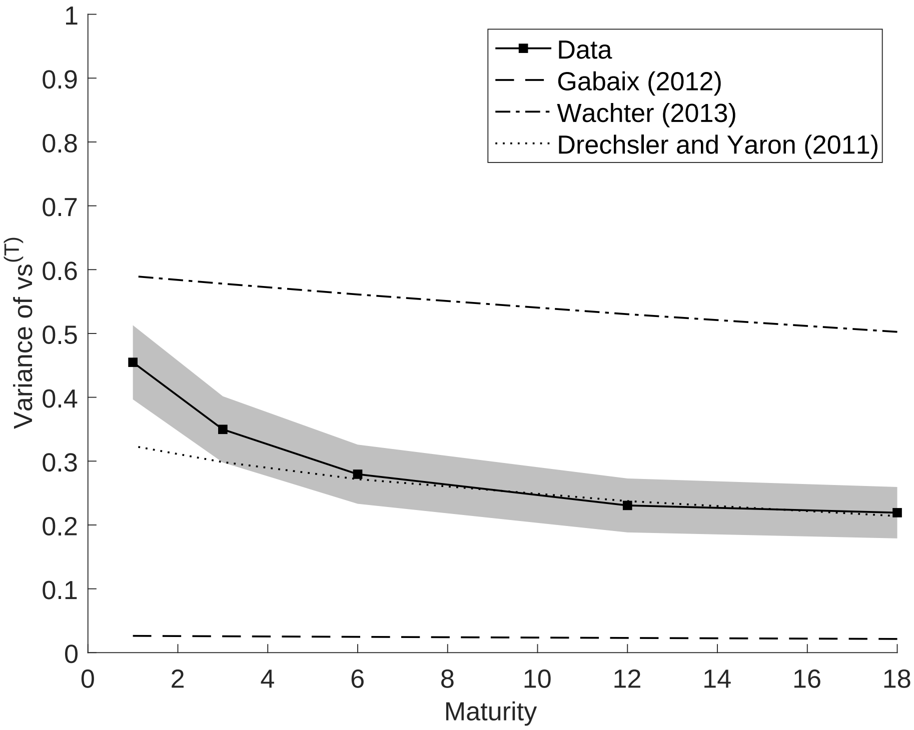


Figure 2 plots the term structure of the variance of variance swap prices in the data (solid line) and in each of the asset pricing models. The models are not able to produce the strong downward slope in the term structure, in particular the slope up to six months to maturity. However, in line with the data, the models by Wachter (2013) (dash-dotted line) and Drechsler and Yaron (2011) (dotted line) do produce a downward sloping term structure, whereas the model by Gabaix (2012) (dashed line) predicts a flat term structure of the variance of variance swap prices. The model by Drechsler and Yaron (2011) matches the observed variation in variance swap prices well, and in particular the variation in long-term variance swap prices. This is interesting, because the model was calibrated to match the dynamics of the one-month variance risk premium. The models by Wachter (2013) and Gabaix (2012) severely over- and understate the observed variation in variance swap prices. This variation is overstated in Wachter (2013), because the high persistence of the disaster intensity makes variation in expected stock market variance high. The variation is understated in Gabaix (2012), because the model produces relatively little variation in expected stock market variance.

Figure 3: The left (right) graph plots the term structure of expected returns (return volatility) on variance swaps in the data (solid line), the model by Gabaix (2012) (dashed line), the model by Wachter (2013) (dash-dotted line), and the model by Drechsler and Yaron (2011) (dotted line). The grey area corresponds to a 95% confidence interval. The results are shown for variance swaps with 1, 3, 6, 12, and 18 months to maturity. The y-axis corresponds to monthly returns.

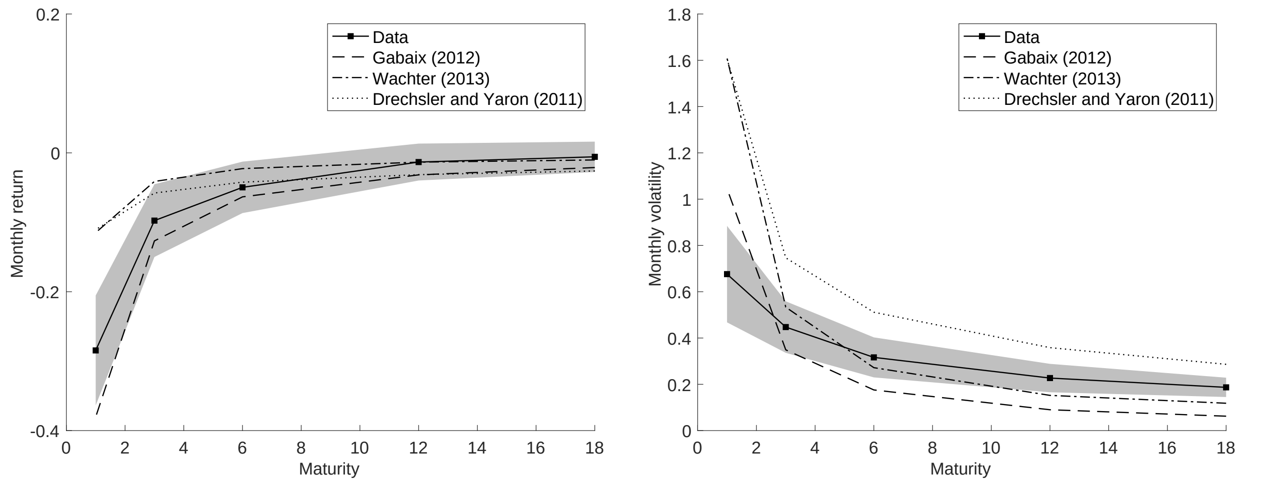


Figure 3 shows the term structure of expected returns (return volatility) on variance swaps in the left (right) graph. I show the results from the data (solid line) as well as the predictions of the models. As seen in the left graph, the model by Gabaix (2012) (dashed line) is able to match the strongly upward sloping term structure of expected returns in the data. However, the model overstates the return volatility of the one-month variance swap and understates the return volatility of the variance swap beyond three months to maturity. The models by Wachter (2013) (dash-dotted line) and Drechsler and Yaron (2011) (dotted line) are not able to match the strong upward sloping term structure of expected returns on variance swaps. In particular the expected return on the one-month variance swap is far off, and this indicates that this short-term risk is not sufficiently severe such that the investor is willing to pay a risk premium which is similar in magnitude as the data. The right graph of Figure 3 shows, however, that the one-month variance swap is sufficiently more risky in the models by Wachter (2013) and Drechsler and Yaron (2011) than in the data, but despite this large risk the models are not able to match the one-month risk premium. Moreover, the model by Wachter (2013) slightly understates the return volatility of long-term variance swaps, whereas the model by Drechsler and Yaron (2011) slightly overstates this return volatility.

In the model by Gabaix (2012), the expected returns in the model are only slightly lower compared to the data, and this could be resolved by adjusting the calibration, for example, decrease the average disaster size. The fact that the model by Gabaix (2012) is able to capture the term structure risk premia on the pricing of variance risk is in line with the findings of Dew-Becker et al. (2017). In the right graph, I show that the model by Gabaix (2012) (dashed line) predicts a stronger downward sloping term structure of return volatility on variance swaps than empirically observed. The model predicts a larger return volatility on variance swaps with one month to maturity and a lower return volatility for variance swaps with maturity beyond one month. An explanation for the large one-month return volatility is the severity of the consumption disaster (-30%) in the current calibration, which makes the return on the one-month variance swap, and thus the return volatility, (extremely) large, if a disaster hits the economy.

The main reason why the model by Wachter (2013) is not able to match the risk premium on the one-month variance swap is that the mean consumption disaster (-15%) in this model is much smaller than in the calibration of the model by Gabaix (2012) (-30%). In the calibration by Wachter (2013), the mean disaster intensity (3.55% p.a.) is increased; however the left graph of Figure 3 shows that increasing the mean disaster size has a much larger effect on the expected returns of variance swaps. Interestingly, the expected returns for long-term variance swaps are quite similar to the data in the model by Wachter (2013). In the right graph, I show that the model by Wachter (2013) (dash-dotted line) predicts a larger return volatility on one-month variance swaps than empirically observed. This result is driven by the extreme realized stock market variance in cases a consumption disaster hits the economy, and this indicates that mean of the consumption disaster or its frequency are too extreme. However, despite this large risk, the model is still not able to match the observed risk premium on one-month variance swaps.

The model by Drechsler and Yaron (2011) is not able to match the empirical risk premium on one-month variance swaps, because the jumps in the long-run risk and the volatility process are not sufficiently severe to capture the empirical magnitude. While at the same, the model predicts a slightly too large risk premium on long-term variance swaps, and, therefore, it is not straightforward how to improve the calibration of this model. In the right graph of Figure 3, I show that the model by Drechsler and Yaron (2011) (dotted line) predicts a larger return volatility than empirically observed, both for short-term and long-term variance swaps. This results from the fact that in order for this model to generate a variance risk premium, the jumps in the volatility process and long-run consumption mean have to be sufficiently frequent and large. If the jumps appear in either of these state variables, volatility of the stock market spikes, and this results in the extremely large return volatility for short-term variance swaps. However, the results of this section show that in order to match the empirical variation in variance discount rates, the model has to feature this type of stochastic volatility in multiple state variables.

In sum, I showed in the Figure 1 that the models I considered struggle to match the empirical term structure on the decomposition of variance swap prices. The data shows a very strong term structure, whereas the models all predict a more or less flat term structure on the decomposition of variance swap prices. Moreover, these results in combination with the pattern of Figure 2 showed that the only model the model by Drechsler and Yaron (2011) is able to match the empirical magnitude of the variation in variance discount rates. Finally, the results of Figure 3 indicated that the model by Gabaix (2012) is best able to match the empirical term structure of expected returns and return volatility of variance swaps. The models by Wachter (2013) and Drechsler and Yaron (2011) severely understate the risk premium on a one-month variance swap, despite the fact that this contract is substantially more risky in the models compared to the data.

In the following section, I show that the empirical results of Section 4.1 are robust to specifying a VAR to obtain expected stock market variance and variance discount rates. Afterward, I show that variance discount rates are related to many important macro-economic variables, and present a joint decomposition of stock market and variance swap return to analyze which components drive its strongly negative correlation.

5 Decomposition of variance swap prices using a VAR

In this section, I show that the variance decomposition of Section 4.1 is robust to specifying a VAR to obtain expectations. Using the VAR it is possible to estimate variance expectations and variance discount rates and analyze its time-series variation. Furthermore, I provide evidence that my VAR specification estimates variance expectations and variance discount rates effectively.

The VAR is used to model variance expectations and obtain variance discount rates as a latent variable from the present value identity (7). It is convenient to model stock market variance with a VAR, because this allows me to obtain variance expectations for each period by iterating forward.

In the benchmark exercise, I focus on the following VAR with four state variables:

$$z_{t+1} = Bz_t + \epsilon_{t+1} \quad \text{and} \tag{13}$$

$$z_t = \left(rv_t \quad pc_t^{(1)} \quad pc_t^{(2)} \quad DEF_t \right)', \tag{14}$$

where $B \in \mathbb{R}^{4\times 4}$ is a matrix with regressor coefficients and $\epsilon_{t+1} \in \mathbb{R}^{4\times 1}$ is a vector with errors. The vector z_t consists of the following variables: rv_t is the log of realized variance, $pc_t^{(1)}$ is the first principal component of the panel of log variance swap prices, $pc_t^{(2)}$ is the second principal component of the panel of log variance swap prices, and DEF_t is the default spread defined as the yield difference between BAA and AAA credits. For simplicity, and standard in the literature, all the variables included in z_t are demeaned such that the intercepts in the VAR are zero.

The first row in the VAR of equation (13) represents the predictive model for stock market variance. In order to decompose the variance swap prices, I calculate variance expectations, and, therefore, the remaining variables in the VAR are included based on its ability to predict stock market variance. $pc_t^{(1)}$ captures the level of the term structure of variance swap prices and predicts future stock market variance well. The level of term structure of variance swap prices is highly correlated (≈ 0.93) with the VIX index, which Drechsler and Yaron (2011) show to be a good predictor of stock market variance. $pc_t^{(2)}$ relates to the slope of the term structure of variance swap prices, which is high during episodes of low stock market variance and vice versa. Finally, the default spread DEF_t is included in the VAR, which Campbell et al. (2018) show to predict variation in long-term variance, and the default spread is a well-known business cycle indicator.

Based on the VAR model, monthly variance expectations for a variance swap with T-months to maturity equals:

$$E_{\text{rv},t}^{(T)} = e_1' \Bigg(\big(1-\rho(T)\big)B + \cdots + \big(1-\rho(1)\big)\rho(T) \times \cdots \times \rho(2)B^T \Bigg) z_t, \tag{15}$$

where $e_1 \in \mathbb{R}^{4\times 1}$ is a unit vector with the first element equal to one and the remaining equal to zero. By the pricing identity (7), variance discount rates are a function of variance expectations from equation (15) and the current variance swap price, as follows:

$$E_{\text{vdr},t}^{(T)} = E_{\text{rv},t}^{(T)} - vs_t^{(T)}. \tag{16}$$

In this way, I obtain an ex ante estimate of variance expectations over the lifetime of the variance swap and an estimate of the variance discount rates that price the variance swap. I use the estimates to decompose variation in the variance swap price to either variance expectations of equation (15) or variance discount rates of equation (16). Before I show the results of this decomposition, I present the estimation results of the VAR in Table 3.

Table 3: This table shows the estimated coefficients of the VAR of equation (13) with t-values in parentheses in Panel A. All variables are normalized to have mean equal to zero, and $pc_t^{(1)}$ and $pc_t^{(1)}$ are additionally standardized to have standard deviation equal to one. Panel B shows the correlation matrix of the residual vector ϵ_t with the standard deviations on the diagonal. Sample period for the dependent variables is January 1996 – June 2019, with 282 monthly data points.

Panel A: Coefficients VAR model

	rv_t	$pc_t^{(1)}$	$pc_t^{(2)}$	DEF_t	R^2
rv_{t+1}	0.056	0.514	-0.344	0.514	0.589
(t-stat.)	(0.73)	(7.43)	(-7.13)	(3.38)	
$pc_{t+1}^{(1)}$	0.049	0.856	0.003	0.105	0.847
(t-stat.)	(1.02)	(20.00)	(0.09)	(1.11)	
$pc_{t+1}^{(2)}$	-0.022	0.151	0.845	-0.022	0.745
(t-stat.)	(-0.35)	(2.70)	(21.77)	(-0.18)	
DEF_{t+1}	0.013	-0.010	-0.004	0.964	0.936
(t-stat.)	(1.34)	(-1.11)	(-0.66)	(48.81)	

Panel B: Correlation/Std Matrix of Residuals.

corr/std	rv_{t+1}	$pc_{t+1}^{(1)}$	$pc_{t+1}^{(2)}$	DEF_{t+1}
rv_{t+1}	0.627	0.627	-0.484	0.354
$pc_{t+1}^{(1)}$	0.627	0.388	-0.598	0.372
$pc_{t+1}^{(2)}$	-0.484	-0.598	0.505	-0.154
DEF_{t+1}	0.354	0.372	-0.154	0.082

The first row in Panel A of Table 3 presents the model for log realized variance each month. In line with expectation, the level of the term structure of variance swap prices $pc_t^{(1)}$ predicts next month's realized variance positively. Variance swap prices rise during episodes of elevated stock market variance. The slope of the term structure of variance swap prices, $pc_t^{(2)}$, predicts next month's realized variance negatively. A result that is also expected, because the slope of the term structure is high (low) during periods of low (high) stock market variance. Finally, DEF_t predicts future realized variance positively and is in line with Campbell et al. (2018). The R^2 of 58.9% to predict next month's variance indicates that most variation is captured. Furthermore, the impulse response functions in Appendix A.5 show that $pc_t^{(1)}$ and $pc_t^{(2)}$ mostly capture variation in short- to mid-term variance, whereas DEF_t captures variation in long-term variance.

The remaining rows in Panel A of Table 3 summarize the dynamics of the explanatory variables in the VAR. The level of the term structure of variance swap prices, $pc_t^{(1)}$, is approximately an AR(1) process with an autoregressive coefficient of 0.86. The slope of the term structure of variance swap prices, $pc_t^{(2)}$, has a similar persistence of 0.85 but is also predicted with a positive coefficient by the level of the term structure of variance swap prices. Finally, the default spread, DEF_t, is more persistent with an autoregressive coefficient of 0.96. The persistence of the variables indicates whether the variables capture variation in short- or long-term variance and the implications are similar to the results of the impulse response functions, as shown in Appendix A.5.

The estimates in Panel A of Table 3 are used to calculate variance expectations, using equation (15), and variance discount rates, using equation (16). Variation in the variance swap price with T-months to maturity is attributed to variation in $E_{\mathrm{rv},t}^{(T)}$, $E_{\mathrm{vdr},t}^{(T)}$ or correlation between $E_{\mathrm{rv},t}^{(T)}$ and $E_{\mathrm{vdr},t}^{(T)}$. This intuition follows from the following equation and is obtained if I calculate the variance of the pricing identity (7) for the variance swap price, as follows:

$$\mathrm{var}\big(vs_t^{(T)}\big) \approx \mathrm{var}\big(E_{\mathrm{rv},t}^{(T)}\big) + \mathrm{var}\big(E_{\mathrm{vdr},t}^{(T)}\big) - 2 \cdot \mathrm{cov}\big(E_{\mathrm{rv},t}^{(T)}, E_{\mathrm{vdr},t}^{(T)}\big). \tag{17}$$

The results of the variance decomposition of equation (17) using the VAR are presented in Table 4.

Table 4: This table shows the results of the variance decomposition of variance swap prices using equation (17). Note that the (co)variances of the third, fourth, and fifth columns are scaled with the variance of the second column such that the sum of the three (co)variances equals one. Standard errors are computed using the Delta method.

T	$\mathrm{var}\big(vs\big)$	$\frac{\mathrm{var}(E_{\mathrm{rv}})}{\mathrm{var}(vs)}$ (s.e.)	$\frac{\mathrm{var}(E_{\mathrm{vdr}})}{\mathrm{var}(vs)}$ (s.e.)	$\frac{-2\cdot\mathrm{cov}(E_{\mathrm{rv}},E_{\mathrm{vdr}})}{\mathrm{var}(vs)}$ (s.e.)
18	0.219	0.285	0.760	-0.045
		(0.204)	(0.294)	(0.330)
12	0.231	0.485	0.606	-0.092
		(0.245)	(0.272)	(0.352)
6	0.280	0.870	0.313	-0.182
		(0.242)	(0.154)	(0.298)
3	0.350	1.128	0.147	-0.275
		(0.189)	(0.070)	(0.219)
1	0.455	1.235	0.069	-0.304
		(0.123)	(0.027)	(0.139)

The second column of Table 4 shows that the variation in variance swap prices decreases in the maturity of the contract. In line with the results of the simple decomposition in

Table 2, the results of Table 4 show that there is strong term structure in the decomposition of variance swap prices. Most of the variation in short-term variance swaps, with at most six months to maturity, is attributed to variance expectations and decreases from 123.5% for one-month variance swaps to 87.0% for six months to maturity. However, short-term variance discount rates also drive a part of the total variation in short-term variance swap prices, which is significantly different from zero. Variation in variance discount rates explains 6.9% of the total variation in one-month variance swap prices up to 31.3% of the variation in six-month variance swap prices. Interestingly, the covariance between expected variance in the next month and the one-month variance discount rate is significantly positive. This result indicates that the variance discount rate increases, rather than decreases, during periods of elevated stock market variance. The positive covariance between expected variance and the variance discount rate is in line with Cheng (2019), who shows that the variance risk premium decreases during periods of elevated stock market variance. The size of the covariance between expected realized variance and variance discount rates decreases in the maturity of the variance swap and is no longer significantly different from zero beyond a horizon of one month.

The variation in long-term variance swap prices with a maturity of 12 or 18 months is mostly attributed to variance discount rates which drives, respectively, 60.6% or 76.0% of the total variation. Moreover, variation in variance expectations only drives 48.5% or 28.5% of the total variation in 12-month or 18-month variance swap prices. Overall, the results from the decomposition using the VAR of Table 4 are remarkably close to the results of the simple decomposition of Table 2. The results from the simple decomposition are model-free and, therefore, the similarity of the results suggests that the VAR is correctly specified. Furthermore, I show in Appendix OA.2 that if I do the decomposition using the predictive regressions or the VAR using a quarterly frequency, the results are highly similar.

Before I close this section, I show that the current specification of the VAR is able to estimate variance expectations and variance discount rates effectively. I test this by estimating the same predictive regressions of the future variance and future returns on the variance swap as before, only in this case variance expectations and variance discount rates obtained from the VAR are used as the predictive variables. The regressions are specified as follows:

$$y_{\mathrm{rv},t+T} = \gamma_{0,\mathrm{rv}} + \gamma_{1,\mathrm{rv}} \cdot E^{(T)}_{\mathrm{rv},t} + \gamma_{2,\mathrm{rv}} \cdot \left(- E^{(T)}_{\mathrm{vdr},t} \right) + u^{\mathrm{rv}}_{t+T} \quad \text{and} \tag{18}$$

$$-y_{\mathrm{vdr},t+T} = \gamma_{0,\mathrm{vdr}} + \gamma_{1,\mathrm{vdr}} \cdot E^{(T)}_{\mathrm{rv},t} + \gamma_{2,\mathrm{vdr}} \cdot \left(- E^{(T)}_{\mathrm{vdr},t} \right) + u^{\mathrm{vdr}}_{t+T}. \tag{19}$$

If variance expectations and variance discount rates are obtained effectively, then variance expectations $E^{(T)}_{\mathrm{rv},t}$ should only predict future stock market variance and variance discount rates $E^{(T)}_{\mathrm{vdr},t}$ should only predict future returns on the variance swap. Moreover, the regressions are specified such that I should find the following: $\gamma_{1,\mathrm{rv}} = 1$ and $\gamma_{2,\mathrm{rv}} = 0$ in the case of regression equation (18) and $\gamma_{1,\mathrm{vdr}} = 0$ and $\gamma_{2,\mathrm{vdr}} = 1$ in the case of regression equation (19). Table 5 presents the results.

Table 5: This table presents the results of the predictive regressions of equations (18) and (19) in which the expected realized variance and expected variance discount rates are the independent variables. t-statistics are represented in parentheses and are computed using Newey-West standard errors with number of lags equal to T.

Dependent variable:	$y_{\mathrm{rv},t+T}$			$-y_{\mathrm{vdr},t+T}$		
Maturity	$\gamma_{1,\mathrm{rv}}$ (t-stat.)	$\gamma_{2,\mathrm{rv}}$ (t-stat.)	R^2	$\gamma_{1,\mathrm{vdr}}$ (t-stat.)	$\gamma_{2,\mathrm{vdr}}$ (t-stat.)	R^2
18	0.943	-0.007	0.127	0.095	0.956	0.267
	(3.61)	(-0.03)		(0.34)	(4.06)	
12	1.104	0.124	0.269	-0.089	0.822	0.191
	(4.48)	(0.58)		(-0.34)	(3.68)	
6	1.016	0.178	0.399	-0.005	0.787	0.118
	(7.92)	(0.74)		(-0.03)	(3.13)	
3	0.964	0.143	0.479	0.030	0.839	0.077
	(13.56)	(0.54)		(0.45)	(3.09)	
1	1.043	0.346	0.592	-0.043	0.653	0.045
	(17.82)	(1.29)		(-0.73)	(2.63)	

Table 5 shows the results of two regressions: a regression to predict future stock market variance in the columns labeled $y_{\mathrm{rv},t+T}$ and a regression to predict future variance swap returns in the columns labeled $-y_{\mathrm{vdr},t+T}$. First, from the regression to predict future variance it follows that, indeed, future variance is predicted by variance expectations and not predicted by variance discount rates. Moreover, the regression coefficient of variance expectations $\gamma_{1,\mathrm{rv}}$ to predict future variance is close to, and not significantly different from, one for each maturity, and this shows that the VAR is able to capture stock market variance beyond one month. Finally, $\gamma_{1,\mathrm{rv}}$ is not exactly equal to one for the regression with one month to maturity, because variance discount rates are implied using the current variance swap prices, which contain information that is not included in the VAR.

Second, from the regression to predict future returns on variance swaps, it follows that, indeed, future returns are predicted by variance discount rates and not predicted by variance expectations. The regression coefficient of variance discount rates to predict future returns $\gamma_{2,\mathrm{vdr}}$ are fairly close, and not significantly different from, one, and this indicates that variance discount rates are able to predict future returns on the variance swap for each maturity. In fact, as indicated by the increasing R^2's, returns on long-term variance swaps are predicted more effectively by variance discount rates than returns on short-term variance swaps. Again, this is in line with the result that variation in long-term variance discount rates is larger than variation in short-term variance discount rates.

In the following section, I show that variance discount rates relate to many important economic variables. This supports the idea that variance discount rates are driven by systematic risk factors, and shows to what extend variance markets are integrated with other financial markets.

6 Time-series variation of variance discount rates

In this section, I analyze the time-series variation in variance discount rates in detail. I show that time-series variation in variance discount rates has a economically sizable effect on average returns of variance swaps. Furthermore, I show that variance discount rates relate to many important macro-economic variables, and I do a joint decomposition of variance swap and stock market returns to analyze the drivers of its negative correlation.

In the next figure, I plot the decomposition of the one-month variance swap price and the decomposition of the 18-month variance swap price.

Figure 4: The figure plots the decomposition of the variance swap price (black line) into variance expectations (dotted line) and variance discount rates (dashed line) over the sample period. The left graph shows the decomposition of the one-month variance swap price, and the right graph the decomposition of the 18-month variance swap price. The variables are represented as six-month moving averages. The shaded area corresponds to the NBER recessions.

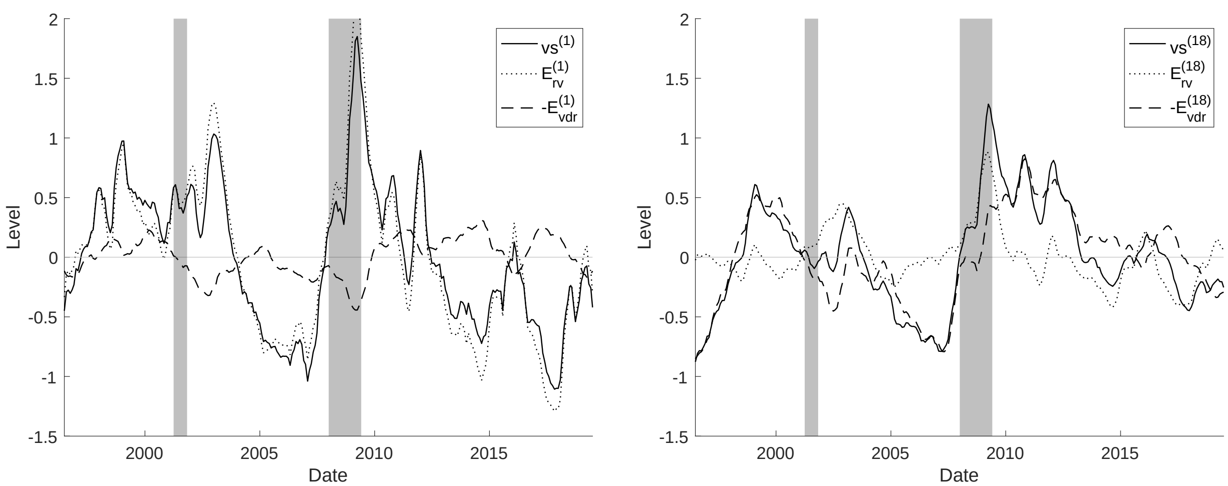


Figure 4 plots the decomposition of the variance swap price (black line) into variance expectations (dotted line) and variance discount rates (dashed line). The left graph shows the demeaned variance swap price with one month to maturity, and it is clearly visible that the monthly variance swap price closely follows expected stock market variance. Moreover, in line with Cheng (2019), there is a negative correlation between expected variance and the variance discount rate, which indicates that the discount rate decreases when expected variance increases.

The right graph of Figure 4 plots the decomposition of the 18-month variance swap price (black line) into variance expectations (dotted line) and variance discount rates (dashed line). Variance discount rates are a more important determinant of the 18-month variance swap price than they are for the one-month variance swap price. It follows from the graph that variance discount rates were relatively high during the period following the financial crisis in 2008, whereas variance discount rates were relatively low during the period leading up to the crisis. Moreover, the variation in the short-term variance discount rates (left graph) and the variation in long-term variance discount rates (right graph) are correlated. In order to analyze this correlation further, I plot in the following figure the variance discount rates obtained from the variance swap prices analyzed in the benchmark exercise.

Figure 5: The figure plots six-month moving averages of the variance discount rates obtained from the variance swap price with one month to maturity (solid grey line), three months to maturity (dotted black line), six months to maturity (dash-dotted black line), 12 months to maturity (dashed black line), and 18 months to maturity (solid black line). The grey area corresponds to NBER recessions.

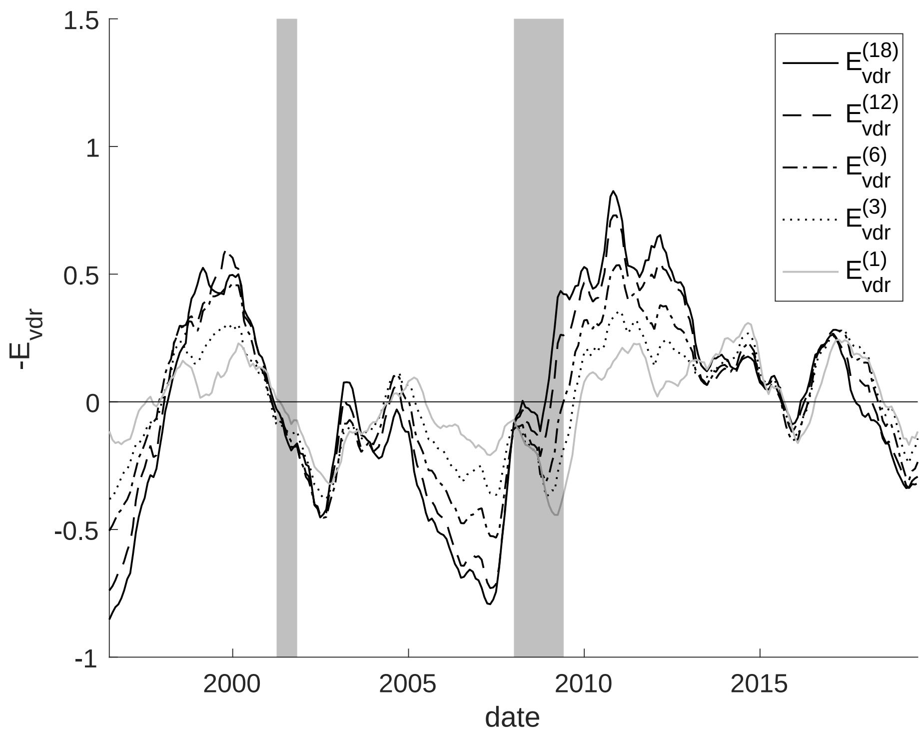


Figure 5 plots the variance discount rates obtained from the variance swap prices ranging from one to 18 months to maturity. The main result from Figure 5 is that the time-variation in term structure of variance discount rates is strongly correlated, such that short-term variance discount rates move in the same direction as long-term variance discount rates do. A finding that suggests that short- and long-term variance discount rates are driven by similar state variables. Overall, the variation in long-term variance discount rates is larger than variation in short-term discount rates. Interestingly, during the financial crisis of 2008, short-term variance discount rates decreased, whereas long-term variance discount rates increased. Therefore, only the price to hedge short-term variance risk decreased during the financial crisis.

I show in Figure 5 that variance discount rates are correlated for each of the considered maturities. In the following, I show that this time variation implies economically sizable differences in the average returns for investing in a variance swap contract. To show this, I run the following regressions:

$$R_{t+1}^{(T)} - 1 = \mu_1 + \mu_2 \cdot E_{\text{vdr},t}^{(T)} + \epsilon_{t+1}, \tag{20}$$

where, $R_{t+1}^{(T)} - 1$ correspond to a monthly simple return on a variance swap with T-months to maturity, and $E_{\text{vdr},t}^{(T)}$ are the variance discount rates obtained from the VAR. Because $E_{\text{vdr},t}^{(T)}$ has a mean of zero, μ_1 equals the average return on a variance swap, and μ_2 equals by how much the average return increases in $E_{\text{vdr},t}^{(T)}$. Table 6 shows the results of regression (20).

Table 6: This table presents the estimates of regression equation (20). The row labeled μ_1 shows the average simple return on a variance swap with maturity T, and the row labeled μ_2 shows the estimate for the variable $E_{\text{vdr},t}^{(T)}$. Note, $E_{\text{vdr},t}^{(T)}$ is scaled to have a standard deviation of one. t-statistics are represented in parentheses.

Maturity	1	3	6	12	18
μ_1	-0.285	-0.098	-0.050	-0.013	-0.006
(t-stat.)	(-7.16)	(-3.77)	(-2.72)	(-1.01)	(-0.52)
μ_2	0.123	0.110	0.082	0.056	0.041
(t-stat.)	(3.09)	(4.24)	(4.46)	(4.24)	(3.77)
R^2	0.033	0.061	0.067	0.061	0.049

The first result of Table 6 is that the average simple return on variance swaps strongly increases in the maturity. The second result, and main takeaway of Table 6 is that the variation in variance discount rates is economically sizable. A one-standard deviation increase in $E_{\text{vdr},t}^{(T)}$ results in an increase in average return of 4.1% for the 18-month variance swap up to an increase of 12.3% for the one-month variance swap. The results suggest that average simple returns on a variance swap with maturities beyond one month are positive, rather than negative, during periods in which $E_{\text{vdr},t}^{(T)}$ exceeds its standard deviation. This puzzling result was noted before by Johnson (2017) and indicates that during some periods investors want to receive a risk premium for holding a variance swap. Furthermore, it is interesting to note that long-term variance swaps do not have an average risk premium significantly different from zero, while at the same time exhibiting sizable time-series variation.

In the following, I run regressions to explain variance discount rates with several macroeconomic variables. This helps in the understanding of the drivers of variation in variance discount rates. In Table 7, I show the estimation results from the regression of some macroeconomic variables to explain the variation in variance discount rates for the benchmark maturities.

Table 7: This table presents the estimation results of a regression with several macro-economic variables as the predictors for variance discount rates ($E^{(T)}_{\text{vdr},t}$). The following variables are included: PVS_t corresponds to the relative price of volatile stocks which Pflueger et al. (2020) show to capture financial market risk perceptions, $S^{\mathbb{Q}}_t$ corresponds to risk-neutral skewness implied from option prices with the S&P 500 as the underlying, DEF_t corresponds to the default spread, TED_t corresponds to the TED-spread defined as the difference between the three-month libor and three-month treasury yield, $y^{(3)}_t$ corresponds to the three-month treasury yield, and $y^{(12)}_t - y^{(3)}_t$ correponds to difference between the 12-month and three-month treasury yield. Note each of the variables included in the regressions are scaled to have a mean of zero and volatility of one.

	$E^{(1)}_{\text{vdr},t}$	$E^{(3)}_{\text{vdr},t}$	$E^{(6)}_{\text{vdr},t}$	$E^{(12)}_{\text{vdr},t}$	$E^{(18)}_{\text{vdr},t}$
PVS_t	0.379	0.381	0.379	0.288	0.146
(t-stat.)	(5.45)	(5.21)	(5.20)	(4.28)	(2.63)
$S^{\mathbb{Q}}_t$	-0.015	0.006	0.068	0.120	0.107
(t-stat.)	(-0.27)	(0.11)	(1.18)	(2.26)	(2.46)
DEF_t	0.384	0.481	0.684	0.812	0.972
(t-stat.)	(4.70)	(5.60)	(7.99)	(10.29)	(14.92)
TED_t	-0.435	-0.412	-0.413	-0.361	-0.291
(t-stat.)	(-6.43)	(-5.79)	(-5.83)	(-5.53)	(-5.40)
$y^{(3)}_t$	0.509	0.486	0.467	0.515	0.474
(t-stat.)	(7.05)	(6.41)	(6.17)	(7.39)	(8.24)
$y^{(12)}_t - y^{(3)}_t$	0.195	0.143	0.115	0.107	0.131
(t-stat.)	(3.83)	(2.68)	(2.16)	(2.19)	(3.23)
R^2	0.361	0.295	0.299	0.405	0.594

The main point of Table 7 is that variance discount rates relate to many other macroeconomic variables which shows that my estimates of variance discount rates are not just noise. The financial market risk perceptions measure (PVS_t)[11] defined in Pflueger et al. (2020) predicts high variance discount rates which indicates that hedging variance risk is relatively cheap during periods of low perceived risk. The third row indicates that a high default spread predicts high variance discount rates and this result is not in line with economic intuition which would predict that during periods of market turmoil (when the default spread is high) hedging variance risk is relatively expensive. However, this result could be driven by the fact

[11] PVS_t is defined as the book-to-market ratio of low-volatility stocks minus the book-to-market ratio of high volatility stocks. This measure is high when perceived risk in the stock market is low.

that the default spread is included in the VAR and, moreover, the default spread is correlated with other variables in the regression. This is something to be explored more into the future. Interestingly, the TED-spread is a good predictor of variance discount rates and the results indicate that during periods of elevated systemic risk (indicated by a high TED-spread), hedging variance risk is expensive. This result could be driven by the fact that during those periods dealers of variance swaps are reluctant to sell these contracts or that banks are willing to pay a large premium to hedge variance risk. In the following, I present evidence that, under the current VAR specification, variance expectations and variance discount rates are obtained effectively.

6.1 Decomposition of variance swap and stock market returns

In this subsection, I decompose returns on variance swaps and the stock market into news about cashflows and news about risk premia. A salient feature of stock market returns is the negative correlation between realized returns and realized stock market variance, and, therefore, returns on variance swaps and stock market returns are negatively correlated. Therefore, the stock market beta of variance swaps is also negative. Using the decomposition of variance swap and stock market returns, I analyze which parts of the variance swap return and stock market return drive the low stock market beta.

In order to derive an expression to decompose variance swap returns, I calculate changes in expectation of identity (7), which yields the following:

$$r_{t+1}^{(T)} - \mathbb{E}_t r_{t+1}^{(T)} \approx \big(\mathbb{E}_{t+1} - \mathbb{E}_t\big) \left[\sum_{i=1}^{T} \big(1 - \rho(T-i+1)\big) \Big(\prod_{j=1}^{i-1} \rho(T-j+1) \Big) rv_{t+i} \right] - \big(\mathbb{E}_{t+1} - \mathbb{E}_t\big) \left[\sum_{i=2}^{T} \Big(\prod_{j=1}^{i-1} \rho(T-j+1) \Big) r_{t+i}^{(T-i+1)} \right] = N_{\mathrm{rv},t+1}^{(T)} - N_{\mathrm{vdr},t+1}^{(T)}. \quad (21)$$

As previously, equation (21) indicates that unexpected returns on a variance swap with maturity T are associated with news about future variance and news about variance discount rates. An increase in expected variance is associated with a positive return on the variance swap, whereas an increase in variance discount rates is associated with a negative return on the variance swap.

I start with the decomposition of returns on variance swaps. Using the VAR of Section 5, I show how to obtain the unexpected shock towards variance swap return ($r_{t+1}^{(T)} - \mathbb{E}_t r_{t+1}^{(T)}$) and news about future stock market variance ($N_{\mathrm{rv},t+1}^{(T)}$). In Appendix A.6, I show, under the assumption that the first two principle components span all variation in variance swap prices, how to obtain the unexpected shock towards variance swap returns. The return realization is defined as follows:

$$r^{(T)}_{t+1} - \mathbb{E}_t r^{(T)}_{t+1} = e_L(T)' \cdot \epsilon_{t+1},$$

where,

$$e_L(T) = \Big(1-\rho(T) \quad \rho(T)\beta_1(T-1) \quad \rho(T)\beta_2(T-1) \quad 0\Big)'.$$

$e_L(T)$ depends on the log-linearization coefficient $\rho(T)$, and the exposure of the variance swap price towards first and second principle component. Note, the return realization of a variance swap is a linear combination of the realization toward stock market variance, the realization toward the first principal component and the realization toward the second principal component.

News about future stock market variance follows directly from the VAR and is defined as follows:

$$N^{(T)}_{\text{rv},t+1} = e_1'\left(\big(1-\rho(T)\big) + \cdots + \big(1-\rho(1)\big)\rho(T)\times\cdots\times\rho(2)B^{T-1}\right)\epsilon_{t+1}. \tag{22}$$

This definition is very similar to definition of $E^{(T)}_{\text{rv},t}$; however, $N^{(T)}_{\text{rv},t}$ depends on the error term rather than the current level of the state variables. The return identity (21) implies that returns on variance swaps are fully explained by news about variance and news about variance discount rates. Therefore, news about variance discount rates is defined in the following way:

$$N^{(T)}_{\text{vdr},t+1} = N^{(T)}_{\text{rv},t+1} - e_L(T)' \cdot \epsilon_{t+1}. \tag{23}$$

I use news about stock market variance of equation (22) and news about variance discount rates of equation (23) to decompose variation in returns on variance swaps.

In the following, I discuss how returns on the stock market are decomposed, and the discussion closely follows Campbell and Vuolteenaho (2004). The identity to decompose stock market return is given by:

$$\begin{aligned} r_{t+1} - \mathbb{E}_t r_{t+1} &\approx \big(\mathbb{E}_{t+1} - \mathbb{E}_t\big)\Big[\sum_{i=0}^{\infty}\rho^i \Delta d_{t+1+i}\Big] - \big(\mathbb{E}_{t+1} - \mathbb{E}_t\big)\Big[\sum_{i=1}^{\infty}\rho^i r_{t+1+i}\Big] \\ &= N_{\text{cf},t+1} - N_{\text{dr},t+1}, \end{aligned} \tag{24}$$

where r_{t+1} is the stock market return, Δd_{t+1} dividend growth, and ρ the log-linearization coefficient. The stock market is a perpetual cash flow, and, therefore, is ρ is constant. In order to decompose the realizations to stock market returns, I follow Campbell and Vuolteenaho

(2004) and model discount rates using a VAR.

The VAR of Campbell and Vuolteenaho (2004) include the following state variables:

$$z_t = \Big(r_t \quad TY_t \quad PE_t \quad VS_t\Big)', \tag{25}$$

where TY_t is the term yield spread, PE_t is the log price earning ratio, and VS_t is the small-stock value spread. The model is estimated on a monthly sample from January 1929 to December 2018. The estimates of the VAR model in Table 13 of Appendix A.7 are very similar to the estimates reported in Campbell and Vuolteenaho (2004). Using these estimates and $\rho = 0.95^{1/12}$, as in Campbell and Vuolteenaho (2004), news about dividends $N_{\mathrm{cf},t+1}$ and news about stock market discount rates $N_{\mathrm{dr},t+1}$ are defined as follows:

$$N_{\mathrm{cf},t+1} = \Big(e_1' + e_1'\rho B\big(I - \rho B\big)^{-1}\Big)\epsilon_{t+1} \quad \text{and} \tag{26}$$

$$N_{\mathrm{dr},t+1} = e_1'\rho B\big(I - \rho B\big)^{-1}\epsilon_{t+1}. \tag{27}$$

Using news about future stock market variance $N^{(T)}_{\mathrm{rv},t+1}$, news about future variance discount rates $N^{(T)}_{\mathrm{vdr},t+1}$, news about future cashflows $N_{\mathrm{cf},t+1}$, and news about future stock market discount rates $N_{\mathrm{dr},t+1}$, the correlation between variance swap and stock market returns can be decomposed.

I define the stock market beta for variance swaps in the following way:

$$N_{\mathrm{rv},t+1} - N_{\mathrm{vdr},t+1} = \alpha + \beta_M\big(N_{\mathrm{cf},t+1} - N_{\mathrm{dr},t+1}\big) + u_{t+1}. \tag{28}$$

The stock market beta β_M of a variance swap is decomposed into four covariances, all divided by $\mathrm{var}\big(N_{\mathrm{cf},t+1} - N_{\mathrm{dr},t+1}\big)$, as follows:

$$\begin{aligned}\beta_M = {} & \frac{\mathrm{cov}\big(N_{\mathrm{rv},t+1}, N_{\mathrm{cf},t+1}\big)}{\mathrm{var}\big(N_{\mathrm{cf},t+1} - N_{\mathrm{dr},t+1}\big)} + \frac{\mathrm{cov}\big(N_{\mathrm{rv},t+1}, -N_{\mathrm{dr},t+1}\big)}{\mathrm{var}\big(N_{\mathrm{cf},t+1} - N_{\mathrm{dr},t+1}\big)} + \frac{\mathrm{cov}\big(-N_{\mathrm{vdr},t+1}, N_{\mathrm{cf},t+1}\big)}{\mathrm{var}\big(N_{\mathrm{cf},t+1} - N_{\mathrm{dr},t+1}\big)} \\ & + \frac{\mathrm{cov}\big(N_{\mathrm{vdr},t+1}, N_{\mathrm{dr},t+1}\big)}{\mathrm{var}\big(N_{\mathrm{cf},t+1} - N_{\mathrm{dr},t+1}\big)} = \beta_{\mathrm{rv,cf}} + \beta_{\mathrm{rv,dr}} + \beta_{\mathrm{vdr,cf}} + \beta_{\mathrm{vdr,dr}}.\end{aligned}$$

Therefore, I decompose β_M into four covariances: the covariance between news about variance and news about dividends, the covariance between news about variance and news about stock market discount rates, the covariance between news about variance discount rates and news about dividends, and the covariance between news about variance discount rates and news about stock market discount rates. Table 8 presents the result of this decomposition.

Table 8: This table shows the results of β_M, estimated using regression equation (28) in the second column for several maturities. The remaining columns show how this β_M is decomposed into covariance between realized variance and dividends $\beta_{\text{rv,cf}}$, realized variance and market discount rates $\beta_{\text{rv,dr}}$, variance discount rates and dividends $\beta_{\text{vdr,cf}}$ and variance discount rates, and stock market discount rates $\beta_{\text{vdr,dr}}$.

T	β_M (t-stat.)	$\beta_{\text{rv,cf}}$ (t-stat.)	$\beta_{\text{rv,dr}}$ (t-stat.)	$\beta_{\text{vdr,cf}}$ (t-stat.)	$\beta_{\text{vdr,dr}}$ (t-stat.)
18	-2.210	-0.230	-1.446	-0.072	-0.462
	(-14.02)	(-2.35)	(-9.01)	(-0.74)	(-1.75)
12	-2.759	-0.352	-2.200	-0.029	-0.178
	(-15.89)	(-2.42)	(-9.79)	(-0.34)	(-0.75)
6	-4.004	-0.562	-3.656	0.017	0.196
	(-16.71)	(-2.36)	(-10.29)	(0.28)	(1.16)
3	-5.237	-0.639	-4.996	0.031	0.366
	(-15.78)	(-2.00)	(-10.13)	(0.72)	(3.47)
1	-7.183	0.097	-7.281	0.000	0.000
	(-9.79)	(0.20)	(-7.76)	(-)	(-)

First, Table 8 shows that β_M increases in the maturity of the variance swap and increases from -7.18 for a variance swap with one month to maturity to -2.21 for a variance swap with 18 months to maturity. Second, this negative β_M is mostly attributed for each maturity to $\beta_{\text{rv,dr}}$ and corresponds to the correlation between news about variance and news about stock market discount rates. This result indicates that during periods of large stock market variance, stock market discount rates are revised upward to compensate for the increased risk of investing in the stock market. Third, the β_M of variance swaps with a maturity beyond one month is driven with a coefficient significantly different from zero by $\beta_{\text{rv,cf}}$ and corresponds to the negative correlation between news about variance and news about dividends. This result shows that during episodes of increased stock market variance, dividends expectations are revised downward. Finally, the sign of the correlation between news about variance discount rates and news about stock market discount rates changes from positive for short-term variance discount rates to negative for long-term variance discount rates.

7 Conclusion

I show that variance discount rates vary over time, which indicates that during some periods investors worry more about variance risk than during others. Moreover, the variation in variance discount rates drives a significant part of the variation in prices in the market for variance and, in particular, for longer horizons. I document a strong term structure in the decomposition of variance swap price variation. Short-term variance swap prices are driven by variation in variance expectations, whereas long-term variance swap prices are mostly driven by variation in variance discount rates. This newly documented stylized fact provides a new challenge to existing asset pricing models, because the model considered in this paper predict a flat term structure. In particular, the disaster model by Gabaix (2012) predicts that all variation in variance swap prices is attributed to variation in variance discount rates. On the other hand, the disaster model by Wachter (2013) predicts that all variation is attributed to variance expectations, and this is driven by the fact that this model incorporates a strong persistence in stock market variance; a feature that is not present in the model by Gabaix (2012). The long-run risk model by Drechsler and Yaron (2011) matches the decomposition of long-term variance swap relatively well, and the model is able to match the overall variation in variance discount rates. However, due to the large variation in short-term disaster risk, short-term variance discount rates move more of the variation in short-term variance swap prices than empirically observed.

In sum, this paper presents new key stylized facts about the market for variance risk. I show that these stylized facts pose a challenge for state-of-the-art asset pricing models, and augmenting the asset pricing models to better describe the pricing of variance risk is thus an interesting avenue for future research.

References

Aït-Sahalia, Y., Karaman, M. and Mancini, L. (2020). The term structure of equity and variance risk premia. *Journal of Econometrics*.

Andersen, T. and Bondarenko, O. Dissecting the market pricing of return volatility. Working paper, (2009).

Baele, L., Driessen, J., Ebert, S., Londono, J. and Spalt, O. (2019). Cumulative prospect theory, option returns, and the variance premium. *The Review of Financial Studies*, 32(9): 3667–3723.

Bakshi, G. and Madan, D. (2000). Spanning and derivative-security valuation. *Journal of Financial Economics*, 55(2):205–238.

Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509.

Barro, R. (2006). Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3):823–866.

Bekaert, G. and Engstrom, E. (2017). Asset return dynamics under habits and bad environment–good environment fundamentals. *Journal of Political Economy*, 125(3):713–760.

Bekaert, G., Engstrom, E. and Ermolov, A. The variance risk premium in equilibrium models. Working paper, (2020).

Bollerslev, T. and Todorov, V. (2011). Tails, fears, and risk premia. *The Journal of Finance*, 66(6):2165–2211.

Bollerslev, T., Tauchen, G. and Zhou, H. (2009). Expected stock returns and variance risk premia. *The Review of Financial Studies*, 22(11):4463–4492.

Campbell, J. and Cochrane, J. (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. and Ammer, J. (1993). What moves the stock and bond markets? A variance decomposition for long-term asset returns. *The Journal of Finance*, 48(1):3–37.

Campbell, J. Y. and Shiller, R. J. (1988a). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228.

Campbell, J. Y. and Shiller, R. J. (1988b). Stock prices, earnings, and expected dividends. *the Journal of Finance*, 43(3):661–676.

Campbell, J. Y. and Vuolteenaho, T. (2004). Bad beta, good beta. *The American Economic Review*, 94(5):1249–1275.

Campbell, J. Y., Giglio, S., Polk, C. and Turley, R. (2018). An intertemporal CAPM with stochastic volatility. *Journal of Financial Economics*, 128(2):207–233.

Cheng, I. H. (2019). The VIX premium. *The Review of Financial Studies*, 32(1):180–227.

Cochrane, J. H. (1992). Explaining the variance of price–dividend ratios. *The Review of Financial Studies*, 5(2):243–280.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, 21(4):1533–1575.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4): 1047–1108.

Dew-Becker, I., Giglio, S., Le, A. and Rodriguez, M. (2017). The price of variance risk. *Journal of Financial Economics*, 123(2):225–250.

Drechsler, I. and Yaron, A. (2011). What's vol got to do with it. *The Review of Financial Studies*, 24(1):1–45.

Eraker, B. and Wu, Y. (2017). Explaining the negative returns to volatility claims: An equilibrium approach. *Journal of Financial Economics*, 125(1):72–98.

Fama, E. F. and Bliss, R. R. (1987). The information in long-maturity forward rates. *The American Economic Review*, 77(4):680–692.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics*, 127(2):645–700.

Giglio, S. and Kelly, B. (2018). Excess volatility: Beyond discount rates. *The Quarterly Journal of Economics*, 133(1):71–127.

Johnson, T. L. (2017). Risk premia and the VIX term structure. *Journal of Financial and Quantitative Analysis*, 52:2461–2490.

Kilic, M. and Shaliastovich, I. (2019). Good and bad variance premia and expected returns. *Management Science*, 65(6):2522–2544.

Kozhan, R., Neuberger, A. and Schneider, P. (2013). The skew risk premium in the equity index market. *The Review of Financial Studies*, 26(9):2174–2203.

Lochstoer, L. A. and Muir, T. Volatility expectations and returns. Working paper, (2019).

Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of Economics*, 132(1):367–433.

Nozawa, Y. (2017). What drives the cross-section of credit spreads? A variance decomposition approach. *The Journal of Finance*, 72(5):2045–2072.

Pflueger, C., Siriwardane, E. and Sunderam, A. (2020). Financial market risk perceptions and the macroeconomy. *The Quarterly Journal of Economics*, 135(3):1443–1491.

Rietz, T. A. (1988). The equity risk premium: A solution. *Journal of Monetary Economics*, 22(1):117–131.

Seo, S. B. and Wachter, J. A. (2018). Do rare events explain cdx tranche spreads? *The Journal of Finance*, 73(5):2343–2383.

Seo, S. B. and Wachter, J. A. (2019). Option prices in a model with stochastic disaster risk. *Management Science*, 65(8):3449–3469.

Shiller, R. J. (1979). The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy*, 87(6):1190–1219.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3):421–436.

Wachter, J. (2013). Can time varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, 68(3):987–1035.

A Appendix: Additional empirical results

A.1 Variance swaps in Kozhan et al. (2013).

The realized variance of a variance swap entered at time t with maturity T is caluclated in the following way:

$$RV_t^{(T)} = \sum_{j=1}^{T} \Big[2\big(e^{r_{t+j}} - 1 - r_{t+j}\big) \Big], \tag{29}$$

where r_{t+j} is daily log return realized on day $t+j$. Note that equation (29) is similar to the sum of squared daily returns as $r^2 \approx 2\big(e^r - 1 - r\big)$. The variance swap rate is defined as the the risk-neutral expectation of the realized variance specified in equation (29). Kozhan et al. (2013) show how to calculate the variance swap rate with maturity T at time t from option prices, as follows:

$$VS_t^{(T)} = \frac{2}{B_t^{(T)}} \left[\int_0^{F_t^{(T)}} \frac{P_t^{(T)}(K)}{K^2} dK + \int_{F_t^{(T)}}^{\infty} \frac{C_t^{(T)}(K)}{K^2} dK \right], \tag{30}$$

where $B_t^{(T)}$ is the risk-free bond price at time t with maturity T, $F_t^{(T)}$ is the forward price at time t with maturity T and $P_t^{(T)}(K)$ and $C_t^{(T)}(K)$ are prices of European put and call options at time t with maturity T and strike price K.

Kozhan et al. (2013) show how to approximate equation (30) using a finite number of available put and call options. Given the set of available option prices $P_t^{(T)}(K_i)$ and $C_t^{(T)}(K_i)$ for $0 \le i \le N$ where prices are mid points from bid and ask quotes, Kozhan et al. (2013) compute variance swap rates as follows. Define the following function:

$$\Delta I(K_i) = \begin{cases} \frac{K_{i+1} - K_{i-1}}{2}, & \text{for } 0 \le i \le N \text{ (with } K_{-1} := 2K_0 - K_1,\ K_{N+1} := 2K_N - K_{N-1}) \\ 0, & \text{otherwise.} \end{cases}$$

Then the variance swap rate is computed as follows:

$$VS_t^{(T)} \approx 2 \sum_{K_i \le F_t^{(T)}} \frac{P_t^{(T)}(K_i)}{B_t^{(T)} K_i^2} \Delta I(K_i) + 2 \sum_{K_i > F_t^{(T)}} \frac{C_t^{(T)}(K_i)}{B_t^{(T)} K_i^2} \Delta I(K_i). \tag{31}$$

Options on the S&P 500 expire every month on the third Friday. Using linear interpolation, I calculate variance swap rates that expire on the last trading day of each month. The linear interpolation works in the following way: Variance swap rates with maturity $T_1 < T$ and $T_2 > T$ are calculated by equation (31); then the variance swap rate with maturity T is constructed as follows:

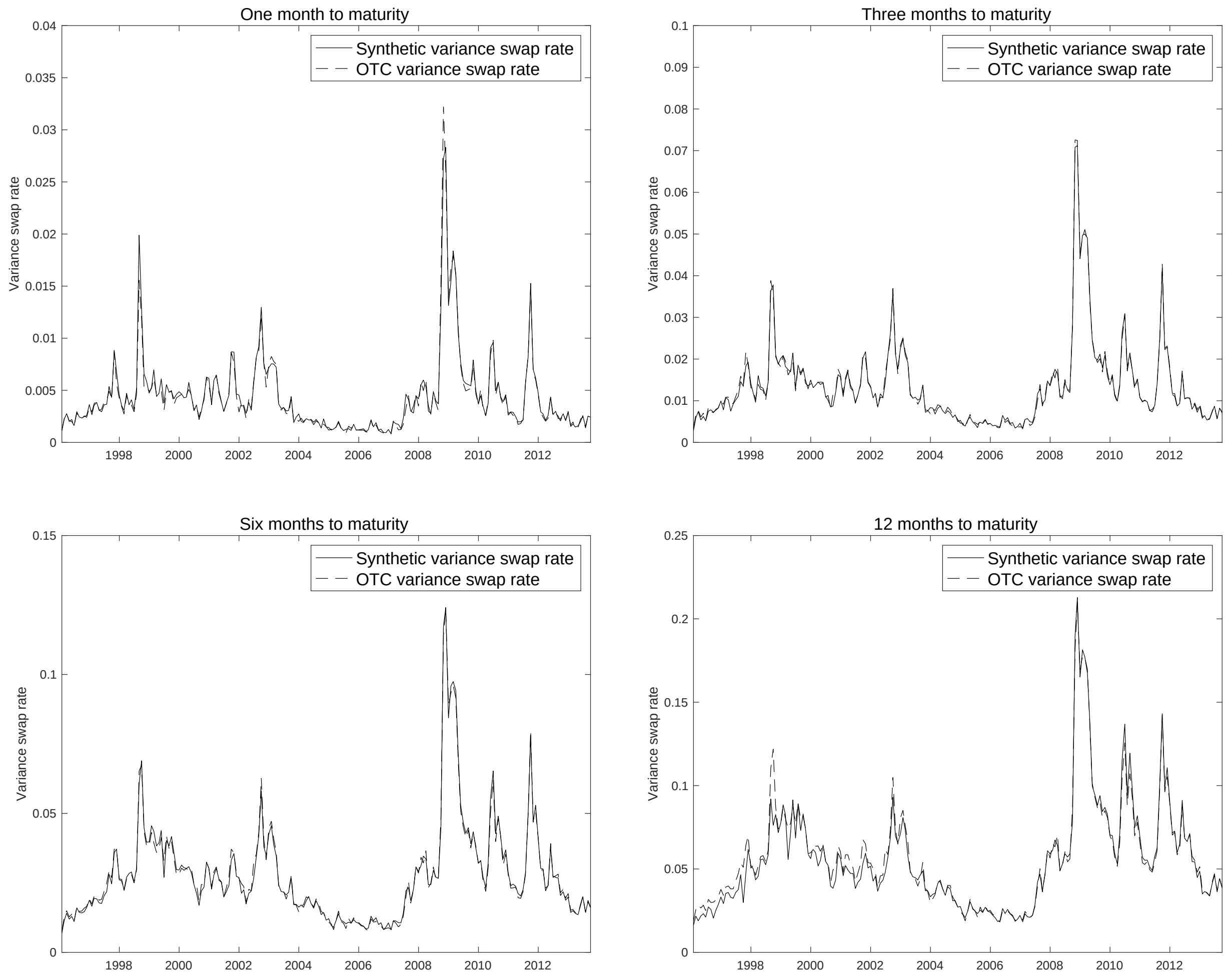
$$VS_t^{(T)} = \alpha VS_t^{(T_1)} + (1-\alpha)VS_t^{(T_2)},$$

where $T = \alpha T_1 + (1-\alpha)T_2$. With the data from OptionMetrics I calculate a panel of variance swap rates with one month up to 18 months to maturity.

A.2 Compare synthetic variance swaps to OTC variance swaps

In this section, I compare the data on synthetic variance swaps that are obtained from option pricing to the data on variance swaps from the OTC market. The data on variance swaps from the OTC market is from Dew-Becker et al. (2017). Their sample covers the period from December 1995 to September 2013 and variance swap rates up to a maturity of 12 months. During the period from January 1996 to September 2013, I observe a synthetic variance swap rate obtained using my methodology and a variance swap rate from the actual OTC data. I plot these rates in following graphs for one, three, six, and 12 months to maturity, which are the maturities of my benchmark analysis.

Figure 6: This figure plots the synthetic variance swap rate and OTC swap rate from Dew-Becker et al. (2017) for four maturities. The top-left graph plots the one-month swap rate, the top-right graph plots the three-month swap rate, the bottom-left graph plots the six-month swap rate, and the bottom-right graph plots the 12-month swap rate.



Overall, Figure 6 provides strong evidence that the synthetic variance swap rate is very similar to the swap rate in the OTC market. This indicates that the option market and the variance swap market are integrated markets and contain the same information regarding the pricing of variance risk. Notable differences include the difference in the one-month swap rate during the financial crisis and the difference in the 12-month swap rates during the first years of my sample. Furthermore, the average correlations between synthetic and OTC swap rate of the four maturities equals 0.991.

In the following table, I present sample statistics from the panel of variance swap returns. Note that I also present sample statistics on the realized variance of the S&P 500, because realized variance plays the role of the dividend payment in the calculation of the return. Moreover, I calculate simple returns and log-returns in order to quantify its differences in the case of variance swaps. The sample statistics of the realized variance and variance swap returns are shown in Table 9.

Table 9: The table shows sample statistics of realized variance in Panel A, simple variance swap returns in Panel B and log variance swap returns in Panel C. The sample statistics of monthly realized variance in Panel A are scaled to represent the yearly standard deviation. The mean, standard deviation, yearly Sharpe ratio and the 5%, 25%, 50%, 75% and 95% quantiles are presented.

Panel A: Realized variance								
Maturity	Mean	SD	SR	5%	25%	Median	75%	95%
-	0.162	0.095	-	0.068	0.101	0.142	0.188	0.323
Panel B: Simple returns on variance swaps								
18	-0.006	0.187	-0.106	-0.231	-0.123	-0.037	0.071	0.317
12	-0.013	0.227	-0.202	-0.269	-0.151	-0.056	0.081	0.346
6	-0.050	0.316	-0.544	-0.363	-0.232	-0.110	0.058	0.480
3	-0.098	0.447	-0.756	-0.477	-0.353	-0.204	0.002	0.581
1	-0.285	0.676	-1.458	-0.779	-0.630	-0.456	-0.193	0.838
Panel C: Log-returns on variance swaps								
18	-0.021	0.167	-	-0.262	-0.131	-0.037	0.069	0.275
12	-0.034	0.195	-	-0.313	-0.164	-0.057	0.078	0.297
6	-0.090	0.264	-	-0.451	-0.265	-0.117	0.057	0.392
3	-0.181	0.365	-	-0.649	-0.435	-0.228	0.002	0.458
1	-0.572	0.640	-	-1.508	-0.995	-0.608	-0.214	0.609

In Table 9, sample statistics of monthly realized variance on the S&P 500 and returns on variance swaps with different maturities ranging from one to 18 months are presented. Panel A presents sample statistics of realized variance scaled to yearly standard deviation, Panel B presents simple returns on variance swaps, and Panel C presents log-returns on variance swaps. The mean monthly realized variance over the sample is equal to 16.2% p.a., with a standard deviation of 9.5%. Furthermore, the distribution of monthly realized variance is right-skewed, indicated by the quantiles of the distribution.

The first observation from Panel B of Table 9 is that, on average, returns on variance swap returns are negative. This result is in line with a positive variance risk premium for the market portfolio as in Bollerslev et al. (2009) and Drechsler and Yaron (2011). Economically, a negative expected return on a variance swap indicates that if an investor wants to hedge variance risk, she pays a risk premium. The premium the investor pays for holding a variance swap is decreasing in the maturity of the variance swap. A variance swap with maturity longer than one month has exposure toward realized variance in the next month and toward expected variance for the remaining of the contract. Dew-Becker et al. (2017) show that the premium for hedging realized variance the next month is much larger than for hedging expected variance, and, therefore, the premium for holding a variance swap is decreasing in the maturity. Moreover, returns on variance swaps are volatile, as seen in the third column of Panel B in Table 9 and the volatility is decreasing in the maturity of the variance swap. However, the yearly Sharpe ratio is strongly increasing in maturity, and the yearly Sharpe ratio for investing in variance swaps with one month to maturity is very low (≈ -1.46), and similar to what Dew-Becker et al. (2017) find. Furthermore, the distributions of variance swap returns are right-skewed, indicated by the quantiles of the return distribution.

The sample average of the log-returns on variance swaps in Panel C of Table 9 are lower than the sample average for simple returns in Panel B of Table 9. This is driven by the fact that log-return distributions are less right-skewed than the simple return distributions and, therefore, is the sample average lower. The distribution of one-month returns is affected the most by the log-transformation. This result derives from the fact that the approximation of log-returns is equal to simple returns is close if the volatility of the return is low.

A.3 Decomposition of variance swap prices in a subsample

In this section, I show that the results of Section 4.1 are robust to running the predictive regressions on a subsample. The analysis is done on the sample period starting in September 2005, and the reason is that the CBOE introduced Long-Term Equity Anticipation Securities (Leaps) for the S&P 500 in this month. Leaps are option contract with the same specifications as before, only with a maturity up to three years. The CBOE lists these options with once every year, and the option expire on the third Friday in January. For this reason, I can obtain a balanced panel of variance swap prices up to a maturity of 24 months using interpolation on the sample period from September 2005 up to June 2019. To decompose the variance swap prices on this subsample, I run the predictive regressions of equations (10) and (11).

Table 10: This table shows the results of the predictive regressions of equations (10) and (11), in which the variance swap price is the independent variable. t-statistics are represented in brackets and are computed using Newey-West standard errors with number of lags equal to T.

Dependent variable:	$y_{\text{rv},t+T}$		$y_{\text{vdr},t+T}$	
Maturity	b_{rv} (t-stat.)	R^2	b_{vdr} (t-stat.)	R^2
24	-0.205	0.023	-1.239	0.454
	(-0.61)		(-3.88)	
18	0.161	0.014	-0.844	0.263
	(0.65)		(-3.68)	
12	0.502	0.120	-0.495	0.109
	(2.81)		(-2.81)	
6	0.772	0.274	-0.231	0.030
	(4.93)		(-1.42)	
3	0.951	0.414	-0.048	0.002
	(9.18)		(-0.44)	
1	1.130	0.565	0.130	0.017
	(14.55)		(1.67)	

Table 10 has two main takeaways. First, the results on this subsample are highly similar to the results over the total sample presented in 4.1. Second, the pattern documented in the term structure of the variance price decomposition continues to hold for longer maturities. That is, the variance swap price for contract with 24 months to maturity is solely driven by variance discount rates. Notably, the coefficient b_{vdr} is less than -1 which indicates the variation in the variance swap price is larger than the variation in variance discount rates. However, the coefficient is not significantly different from -1.

A.4 Variation in upside and downside variance swap prices

In this section, I show how the price of a variance swap can be decomposed as the sum of an upside and downside variance swap. In addition, I show that most of the variation in variance swap prices is attributed to variation in downside variance swap prices.

I start by showing how to decompose the variance swap as in Kozhan et al. (2013), into the sum of an upside and downside variance swap. As shown in Kozhan et al. (2013), the payoff of the variance swap is given by:

$$g(r_{t,T}) = 2(e^{r_{t,T}} - 1 - r_{t,T}),$$

where $r_{t,T}$ is the log-return on a forward contract from time t to maturity T. The payoff of the upside and downside variance swap are defined as follows:

$$g_u(r_{t,T}) = 2(e^{r_{t,T}} - 1 - r_{t,T}) \cdot \mathbb{1}(r_{t,T} \geq 0) \text{ and} \tag{32}$$

$$g_d(r_{t,T}) = 2(e^{r_{t,T}} - 1 - r_{t,T}) \cdot \mathbb{1}(r_{t,T} < 0), \tag{33}$$

where $\mathbb{1}(\cdot)$ is an indicator function. Using the formula from Bakshi and Madan (2000), I show that the prices of the payoff functions (32) and (33) are given by:

$$VS_{u,t}^{(T)} = \frac{2}{B_t^{(T)}} \left[\int_{F_t^{(T)}}^{\infty} \frac{C_t^{(T)}(K)}{K^2} dK \right] \text{ and}$$

$$VS_{d,t}^{(T)} = \frac{2}{B_t^{(T)}} \left[\int_{0}^{F_t^{(T)}} \frac{P_t^{(T)}(K)}{K^2} dK \right],$$

such that $VS_t^{(T)} = VS_{u,t}^{(T)} + VS_{d,t}^{(T)}$. Note, by Proposition 1 of Kozhan et al. (2013) there exists a unique trading strategy that perfectly hedges the payoff functions (33) and (32). However, the inclusion of the indicator function in the payoff functions makes this trading strategy infeasible to implement without strong assumptions on the process of the underlying. Therefore, I will only focus in this analysis on the prices of the upside and downside variance swap which is a combination of the cashflow and the discount rate.

My definition of the upside and downside variance swap is similar to Andersen and Bondarenko (2009) and Baele et al. (2019). While these papers focused on the average upside and downside variance premium for a maturity of one month, my analysis focuses on time-series variation in upside and downside variance swap prices and I include maturities from one-month up to 18 months. Moreover, in some other studies, the conditioning is different in the pricing equation than in the calculation of the realized payoff. For instance, Kilic and Shaliastovich (2019) and Dew-Becker et al. (2017) use a similar specification for the upside and downside variance swap price, but in order to calculate realized payoff they condition on intraday or daily returns. Hence, their realized payoff conditions on a different set of events compared to the pricing equation, for which the condition is whether the stock price at maturity is above or below the current forward price. This blurs the comparison of the realized payoff and price of the upside and downside variance swap, and, therefore, does not identify the upside and downside variance discount rate well.

I derive an identity to decompose variation in the variance swap price into variation due to the upside and downside variance swap price, in the following way (I suppressed the constants):

$$\begin{aligned}
vs_t^{(T)} &= \log\left(VS_{\mathrm{u},t}^{(T)} + VS_{\mathrm{d},t}^{(T)}\right) \approx \rho_1(T) vs_{\mathrm{u},t}^{(T)} + (1-\rho_1(T)) vs_{\mathrm{d},t}^{(T)}, \\
\iff 1 &\approx \frac{\mathrm{cov}(vs_t^{(T)}, \rho_1(T) vs_{\mathrm{u},t}^{(T)})}{\mathrm{var}(vs_t^{(T)})} + \frac{\mathrm{cov}(vs_t^{(T)}, (1-\rho_1(T)) vs_{\mathrm{d},t}^{(T)})}{\mathrm{var}(vs_t^{(T)})} \\
&=: b_u + b_d.
\end{aligned} \tag{34}$$

The coefficients b_u and b_d are estimated in two stages, in the first stage I estimate $\rho_1(T)$ using a simple regression and in the second stage b_u and b_d are estimated with a simple regression using $\rho_1(T)$ from the first stage. Similar to before, the sum of the coefficients should be close to one, and if this is the case it indicates that this log-linear approximation is, in fact, a good approximation. The results these second-stage regressions are presented in Table 11.

Table 11: In this table the results of the decomposition of the variance swap price into the upside and downside variance swap price are presented. The definition of the coefficients b_u and b_d are given in equation (34). The t-statistics of the coefficients are given in parentheses and are calculated using Newey-West standard errors with 50 lags.

Dependent variable:	$\rho_1(T) vs_{\mathrm{u},t}^{(T)}$		$(1-\rho_1(T)) vs_{\mathrm{d},t}^{(T)}$	
Maturity	b_u (t-stat.)	R^2	b_d (t-stat.)	R^2
18	0.256	0.859	0.742	0.980
	(19.18)		(57.96)	
12	0.275	0.865	0.722	0.978
	(16.58)		(45.13)	
6	0.368	0.886	0.634	0.956
	(20.74)		(50.42)	
3	0.371	0.923	0.635	0.970
	(22.96)		(56.47)	
1	0.314	0.957	0.687	0.991
	(27.06)		(64.09)	

The main result from Table 11 is that the main driver of variation in the variance swap price is the downside variance swap price. This result makes sense, because an important determinant of the variance swap price is crash risk. Overall, the importance of the downside variance swap price increases in the maturity of the variance swap (except for the one-month variance swap). Finally, the sum of the coefficients b_u and b_d is very close to one, which indicates that the log-linear approximation is a good approximation for the log-price of the variance swap.

A.5 Impulse response functions of the VAR

In this subsection, I show the impulse response function of stock market variance in response to a change of each of the other variables in the VAR of equation (13). The impulse responses are presented in Figure 7.

Figure 7: This figure plots the monthly impulse response functions of stock market variance in response to a change of the variables in the VAR. The scale of the y-axis is in standard deviation of stock market variance, where each of the variables increases by one standard deviation at time 0. The top-left graph plots the responses to a change in rv, the top-right graph plots the responses to a change in $pc^{(1)}$, the bottom-left graph plots the responses to a change in $pc^{(2)}$, and the bottom-right graph plots the responses to a change in DEF.

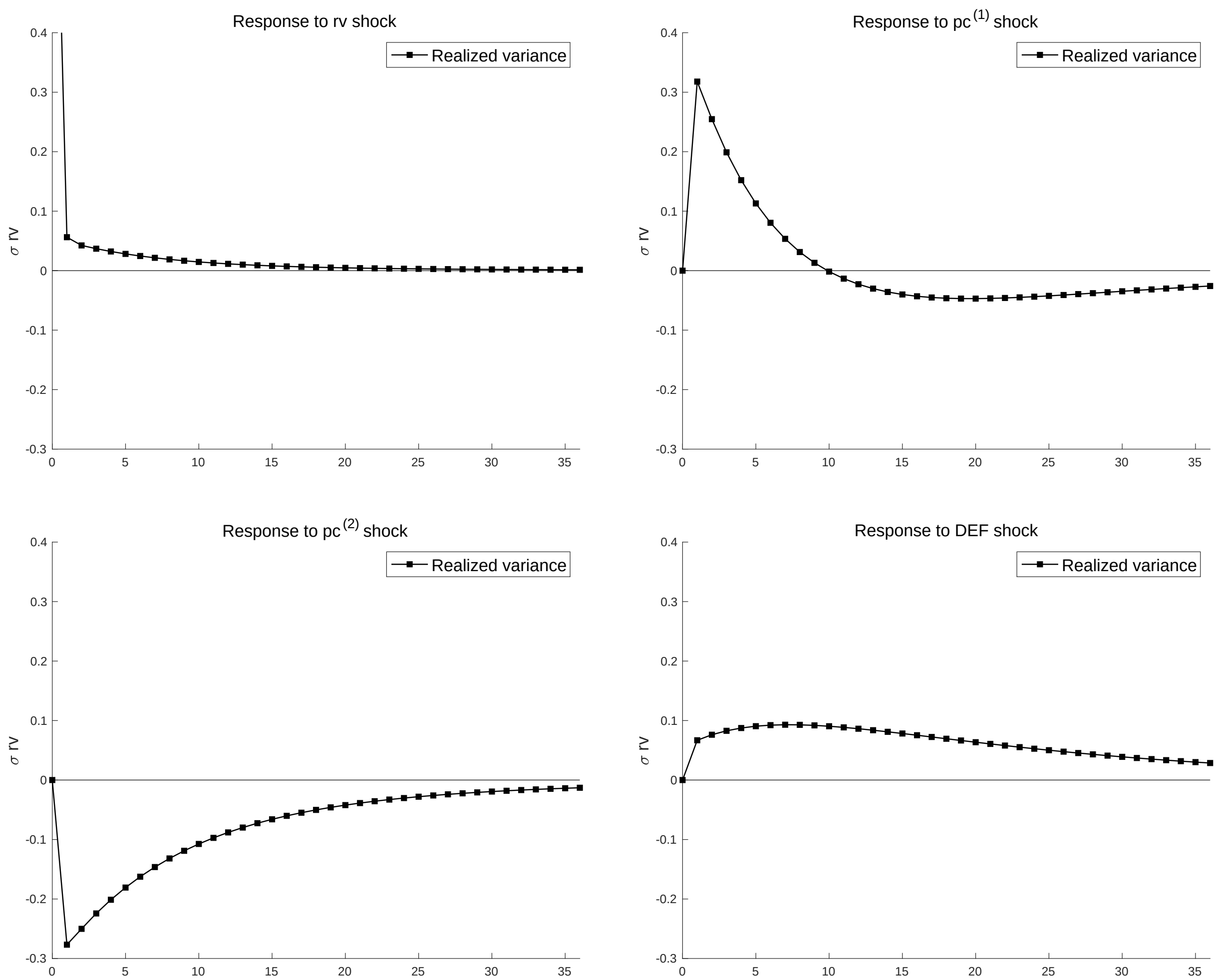


Figure 7 plots the impulse response functions of stock market variance for a horizon up to 36 months. The top-left graph shows that there is very little persistence in stock market variance, if the current level increases. The top-right graph shows that an increase in $pc^{(1)}$ increases future stock market variance up to 10 months forward. The bottom-left graph shows that an increase in $pc^{(2)}$ decreases future stock market variance and the shock is more persistent than a shock in $pc^{(1)}$. Finally, the bottom-right graph shows that shocks towards DEF are the most persistent and, therefore, affect long-term stock market variance.

A.6 Decomposition of variance swap returns

In this section, I show how to use the VAR of equation (13) in Section 5 to decompose returns on variance swaps using identity (21). The left-hand side of identity (21) corresponds to realizations of variance swap returns. In order to calculate return realizations using the VAR of equation (13), I assume that all variation in variance swap prices is captured by the first two principal components of the panel of variance swap rates. This two-factor assumption implies the following:

$$vs_t^{(T)} = \beta_0(T) + \beta_1(T) \cdot pc_t^{(1)} + \beta_2(T) \cdot pc_t^{(2)}. \tag{35}$$

Simple regressions show that equation (35) captures between 99.1% and 99.8% of the variation, with an average of 99.5% for all variance swap rates with one up to 18 months to maturity. Therefore, the variation in variance swap rates is accurately captured by the first two principal components. Under this assumption, realizations of variance swap returns are defined as follows:

$$\begin{aligned} & r_{t+1}^{(T)} \approx k(T) + \rho(T) \cdot vs_{t+1}^{(T-1)} + \big[1 - \rho(T)\big] rv_{t+1} - vs_t^{(T)} \\ \Longleftrightarrow \quad & r_{t+1}^{(T)} - \mathbb{E}_t r_{t+1}^{(T)} \approx \Big(\mathbb{E}_{t+1} - \mathbb{E}_t\Big)\Big[\rho(T) vs_{t+1}^{(T-1)} + \big[1 - \rho(T)\big] rv_{t+1}\Big] \\ & \qquad = e_L(T)' \cdot \epsilon_{t+1}, \end{aligned} \tag{36}$$

where,

$$e_L(T) = \Big(1 - \rho(T) \quad \rho(T)\beta_1(T-1) \quad \rho(T)\beta_2(T-1) \quad 0\Big)'.$$

It follows from the definition of $e_L(T)$ that the return realization of a variance swap is a linear combination of the realization toward stock market variance, the realization toward the first principal component and the realization toward the second principal component. Vector $e_L(T)$ depends on the maturity of the variance swap, because the relative importance of each of the individual realizations toward the variables in the VAR depends on the maturity and is captured by the constants $\rho(T)$, $\beta_1(T-1)$, and $\beta_2(T-1)$. In sum, the realizations of variance swap returns follow from the two-factor assumption of equation (35) and from the assumption that log-linear approximation of the return holds with an equality.

I now show how to decompose the variance of variance swap returns. By calculating the variance of identity (21), I obtain the following identity to decompose the variation in returns, as follows:

$$\text{var}\big(\bar{r}_{t+1}^{(T)}\big) \approx \text{var}\big(N_{\text{rv},t+1}^{(T)}\big) + \text{var}\big(N_{\text{vdr},t+1}^{(T)}\big) - 2 \cdot \text{cov}\big(N_{\text{rv},t+1}^{(T)}, N_{\text{vdr},t+1}^{(T)}\big), \tag{37}$$

where $\bar{r}_{t+1}^{(T)} := e_L(T)' \cdot \epsilon_{t+1}$ is the return realization of a variance swap with maturity T at time $t+1$. Equation (37) implies that variation in variance swap returns is attributed to news about variance, news about variance discount rates, or the covariance between the two. Table 12 presents the results of this decomposition.

Table 12: This table shows the results of the variance decomposition of variance swap prices using equation (37). Note that the (co)variances of the third, fourth, and fifth columns are scaled with the variance of the second column such that the sum of the three (co)variances equals one. Standard errors are computed using the Delta method.

T	$\text{var}(\bar{r})$	$\frac{\text{var}(N_{\text{rv}})}{\text{var}(\bar{r})}$ (s.e.)	$\frac{\text{var}(N_{\text{vdr}})}{\text{var}(\bar{r})}$ (s.e.)	$\frac{-2\text{cov}(N_{\text{rv}},N_{\text{vdr}})}{\text{var}(\bar{r})}$ (s.e.)
18	0.023	0.630	0.674	-0.304
		(0.274)	(0.274)	(0.415)
12	0.032	0.966	0.414	-0.380
		(0.325)	(0.192)	(0.420)
6	0.063	1.224	0.123	-0.346
		(0.246)	(0.063)	(0.286)
3	0.115	1.217	0.033	-0.250
		(0.124)	(0.017)	(0.138)
1	0.393	1.000	0.000	0.000
		(-)	(-)	(-)

The first observation of Table 12 is that the variance of variance swap returns is strongly decreasing in the maturity of the variance swap and equals 39.3% for a variance swap with one month to maturity and equal to 2.3% for a variance swap with 18 months to maturity. The return on a one-month variance swap is solely driven by news about stock market variance, because the variance swap matures in the next periods and, therefore, is independent of variance discount rates by definition. Variation in returns on short-term variance swaps is driven by news about stock market variance and drives, respectively, 122.4% and 121.7% of the total variation in variance swap returns with six and three months to maturity. This result indicates that variance expectations change more quickly over time than variance discount rates do. However, even variation in short-term variance swap returns is driven by news about variance discount rates with a coefficient significantly different from zero and accounts for 12.3% of the variation in six-month variance swap returns and 3.3% of the variation in three-month variance swap returns.

Variation in returns of long-term variance swaps is also largely driven by news about variance and drives 96.6% of the total variation in 12-month variance swap returns and 63.0% of the total variation in 18-month variance swap returns. However, variation in long-term variance swap prices is driven by a larger fraction due to news about variance discount rates and drives 67.4% of the total variation in 18-month variance swap returns and 41.4% of the total variation in 12-month variance swap returns. Interestingly, the covariance between news about stock market variance and news about variance discount rates is positive for each maturity and indicates that hedging variance risk becomes less expensive if expected stock market variance increases.

In Appendix OA.3, I show that the results of obtaining the realizations to variance swap return using equation (36) or by including the variance swap return $r_{t+1}^{(T)}$ as an additional state variable in the VAR yield very similar results. If $r_{t+1}^{(T)}$ is an additional state variable in the VAR, realizations toward variance swap returns are obtained directly from the VAR.

A.7 Decomposition of stock market returns

In this section, I discuss how to decompose stock market returns into news about dividends and news about stock market discount rates, following Campbell and Vuolteenaho (2004).

Table 13 presents the estimation results of the VAR to decompose stock market returns.

Table 13: This table shows the coefficients of the VAR of equation (13) with state variables of equation (25), and a vector with intercepts c. The t-values of the coefficients are denoted in parentheses, sample period for the dependent variables is January 1929 – December 2018, with 1080 monthly data points.

Coefficients VAR model						
	c	r_t	TY_t	PE_t	VS_t	R^2
r_{t+1}	0.064	0.101	0.004	-0.015	-0.012	0.025
(t-stat.)	(3.59)	(3.35)	(2.06)	(-3.26)	(-2.27)	
TY_{t+1}	-0.026	0.067	0.938	-0.004	0.053	0.897
(t-stat.)	(-0.29)	(0.44)	(90.28)	(-0.16)	(2.04)	
PE_{t+1}	0.025	0.511	0.001	0.992	-0.003	0.991
(t-stat.)	(2.10)	(25.07)	(0.76)	(320.99)	(-0.83)	
VS_{t+1}	0.019	0.011	0.000	-0.001	0.989	0.979
(t-stat.)	(1.14)	(0.40)	(0.14)	(-0.12)	(203.71)	

The estimates in Table 13 are very similar to the estimates in Campbell and Vuolteenaho (2004). Using the estimates of the VAR, cashflow news and discount rate news are defined as in equations (26) and (27).

Over the sample period January 1996 until December 2018, I estimate $N^{(T)}_{\text{rv},t+1}$, $N^{(T)}_{\text{vdr},t+1}$, $N_{\text{cf},t+1}$, and $N_{\text{dr},t+1}$. In Figure 8, I plot each standardized news term over the sample period from stock market returns and returns on a 12-month variance swap. Furthermore, Table 14 presents the correlation coefficients of the variables over the sample period.

Figure 8: The figure plots yearly moving averages of $N^{(12)}_{\text{rv},t}$, $N^{(12)}_{\text{vdr},t}$, $N_{\text{cf},t}$ and $N_{\text{dr},t}$ over the sample period from January 1996 – December 2018. Each of the news terms is standardized with its standard deviation. The shaded area corresponds to the NBER recessions.

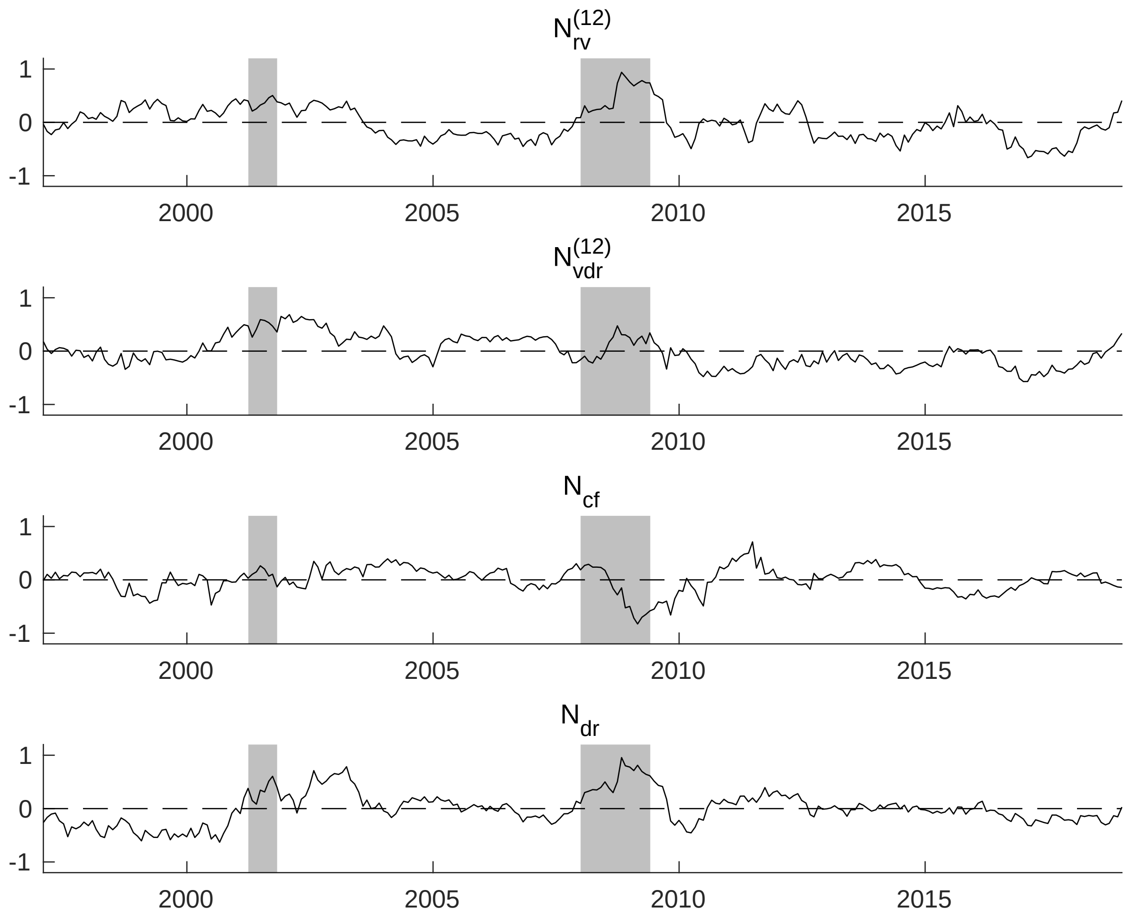


Table 14: This table represent the estimated correlation matrix of $N^{(12)}_{\mathrm{rv},t}$, $N^{(12)}_{\mathrm{vdr},t}$, $N_{\mathrm{cf},t}$, and $N_{\mathrm{dr},t}$ over the sample period from January 1996 – December 2018.

Correlation Matrix				
	$N^{(12)}_{\mathrm{rv},t}$	$N^{(12)}_{\mathrm{vdr},t}$	$N_{\mathrm{cf},t}$	$N_{\mathrm{dr},t}$
$N^{(12)}_{\mathrm{rv},t}$	1.000	0.302	-0.166	0.566
$N^{(12)}_{\mathrm{vdr},t}$	0.302	1.000	0.021	-0.069
$N_{\mathrm{cf},t}$	-0.166	0.021	1.000	0.266
$N_{\mathrm{dr},t}$	0.566	-0.069	0.266	1.000

The first graph of Figure 8 plots news about variance, the second graph plots news about variance discount rates, the third graph plots news about dividends, and the fourth graph plots news about stock market discount rates. The first and fourth graphs show that during episodes of large stock market variance, stock market discount rates tend to increase. This result is supported by the evidence in the correlation matrix of Table 14, which shows that the correlation coefficient between news about variance and news about stock market discount rates is positive.

Online Appendix to
Variance Discount Rates: What Drives Preferences over Variance Risk?

Joren Koëter*

December 23, 2022

Abstract

This online appendix (OA) contains additional material and results. Section OA.1 shows that the results of the VAR are robust to the inclusion of different state variables than studied in the main paper. Section OA.2 presents that the results of the main paper using a quarterly rather than monthly frequency. Section OA.3 specifies a VAR in which the variance swap return is included, and shows that the variance swap return decomposition are very similar to before. Finally, Section 4.2 discusses the various asset pricing model in detail.

OA.1 Different specifications of the VAR

In this section, I show that the results of the decomposition of variance swap prices are robust to a variety of different specifications of the VAR. Many of the variables included have been shown by prior studies to be important state variables that predict stock market returns or stock market variance. In Table OA.1, I show the results of the decomposition of variance swap prices for the following specifications:

1. $z_t = \begin{pmatrix} rv_t & pc_t^{(1)} & pc_t^{(2)} \end{pmatrix}'$. In this specification the default spread is excluded in order to alleviate any concerns induced by the persistence of the default spread.

2. $z_t = \begin{pmatrix} rv_t & pc_t^{(1)} & pc_t^{(2)} & rv_{t-13,t-1} \end{pmatrix}'$. In this specification lagged long-term stock market variance is included. $rv_{t-13,t-1}$ is defined as the log of realized variance in the past year, and lagged one month additionally such that the variable does not overlap with rv_t.

*Department of Finance at Rotterdam School of Management, Erasmus University. Email: koeter@rsm.nl.

3. $z_t = \left(rv_t \quad pc_t^{(1)} \quad pc_t^{(2)} \quad pe_t \right)'$. In this specification the log of smoothed price-earnings ratio is included, because Campbell et al. (2018) show that it predicts future stock market variance.

4. $z_t = \left(rv_t \quad pc_t^{(1)} \quad pc_t^{(2)} \quad def_t \quad pe_t \quad rv_{t-13,t-1} \right)'$. In this specification all state variables used in the previous analyses are included.

Table OA.1: This table shows the results of the variance decomposition of variance swap prices using equation (17). Note that the (co)variances of the third, fourth, and fifth columns are scaled with the variance of the second column such that the sum of the three (co)variances equals one. Standard errors are computed using the Delta method.

T	$\frac{\sigma^2_{\text{rv}}}{\sigma^2_{\text{vs}}}$ (s.e.)	$\frac{\sigma^2_{\text{vdr}}}{\sigma^2_{\text{vs}}}$ (s.e.)	$\frac{-2\sigma_{\text{rv,vdr}}}{\sigma^2_{\text{vs}}}$ (s.e.)	$\frac{\sigma^2_{\text{rv}}}{\sigma^2_{\text{vs}}}$ (s.e.)	$\frac{\sigma^2_{\text{vdr}}}{\sigma^2_{\text{vs}}}$ (s.e.)	$\frac{-2\sigma_{\text{rv,vdr}}}{\sigma^2_{\text{vs}}}$ (s.e.)
	Specification 1			Specification 2		
18	0.148	0.678	0.174	0.423	0.604	-0.027
	(0.102)	(0.297)	(0.229)	(0.291)	(0.233)	(0.335)
6	0.750	0.201	0.049	0.855	0.223	-0.078
	(0.217)	(0.126)	(0.213)	(0.258)	(0.126)	(0.286)
1	1.200	0.035	-0.235	1.200	0.035	-0.236
	(0.124)	(0.018)	(0.136)	(0.124)	(0.018)	(0.136)
	Specification 3			Specification 4		
18	0.227	0.800	-0.028	0.564	1.016	-0.580
	(0.176)	(0.360)	(0.429)	(0.306)	(0.353)	(0.540)
6	0.774	0.219	0.007	1.040	0.443	-0.482
	(0.223)	(0.135)	(0.239)	(0.272)	(0.175)	(0.369)
1	1.200	0.036	-0.236	1.200	0.035	-0.236
	(0.124)	(0.019)	(0.136)	(0.124)	(0.018)	(0.136)

The results in Table OA.2 show that the main result of the paper holds in each of the specifications; that is, short-term variance swap prices mainly driven by variance expectations, whereas long-term variance swap prices are mainly driven by variance discount rates. Moreover, the decomposition of the one-month variance swap price shows that the variables other than $pc_t^{(1)}$ and $pc_t^{(2)}$ add little in terms of predicting the one-month variance. The variables mostly add something for long-term stock market variance, due to the persistence of each of the variables. To test how well each of the specifications estimates variance expectations and variance discount rates, I run the following regressions(the same as equations (18) and (19) in the main paper):

$$y_{\text{rv},t+T} = \gamma_{0,\text{rv}} + \gamma_{1,\text{rv}} \cdot E^{(T)}_{\text{rv},t} + \gamma_{2,\text{rv}} \cdot \left(- E^{(T)}_{\text{vdr},t}\right) + u^{\text{rv}}_{t+T} \quad \text{and}$$
$$-y_{\text{vdr},t+T} = \gamma_{0,\text{vdr}} + \gamma_{1,\text{vdr}} \cdot E^{(T)}_{\text{rv},t} + \gamma_{2,\text{vdr}} \cdot \left(- E^{(T)}_{\text{vdr},t}\right) + u^{\text{vdr}}_{t+T}.$$

In case variance expectations and variance discount rates are estimated effectively, I should find $\gamma_{1,\text{rv}} = \gamma_{2,\text{vdr}} = 1$ and $\gamma_{2,\text{rv}} = \gamma_{1,\text{vdr}} = 0$.

Table OA.2: This table presents the results of the predictive regressions of equations (18) and (19) in which the expected realized variance and expected variance discount rates are the independent variables. t-statistics are represented in parentheses and are computed using Newey-West standard errors with number of lags equal to T.

T	$\gamma_{1,\text{rv}}$ (t-stat.)	$\gamma_{2,\text{rv}}$ (t-stat.)	$\gamma_{1,\text{vdr}}$ (t-stat.)	$\gamma_{2,\text{vdr}}$ (t-stat.)	$\gamma_{1,\text{rv}}$ (t-stat.)	$\gamma_{2,\text{rv}}$ (t-stat.)	$\gamma_{1,\text{vdr}}$ (t-stat.)	$\gamma_{2,\text{vdr}}$ (t-stat.)
	Specification 1				Specification 2			
18	2.184	-0.341	-1.163	1.299	1.026	-0.284	-0.004	1.223
	(5.94)	(-1.86)	(-3.07)	(7.11)	(3.92)	(-1.39)	(-0.01)	(5.81)
6	1.072	0.014	-0.057	0.939	1.024	-0.025	-0.011	0.973
	(8.23)	(0.06)	(-0.43)	(3.80)	(8.27)	(-0.10)	(-0.08)	(4.12)
1	1.075	0.765	-0.075	0.234	1.075	0.760	-0.075	0.239
	(16.23)	(1.91)	(-1.14)	(0.58)	(16.17)	(1.90)	(-1.12)	(0.60)
	Specification 3				Specification 4			
18	1.723	-0.161	-0.780	1.143	0.985	-0.058	-0.015	1.032
	(6.77)	(-0.81)	(-3.18)	(6.06)	(5.55)	(-0.32)	(-0.08)	(5.64)
6	1.076	-0.016	-0.067	0.988	1.013	0.075	-0.009	0.915
	(8.72)	(-0.07)	(-0.53)	(4.32)	(9.98)	(0.42)	(-0.13)	(4.87)
1	1.072	0.754	-0.075	0.245	1.037	0.257	-0.038	0.743
	(16.36)	(1.91)	(-1.13)	(0.62)	(18.51)	(1.17)	(-0.67)	(3.39)

Table OA.2 shows that many VAR specifications are not able to obtain the one-month variance discount rate, and long-term variance expectations. In particular, specifications 1 through 3 fail in obtaining the one-month variance discount rate effectively. It follows that the default spread is an important variable in order to obtain this discount rate. Furthermore, specification 1 and 3 predict long-term variance expectations which exceed the realized variance. Finally, specification 4 does a good job in obtaining short-term variance discount rates as well as long-term variance expectations, and the results are very similar to the results of the benchmark specification of the VAR.

OA.2 VAR using quarterly data

In this section, I show that the results continue to hold if the analyses are done at the quarterly rather than monthly frequency. This is done to alleviate concerns that the estimates based on monthly frequency overstate the persistence of the variables and, therefore, create a bias in the variance expectations. The VAR is used to calculate quarterly stock market variance expectations and for this reason the variance swap rate with three months to maturity is the shortest maturity considered in this exercise. Table OA.3 presents the estimation results.

Table OA.3: This table shows the estimated coefficients of the VAR of equation (13) with t-values in parentheses. All variables are normalized to have mean equal to zero, and $pc_t^{(1)}$ and $pc_t^{(1)}$ are additionally standardized to have standard deviation equal to one. The sample period for the dependent variables is March 1996 to June 2019, with 94 quarterly data points.

Coefficients VAR model					
	rv_t	$pc_t^{(1)}$	$pc_t^{(2)}$	DEF_t	R^2
rv_{t+1}	-0.031	0.468	-0.267	0.448	0.459
(t-stat.)	(-0.17)	(3.07)	(-3.28)	(1.66)	
$pc_{t+1}^{(1)}$	-0.013	0.750	0.003	0.104	0.585
(t-stat.)	(-0.07)	(5.02)	(0.04)	(0.40)	
$pc_{t+1}^{(2)}$	-0.108	0.294	0.696	0.140	0.619
(t-stat.)	(-0.60)	(1.99)	(8.82)	(0.54)	
DEF_{t+1}	0.024	-0.025	-0.007	0.847	0.709
(t-stat.)	(0.46)	(-0.59)	(-0.33)	(11.50)	

Overall, the estimation results based on quarterly frequency are very similar to the results based on monthly frequency. Variance expectations and variance discount rates are calculated using these estimates and by adjusting equations (15) and (16) accordingly. Table OA.4 presents the results.

Table OA.4: This table shows the results of the variance decomposition of variance swap rates using equation (17), based on the VAR estimated on quarterly data. Note that the (co)variances of the third, fourth, and fifth columns are scaled with the variance of the second column such that the sum of the three (co)variances equals one.

T	$\text{var}(vs)$	$\frac{\text{var}(E_{\text{rv}})}{\text{var}(vs)}$	$\frac{\text{var}(E_{\text{vdr}})}{\text{var}(vs)}$	$\frac{-2\cdot\text{cov}(E_{\text{rv}},E_{\text{vdr}})}{\text{var}(vs)}$
18	0.220	0.218	0.665	0.117
12	0.227	0.434	0.505	0.060
6	0.274	0.824	0.206	-0.030
3	0.348	1.021	0.085	-0.107

The results of the decomposition of variance swap rates in Table OA.3 are remarkably close to the results of Table 4. Therefore, my results are robust whether the frequency of the VAR is monthly or quarterly. Finally, I also decompose the variance swap rate using the predictive regressions of equations (10) and (11). Table OA.5 presents the results.

Table OA.5: This table shows the results of the predictive regressions of equations (10) and (11) in which the variance swap price is the independent variable. The frequency of the data is quarterly. t-statistics are represented in parentheses and are computed using Newey-West standard errors with number of lags equal to $\frac{1}{3} \cdot T$.

Dependent variable:	$y_{\text{rv},t+T}$		$y_{\text{vdr},t+T}$	
Maturity	b_{rv} (t-stat.)	R^2	b_{vdr} (t-stat.)	R^2
18	0.240	0.029	-0.741	0.218
	(1.21)		(-3.76)	
12	0.540	0.129	-0.452	0.092
	(3.32)		(-2.71)	
6	0.836	0.303	-0.162	0.016
	(5.48)		(-1.04)	
3	0.964	0.417	-0.036	0.001
	(8.05)		(-0.30)	

The similarity between the results of Table 2 and Table OA.5 indicate that the results of predictive regressions are robust to decreasing the frequency to the quarterly level.

OA.3 Decomposition of variance swap returns

In this section, I show that the results in Section 6.1 for the variance swap return decomposition are robust to directly adding the variance swap return to the VAR. The VAR of equation (13) is estimated using the following state variables:

$$z_t = \left(r_t^{(T)} \quad rv_t \quad pc_t^{(1)} \quad pc_t^{(2)} \quad DEF_t \right)',$$

where $r_t^{(T)}$ are the returns on a variance swap with T-periods to maturity. Therefore, to decompose the returns of different maturities using identity (21), the VAR has to be re-estimated for each maturity. In this appendix, I show the estimation results of the VAR with returns on 12-month variance swaps $r_t^{(12)}$ and the results of the decomposition. The estimation results of the VAR are in Table OA.6.

Table OA.6: This table shows the estimated coefficients of the VAR of equation (13) with t-values in parentheses. All variables are normalized to have the mean equal to zero, and $pc_t^{(1)}$ and $pc_t^{(1)}$ are additionally standardized to have the standard deviation equal to one. The sample period for the dependent variables is January 1996 to June 2019, with 282 monthly data points.

Coefficients VAR model						
	$r_t^{(12)}$	rv_t	$pc_t^{(1)}$	$pc_t^{(2)}$	DEF_t	R^2
$r_{t+1}^{(12)}$	0.018	0.017	-0.043	-0.047	0.064	0.085
	(0.26)	(0.71)	(-2.02)	(-3.13)	(1.42)	
rv_{t+1}	0.357	0.029	0.511	-0.308	0.523	0.594
	(1.53)	(0.38)	(7.24)	(-6.15)	(3.49)	
$pc_{t+1}^{(1)}$	-0.039	0.037	0.863	-0.002	0.106	0.848
	(-0.27)	(0.76)	(19.76)	(-0.07)	(1.13)	
$pc_{t+1}^{(2)}$	0.260	-0.016	0.133	0.869	-0.017	0.746
	(1.39)	(-0.25)	(2.34)	(21.64)	(-0.14)	
DEF_{t+1}	0.092	0.008	-0.012	0.002	0.967	0.938
	(3.06)	(0.78)	(-1.28)	(0.36)	(46.63)	

The inclusion of $r_t^{(12)}$ into the VAR does not alter the models of the four other variables much. Only $r_t^{(12)}$ positively predicts the default spread in the next period. Low returns on 12-month variance swaps are predicted by a large level of the term structure of variance swap rates $pc_t^{(1)}$ and a large slope of the term structure of variance swap rates $pc_t^{(2)}$.

Using the estimates of this VAR, $\bar{N}^{(T)}_{\mathrm{rv},t}$ is obtained, as follows:

$$\bar{N}^{(T)}_{\mathrm{rv},t} = e_2'\Bigg(\big(1-\rho(T)\big) + \cdots + \big(1-\rho(1)\big)\rho(T) \times \cdots \times \rho(2)B^{T-1}\Bigg)\epsilon_t,$$

where e_2 is vector of zeros and the second element a one (as rv_t is the second element in z_t). Furthermore, $\bar{N}^{(T)}_{\mathrm{vdr},t}$ is using this VAR obtained in the following way:

$$\bar{N}^{(T)}_{\mathrm{vdr},t} = \bar{N}^{(T)}_{\mathrm{rv},t} - e_1' \cdot \epsilon_t.$$

The results of the decomposition of variance swap returns are shown in Table OA.7.

Table OA.7: This table shows the results of the variance decomposition of variance swap rates using equation (37). Note that the (co)variances of the third, fourth, and fifth columns are scaled with the variance of the second column such that the sum of the three (co)variances equals one. Standard errors are computed using the Delta method.

Maturity	$\mathrm{var}\big(e_1'\epsilon_t\big)$	$\mathrm{var}\big(\bar{N}^{(T)}_{\mathrm{rv},t}\big)$	$\mathrm{var}\big(\bar{N}^{(T)}_{\mathrm{vdr},t}\big)$	$-2\mathrm{cov}\big(\bar{N}^{(T)}_{\mathrm{rv},t}, \bar{N}^{(T)}_{\mathrm{vdr},t}\big)$
18	0.026	0.593	0.542	-0.134
		(0.265)	(0.224)	(0.348)
12	0.035	0.893	0.369	-0.262
		(0.307)	(0.165)	(0.379)
6	0.064	1.172	0.152	-0.324
		(0.236)	(0.068)	(0.276)
3	0.123	1.151	0.037	-0.188
		(0.147)	(0.014)	(0.156)
1	0.378	1.000	0.000	0.000

Note that the decomposition of the variance swap returns with 12 months to maturity is obtained using the estimates of the VAR of Table OA.6. To decompose the returns on variance swaps with T months to maturity, the VAR is re-estimated with $r_t^{(T)}$ as a state variable. Overall, the results of Table OA.7 are very similar to the results of the decomposition using only four state variables represented in Table 12.

Furthermore, each of the decomposition objects in identity 21 are directly compared using the methodology with five variables in the VAR and the method with only four variables. The results of the comparison are represented in the next table.

Table OA.8: This table shows the results of the comparison between the method in which the variance swap return is modeled directly in the VAR and the method in which the realization is obtain using equation (36). $e_1'\epsilon_t$, $\bar{N}_{\text{rv},t}^{(T)}$, and $\bar{N}_{\text{vdr},t}^{(T)}$ correspond to the variables obtained using the VAR with five variables. $e_L(T)'\epsilon_t$, $N_{\text{rv},t}^{(T)}$, and $N_{\text{vdr},t}^{(T)}$ correspond to the variables obtained using the VAR with four variables.

Maturity	$\text{corr}\big(e_1'\epsilon_t, e_L(T)'\epsilon_t\big)$	$\text{corr}\big(\bar{N}_{\text{rv},t}^{(T)}, N_{\text{rv},t}^{(T)}\big)$	$\text{corr}\big(\bar{N}_{\text{vdr},t}^{(T)}, N_{\text{vdr},t}^{(T)}\big)$
18	0.963	0.998	0.940
12	0.981	0.998	0.952
6	0.986	0.999	0.917
3	0.990	0.999	0.767
1	0.987	0.987	-

The correlations in Table OA.8 are all very high except for the correlation of the revised discount rate expectations of variance swap returns with a maturity of three months. However, the variance of this object is very small, and, therefore, a tiny deviation yields large changes in correlation as indicated by the large correlations in the other columns.

OA.4 Appendix asset pricing models

In the following subsections, I discuss more the models considered in this paper in detail. Section OA.4.1 discusses the model by Gabaix (2012), Section OA.4.2 discusses Wachter (2013), and Section OA.4.3 discusses Drechsler and Yaron (2011).

OA.4.1 Variable disaster risk and CRRA preferences

In this subsection, I discuss the variable rare disaster model of Gabaix (2012). I use the following specification from Dew-Becker et al. (2017):

$$\begin{aligned}\Delta c_{t+1} &= \mu_c + \sigma_c \epsilon_{c,t+1} + J_{c,t+1},\\ L_{t+1} &= \big(1-\rho_L\big)\bar{L} + \rho_L L_t + \sigma_L \epsilon_{L,t+1} \quad \text{and}\\ \Delta d_{t+1} &= \eta\sigma_c\epsilon_{c,t+1} - L_t \cdot \mathbb{1}_{J_{c,t}\neq 0},\end{aligned}$$

where $\epsilon_{c,t+1}, \epsilon_{L,t+1} \sim N(0,1)$ and $J_{c,t+1}$ is the jump process (rare disaster). The state variable L_t captures the exposure of the dividend process toward the rare disaster, and this exposure varies over time. During times when L_t is large, the stock market is affected more by consumption disasters than when L_t is low. The rare disaster process is modeled as a compound

Poisson process and is defined as follows:

$$J_t = \sum_{i=1}^{N_t} \xi_{i,t}, \text{ where } N_t \sim \text{Poisson}(\lambda_t) \text{ and } \xi_{i,t} \sim N\big(\mu_d, \sigma_d\big). \tag{OA.1}$$

Note that in the model by Gabaix (2012) $\lambda_t = \lambda$; that is the jump intensity does not vary over time. The representative agent in the model has power utility preferences with risk aversion parameter γ, which yields the following stochastic discount factor:

$$M_{t+1} = \delta\Big(\frac{C_{t+1}}{C_t}\Big)^{-\gamma},$$

where δ is the utility discount rate. I use the calibration from Dew-Becker et al. (2017), which is calibrated to match the risk premium on one-month variance swaps, and is given in Table OA.12 of Appendix OA.4.1.

In order to obtain an equation for the realized stock market variance, I use the following log-linear stock market return approximation:

$$r_{m,t+1} \approx \kappa_0 + \kappa_1 pd_{t+1} - pd_t + \Delta d_{t+1}, \tag{OA.2}$$

where κ_0, κ_1 are log-linearization constants and I use the follow approximation: $pd_t \approx z_0 + z_1 L_t$. The stock market return is then driven by two Gaussian shocks $\big(\epsilon_{c,t+1}$ and $\epsilon_{L,t+1}\big)$ and a jump shock $\big(L_t \cdot \mathbb{1}_{J_{c,t}\neq 0}\big)$. I follow Dew-Becker et al. (2017), who assume that, in the absence of a disaster, the shocks to consumption and the variable disaster have a deterministic variance. In the case of a disaster occurring, Dew-Becker et al. (2017) assume that the largest daily decline in the value of the stock market is $F\%$. Under these assumptions, realized variance over the next period equals the following:

$$RV_{t+1} = \kappa_1^2 z_1^2 \sigma_L^2 + \eta^2 \sigma_c^2 + F \cdot L_t \mathbb{1}_{J_{c,t}\neq 0}, \tag{OA.3}$$

where the first two summands correspond to the variance from the consumption process and variable rare disaster process, and the last summand is the realized variance from the jump process.

Equation (OA.3) offers a first insight into the drivers of variance risk in the model by Gabaix (2012). Realized variance depends on whether the consumption disaster hits the economy. Given that the consumption disaster is a (very) undersirable outcome of the agent, she is willing to pay a large price to hedge this risk. Furthermore, it follows from equation (OA.3) that the disaster size L_t drives variation in expected stock market variance.

The variance swap rate at time t with T months to maturity is computed as the sum of risk-neutral realized variances of equation (OA.3), as follows:

$$VS_t^{(T)} = \sum_{t=1}^{T} \mathbb{E}_t^{\mathbb{Q}}\big(RV_{t+i}\big) = T \cdot v_0 + v_1 \sum_{i=i}^{T} \mathbb{E}_t\big(L_{t+i-1}\big), \tag{OA.4}$$

where $v_0 = \kappa_1^2 z_1^2 \sigma_L^2 + \eta^2 \sigma_c^2$ is the diffusive variance and $v_1 = F \cdot \mathbb{E}^{\mathbb{Q}}\big(\mathbb{1}_{J_c \neq 0}\big)$. These equations show that the size of the disaster L_t also drives the risk premium embedded in the variance swap rate.

The calibration of the model by Gabaix (2012) is from Dew-Becker et al. (2017) and given in the following table.

Table OA.9: Calibration of the model by Gabaix (2012).

Parameter	Value	Parameter	Value
μ_c	$0.01/12$	σ_c	$0.02/\sqrt{12}$
μ_d	-0.3	σ_d	0.15
$\bar{L}$	$-\log(0.5)$	σ_L	0.04
ρ_L	$0.87^{1/12}$	η	5
β	$0.96^{1/12}$	γ	7
λ	$\frac{0.01}{12}$		

Note that in the calibration of Dew-Becker et al. (2017) the risk-aversion is raised to 7 in order to match the Sharpe ratio on one-month variance swaps.

In the following, I present more details of the results from the simulation study for the model by Gabaix (2012). First, I present sample statistics of realized variance and variance discount rates in the model. The results of this simulation study are represented in the following table.

Table OA.10: This table presents sample statistics of the realized variance and variance discount rates in the model by Gabaix (2012). The mean, standard deviation and Sharpe Ratio of the 18-, 12-, 6-, 3-, and 1-month simple variance discount rates are presented. The second column consists of the empirical result, and the third, fourth, and fifth columns represent the 5%, 50%, and 95% quantile of the simulation study, respectively.

Statistic	Data	Model		
	Est.	5%	50%	95%
Realized variance				
$\mathbb{E}(RV)$	0.162	0.115	0.116	0.117
$\sigma(RV)$	0.095	0.000	0.021	0.043
Variance discount rates				
$\mathbb{E}(r^{(18)})$	-0.006	-0.024	-0.021	-0.017
$\sigma(r^{(18)})$	0.187	0.021	0.056	0.109
$SR(r^{(18)})$	-0.106	-4.091	-1.278	-0.572
$\mathbb{E}(r^{(12)})$	-0.013	-0.037	-0.032	-0.025
$\sigma(r^{(12)})$	0.227	0.021	0.080	0.161
$SR(r^{(12)})$	-0.202	-5.976	-1.326	-0.578
$\mathbb{E}(r^{(6)})$	-0.050	-0.075	-0.063	-0.050
$\sigma(r^{(6)})$	0.316	0.023	0.155	0.319
$SR(r^{(6)})$	-0.544	-11.028	-1.362	-0.584
$\mathbb{E}(r^{(3)})$	-0.098	-0.150	-0.127	-0.099
$\sigma(r^{(3)})$	0.447	0.028	0.308	0.636
$SR(r^{(3)})$	-0.756	-17.644	-1.370	-0.585
$\mathbb{E}(r^{(1)})$	-0.285	-0.451	-0.380	-0.296
$\sigma(r^{(1)})$	0.676	0.068	0.926	1.905
$SR(r^{(1)})$	-1.458	-22.305	-1.366	-0.585

Table OA.10 confirms the finding of Figure 3 that the model by Gabaix (2012) is able to capture the strongly increasing term structure of expected variance swap returns documented in the data. Moreover, Table OA.10 shows that the volatility of variance swap returns varies a lot across simulation sets, and this results from the fact that the probability of a disaster is small (1% p.a.). If no disasters occur in a simulation set, the volatility of variance swap returns is very low. Finally, I conclude from Table OA.10 that the model by Gabaix (2012) is not able to capture the dynamics of empirical stock market volatility.

In the following, I decompose variance swap rates in the model by Gabaix (2012) for each simulation set seperately. Table OA.11 presents the results.

Table OA.11: This table presents the results of the simple variance decomposition of variance swap rates in the data and in the model by Gabaix (2012). The results of the data are from Table 2, with standard errors in parentheses. The regression coefficients of the model are estimated for each simulation set, and the mean and standard deviation of the regression coefficients are represented in the table.

Maturity	Data		Model	
	b_{rv}	b_{vdr}	b_{rv}	b_{vdr}
18	0.245	-0.728	0.002	-0.985
	(0.185)	(0.188)	(0.028)	(0.125)
12	0.558	-0.419	0.002	-0.989
	(0.151)	(0.158)	(0.028)	(0.101)
6	0.833	-0.168	0.002	-0.994
	(0.119)	(0.122)	(0.028)	(0.069)
3	0.957	-0.040	0.002	-0.997
	(0.083)	(0.086)	(0.027)	(0.047)
1	1.101	0.101	0.002	-0.998
	(0.056)	(0.056)	(0.027)	(0.027)

Table OA.11 shows that the result of Figure 1 is stable across the simulation sets. In particular, short-term variance swap rates are solely driven by variance discount rates, and this number is very similar across simulations, and, therefore, it is strong evidence that the model is not in line with the data.

In the following subsection, I discuss the model by Wachter (2013).

OA.4.2 Time-varying disaster risk and Epstein-Zin preferences

In this subsection, I discuss a discrete version of the model by Wachter (2013). Similar to Gabaix (2012), the consumption disaster risk varies over time. However, in the model by Wachter (2013), the disaster intensity, rather than the disaster size, varies over time. Furthermore, the agent in the model has preferences as in Epstein and Zin (1989), rather than CRRA preferences.

Consumption and dividend growth in the model are given by

$$\Delta c_{t+1} = \mu_c + \sigma_c \epsilon_{c,t+1} + J_{t+1} \quad \text{and}$$
$$\Delta d_{t+1} = \eta \Delta c_{t+1},$$

where $\epsilon_c \sim N(0,1)$ and J_t is a compound-Poisson as in equation (OA.1) of the model by Gabaix (2012). However, in this model the intensity of the consumption disaster is time-varying and follows the following square-root process:

$$\lambda_{t+1} = \phi\lambda_t + (1-\phi)\mu_\lambda + \sigma_\lambda\sqrt{\lambda_t}\epsilon_{\lambda,t+1},$$

where $\epsilon_{\lambda,t} \sim N(0,1)$. The investor has Epstein-Zin utility with elasticity of intertemporal substitution (EIS) equal to one and, therefore, is the log-utility given by

$$v_t = (1-\beta)c_t + \frac{\beta}{1-\alpha}\log\mathbb{E}_t\exp\big(v_{t+1}(1-\alpha)\big),$$

where β is the utility discount rate and $\gamma = 1-\alpha$ is the risk aversion parameter. The calibration of the model is from Dew-Becker et al. (2017) and is given in Table OA.12 of Appendix OA.4.2.

An equation for the realized variance in the model by Wachter (2013) follows from the log-linear market return which is given by

$$r_{m,t+1} \approx \kappa_0 + \kappa_1 pd_{t+1} - pd_t + \Delta d_{t+1},$$

where κ_0, κ_1 are log-linearization constants for the log-market return and pd_t is the log price-dividend ratio and is approximately linear in the state variable: $pd_t \approx z_0 + z_1\lambda_t$. Under these assumptions, realized variance in this model given by

$$RV_{t+1} = \eta^2\sigma_c^2 + \kappa_1^2 z_1^2\sigma_\lambda^2\lambda_t - F\eta J_{t+1}, \tag{OA.5}$$

where the first two summands correspond the variances of the diffusive shocks $\epsilon_{c,t+1}$ and $\epsilon_{\lambda,t+1}$ and the last summand corresponds to the realized variance from the consumption disaster.

Equation OA.5 offers a first insight into the pricing of variance risk in the model by Wachter (2013). The first summand of equation (OA.3) is constant over time; however, the second summand scales with the level of the intensity of the consumption disaster. This results from the fact that the disaster intensity follows a square-root process, which indicates that future variance of the disaster intensity scales with the current level of the disaster intensity. Therefore, even in the absence of consumption disasters, stock market variance is time-varying in this model. This result is different from the model by Gabaix (2012) in which the variance of the stock market only varies if a disaster hits the economy. The third summand of equation (OA.5) corresponds to the stock market variance that follows from the disaster process.

Variance swap rates are computed as the risk-neutral expectation of the sum of realized variances of equation (OA.5) of period $t+1$ until $t+T$, as follows:

$$VS_t^{(T)} = \sum_{t=1}^{T} \mathbb{E}_t^{\mathbb{Q}}\big(RV_{t+i}\big) = T \cdot v_0 + v_1 \sum_{i=1}^{T} \mathbb{E}^{\mathbb{Q}}(\lambda_{t+i-1}), \tag{OA.6}$$

where $v_0 = \eta^2\sigma_c^2$ and $v_1 = \kappa_1^2 z_1^2 \sigma_\lambda^2 - F\eta \exp\big(-\alpha\mu_d + \frac{1}{2}\alpha\sigma_d^2\big)(\mu_d - \alpha\sigma_d^2)$. Due to the Epstein-Zin preferences of the agent, the risk-neutral dynamics of the disaster intensity are different from the real-world dynamics in the sense that states with low lifetime utility, which correspond to states with high disaster intensity, receive a larger risk-neutral probability. This yields the agent a premium for instruments that offer protection against states in which disaster intensity is high, and this feature is not present in a model with CRRA preferences.

The calibration of the model is given in Table OA.12.

Table OA.12: This table shows the calibration of the model by Wachter (2013).

Parameter	Value	Parameter	Value
μ_c	$0.0252/12$	σ_c	$0.02/\sqrt{12}$
μ_d	-0.15	σ_d	0.10
μ_λ	$0.0355/12$	σ_λ	$0.067/12$
ϕ	$\exp(-0.08/12)$	β	$\exp(-0.012/12)$
η	2.6	γ	$4.9 = 1-\alpha$

Note that in the calibration of Dew-Becker et al. (2017) the risk-aversion is raised to 4.9 in order to match the Sharpe ratio on one-month variance swaps as closely as possible.

In the following, I present more details of the results from the simulation study for the model by Wachter (2013). First, I present sample statistics of realized variance and variance discount rates in the model. The results of this simulation study are represented in the following table.

Table OA.13: This table presents sample statistics of the realized variance and variance discount rates in the model by Wachter (2013). The mean, standard deviation, and Sharpe ratio of the 18-, 12-, 6-, 3-, and 1-month simple variance discount rates are presented. The second column consists of the empirical result, and the third, fourth, and fifth columns represent the 5%, 50%, and 95% quantile of the simulation study, respectively.

Statistic	Data	Model		
	Est.	5%	50%	95%
Realized variance				
$\mathbb{E}(RV)$	0.162	0.094	0.123	0.168
$\sigma(RV)$	0.095	0.028	0.048	0.073
Variance discount rates				
$\mathbb{E}(r^{(18)})$	-0.006	-0.015	-0.010	-0.005
$\sigma(r^{(18)})$	0.187	0.080	0.101	0.179
$SR(r^{(18)})$	-0.106	-0.556	-0.357	-0.097
$\mathbb{E}(r^{(12)})$	-0.013	-0.019	-0.014	-0.005
$\sigma(r^{(12)})$	0.227	0.084	0.118	0.256
$SR(r^{(12)})$	-0.202	-0.721	-0.412	-0.075
$\mathbb{E}(r^{(6)})$	-0.050	-0.032	-0.024	-0.007
$\sigma(r^{(6)})$	0.316	0.086	0.183	0.502
$SR(r^{(6)})$	-0.544	-1.284	-0.464	-0.048
$\mathbb{E}(r^{(3)})$	-0.098	-0.060	-0.045	-0.009
$\sigma(r^{(3)})$	0.447	0.074	0.338	1.008
$SR(r^{(3)})$	-0.756	-2.674	-0.459	-0.031
$\mathbb{E}(r^{(1)})$	-0.285	-0.173	-0.127	-0.018
$\sigma(r^{(1)})$	0.676	0.045	0.993	3.058
$SR(r^{(1)})$	-1.458	-12.294	-0.442	-0.019

Table OA.13 confirms the finding of Figure 3 that the model by Wachter (2013) is not able to capture the strongly increasing term structure of expected variance swap returns documented in the data. Moreover, Table OA.13 shows that also in the model by Wachter (2013) the volatility of variance swap returns varies a lot across simulation sets, and this results from the fact that the probability of a disaster is, on average, small (3.55% p.a.). If no disasters occur in a simulation set, the volatility of variance swap returns is very low. Finally, I conclude from Table OA.10 that the model by Wachter (2013) does a better job than the model by Gabaix (2012) of capturing the empirical dynamics of stock market volatility.

In the following, I decompose variance swap rates in the model by Wachter (2013) for each simulation set separately. Table OA.14 presents the results.

Table OA.14: This table presents the results of the simple variance decomposition of variance swap rates in the data and in the model by Wachter (2013). The results of the data are from Table 2, with standard errors in parentheses. The regression coefficients of the model are estimated for each simulation set, and the mean of the regression coefficients is represented in the table with the standard deviation in parentheses.

Maturity	Data		Model	
	b_{rv}	b_{vdr}	b_{rv}	b_{vdr}
18	0.245	-0.728	0.973	-0.037
	(0.185)	(0.188)	(0.040)	(0.040)
12	0.558	-0.419	0.963	-0.033
	(0.151)	(0.158)	(0.029)	(0.032)
6	0.833	-0.168	0.954	-0.033
	(0.119)	(0.122)	(0.021)	(0.024)
3	0.957	-0.040	0.950	-0.039
	(0.083)	(0.086)	(0.019)	(0.020)
1	1.101	0.101	0.948	-0.052
	(0.056)	(0.056)	(0.018)	(0.018)

Table OA.14 confirms the finding of Figure 1 that variance swap rates are driven by variance expectations in the model by Wachter (2013). Moreover, this result is very stable across the simulation sets, as indicated by the low standard deviation of b_{rv}. Therefore, this is strong evidence that the model is not in line with the data because my analysis shows that long-term variance swaps are mostly driven by variance discount rates.

In the following subsection, I discuss long-run risk model by Drechsler and Yaron (2011).

OA.4.3 Long-run risk

In this subsection, I discuss the long-run risk model by Drechsler and Yaron (2011). This model is a generalization of the long-run risk model by Bansal and Yaron (2004) in order to incorporate stylized facts regarding the variance risk premium. The model is generalized in the sense that the long-run mean consumption growth and the stochastic volatility incorporate jump shocks. Moreover, the long-run mean of the stochastic volatility process varies over time. The agent in the model has Epstein-Zin preferences, as is standard in long-run risk models. An important difference between the long-run risk model and the previously discussed consumption disaster models is that there are no consumption disasters in the long-run risk model. However, the state variables, which govern the future consumption growth rate and future consumption volatility, are exposed to jump risk.

Drechsler and Yaron (2011) specify the state vector of the economy as a VAR with Gaussian and jump shocks, as follows:

$$Y_{t+1} = \begin{pmatrix} \Delta c_{t+1} \\ x_{t+1} \\ \bar{\sigma}^2_{t+1} \\ \sigma^2_{t+1} \\ \Delta d_{t+1} \end{pmatrix} = \mu + FY_t + G_t z_{t+1} + J_{t+1}, \tag{OA.7}$$

where, μ is a vector with the means of each state variable, F is specified as follows:

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & \rho_x & 0 & 0 & 0 \\ 0 & 0 & \rho_{\bar{\sigma}} & 0 & 0 \\ 0 & 0 & (1-\tilde{\rho}_\sigma) & \rho_\sigma & 0 \\ 0 & \phi & 0 & 0 & 0 \end{pmatrix}, \tag{OA.8}$$

$G_t G_t'$ is the variance-covariance matrix, $z_{t+1} \sim N(0, I)$ is a vector of Gaussian shocks, and J_{t+1} is a vector of jump shocks. Jumps are compound-Poisson as in equation (OA.1) with intensity λ_t, which can vary over time, similar to the model by Wachter (2013). Drechsler and Yaron (2011) consider a specification with jumps in x_t and σ^2_t, where $J_{x,t}$ is compound normal distributed and $J_{\sigma,t}$ is compound gamma distributed.

The first and last element of Y_t are the consumption and dividend growth, respectively. These processes have a time-varying mean, which is driven by the persistent process x_t, the second element of Y_t. The third element of Y_t is the long-run mean $\bar{\sigma}^2_t$ of the stochastic volatility process σ^2_t, the fourth element of Y_t.

The variance-covariance matrix, $G_t G_t'$, which governs the stochastic volatility of the model, and the jump intensity, λ_t, are affine in the state variable σ^2_t:

$$\begin{aligned} G_t G_t' &= h + H_\sigma \sigma^2_t \quad \text{and} \\ \lambda_t &= l_0 + l_1 \sigma^2_t, \end{aligned}$$

and, therefore, all variation in either the jump intensity or stochastic volatility is driven by σ^2_t.

The representative agent in the model has Epstein-Zin utility for which the stochastic discount factor is given by

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{c,t+1},$$

where $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, δ is the utility discount rate, γ is the risk aversion, ψ is the EIS, and $r_{c,t+1}$, the return on wealth. Drechsler and Yaron (2011) solve a log-linear version of the model and use $pd_{t+1} \approx A_{0,m} + A'_m Y_{t+1}$, which says that the log price-dividend ratio is linear in the state variables. Under these conditions is the log-linearized market return, written as follows:

$$r_{m,t+1} = r_0 + \left(B'_r F - A'_m\right) + B'_r G_t z_{t+1} + B'_r J_{t+1}, \tag{OA.9}$$

where $A_{0,m}$, A'_m, r_0 and B'_r are given in equations (8) and (9) of Drechsler and Yaron (2011). Realized variance during period $t+1$ is equal to

$$RV_{t+1} = B'_r h B_r + B'_r H_\sigma \sigma_t^2 B_r + B'_r J_{t+1} J'_{t+1} B_r. \tag{OA.10}$$

The assumption underlying this realized variance equation is that the Gaussian shocks z_{t+1} occur diffusively during period $t+1$, while jumps happen on a single day.

Equation (OA.10) offers a first insight into the pricing of variance risk in the model by Drechsler and Yaron (2011). The first summand corresponds to the constant variance coming from the Gaussian shocks in the model. The second summand corresponds to the stochastic variance coming from the Gaussian shocks for which the variance is governed by the state variable σ_t^2. Finally, the third summand corresponds to the realized variance coming from the jump realizations in the state variables x_t and σ_t^2. Similar to the model by Wachter (2013) is time-variation in the realized variance coming from stochastic variance of Gaussian shocks and from the jump shocks.

The variance swap rate at time t with maturity T is computed as the risk-neutral expectation of the realized variance of equation (OA.10), as follows:

$$VS_t^{(T)} = \sum_{t=1}^{T} \mathbb{E}_t^{\mathbb{Q}}\left(RV_{t+i}\right) = T \cdot v_0 + v_1 \sum_{t=1}^{T} \mathbb{E}_t^{\mathbb{Q}}\left(\sigma_{t+i-1}^2\right),$$

where,

$$v_0 = B'_r h B_r \text{ and } v_1 = B'_r H_\sigma B_r + l_1 \cdot B'_r \Psi^{\mathbb{Q}} B_r.$$

In the last equation, $\Psi^{\mathbb{Q}}$ is a matrix that has the risk-neutral variance of the disaster realization on the diagonal and corresponds to equation (21) of Drechsler and Yaron (2011). In order to derive these equations, I used $l_{0,x} = l_{0,\sigma} = 0$ and $l_{1,x} = l_{1,\sigma}$ from the calibration of Drechsler and Yaron (2011). The full calibration of the model is from Drechsler and Yaron (2011) and presented in Table 5 of their paper, and I use the calibration in which jump shocks in the x_t process follow a compound-Poisson in combination with a normal distribution.

In the following, I present more details of the results from the simulation study for the model by Drechsler and Yaron (2011).[1] First, I present sample statistics of realized variance and variance discount rates in the model. The results of this simulation study are represented in the following table.

Table OA.15: This table presents sample statistics of the realized variance and variance discount rates in the model by Drechsler and Yaron (2011). The mean, standard deviation, and Sharpe ratio of the 18-, 12-, 6-, 3-, and 1-month simple variance swap returns are presented. The second column consists of the empirical result, and the third, fourth, and fifth columns represent the 5%, 50%, and 95% quantile of the simulation study, respectively.

Statistic	Data	Model		
	Est.	5%	50%	95%
Realized variance				
$\mathbb{E}(RV)$	0.162	0.157	0.169	0.187
$\sigma(RV)$	0.095	0.051	0.087	0.134
Variance discount rates				
$\mathbb{E}(r^{(18)})$	-0.006	-0.036	-0.026	-0.014
$\sigma(r^{(18)})$	0.187	0.191	0.276	0.387
$SR(r^{(18)})$	-0.106	-0.654	-0.333	-0.128
$\mathbb{E}(r^{(12)})$	-0.013	-0.045	-0.032	-0.015
$\sigma(r^{(12)})$	0.227	0.232	0.343	0.488
$SR(r^{(12)})$	-0.202	-0.659	-0.326	-0.109
$\mathbb{E}(r^{(6)})$	-0.050	-0.063	-0.043	-0.017
$\sigma(r^{(6)})$	0.316	0.304	0.477	0.734
$SR(r^{(6)})$	-0.544	-0.686	-0.314	-0.084
$\mathbb{E}(r^{(3)})$	-0.098	-0.089	-0.060	-0.020
$\sigma(r^{(3)})$	0.447	0.394	0.671	1.144
$SR(r^{(3)})$	-0.756	-0.736	-0.309	-0.064
$\mathbb{E}(r^{(1)})$	-0.285	-0.176	-0.116	-0.027
$\sigma(r^{(1)})$	0.676	0.708	1.352	2.697
$SR(r^{(1)})$	-1.458	-0.820	-0.292	-0.036

[1] I thank Friedrich Lorenz for sharing the codes to solve the model.

Table OA.15 confirms the finding of Figure 3 that the model by Drechsler and Yaron (2011) is not able to capture the strongly increasing term structure of expected variance swap returns documented in the data. Moreover, it shows that the model predicts, for each maturity, a volatility of variance swap returns, which is larger than observed empirically. Finally, I conclude from Table OA.10 that the model by Drechsler and Yaron (2011) does a good job of capturing the empirical dynamics of stock market volatility.

In the following, I decompose variance swap rates in the model by Drechsler and Yaron (2011) for each simulation set separately. Table OA.16 presents the results.

Table OA.16: This table presents the results of the simple variance decomposition of variance swap rates in the data and in the model by Drechsler and Yaron (2011). The results of the data are from Table 2, with standard errors in parentheses. The regression coefficients of the model are estimated for each simulation set and the mean of the regression coefficients are represented in the table with the standard deviation in parentheses.

Maturity	Data		Model	
	b_{rv}	b_{vdr}	b_{rv}	b_{vdr}
18	0.245	-0.728	0.349	-0.560
	(0.185)	(0.188)	(0.105)	(0.108)
12	0.558	-0.419	0.412	-0.502
	(0.151)	(0.158)	(0.105)	(0.107)
6	0.833	-0.168	0.506	-0.418
	(0.119)	(0.122)	(0.097)	(0.098)
3	0.957	-0.040	0.567	-0.379
	(0.083)	(0.086)	(0.087)	(0.086)
1	1.101	0.101	0.615	-0.385
	(0.056)	(0.056)	(0.075)	(0.075)

Table OA.16 confirms the finding of Figure 1 that short-term variance swap rates are driven by variance expectations and long-term variance swap rates by variance discount rates. Moreover, this result is stable across the simulation sets, as indicated by the low standard deviations of b_{rv} and b_{vdr}. Therefore, this is strong evidence that the model predicts a variation in short-term variance discount rates, which is substantially larger than observed empirically.

References

Bansal, R. and Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509.

Campbell, J. Y., Giglio, S., Polk, C. and Turley, R. (2018). An intertemporal CAPM with stochastic volatility. *Journal of Financial Economics*, 128(2):207–233.

Dew-Becker, I., Giglio, S., Le, A. and Rodriguez, M. (2017). The price of variance risk. *Journal of Financial Economics*, 123(2):225–250.

Drechsler, I. and Yaron, A. (2011). What's vol got to do with it. *The Review of Financial Studies*, 24(1):1–45.

Epstein, L. and Zin, S. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.

Gabaix, X. (2012). Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. *The Quarterly Journal of Economics*, 127(2):645–700.

Wachter, J. (2013). Can time varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, 68(3):987–1035.