Volatility and the Pricing Kernel*

David Schreindorfer
W.P. Carey School of Business
Arizona State University

Tobias Sichert
Department of Finance
Stockholm School of Economics

August 28, 2023

Abstract

Negative stock market returns are significantly more painful to investors when they occur in periods of low volatility. To establish this fact, we show that a drop in volatility makes the pricing kernel steeper and average put option returns more negative. Asset pricing theories based on habits and long-run risks imply that the pricing of stock market risk does not vary with volatility, or that it moves in the opposite direction. An explanation of our finding that is consistent with prior empirical evidence, as well as the model of Gabaix (2012), is that volatility evolves independently of the pricing kernel.

JEL Classification: G12, G13, G33

Keywords: Pricing kernel, volatility, equity index options, tail risk, risk-return trade-off, recovery, habits, long-run risks, rare disasters.

*Contact: david.schreindorfer@asu.edu and tobias.sichert@hhs.se. We thank our discussants Ian Dew-Becker, Jens Christensen, Kris Jacobs, Christian Skov Jensen, Mete Kilic, Dmitriy Muravyev, and Paola Pederzoli, as well as Caio Almeida, Tyler Beason, Oliver Boguth, Christian Julliard, Lars-Alexander Kuehn, Jonathan Payne, Seth Pruitt, Robert Ready, Guillaume Roussellet, Ivan Shaliastovich, Per Strömberg, and seminar participants at Arizona State, Stockholm School of Economics, McGill, Princeton, Carnegie Mellon, University of Oregon, Goethe University, ITAM, and the University of Georgia, the Virtual Derivatives Workshop, the JEF seminar, the BI-SHoF Conference 2022, the CICF 2022, the SoFiE Conference 2022, the EFA 2022, the MFA 2023, and the SGF Conference 2023 for helpful comments and suggestions. Tobias Sichert acknowledges support from the Jan Wallanders and Tom Hedelius Foundation and the Tore Browaldhs Foundation, grant no. Fh21-0026.

The pricing of stock market risks is central for our understanding of optimal investment decisions. In this paper, we document that negative stock market returns are significantly more painful to investors when they occur in periods of low volatility. Specifically, low volatility is associated with (i) a steeper pricing kernel and (ii) larger risk premia on out-of-the-money put options. We show that this fact is inconsistent with most theoretical explanations for time-varying volatility.

To illustrate our finding, it is essential to distinguish between the pricing kernel, M_{t+1}, which is potentially a function of many different shocks, and its projection onto stock market returns, $E_t[M_{t+1}|R_{t+1}]$, which is only a function of returns and investors' time-t information set. The "projected pricing kernel" reveals how marginal utility is expected to vary with returns, and it has the same pricing implications as M_{t+1} for any claim on the market. We propose a maximum likelihood estimator of the projection based on options and return data. The estimator requires no distributional assumptions about returns and allows us to condition on ex ante stock market volatility. Figure I shows our estimate for the 10th and 90th percentile of volatility. The steeper curve in periods of low volatility implies that negative returns are significantly more painful to investors in calm markets.

This time-varying pricing of stock market risks is also evident in index option returns. While it is well-known that out-of-the-money put options earn very large negative returns on average (Rubinstein 1994), we document the complimentary fact that put returns become even more negative in calm markets. Like our estimates of the projected pricing kernel, this finding suggests that investors are more concerned about negative returns in times of low volatility.

The finding in Figure I is surprising from the perspective of prior microeconomic evidence, which shows that individuals become more risk averse in recessions (Cohn et al. 2015; Guiso et al. 2018). All else equal, higher risk aversion in volatile markets makes the pricing kernel (and its projection onto returns) steeper, rather than flatter. Our finding may also appear surprising based on the fact that options become more valuable when volatility rises. All else equal, higher put prices map to lower average

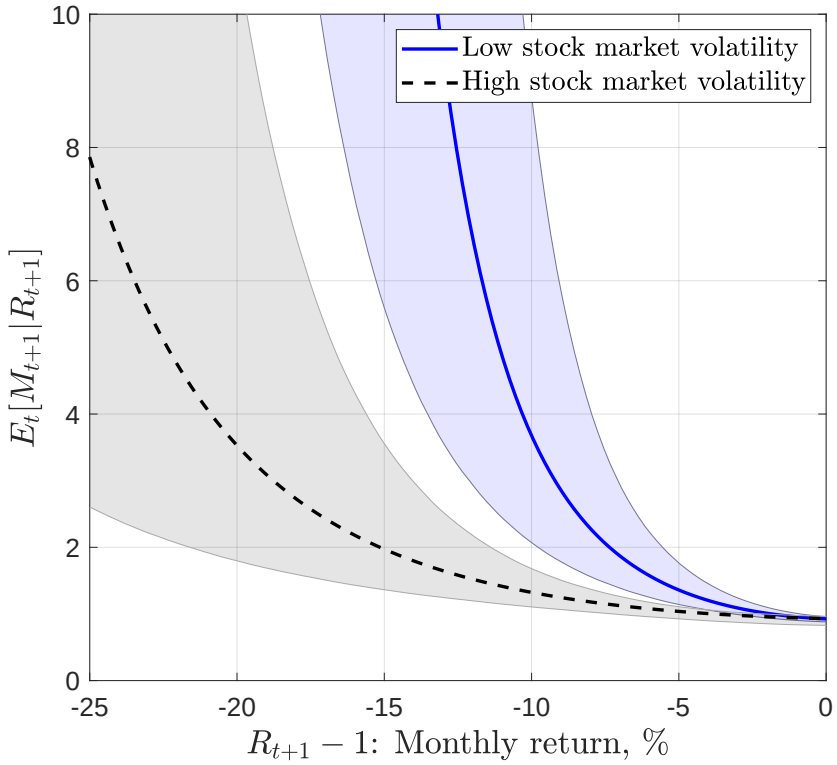

Figure I: Volatility and the projected pricing kernel. We plot the projected pricing kernel, $E_t[M_{t+1}|R_{t+1}]$, for the 10th and 90th percentile of conditional stock market volatility. The pricing kernel is measured at a monthly horizon, parameterized by equations (4) and (5) with a polynomial order of $N = 2$, and estimated over the 1990-2019 sample. Shaded areas represent pointwise 90% confidence bounds.

put returns, yet we document empirically that the opposite relation hold true. On the other hand, our finding is consistent with the absence of a risk-return trade-off in the time series of stock market returns (Glosten et al. 1993). If a rise in risk (stock market volatility) is accompanied by a sufficiently large drop in risk prices (a flatter projected pricing kernel), expected returns will not vary with volatility.

To understand how the time-varying pricing of stock market risks relates to the joint dynamics of volatility and the pricing kernel (before projecting it onto returns), we derive $E_t[M_{t+1}|R_{t+1}]$ under the assumption that returns and the pricing kernel are conditionally jointly lognormal. This setting is stylized, but it encompasses the models of Campbell and Cochrane (1999) and Bansal and Yaron (2004) and

the intuition it provides carries over to models with non-normal shocks. We show analytically that a rise in stock market volatility makes the projected pricing kernel flatter, as in the data, *if it is not accompanied by a rise in the conditional volatility of the pricing kernel*. This finding has important implications for structural models.

We show that, in contrast to Figure I, the habit model of Campbell and Cochrane (1999) implies that negative returns are less painful in calm markets, whereas the long-run risks model of Bansal and Yaron (2004) implies that investors are indifferent to the timing of negative returns. The same holds true for other models with recursive utility and persistent state variables, including Drechsler and Yaron (2011), Wachter (2013), and Constantinides and Ghosh (2017), which closely resemble the original long-run risks model in terms of their implications for the projected pricing kernel. We argue that the models' counterfactual implications for the pricing kernel are an inherent feature of their economic mechanisms, which rely on positive comovement between the conditional volatilities of returns and the pricing kernel to rationalize asset price dynamics. Without this tight connection, neither model would be able to explain the countercyclical nature of volatility, the leverage effect of Black (1976), or the long-horizon predictability of excess returns.

More encouragingly, we find that the model of Gabaix (2012) is consistent with our empirical evidence. In his model, stocks' exposure to consumption disasters varies over time and the dynamics of this exposure are specified to be independent of the pricing kernel. As a result, the volatility of returns evolves independent of the pricing kernel's volatility and the projection becomes steeper when volatility falls. One explanation for our finding is therefore that stock market volatility evolves independently of the pricing kernel. This explanation is consistent with empirical evidence in Dew-Becker et al. (2017), who show that shocks to expected (as opposed to realized) stock market variance command no risk premium, Jurado et al. (2015), who show that stock market volatility is only weakly correlated with macroeconomic uncertainty (a plausible driver of pricing kernel volatility), and (Glosten et al. 1993), who show that volatility does not predict stock market returns in the time series.

Related Literature. We build on a large literature that studies the pricing kernel's projection onto stock market returns based on index options, starting with Aït-Sahalia and Lo (2000), Jackwerth (2000), and Rosenberg and Engle (2002). The central finding of these studies is that the pricing kernel is a non-monotonic function of stock market returns – an observation dubbed the "pricing kernel puzzle" due to its inconsistency with standard models.[1] With two exceptions that we discuss below, however, the literature has not systematically examined time-variation in the pricing kernel's conditional distribution, and it has not connected properties of the projected pricing kernel to macro-finance puzzles.

Methodologically, we build on estimation approaches that take the risk-neutral distribution implied by option prices as given and then select parameters of $E[M|R]$ via a criterion function based on realized returns. Prior papers have implemented this idea in somewhat unconventional estimation frameworks: Bliss and Panigirtzoglou (2004) maximize the p-value of a Berkowitz test for uniformity and independence of returns, whereas Linn et al. (2018) minimize a generalized method of moments criterion for moments of the conditional CDF of returns. We follow the same general idea, but recognize that estimation can be performed via standard maximum likelihood. In particular, once a functional form has been specified for $E[M|R]$, the likelihood function of returns can be computed without additional parametric assumptions about the density of returns.

In terms of results, Bliss and Panigirtzoglou (2004) focus on estimating the representative investor's risk aversion, but show as an auxiliary result that estimates are higher in subsamples with low volatility. Our finding is consistent with this result, but differs along three key dimensions. First, Bliss and Panigirtzoglou's estimates are based on specific utility functions that imply a near-linear projected pricing kernel. In contrast, we model the projection with a flexible polynomial and find that it is strongly convex in returns. Our estimates show that linear

[1] In the online appendix, we show that our estimates are consistent with the projected pricing kernel's non-monotonicity. The main text focuses on the negative return region to draw attention to the novel fact we document – covariation with volatility – and away from the existing fact.

specifications are inconsistent with the large variance risk premium in the data, as well as the central importance of stock market tail events for the equity premium (Beason and Schreindorfer 2022). Second, Bliss and Panigirtzoglou interpret their auxiliary finding as showing that risk aversion rises in times of low volatility. We argue in stark contrast that the data is most consistent with stock market volatility evolving largely independently from the pricing kernel and risk aversion. Third, we use the projected pricing kernel to shed new light on the economic mechanisms of equilibrium models, whereas Bliss and Panigirtzoglou's study is purely empirical.

A contemporaneous and independently developed paper by Kim (2022) also studies time-variation in the projected pricing kernel. Kim shows how the projection varies with many different macroeconomic covariates (including volatility) and analyzes implications for time-variation in conditional risk premia. Similarly, Driessen et al. (2020) study the time-variation of the projected pricing kernel with respect to the CFNAI index and find relatively small effects. We focus on covariation with volatility only, but additionally explore its economic drivers theoretically by linking it to the joint dynamics of returns and the pricing kernel. Empirically, Kim finds qualitatively similar covariation of the projection with volatility as we do, but the effect he estimates is noticeably smaller. We show that this is a result of how volatility-dependence is modelled. Specifically, the nonlinear functional form we propose leads to substantially better fit to returns than the linear specification in Kim (2022), despite being more parsimonious.

Our theoretical result that market volatility must evolve close to independent of the pricing kernel in order to explain time-variation in the projected pricing kernel is consistent with prior empirical work. In particular, Jurado et al. (2015) show that macroeconomic uncertainty is only weakly correlated with stock market volatility and considerably more persistent. Additionally, these authors find that an increase in macroeconomic uncertainty leads to sizable and protracted decline in real activity (production, hours, employment), whereas Berger et al. (2020) show that an increase in expected stock market volatility does not. It is therefore plausible that

stock market volatility is not an important determinant of marginal utility, whereas macroeconomic uncertainty is.

The idea to estimate time-variation in the pricing kernel is complementary to recent "recovery" research, started by Ross (2015). The general idea of this literature is to impose economically motivated restrictions on the pricing kernel and data-generating process in order to recover physical probabilities from risk-neutral probabilities based on options data. Our assumptions are statistical in nature, but implicitly share the goal of recovering conditional physical probabilities from options. The parametric structure we impose on the projected pricing kernel allows us to employ a likelihood-based estimation approach to recover physical densities that provide the best fit to the data.

Our estimation yields conditional return distributions as a byproduct. This allows us to shed new light on a number of findings in related papers that study time-variation in return moments based on options data. First, Gormsen and Jensen (2022) show that stock market returns become more left-skewed and fat-tailed in times of low volatility. We repeat their analysis separately for physical moments and moment risk premia, and find that the co-movement of option-implied moments results almost entirely from co-movement in risk premia, rather than co-movement in physical moments (see Section IA. VIII). This result is a useful complement to Gormsen and Jensen's finding because it provides additional information about its economic drivers. Second, Martin (2017) derives a lower bound for the conditional equity premium that can be computed from option prices. To do so, he assumes the "negative correlation condition", $\text{Cov}_t[M_{t+1}R_{t+1}, R_{t+1}] \leq 0$, and verifies that it holds in a wide range of asset pricing models. We evaluate the condition empirically based on our estimates of the pricing kernel and conditional return distribution and find that it is violated on about 1% of all days (see Section IA. VII). Importantly, the violations occur primarily during the financial crisis of 2008, precisely the time when Martin finds the equity premium to be particularly high. This finding suggests that the equity premium may not be as volatile as implied by Martin's estimates.

I. Estimation

This section explains our approach for estimating the pricing kernel as a function of stock market returns and conditional volatility, discusses data sources, and illustrates the robustness and statistical significance of our estimates. Throughout, the pricing kernel in period $(t+1)$ is denoted by M_{t+1}, the ex-dividend market return by R_{t+1}, and "t"- subscripts indicate moments and probability density functions that condition on investors' information set at time-t.

A. Estimation Approach

In the absence of arbitrage opportunities, the pricing kernel's projection onto stock market returns equals[2]

$$E_t[M_{t+1}|R_{t+1}] = \frac{1}{R_t^f}\frac{f_t^*(R_{t+1})}{f_t(R_{t+1})}, \tag{1}$$

where R_t^f is the risk-free rate and $f_t^*(R_{t+1})$ and $f_t(R_{t+1})$ denote the conditional risk-neutral and physical density of R_{t+1}, respectively. The projection measures the mean of M_{t+1} conditional on investors' information set at time-t and conditional on a (potential) return outcome at time-$(t+1)$. Apart from the market return, (1) therefore averages over all shocks that affect the pricing kernel at $(t+1)$. Importantly, (1) is generally a *nonlinear* conditional expectation function of R_{t+1} for any time-t information set, i.e., it is not a linear projection. Our estimation conditions on volatility as part of investors' time-t information set, as further detailed below.

To estimate $E_t[M_{t+1}|R_{t+1}]$, we extract f_t^* from option prices for each day of the sample based on the classic result of Breeden and Litzenberger (1978). This methodology is fairly standard and we refer interested readers to Appendix A for details. Next, we model the projection with a flexible parametric function of returns and the conditional return volatility, $M(R_{t+1}, \sigma_t; \theta)$, and combine it with (1) to

[2]The fact that the pricing kernel equals the ratio of risk-neutral to physical probabilities (scaled by R^f) is a well-known textbook result – see, e.g., Cochrane (2005), p. 51. We provide a derivation for the projected pricing kernel in the online appendix.

express the conditional physical density as

$$f_t(R_{t+1};\theta) = \frac{f_t^*(R_{t+1})}{R_t^f \times M(R_{t+1},\sigma_t;\theta)}. \tag{2}$$

Given a functional form for $M(R_{t+1},\sigma_t;\theta)$, the unknown parameter vector θ can be estimated by maximizing the log-likelihood of realized returns,

$$LL(\theta) = \sum_{t=1}^{T} \ln f_t(R_{t+1};\theta). \tag{3}$$

Our notation emphasizes f_t's dependence on the parameter vector θ, but it is important to note that the density does not belong to a known parametric family of distributions. Rather, it results from applying a (parametric) change-of-measure to the (nonparametric) risk-neutral distribution f_t^*, whose shape is completely flexible and implied by the market prices of equity index options.

Our maximum likelihood estimator is statistically efficient and it incorporates conditioning information from the entire risk-neutral distribution. Both features represent important advantages over moment-based estimation approaches. Furthermore, our estimator makes it straightforward to incorporate information from additional time-t conditioning variables, which we utilize to illustrate the robustness of our findings in Section I.E.

B. Parameterizing the Projected Pricing Kernel

We model the projection as an exponential polynomial,[3]

$$M(R_{t+1},\sigma_t;\theta) = \exp\left\{\delta_t + \sum_{i=1}^{N} c_{it} \times (\ln R_{t+1})^i\right\}, \tag{4}$$

where the polynomial coefficients c_{it} vary with volatility according to

$$c_{it} = \frac{c_i}{\sigma_t^{b\times i}}, \tag{5}$$

δ_t is a time-varying intercept, and $\theta = (b, c_1, ..., c_N)$. The intercept is calculated for each day of the sample to satisfy the theoretical restriction that $f_t(R_{t+1};\theta)$ integrates

[3]Prior papers that have modelled the pricing kernel as a polynomial include Chapman (1997), Dittmar (2002), Rosenberg and Engle (2002), and Jones (2006).

to one, i.e., δ_t does not represent a free parameter.[4,5] The conditional volatility σ_t is estimated with the heterogeneous autoregressive (HAR) model of Corsi (2009) based on intradaily return data– see Appendix B for details.

We experimented with different functional forms for the time-varying polynomial coefficients c_{it}, and found that (5) provides a very good fit (in terms of log-likelihood) despite its parsimony. Additionally, when we estimated a more flexible functional form for the relationship between c_{it}'s and σ_t with more free parameters, we found that its shape closely resembles the one in (5) – see Section I.B of the online appendix for details. This alternative specification for c_{it}'s is used to illustrate the robustness of our results in Section I.F. Lastly, (5) nests two interesting special cases. For $b = 0$, the projected pricing kernel equals a time-invariant function of *returns*,

$$M(R_{t+1}, \sigma_t; \theta) = \exp\left\{\delta_t + \sum_{i=0}^{N} c_i \times (\ln R_{t+1})^i\right\}, \tag{6}$$

i.e., the graph of $E[M|R]$ does not vary with volatility, apart from a small vertical shift induced by δ_t. For $b = 1$, the projected pricing kernel equals a time-invariant function of *standardized returns* (up to a vertical shift due to δ_t),

$$M(R_{t+1}, \sigma_t; \theta) = \exp\left\{\delta_t + \sum_{i=0}^{N} c_i \times \left(\frac{\ln R_{t+1}}{\sigma_t}\right)^i\right\}. \tag{7}$$

In this case, the graph of $E[M|R]$ scales horizontally and proportionally with volatility. Intermediate values of b allow $E[M|R]$ to change with volatility to varying degrees. To formally test whether $E[M|R]$ varies with volatility, we evaluate the hypothesis $H_0 : b = 0$.

[4]The intercept equals $\delta_t = -\ln R_t^f + \ln\left(\int_0^\infty f^* \times \exp\left\{-\sum_{i=1}^{N} c_{it} \times (\ln R_{t+1})^i\right\} dR_{t+1}\right)$, i.e., its value is implied by R_t^f, f_t^*, and the polynomial coefficients $(b, c_1, ..., c_N)$. We find δ_t for each date by evaluating this integral numerically. By substituting the expression for δ_t into (4) and then (4) into (2), it can be verified that f_t integrates to one.

[5]Instead of computing δ_t based on the theoretical restriction $\int f = 1$, one could add a time-varying intercept c_{0t} to polynomial (4) and model c_{0t} as a function of volatility. Since this approach does not guarantee $\int f = 1$, however, it becomes necessary to add a penalty for violations of the restriction to the objective function. In turn, doing so requires the researcher to make a (necessarily subjective) choice on the relative importance of the restriction and the fit to realized returns. Kim (2022) does so in the context of a moment-based estimation of the pricing kernel.

C. *Parameter Identification*

The pricing kernel controls the extent to which conditional real world probabilities differ from their risk-neutral counterparts. Specifically, (2) shows that $f_t(R)$ takes on smaller values than $f_t^*(R)$ for return regions where $M(R_{t+1}, \sigma_t; \theta) > 1/R_t^f$ and higher values where $M(R_{t+1}, \sigma_t; \theta) < 1/R_t^f$. Individual elements of $\theta = \{c_1, ..., c_N, b\}$ are therefore identified if they alter the shape of $E[M|R]$ in such a way that it better explains the relative likelihood of different return realizations. Since f_t^* does not vary with θ, one can equivalently think of parameters as being identified by risk premia: An increase in the mean of f_t is equivalent to a higher equity premium, an increase in the variance of f_t is equivalent to a higher (less negative) variance premium, etc.

Most elements of θ alter the shape of f_t in multiple ways relative to that of f_t^*. Nevertheless, it is useful to discuss the main sources of parameter identification. c_1, the slope of $E[M|R]$, controls the relative probabilities of negative and positive returns. If the slope is negative, for example, the left tail of f_t^* gets downweighted in computing f_t, whereas the right tail gets upweighted. c_1 is therefore identified by the mean of f_t and the likelihood of negative returns. c_2, the curvature of $E[M|R]$, controls the relative probabilities of small and large absolute returns. If the curvature is positive, both extreme tails of f_t get downweighted relative to the tails of f_t^*, whereas the center of the distribution gets upweighted. Hence, c_2 is identified by the variance of f_t and the likelihood of extreme returns. c_2 is also negatively related to the mean of f_t because f_t^* is left-skewed, so that the equity premium further aids in its identification. Similarly, c_3, c_4, etc. are identified by higher order moments of f_t. The scaling parameter b controls how parameters of $E[M|R]$ vary with volatility, and therefore the amount of time-variation in the probabilities of different returns. For $b > 0$, an increase in volatility makes the slope of $E[M|R]$ less negative and its curvature less positive. b is therefore identified by the amount of time-variation in the moments of f_t, relative to time-variation in the corresponding f_t^* moments. We illustrate these channels quantitatively in Table IA.I of the online appendix by showing the sensitivity of moments of f_t to individual parameters.

D. *Data*

We use the S&P 500 index as a proxy for the aggregate stock market and focus on a return horizon of one month (30 calendar days). Return data comes from the Center for Research in Security Prices (CRSP). Option price quotes for the estimation of f_t^* come from the Chicago Board Options Exchange (CBOE). Because this data limits our sample to the 30-year period from 1990 to 2019, we sample daily to maximize the efficiency of our estimates, i.e., we work with a daily sample of $T = 7,556$ overlapping monthly returns. The estimation of conditional return volatilities (detailed in Appendix A) relies on intra-daily price quotes for S&P 500 futures, which were purchased from TickData. We use quotes for the large futures contract (ticker "SP") from 1990 to 2002, and for the E-Mini Futures contract (ticker "ES") from 2003 to 2019, i.e., we use data for the more actively traded futures contract in each part of the sample. Lastly, we use interest rates data from the Federal Reserve Bank of St. Louis' FRED database for robustness tests.

E. *Estimation Results*

Table I shows estimates for the parameterized pricing kernel in (4) and (5), and polynomial orders between $N = 1$ and $N = 5$. To account for autocorrelation that results from the use of overlapping return data, we determine the statistical significance of our estimates based on a block bootstrap with a block length of 21 trading days.[6]

The estimation results are easily summarized. The volatility-scaling parameter b is positive and significantly different from zero for all polynomial orders, and its significance grows in N. The observation that the shape of $E[M|R]$ varies with volatility is therefore not sensitive to the assumed polynomial order. In fact, $E[M|R]$ is well-described as scaling proportionally with volatility since the point estimate of b is close to one for all $N > 1$.

[6] Volatility is also persistent, but this fact does not require a standard error adjustment because it does not induce autocorrelation into the observations that enter the objective function (3).

Table I: Estimation results

We estimate the projected pricing kernel in (4) and (5) for different polynomial orders N by maximizing the log likelihood of realized returns, (3). Statistical inference is based on a block bootstrap with a block length of 21 trading days. *, **, and *** denote significance at the 10%, 5% and 1% levels.

N	1	2	3	4	5
Log-likelihood	14,275	14,370	14,370	14,384	14,384
$\hat{c}_1$	-0.017***	-0.067	-0.068	-0.095**	-0.095**
$\hat{c}_2$	–	0.100***	0.103***	0.020	0.020
$\hat{c}_3$	–	–	0.000	0.011*	0.011*
$\hat{c}_4$	–	–	–	0.003***	0.003***
$\hat{c}_5$	–	–	–	–	0.000
$\hat{b}$	1.600*	0.976**	0.973**	1.098***	1.097***

The log-likelihood increases substantially when the polynomial order is increased from $N = 1$ to $N = 2$, but only moderately thereafter. A likelihood ratio test rejects $N = 1$ in favor of $N = 2$ with a p-value of 0.15%, but fails to reject $N = 2$ in favor of any $N > 2$ at the 10% level (untabulated).[7] We show below that the reason for the log-linear pricing kernel's poor fit lies in its inability to match the sample variance premium. Hence, the data clearly favors specifications for which the logarithm of $E[M|R]$ is convex. Since the parsimonious quadratic ($N = 2$) kernel is not rejected in favor of more flexible specifications, we use it as our benchmark case. All subsequent results are based on this estimate, unless otherwise mentioned.

Figure I in the introduction illustrates graphically how $E[M|R]$ varies with volatility by plotting it for the 10th and 90th percentile of σ_t (p_{10} and p_{90}). The figure shows that the pricing kernel is considerably steeper when volatility is low. For example, for a monthly return of -10%, the projected pricing kernel equals $M(R_{t+1} = -0.1, \sigma_t = p_{10}; \theta) = 3.68$ when volatility is low and $M(R_{t+1} = -0.1, \sigma_t = p_{90}; \theta) = 1.32$ when volatility is high.

[7] There is no established method for dealing with overlapping data in likelihood ratio tests. We therefore rely on an ad-hoc sub-sampling approach. Specifically, we use observations $1, 22, 43, ...$, as the first subsample, observations $2, 23, 44, ...$, as the second subsample, and so on, up to observations $21, 42, 63, ...$, as the last subsample. We then estimate the two nested specifications of $E[M|R]$ in each subsample, compute their likelihood ratio, and average the individual likelihood-ratio statistics across the 21 subsamples. Finally, we compute critical values based on the statistic's asymptotic χ^2-distribution.

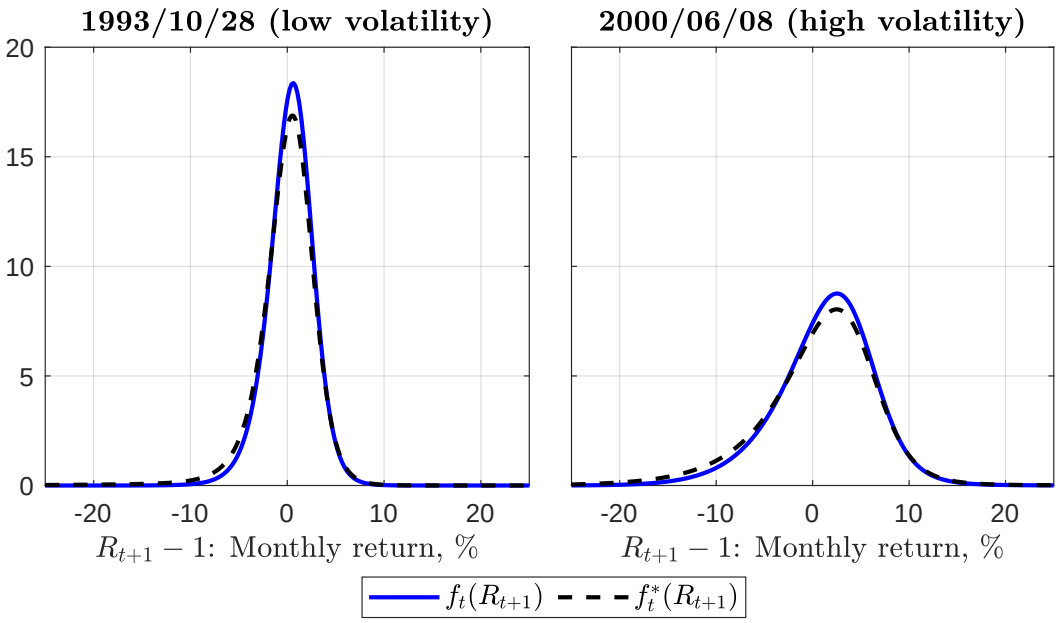

Figure II: Conditional density estimates for select days. We plot the estimated physical and risk-neutral return density for days on which conditional volatility is close to its 10th (left panel) or 90th (right panel) percentile. Estimates are based on the $E[M|R]$ specification in equations (4) and (5) and a polynomial order of $N = 2$.

Figure II shows the resulting conditional return densities for two dates. For comparability with Figure I, we choose days on which conditional volatility is close to its 10th percentile and 90th percentile, respectively. Because our parameterization of $E[M|R]$ implies a smooth change-of-measure, f_t inherits many of f_t^*'s properties. It is unimodal, roughly bell-shaped, and its conditional volatility moves with that of f_t^*. Relative to f_t^*, however, f_t has more probability mass in the center and less mass in the left tail. As a result, the physical density is less left-skewed and leptokurtic than its risk-neutral counterpart, the equity premium is positive, and the variance premium is negative.

Across the 7,556 trading days in our sample, the conditional physical (risk-neutral) density has an average mean of 9.06% (0.98%) p.a., standard deviation of 13.83% (17.97%) p.a., skewness of -0.61 (-1.48), and kurtosis of 4.43 (10.47). We show the time series of these moments in Figure IA.III of the online appendix. Our density estimates imply that the conditional equity premium $E_t[R_{t+1}] - E_t^*[R_{t+1}]$

has an average of 8.1% p.a., which closely matches the average excess return on the S&P 500 of 8.0% over the 1990-2019 period. Similarly, our density estimates imply that the conditional variance premium $var_t[R_{t+1}] - var_t^*[R_{t+1}]$ has an average of -12.4%^{2} per month, which closely matches the average $\sigma_t^2 - \left(\frac{VIX_t}{100}\right)^2$ of -13.2%^{2} per month over 1990-2019. The parametric pricing kernel therefore provides a good fit for stock market risk premia in our sample. Additionally, both risk premia are well-identified by $E[M|R]$: Our bootstrap estimates imply 99% confidence intervals of $[2.8\%, 13.3\%]$ per year for the average equity premium and $[-17.0\%^2, -7.5\%^2]$ per month for the average variance premium.

F. Robustness

We perform five robustness tests. First, we model the projected pricing kernel's volatility-dependence with the alternative specification

$$c_{it} = \sum_{k=0}^{K} c_{ik} \times \sigma_t^k, \tag{8}$$

which assumes that coefficients of the $E[M|R]$-polynomial are themselves polynomials of volatility. The combination of (4) and (8) is equivalent to a bivariate polynomial in $\ln R_{t+1}$ and σ_t with a tensor product base. We find that, for $K = 2$ and higher orders, the estimated functional relationship between σ_t and c_{it}'s implied by (8) closely resembles the one in our benchmark specification (5). As a result, the relationship between σ_t and the shape of $M(R_{t+1}, \sigma_t; \theta)$ also closely resembles the one in our benchmark specification. We illustrate this fact for $K = 3$ below and for other polynomial orders in the online appendix.[8] Relative to (8), our benchmark specification (5) has the advantage of being more parsimonious.

Second, we allow the projected pricing kernel to comove with additional macroe-

[8]The linear case ($K = 1$) implies too little time-variation in c_{it}'s and therefore too little time-variation in $M(R_{t+1}, \sigma_t; \theta)$. Relative to our benchmark estimates, the log-likelihood is lower (at 14,352) and implied physical moments are closer to their risk-neutral counterparts. The $K = 1$ case corresponds to the specification in Kim (2022), apart from the fact that he models the pricing kernel's intercept as parametric function of volatility, whereas we choose it such that the implied return densities integrate to one – see footnote 5 for details.

conomic time series by modeling coefficients of the $E[M|R]$-polynomial as

$$c_{it} = \frac{c_{i0} + c_{i1} \times \text{short rate} + c_{i2} \times \text{term spread} + c_{i3} \times \text{credit spread}}{\sigma_t^{b \times i}}. \tag{9}$$

In doing so, we are able to evaluate whether volatility continues to induce variability in $E[M|R]$ once other sources of variation are accounted for, i.e., whether b continues to be significantly different from zero. We measure the short rate by the yield of a 3-month Treasury bill, the term spread by the difference in yields of a 10-year Treasury bond and a 3-month Treasury bill, and the credit spread by the difference in yields of a 10-year corporate bond with Moody's Aaa rating and an equivalent bond with a Baa rating.

Third, instead of modelling $E[M|R]$ as a polynomial, we model $f_t(R_{t+1};\theta)$ as a parametric density and obtain $E[M|R]$ from (1) as the ratio of risk-neutral and physical densities, scaled by the risk-free rate. Specifically, we parameterize the density of standardized log returns, $g_t\left(\frac{\ln R_{t+1}}{\sigma_t};\theta\right)$, with a normal inverse Gaussian (NIG) distribution and compute the distribution of simple returns via a change of variables as $f_t(R_{t+1};\theta) = g_t\left(\frac{\ln R_{t+1}}{\sigma_t};\theta\right)/(\sigma_t \times R_{t+1})$. The NIG distribution is unimodal, bell-shaped, allows for nonzero skewness and excess kurtosis, and depends on four parameters, which we estimate via maximum likelihood. This method for estimating the conditional distribution resembles the popular approach of scaling historical return innovations with an estimate of conditional volatility – see, e.g., Rosenberg and Engle (2002), Barone-Adesi et al. (2008) and Christoffersen et al. (2013) – and shares its limitation that higher conditional moments (beyond volatility) are time-invariant by construction. In contrast, the parameterized pricing kernel in our benchmark specification allows all return moments to vary over time.

Fourth, we re-estimate the benchmark specification in the second half of the sample (2005–2019) to address concerns about a possible segmentation between index option and equity markets. In particular, Dew-Becker and Giglio (2022) argue that the two markets have historically been segmented, but also provide evidence suggesting that they have become well-integrated since about the mid 2000's. If

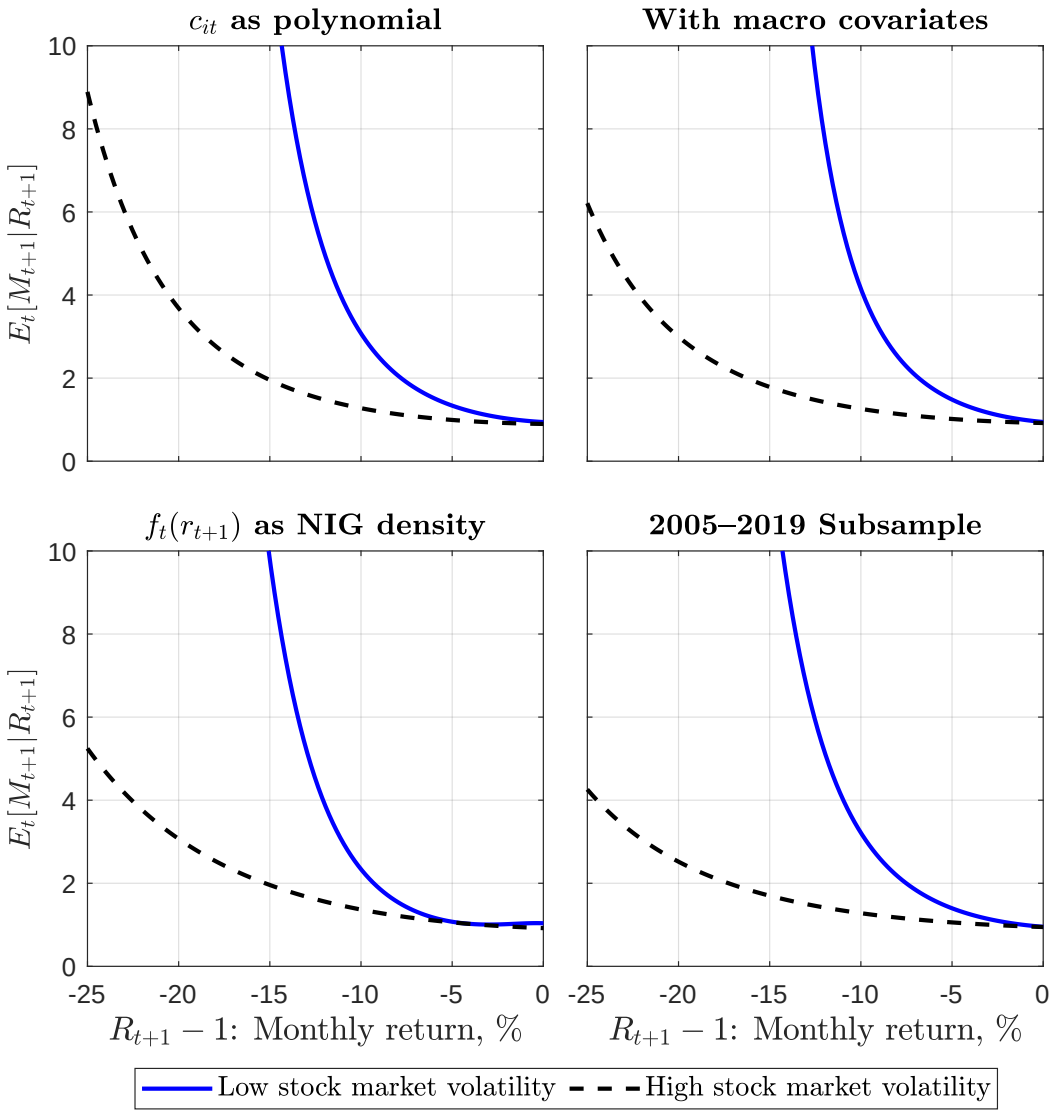

Figure III: Robustness. We plot the projected pricing kernel for the 10th and 90th percentile of conditional stock market volatility. Top-left: $\ln E[M|R]$ is a $N = 2$ polynomial with coefficients that depend on volatility via (8) with $K = 3$. Top-right: $\ln E[M|R]$ is a $N = 2$ polynomial with coefficients that depend on volatility, short-term interest rates, the term spread, and credit spreads via (9). Bottom-left: We model the distribution of standardized log returns $\ln R_{t+1}/\sigma_t$ with a Normal Inverse Gaussian distribution, compute $f_t(R_{t+1})$ via a change-of-variables, and obtain $E[M|R]$ from (1). Bottom-right: $E[M|R]$ is equivalent to the benchmark specification, but estimated over the 2005-2019 subsample.

time-variation in the estimated projected pricing kernel was a result of market segmentation, one would expect it to be substantially weaker in more recent data.

Figure III shows that, for each of the four alternative estimates, the projected pricing kernel's volatility-dependence looks similar to our benchmark estimates in Figure I. Parameter estimates for these specifications are reported in Section I.III of the online appendix. In the bivariate polynomial specification, a likelihood ratio test strongly rejects the hypothesis $H_0 : c_{ik} = 0 \forall i, k > 0$ (time-invariance) with a p-value of 0.2%. In the specification with additional covariates, the estimated volatility-scaling parameter of $\hat{b} = 1.06$ is very close to the benchmark estimate of 0.976, and it remains statistically significant with a p-value of 0.064. The parameterized density approach does not lend itself to a formal statistical test, but the amount of time-variation in $E[M|R]$ is quantitatively similar to that in Figure I. In the 2005–2019 estimation, the corresponding point estimate of b is once again similar to the benchmark at $\hat{b} = 1.01$. Our main result is therefore not sensitive to the way $E[M|R]$ is parameterized, the volatility-dependence of $E[M|R]$ does not reflect comovement between volatility and other state variables, and it can also not be explained by market segmentation.

G. The Economic Importance of Convexity

We saw above that the log-quadratic specification of $E[M|R]$ provides a significantly better fit to return data than its log-linear counterpart. To illustrate what this difference implies economically, we now illustrate its implications for risk premia.

For $N = 1$, the estimates in Table I imply that the conditional physical density has an average standard deviation of 16.9%, skewness of -1.33, and kurtosis of 11.43. All of these moments are much closer to their risk-neutral counterparts than in the benchmark $N = 2$ case, and risk premia on higher moments are smaller as a result. For example, the average variance premium equals $-4.4\%^2$ per month with a 99% confidence interval of $[-7.1\%^2, -1.0\%^2]$. The -13.2%^{2} sample average of $\sigma_t^2 - \left(\frac{VIX_t}{100}\right)^2$ has a bootstrapped p-value of 0.00% under the sampling distribution

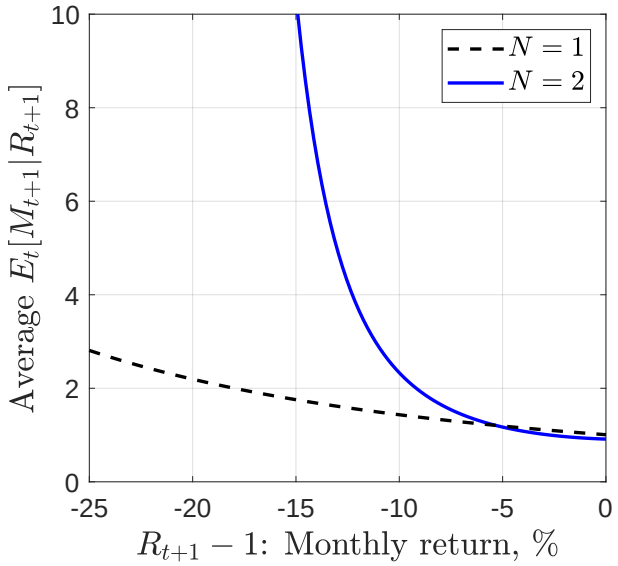

Figure IV: Average $E[M|R]$ for linear and quadratic specifications. $E[M|R]$ is parameterized by (4) and (5) with polynomial orders of either $N = 1$ or $N = 2$.

of the $N = 1$ estimator. Hence, a log-linear projected pricing kernel is inconsistent with the variance premium.

Whereas the log-linear specification provides a good fit to the sample equity premium of 8.0% p.a., with an implied value that also equals 8.0%, we find that it is inconsistent with its sources. Specifically, Beason and Schreindorfer (2022) propose a non-parametric decomposition of the equity premium into contributions of different return regions, and find that monthly returns below -10% account for about 80/100 of the total premium. Based on our parametric estimates, we find the equivalent contribution to be 37/100 for $N = 1$ and 72/100 for $N = 2$.[9] The log-linear specification of the projected pricing kernel therefore substantially understates the importance of stock market tail events for the equity premium, whereas the log-quadratic specification captures it well.

The reason for both shortcomings is illustrated in Figure IV, which shows the

[9]The decomposition is based on the average distributions $f(R) = \frac{1}{T}\sum_{t=1}^{T} f_t(R)$ and $f^*(R) = \frac{1}{T}\sum_{t=1}^{T} f_t^*(R)$. The relative contribution of returns below -10% equals $\left(\int_0^{0.9}(R-1)[f(R) - f^*(R)]dR\right) / \left(\int_0^{\infty}(R-1)[f(R) - f^*(R)]dR\right)$.

average projected pricing kernel for polynomial orders of $N = 1$ and $N = 2$. Relative to the log-quadratic case, the log-linear pricing kernel is substantially flatter, especially in the far left tail of the return distribution. It is therefore misspecified in the sense that it substantially understates investors' aversion against tail events. As a result, it is problematic to estimate investors' risk aversion based on such a specification, as, e.g., in Bliss and Panigirtzoglou (2004).

II. Volatility and Expected Put Option Returns

To address potential concerns about a misspecified parametric functional form of $E[M|R]$, we present a non-parametric alternative via put returns to measure our main finding. We find that returns are increasing in volatility, which confirms that negative returns are less painful to investors in bad times.

For our test, we use a panel of realized out-of-the-money put returns, with target moneyness levels $K_i/S_t - 1$ that range from -24% to 0, in steps of 2%. For each month, we select the put option with 30 days until maturity that expires on the main expiration cycle (third Friday), and has moneyness closest to the target, but is at most $\pm 1\%$ away from the target. Finally, we compute realized hold-until-maturity returns for each put option. To proxy for low and high volatility, we split the sample and calculate average returns if the ex ante σ_t is below its 25th percentile, or above its 75th percentile, respectively.

The circles plotted in Figure V denote the average realized excess put returns, conditional on low and high volatility. It is evident that when volatility is low, average realized put returns decrease rapidly in their moneyness, while the decrease is slow in times of high volatility.

This finding might be counter-intuitive at first sight. Since option prices are higher if volatility is high, one might expect that returns are on average lower. However, the results show that the opposite is true. While option in times of high volatility are expensive in dollar terms, they are relatively cheap in terms of expected returns. In addition, we compute the expected monthly S&P 500 put returns implied

by our benchmark estimation of $E[M|R]$. Figure V illustrates the results, where low and high volatility periods are defined as before. One can clearly see that implied values are well in line with realized ones.[10]

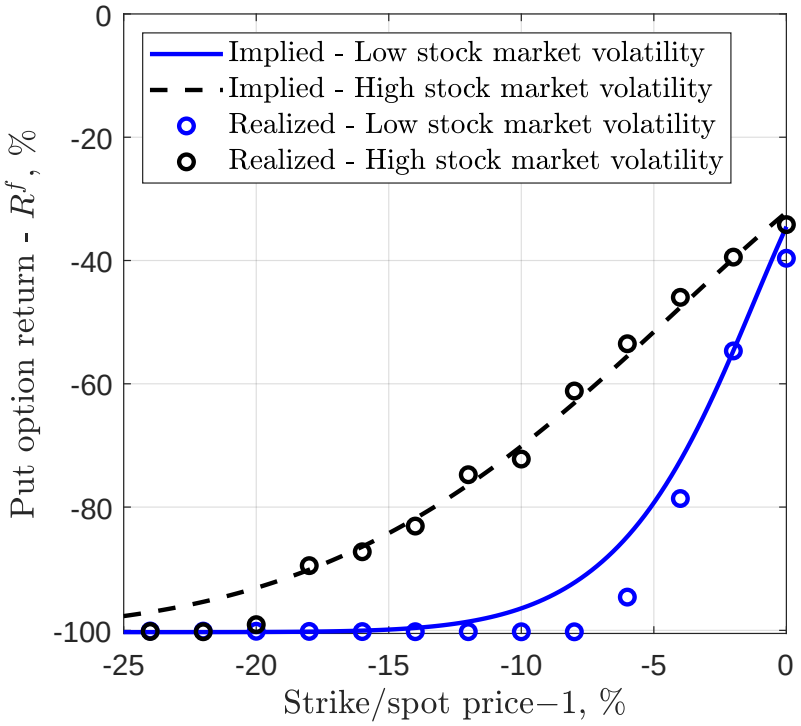

Figure V: Put returns. The plot shows expected and realized monthly S&P 500 put option returns. Implied values are from our benchmark estimate of $E[M|R]$ and the lines correspond to high and low volatility (10th and 90th percentile of σ_t, respectively). The circles denote average realized monthly S&P 500 put option excess returns in our sample. Here, low and high volatility are defined as days on which conditional volatility is below its 25th percentile and above its 75th percentile, respectively.

To understand the relationship of the observed implied pattern to the slope of $E[M|R]$ intuitively, first note that the expected return on an Arrow-Debreu security that pays 1\$ if the return on the index at maturity is exactly equal to R is $1/E[M|R]$, i.e., it is inversely proportional to the level of M. Since $E[M|R]$ is always higher in times of low volatility (see Figure I) relative to high volatility, all Arrow-Debreu

[10] In contrast, specifications without time-variation in $E[M|R]$ ($b = 0$) fail to match the time-variation if expected put returns, and log-linear specifications fail to match the convexity of $E[M|R]$ and imply much too high levels of put returns.

securities that pay off in the domain of negative returns have lower expected returns relative to high volatility times. To make the connection to puts it is helpful to think of a put option as a portfolio of all Arrow-Debreu securities that pay off below the strike of the put. Since all the securities in the portfolio have a lower return, their weighted average does as well, and hence put returns are more negative in low volatility periods.

To confirm the empirical result regarding put returns and to assess the statistical significance of the time-variation, we run then following predictive regression:

$$R^{Put}_{t+1,i} - R_{f,t} = \alpha + \beta \ln(\sigma_t) + \gamma K_i/S_t + \epsilon_{t+1,i}. \tag{10}$$

Based on the graphical evidence in Figure V we expect that the β coefficients depends on moneyness in a non-trivial way, and we hence run the regression separately for each moneyness bin. We use a log-transform of the volatility, since this is more consistent with the theoretical result that put returns approach the risk-free rate as $\sigma \to \infty$ (Hu and Jacobs 2020). We use the options' initial moneyness K_i/S_t as control, since it clearly relates to the expected return, even when holding moneyness in a tight interval.

The results in Table II show that put returns are indeed predictable by ex ante expected return volatility for a large range of moneyness levels. For intermediate moneyness levels, the effect is statistically and economically significant. To illustrate the magnitude, note that σ_t increase by a factor of about 2.2 between p_{10} and p_{90}. Hence a $\beta = 0.5$ implies an increase in expected put returns of $0.5 \times \ln(2.2) = 39\%$ on a monthly basis from good to bad times.[11] For both very high and very low moneyness levels Figure V shows there is little difference in realized returns, and hence it is not surprising that the predictive relationship is not significant everywhere.

[11] This finding on the time-series predictability of S&P 500 put returns is consistent with findings on the cross-section of singe stock put returns (across different underlings) in Hu and Jacobs (2020) and Aretz et al. (2022). Both papers show empirically that single stock put returns are increasing in (idiosyncratic) volatility. Importantly, however, their finding is a cross-sectional result, and is driven by different levels of volatility of single stocks, and not by time-series variation in (systematic) volatility.

Table II: Predicting out-of-the-money put returns

The table shows β regression coefficients from (10) for different levels of initial moneyness. The description of the option panel construction is provided in the text. *, **, and *** denote significance at the 10%, 5% and 1% levels. t-statistics (in parenthesis) are adjusted for heteroscedasticity and autocorrelation (Newey and West 1987).

$K/S-1,\%$	-20	-18	-16	-14	-12	-10	-8	-6	-4	-2	0
$\hat{\beta}$	0.02	0.24	0.34^{**}	0.43^{**}	0.45^{**}	0.49^{**}	0.49^{**}	0.52^{**}	0.41^{*}	0.29	0.05
	(1.27)	(1.63)	(2.01)	(2.17)	(2.36)	(2.19)	(2.45)	(2.50)	(1.86)	(1.31)	(0.25)

Overall, these results show nonparametrically that negative returns are less painful to investors in bad times. In low volatility times, they are willing to incur a much more negative returns on average on an insurance via a put option for a given moneyness level relative to high volatility times. Hence we confirm our main finding using a economically potentially more intuitive and natural unit.

Moreover, put returns are heavily studied in the options literature. One of the main puzzles are their high prices and low returns, which Gabaix (2012) calls one of the ten major puzzles in finance. We add to that puzzle, by showing that put returns are particularly low in low volatility times, or conversely, their prices are relatively high.

III. Economic Interpretation

Which economic forces cause time-variation in $E_t[M_{t+1}|R_{t+1}]$? What does the time-varying pricing of stock market risks reveal about the joint dynamics of volatility and the pricing kernel (before projecting it onto returns)? What connections, if any, are there to other stock market phenomena, such as the average equity premium, return predictability, or the risk-return trade-off? This section attempts to answer these questions by examining the projected pricing kernel in a number of structural asset pricing models. We relegate all computational details to the online appendix.

A. Intuition and Overview of Results

Figure VI shows $E_t[M_{t+1}|R_{t+1}]$ in the models of Campbell and Cochrane (1999), Bansal and Yaron (2004), and Gabaix (2012). A number of additional models are analyzed in the online appendix and briefly discussed below. In contrast to the data, the habit model (top-left panel) implies that the projection becomes steeper when volatility rises, whereas the long-run risks model (top-right panel) implies that its shape does not vary with volatility. In line with the data, the time-varying disaster risk model (bottom panel) implies that the projection becomes flatter when volatility rises. To understand the source of this behavior and link it to the models' economic mechanisms, we first characterize the shape of $E_t[M_{t+1}|R_{t+1}]$ in a general lognormal setting, which encompasses two of the aforementioned models.

PROPOSITION 1. Assume that the log pricing kernel and log return follow

$$\begin{aligned}\ln M_{t+1} =& \mu_{m,t} - \sigma_{m,t}\varepsilon_{s,t+1}\\ \ln R_{t+1} =& \mu_{r,t} + \sigma_{s,t}\varepsilon_{s,t+1} + \sigma_{i,t}\varepsilon_{i,t+1},\end{aligned}$$

where the systematic shock $\varepsilon_{s,t}$ and the idiosyncratic shock $\varepsilon_{i,t}$ are IID standard normal. Then time-variation in the shape of $E_t[M_{t+1}|R_{t+1}]$ can be characterized by

$$\frac{\partial \ln E_t[M_{t+1}|R_{t+1}]}{\partial \ln R_{t+1}} = -\frac{\sigma_{m,t}}{\sigma_{s,t}} \times \frac{\sigma_{s,t}^2}{\sigma_{s,t}^2 + \sigma_{i,t}^2}. \tag{11}$$

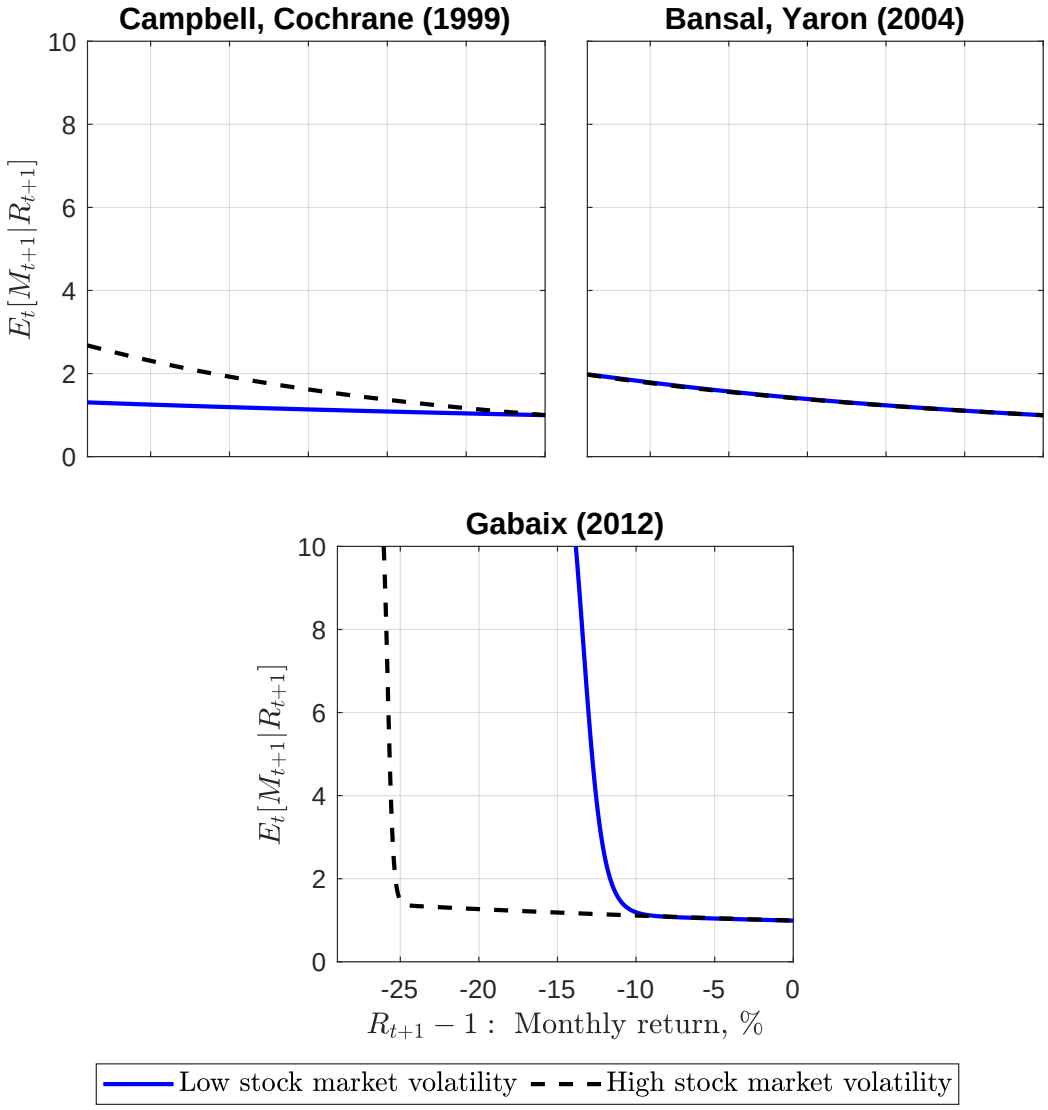

Figure VI: $E[M|R]$ **in Macrofinance models.** We plot the projected pricing kernel for the 10-th and 90-th percentile of the conditional volatility of returns. The models of Campbell and Cochrane (1999) and Bansal and Yaron (2004) are calibrated as in the original studies, whereas the model of Gabaix (2012) is calibrated as in Dew-Becker et al. (2017).

Proof. *Joint normality implies that the log pricing kernel, conditional on the log return, is distributed as* $\ln M_{t+1}|\ln R_{t+1} \sim N\left(E_t[\ln M_{t+1}|\ln R_{t+1}], \mathrm{Var}_t[\ln M_{t+1}|\ln R_{t+1}]\right)$*, where* $E_t[\ln M_{t+1}|\ln R_{t+1}] = \mu_{m,t} - \frac{\sigma_{m,t}\sigma_{s,t}}{\sigma_{s,t}^2+\sigma_{i,t}^2}\left(\ln R_{t+1} - \mu_r\right)$ *and* $\mathrm{Var}_t[\ln M_{t+1}|\ln R_{t+1}] = \sigma_{m,t}^2\left(1 - \frac{\sigma_{s,t}^2}{\sigma_{s,t}^2+\sigma_{i,t}^2}\right)$*. Using the moment generating function of a normal random variable, the conditional expectation of the pricing kernel equals*

$$\mathbb{E}_t[M_{t+1}|R_{t+1}] = \exp\left\{\mu_{m,t} - \frac{\sigma_{m,t}\sigma_{s,t}}{\sigma_{s,t}^2+\sigma_{i,t}^2}\left(\ln R - \mu_r\right) + \sigma_{m,t}^2\left(1 - \frac{\sigma_{s,t}^2}{\sigma_{s,t}^2+\sigma_{i,t}^2}\right)/2\right\}.$$

Taking logs and differentiating with respect to $\ln R_{t+1}$ *yields* (11). ■

To understand the intuition behind Proposition 1, note that $-\frac{\sigma_{m,t}}{\sigma_{s,t}}$ equals the slope of $\ln E_t[M_{t+1}|R_{t+1}]$ in a world without idiosyncratic risk ($\sigma_{i,t} = 0$). In this world, log returns and the log pricing kernel are perfectly correlated, the market represents the only risk factor and, all else equal, an increase in return volatility (due to $\sigma_{s,t}$) makes the projection flatter. The intuitive reason is that a drop in the market is more indicative of high marginal utility (deteriorating macroeconomic conditions) if it occurs during calm markets, rather than the middle of a recession.[12] The second factor in (11), $\frac{\sigma_{s,t}^2}{\sigma_{s,t}^2+\sigma_{i,t}^2}$, equals the fraction of systematic return risk or, equivalently, the squared conditional correlation between $\ln M_{t+1}$ and $\ln R_{t+1}$. It adjusts for the fact that, in reality, the pricing kernel involves additional factors and is therefore imperfectly correlated with the market. Intuitively, an increase in the fraction of systematic return risk makes returns more informative about the pricing kernel and therefore results in a steeper projection.

Proposition 1 shows that time-variation in the shape of $E_t[M_{t+1}|R_{t+1}]$ reflects

[12] Suppose, for example, that a -10% return corresponds to a two standard deviation event in normal times and, as such, tends to coincide with a two standard deviation shock to the pricing kernel – a fairly large macroeconomic shock. If return volatility doubles in recessions while pricing kernel volatility remains unchanged, then a return of -10% corresponds to a one standard deviation event and, as such, tends to coincide with a one standard deviation shock in the pricing kernel – a smaller macroeconomic shock.

the joint evolution of $(\sigma_{m,t}, \sigma_{s,t}, \sigma_{i,t})$ over time. These dynamics can take many different forms, but several stylized cases are noteworthy. First, suppose that idiosyncratic risk is time-invariant (constant $\sigma_{i,t}$), while systematic risk increases less than proportionally with $\sigma_{m,t}$.[13] In this case, which we'll find to describe the habit model, both fractions in equation (11) increase and $E_t[M_{t+1}|R_{t+1}]$ steepens when return volatility rises (due to $\sigma_{s,t}$). Second, suppose that $\sigma_{m,t}$, $\sigma_{s,t}$, and $\sigma_{i,t}$ are all proportional to one another. In this case, which we'll find to closely approximate the long-run risks model, both fractions in equation (11) are time-invariant and the shape of $E_t[M_{t+1}|R_{t+1}]$ remains unchanged when return volatility rises. Third, suppose that the volatility of returns evolves independently from $\sigma_{m,t}$. Then an increase in return volatility due to $\sigma_{i,t}$ will always make $E_t[M_{t+1}|R_{t+1}]$ flatter, whereas an increase due to $\sigma_{s,t}$ will make $E_t[M_{t+1}|R_{t+1}]$ flatter as long as $\sigma_{s,t} > \sigma_{i,t}$. Even though the model of Gabaix (2012) is not lognormal, we show below that its implications for time-variation in $E_t[M_{t+1}|R_{t+1}]$ are well-described by this scenario.

In what follows, we briefly review each of the aforementioned models and discuss their implications for the dynamics of $(\sigma_{m,t}, \sigma_{s,t}, \sigma_{i,t})$.

B. Habits

B.1. Campbell and Cochrane (1999): Model statement

Aggregate consumption and dividends follow homoscedastic random walks,

$$\begin{aligned} \Delta c_{t+1} =& g + \sigma \varepsilon_{c,t+1} \\ \Delta d_{t+1} =& g + \sigma_w \left(\rho \varepsilon_{c,t+1} + \sqrt{1-\rho^2} \varepsilon_{d,t+1} \right) \end{aligned} \tag{12}$$

where ε_c and ε_d are IID standard normal. Equation (12) implies that the correlation between consumption and dividend growth rates equals ρ. Equity represents a claim to the dividends in all future periods. The representative agent's utility function is

$$E_t \left[\sum_{h=0}^{\infty} \delta^h \frac{(C_{t+h} - X_{t+h})^{1-\gamma} - 1}{1-\gamma} \right], \tag{13}$$

[13] This is the case, for example, if $\sigma_{s,t}$ is an increasing and concave function of $\sigma_{m,t}$ that goes through the origin.

where $\delta > 0$ controls time preference and $\gamma > 0$ controls risk preference. Time variation in the habit, X_t, is modelled via the surplus consumption ratio

$$S_t = \frac{C_t - X_t}{C_t}, \tag{14}$$

whose logarithm, $s_t = \ln S_t$, evolves via the heteroscedastic AR(1) process,

$$\begin{aligned} s_{t+1} =& (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)\sigma\varepsilon_{c,t+1} \\ \lambda(s_t) =& \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})} - 1 & , s_t < s_{max} \\ 0 & , s_t \geq s_{max} \end{cases}, \end{aligned} \tag{15}$$

where $\bar{S} = \sigma\sqrt{\gamma/(1-\phi)}$ and $s_{max} = \bar{s} + (1-\bar{S}^2)/2$. The parameters $\bar{s} = \ln \bar{S}$ and ϕ control the unconditional mean and persistence of s_t, whereas the function $\lambda(s_t)$ controls its sensitivity to consumption innovations.

B.2. Volatility Dynamics and $E[M|R]$ in the Habit Model

The conditional variance of idiosyncratic shocks ($\varepsilon_{d,t+1}$) is time-invariant and given by $\sigma^2_{i,t} = \sigma^2_w(1-\rho^2)$. Systematic shocks ($\varepsilon_{c,t+1}$) are reflected in dividend growth rates, and in time-variation in the price-dividend (P/D) ratio. Because the P/D ratio has to be computed numerically, its conditional variance cannot be expressed analytically. Schreindorfer (2023) shows that, for a numerically accurate solution, the conditional variance of P/D looks drastically different from its counterpart in Campbell and Cochrane (1999). We first explain the model mechanism for the original (inaccurate) solution, which most readers are likely familiar with, and then point out what changes when the model is solved accurately. The dashed line in the top-left panel of Figure VII shows that, for the original solution, the conditional variance of systematic shocks, $\sigma^2_{s,t}$, is a concave and monotonically increasing function of $\sigma^2_{m,t}$.[14] Proposition 1 implies that, as a result, a rise in the volatility of returns (due to $\sigma_{s,t}$) makes $E[M|R]$ steeper, rather than flatter, as in the data.

[14] Because the pricing kernel, $\ln M_{t+1} = \ln\delta - \gamma(\Delta c_{t+1} + \Delta s_{t+1})$ is exogenous, its conditional variance is not affected by the numerical solution. It is straightforward to show that $\sigma^2_{m,t} = \gamma(1-\phi)(1+2\bar{s}) - 2\gamma(1-\phi)s_t$.

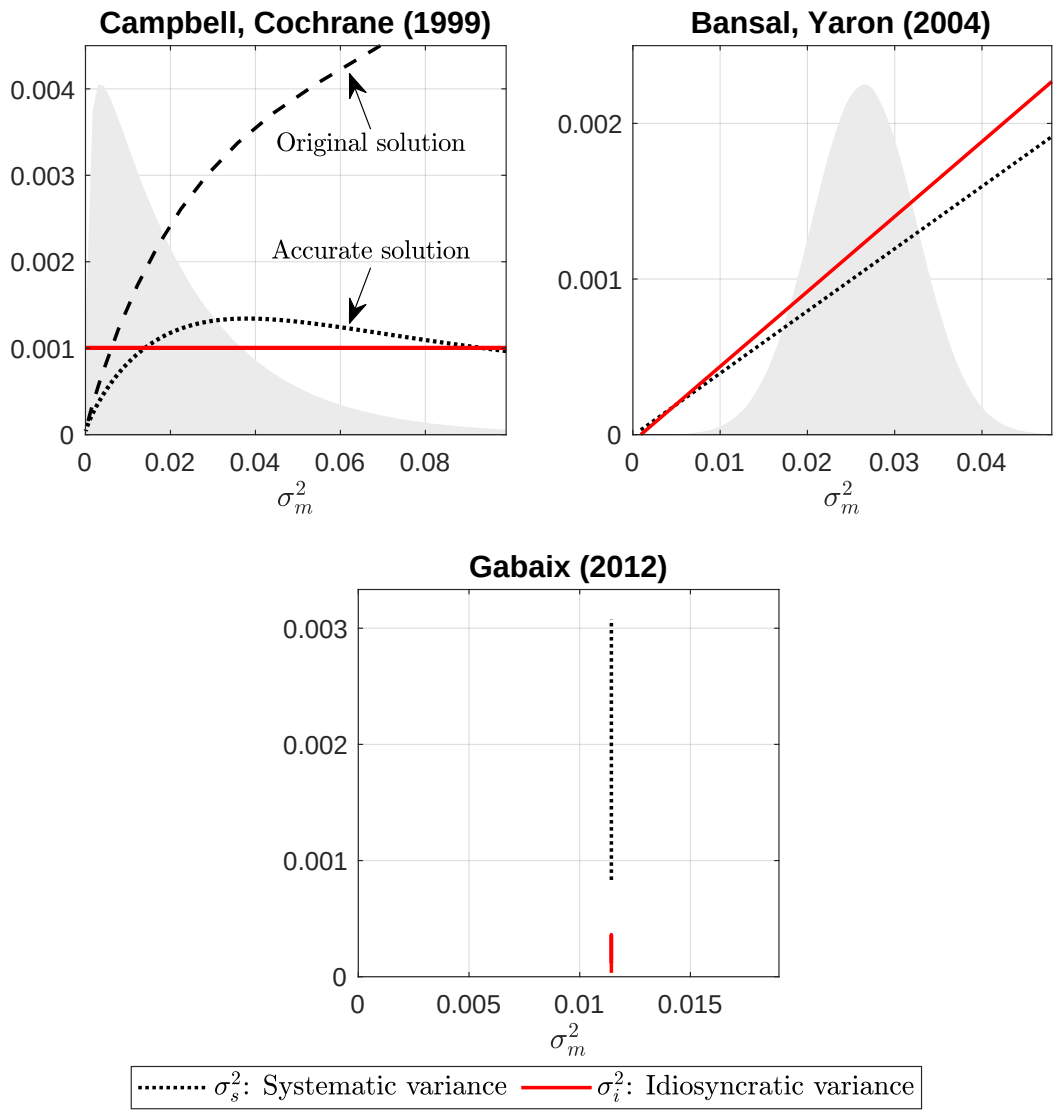

Figure VII: Sources of time-variation in $E[M|R]$. We plot the conditional variances of systematic and idiosyncratic return shocks as functions of the pricing kernel's conditional variance. Proposition 1 connects these moments to the shape of $E[M|R]$. The models of Campbell and Cochrane (1999) and Bansal and Yaron (2004) are calibrated as in the original studies, whereas the model of Gabaix (2012) is calibrated as in Dew-Becker et al. (2017).

To see why these volatility dynamics are essential for the habit mechanism, note that a sequence of adverse consumption shocks leads to a decrease in s_t and an increase in $\sigma_{m,t}$. This raises risk aversion and decreases the P/D ratio. Since $\sigma_{s,t}$ increases at the same time, the model captures (i) the countercyclical nature of stock market volatility and (ii) the "leverage effect", i.e., the negative correlation between changes in volatility and contemporaneous returns (Black 1976), and (iii) because the simultaneous rise in return volatility and $\sigma_{m,t}$ induces a large rise in risk premia, the model rationalizes the long-horizon predictability of excess returns. The habit model's counterfactual implications for time-variation in $E[M|R]$ therefore result directly from its core mechanism for dynamic asset pricing facts.

For the accurate model solution, $\sigma^2_{s,t}$ is a hump-shaped, rather than monotonically increasing, function of $\sigma^2_{m,t}$, as shown in Figure VII. This alters the model's predictions in several ways relative to the original solution. First, during recessions (states of high $\sigma^2_{m,t}$), a further increase in $\sigma_{m,t}$ decreases $\sigma_{s,t}$ and therefore decreases the conditional volatility of returns. When solved accurately, the habit model is therefore no longer consistent with the fact that stock market volatility spikes during recessions or with the leverage effect (see Schreindorfer 2023 for more detail). Second, states of low return volatility (low $\sigma_{s,t}$) can be associated with either low or high values of $\sigma_{m,t}$ for the accurate model solution, and hence with either a flat or steep $E[M|R]$ curve. Probabilistically, however, values around the 10-th percentile of return volatility are more than 100,000 times as likely to coincide with low values of σ_m and a flat $E[M|R]$ curve than with high values of σ_m and a steep $E[M|R]$ curve. Even for the accurate solution, for which we depict $E[M|R]$ in Figure VI, the habit mechanism for time-varying volatility therefore remains inconsistent with our empirical evidence.

C. Recursive Utility and Long-run Risks

C.1. Bansal and Yaron (2004): Model statement

The representative agent has Epstein-Zin utility, calibrated to imply a preference for the early resolution of uncertainty (EIS>1/RRA), and is endowed with

$$\begin{aligned}
\Delta c_{t+1} &= g + x_t + \sigma_t \varepsilon_{t+1} \\
\Delta d_{t+1} &= g + \phi x_t + \varphi_d \sigma_t \varepsilon_{d,t+1} \\
x_{t+1} &= \rho x_t + \varphi_e \sigma_t \varepsilon_{x,t+1} \\
\sigma^2_{t+1} &= \sigma^2 + \nu_1(\sigma^2_t - \sigma^2) + \sigma_w \varepsilon_{\sigma,t+1}
\end{aligned}$$

where $(\varepsilon_{t+1}, \varepsilon_{d,t+1}, \varepsilon_{x,t+1}\varepsilon_{\sigma,t+1}) \overset{\text{IID}}{\sim} N(0,1)$. The "long run risks" x_t and σ_t generate persistent variation in the conditional mean and volatility of consumption and dividend growth rates. In the following, we illustrate the model's implications for volatility dynamics based on the log-linearized solution, which implies conditional lognormality. We refer interested readers to Bansal and Yaron (2004) for details on the solution coefficients (A_0, A_1, A_2) and (A_{0m}, A_{1m}, A_{2m}).

C.2. Volatility Dynamics and $E[M|R]$ in the Long-run Risks Model

The conditional variance of idiosyncratic shocks $(\varepsilon_{d,t+1})$ is proportional to σ^2_t and given by $\sigma^2_{i,t} = \varphi^2_d \sigma^2_t$. Systematic return shocks ($\varepsilon_{x,t+1}$ and $\varepsilon_{\sigma,t+1}$) are reflected in the P/D ratio. Based on the model's log-linearized solution $\ln(P_t/D_t) = A_{0m} + A_{1m} x_{t+1} + A_{2m}\sigma^2_{t+1}$, it is straightforward to show that the conditional variance of systematic risk is linear in σ^2_t,

$$\sigma^2_{s,t} = A^2_{2m}\sigma^2_w + A^2_{1m}\varphi^2_e \sigma^2_t \tag{16}$$

as is the conditional variance of the log pricing kernel,[15]

$$\sigma^2_{m,t} = (\theta-1)^2 \kappa^2_1 A^2_2 \sigma^2_w + \left[((\theta-1) - \theta/\psi)^2 + ((\theta-1)\kappa_1 A_1 \varphi_e)^2\right] \sigma^2_t \tag{17}$$

[15]The log pricing kernel equals $\ln M_{t+1} = \theta \ln \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1) r_{a,t+1}$. To derive (17), we substitute for the log-linearized return on wealth, $r_{a,t+1} = \kappa_0 + \kappa_1 w c_{t+1} - w c_t + \Delta c_{t+1}$, replace the log wealth consumption ratio by $w c_{t+1} = A_0 + A_1 x^2_{t+1} + A_2 \sigma^2_{t+1}$, and then take the conditional variance.

For Bansal and Yaron's calibration, the intercepts of σ_s^2 and σ_m^2 are both close to zero, and $\sigma_{m,t}^2$, $\sigma_{s,t}^2$, and $\sigma_{i,t}^2$ are approximately proportional to one another. This proportionality is illustrated in the top-right panel of Figure VII and it explains, via Proposition 1, why the shape of $E[M|R]$ is time-invariant in the model.

The positive comovement between $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$ is central to the long-run risks mechanism. As in the habit model, it implies that elevated stock market volatility (due to $\sigma_{s,t}$) coincides with a large equity premium and a low P/D ratio. The long-run risks model therefore explains the negative correlation between volatility innovations and contemporaneous returns, as well as the long-horizon predictability of excess returns. Additionally, because Epstein-Zin preferences imply aversion against heteroscedasticity in x_t, a positive realization of $\varepsilon_{\sigma,t}$ increases the pricing kernel and lowers P/D (contributing to the equity premium), while also increasing $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$. Covariation between $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$ is therefore a byproduct of how the model generates the equity premium. Table 5 in Bansal and Yaron (2004) shows that about 40% of the pricing kernel's variance, and hence about 40% of the equity premium, is attributable to the time-varying volatility channel. Positive comovement between $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$ is therefore essential for the long-run risks model's main empirical predictions.

The fact that $\sigma_{i,t}^2$ covaries with $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$ is not important, because the long-run risks factor σ_t^2 affects P/D primarily by controlling the conditional volatility of x_t, not by controlling the conditional volatility of consumption and dividend growth rates. If consumption and dividend growth rates were homoscedastic, $\sigma_{i,t}^2$ would be time-invariant and $E[M|R]$ would behave as in the Campbell-Cochrane model, but the model's asset pricing implications would remain nearly unchanged.[16] Obviously, this change would remove the model further from explaining our empirical evidence.

[16]To confirm this assessment, we solved the model with homoscedastic consumption and dividend growth rates by replacing σ_t by σ (the square root of the unconditional mean of σ_t^2) in the equations for Δc_{t+1} and Δd_{t+1}, while maintaining the time-varying σ_t in the equation for x_{t+1}. Based on a long model simulation, the mean/std/AC1 of the annual (time-aggregated) log P/D ratio equal 3.04/0.18/0.69 for both the heteroscedastic and homoscedastic versions of the model, whereas the annual log equity return has a mean/std/AC1 of 6.68/16.62/0.02 in the heteroscedastic model and 6.69/16.73/0.02 in the homoscedastic model.

C.3. Related Recursive Utility Models

The online appendix examines a number of related models with recursive preferences and persistent state variables. First, we consider Drechsler and Yaron (2011), who augment the long-run risks model with jumps in x_t and σ_t^2 and a second volatility factor. These changes result in skewed returns and a better fit to volatility moments, but don't alter the model's main economic mechanism. We find that they also have little effect on $E[M|R]$, which remains close to time-invariant.

Second, we consider Wachter (2013), who augments the rare disaster model of Rietz (1988) and Barro (2006) with recursive utility and persistent variation in the disaster probability, which is modelled as a CIR process. The CIR process implies that a rise in the probability of disasters not only lowers expected consumption growth, but also increases the conditional volatility of expected consumption growth. Hence, Wachter essentially fuses Bansal and Yaron's x_t and σ_t^2 into a singe state variable. As in the original long-run risks model, $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$ are approximately proportional to one another and the $E[M|R]$ curve is time-invariant as a result.

Lastly, we consider Constantinides and Ghosh (2017), who propose a quantitative implementation of the Constantinides and Duffie (1996) mechanism. Their model features recursive preferences and a persistent state that simultaneously controls time-variation in (i) the higher moments of households' idiosyncratic income shocks and (ii) the conditional mean of aggregate dividend growth. The state is modelled as an autoregressive Gamma process, which implies, similar to the CIR process in Wachter (2013), that the state's volatility increases in its level. When household risk rises and increases $\sigma_{m,t}^2$, the conditional mean of dividends simultaneously becomes more volatile, which increases the volatility of P/D and therefore $\sigma_{s,t}^2$. Once again, this mechanism implies that $\sigma_{m,t}^2$ and $\sigma_{s,t}^2$ are approximately proportional to one another and that the $E[M|R]$ curve is time-invariant as a result.

These examples illustrate that it is the long-run risks mechanism for time-varying volatility, rather than the details of its implementation in Bansal and Yaron (2004), that lead to the models' inconsistency with the observed variation in $E[M|R]$.

D. Time-varying Disaster Resilience

D.1. Gabaix (2012): Model statement

We consider the version of Gabaix's (2012) time-varying disaster risk model in Dew-Becker et al. (2017).[17] The representative agent has time-separable power utility and is endowed with

$$\begin{aligned}\Delta c_{t+1} =& g + \sigma\varepsilon_{t+1} + \xi_{t+1}J_{t+1}\\ \Delta d_{t+1} =& g + \lambda\sigma\varepsilon_{t+1} - L_t\xi_{t+1}\\ L_{t+1} =& (1-\rho_L)\bar{L} + \rho_L L_t + \sigma^L\varepsilon_{L,t+1}\end{aligned}$$

where $\varepsilon_{t+1} \overset{\text{IID}}{\sim} N(0,1)$ captures small consumption shocks, $\xi_{t+1} \overset{\text{IID}}{\sim} \text{Bernoulli}(p_J)$ is a disaster indicator, and $J_{t+1} \overset{\text{IID}}{\sim} N(E_J, V_J^2)$ is the disaster size. The recovery rate of dividends, L_t, follows an autoregressive process with normal innovations, $\varepsilon_{L,t+1} \overset{\text{IID}}{\sim} N(0,1)$. Hence, consumption disasters lead to a large drop in dividends if they occur in periods with high L_t and a small drop in dividends if they occur in periods with high L_t. Different from the model of Wachter (2013), but as in the original disaster risk model of Rietz (1988) and Barro (2006), the probability of disasters, p_J, is assumed to be constant.

D.2. Volatility Dynamics and $E[M|R]$ in the Rare Disaster Model

Conditional on the model's state variable L_t, dividends are only exposed to systematic shocks (ε_{t+1} and ξ_{t+1}). L_t controls the conditional variance of systematic shocks, $\sigma_{s,t}^2 = \text{Var}_t[\Delta d_{t+1}] = \lambda^2\sigma^2 + L_t^2 p_J(1-p_J)$. Because L_t evolves independent from consumption (and the pricing kernel), $\varepsilon_{L,t+1}$ represents an idiosyncratic shock that affects returns via the P/D ratio. The variance of idiosyncratic returns, $\sigma_{i,t}^2 = \text{Var}_t[\ln(P_{t+1}/D_{t+1})]$, has to be computed numerically.

[17] Dew-Becker et al.'s version of the model adds Gaussian innovations to consumption growth (which is constant absent disasters in the model's original version), it assumes a normally distributed (rather than constant) disaster size, and it specifies the recovery rate of dividends as an autoregressive (rather than linearity generating) process. These modifications do not alter the model's basic economics, but result in a pricing kernel that is a continuous (rather than discontinuous) function of the model's state. We follow Dew-Becker et al.'s calibration of the model.

What is unique about Gabaix's model is that the volatility of returns evolves independently from the pricing kernel. This feature implies that changes in return volatility are not accompanied by changes in $\sigma_{m,t}$, which is time-invariant because the pricing kernel is IID. Proposition 1 shows that, under lognormality, an increase in return volatility that is not accompanied by a change in $\sigma_{m,t}$ makes $E[M|R]$ flatter, as in the data. While Gabaix' model is not lognormal, this mechanism nevertheless appears to describe its implications for $E[M|R]$ well.

E. Discussion

As Gabaix' model shows, one explanation for our empirical finding is that stock market volatility evolves independently from the pricing kernel. This explanation is consistent with empirical evidence in Dew-Becker et al. (2017), who show that shocks to expected (as opposed to realized) stock market variance command no risk premium, as well as empirical evidence in Jurado et al. (2015), who show that stock market volatility is only weakly correlated with macroeconomic uncertainty (a plausible driver of pricing kernel volatility).

Two additional pints are worth making. First, the fact that M_{t+1} is IID in Gabaix' model is not important for its ability to rationalize our evidence. What matters is that the pricing kernel's conditional variance $\sigma^2_{m,t}$ does not comove with the conditional variance of returns, as one can see from Proposition 1. Second, what matters for the model's ability to explain our evidence is not that shocks to the variance of returns are independent from M_{t+1} (which is important for replicating the evidence in Dew-Becker et al.), but rather that shocks to the variance of returns are independent from shocks to the variance of the pricing kernel. This is a subtle but important distinction. It is conceivable that shocks to return variance affect the realized value of the pricing kernel without altering its conditional variance, which would rationalize our evidence but not that of Dew-Becker et al. (2017). Assuming that the variance of returns is independent of both M_{t+1} and its conditional moments, as in Gabaix's model, is an intuitive explanation for both findings.

IV. Conclusion

Option markets provide us with valuable information to assess how the pricing of stock market risks varies over time. We show that negative returns are substantially more painful to investors when they occur in periods of low stock market volatility, which is reflected in a steeper projected pricing kernel and larger (negative) risk premia on put options. This evidence provides a useful diagnostic test for asset pricing models, which routinely assume difficult-to-measure dynamics in preferences and fundamentals to rationalize asset prices. We show that many popular models require counterfactual dynamics of the pricing kernel to explain asset price dynamics, including the countercyclical nature of volatility and long-horizon predictability of excess returns.

We trace the source of the discrepancy of models relative to the data back to the joint dynamics of returns and pricing kernel. Our Proposition 1 shows that the observed variation in the projected pricing kernel is consistent with return volatility evolving close to independently from the volatility of the pricing kernel. We find this mechanism to be present in the model of Gabaix (2012), and hence the model is well in line with our empirical evidence. Our theoretical finding is also supported by prior empirical evidence. In particular, Jurado et al. (2015) show that macroeconomic uncertainty is considerably more persistent and only weakly correlated with return volatility. Additionally, they show that an increase in macroeconomic uncertainty is associated with a decline in future economic activity, whereas Berger et al. (2020) show that the same is not true for an increase in expected stock market volatility. It therefore makes sense that market volatility is not related to investors' marginal utility.

Appendix

This appendix explains how we extract risk-neutral distributions from option prices and details the time series model for conditional volatility.

A. Extracting Risk-neutral Densities from Options

We follow the methodology in Beason and Schreindorfer (2022) to extract risk-neutral densities from option prices. Breeden and Litzenberger (1978) show that the risk-neutral PDF of the future price level S_{t+1} is given by

$$f_t^*(S_{t+1}) = R_t^f \times \frac{\partial^2 P_t(K)}{\partial K^2}\bigg|_{K=S_{t+1}}, \tag{A.1}$$

where P is the price of a put option and K the associated strike price. The risk-neutral PDF of ex-dividend returns follows from the change of variables $R_{t+1} = \frac{S_{t+1}}{S_t}$ as $f_t^*(R_{t+1}) = S_t \times f_t^*(S_{t+1})$. To recover risk-neutral densities from options based on (A.1), it is necessary to observe option prices for the desired maturity and a continuum of strikes. We generate these prices via interpolation and extrapolation of observed quotes as follows. For each day in the sample, we use Black's formula (a version of Black and Scholes 1973) to convert observed option prices to implied volatility (IV) units, fit an interpolant to them, evaluate the interpolant at a maturity of 30 calendar days and a fine grid of strike prices, map interpolated IVs back to option prices, and finally compute f_t^* via finite differences based on (A.1). Importantly, this approach does not assume the validity of the Black-Scholes model because Black's formula is merely used to map back-and-forth between two spaces. The mapping relies on LIBOR rates that are linearly interpolated to the options' maturities and forward prices for the underlying. The remainder of this appendix details the interpolation of IVs.

The SVI Method

We interpolate IVs based on Jim Gatheral's SVI method.[18] SVI describes implied variance (the square of IV) for a given maturity τ with the function

$$\sigma^2_{BSM}(x) = a + b\left(\rho(x-m) + \sqrt{(x-m)^2+\sigma^2}\right), \tag{A.2}$$

where $x = \log(K/F_{t,\tau})$ is the option's log-moneyness, $F_{t,\tau}$ the forward price for maturity τ, and a, b, ρ, m, σ are parameters. The method is widely used in financial institutions because it is parsimonious, yet known to provide a very good approximation to IVs, both in the data and in fully-specified option pricing models.

We make two modifications to the basic SVI method to allow for interpolation in the maturity, in addition to the moneyness dimension. First, we parameterize σ^2_{BSM} as a function of standardized moneyness, $\kappa \equiv \frac{\log(K/F_{t,\tau})}{VIX_t/100\times\sqrt{\tau}}$, rather than x, to limit the extent to which the shape of the IV curve varies with maturity. Second, we specify linear functions of τ for the five coefficients, e.g.,

$$a = a_0 + a_1\tau, \tag{A.3}$$

and similarly for (b, ρ, m, σ). Jointly, (A.2) and (A.3) describe IVs as a bivariate function of κ and τ that is parameterized by $\theta \equiv (a_0, a_1, b_0, b_1, \rho_0, \rho_1, m_0, m_1, \sigma_0, \sigma_1)$.

An important criterion for the successful interpolation and extrapolation of IVs is that the corresponding option prices respect theoretical no arbitrage restrictions, i.e. that they are (i) non-negative, (ii) monotonic in K, (iii) convex in K, and (iv) imply (via Equation A.1) a density $f^*_t(R)$ that integrates to one. We impose these constraints in the estimation as further described below.

Data and Implementation

We clean the options data by removing observations that (i) violate the static no-arbitrage bounds $P \leq K/R^f$ or $C \leq S$, (ii) have a bthbid quote of zero, (iii) have the

[18] SVI was devised at Merrill Lynch and disseminated publicly by Gatheral (2004). See Gatheral (2006) for a textbook treatment and Berger et al. (2020) for a recent application in economics.

CBOE's error code 999 for ask quotes or 998 for bid quotes, (iv) have non-positive bid-ask spreads, (v) have midquotes less than \$0.50, (vi) are singles (a call quote without a matching put quote or vice versa), (vii) are PM settled, or (viii) have IVs less than 2% or more than 200%. To detect any additional outliers, we fit a linear function κ and τ to IVs on each date, and remove observations that are highly influential based on their Cook's distance (a common statistical metric for detecting outliers). Finally, we restrict the sample to puts with a standardized moneyness below 0.5, calls with a standardized moneyness above -0.5, and maturities between 8 and 120 calendar days, i.e. we exclude long-term and in-the-money options.

For each day in the sample, we estimate the SVI parameter vector θ by minimizing the root mean squared error between observed IVs and the SVI interpolant,

$$\hat{\theta}_t = \underset{\theta}{\operatorname{argmin}} \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} [\sigma_{BSM,t,i} - \sigma_{BSM}(\kappa_{t,i}, \tau_{t,i}; \theta)]^2}, \tag{A.4}$$

where N_t is the number of observations on day t. We use a particle swarm algorithm to minimize the objective function and discard parameters for which SVI-implied prices violate no arbitrage constraints. The positivity, monotonicity, and convexity of option prices are checked on a bivariate grid for κ and τ.[19] At every maturity in the τ-grid, we integrate f_t^* over the κ-region and discard parameters for which these integrals do not fall within (a numerical error tolerance of) 1 basis point of one.

The fit to IVs results in an average (median) R^2 of 98.8% (99.6%) across the 7,556 trading days in our sample.

B. Conditional Volatility Estimation

Our implementation of the HAR (Heterogeneous AR) model of Corsi (2009) is:

$$RV_t^{(21)} = \alpha + \beta^m RV_{t-21}^{(21)} + \beta^w RV_{t-21}^{(5)} + \beta^d RV_{t-21}^{(1)} + \epsilon_t, \tag{B.1}$$

[19] The κ-grid includes the integers from -20 to -11, 61 equally-spaced points between -10 and 5, and the integers from 6 to 10, for a total of 76 points. The width of this grid ensures that even extrapolated option prices are arbitrage free. The τ-grid is equally-spaced with 12 points between 10 and 120 days to maturity.

where the realized volatility $RV_t^{(1)} = (\sum_{i=1}^{N} r_i^2)^{0.5}$ denotes the square root of the sum of N squared intra-day log returns of day t, and $RV_t^{(h)} = (\frac{1}{h}\sum_{j=0}^{h} RV_{t-j}^{(1)})^{0.5}$. In the model, the past week $RV_t^{(5)}$ and the past month $RV_t^{(21)}$ represent the long-memory feature of the volatility model. We calculate RV based on squared five-minute log returns, which is a popular choice, as it presents a good trade-off between reducing noise (high sampling frequency) and reducing bias due to micro-structure effects (low sampling frequency). We sub-sample our estimator every minute, which reduces the noise without any bias, and add the squared log overnight return to each intra-daily variance estimate.

The intra-day returns are based on high frequency future prices for the S&P 500 index obtained from Tick Data Inc. In 1997, the CME introduced the so-called mini future (symbol: ES). Over time, the standard "large" futures contract (symbol: SP) lost market share to the mini, and eventually was discontinued in 2021. Since the trading volume of the mini (adjusted for the smaller multiplier) overtook the large contract during the year 2002, we switch our RV calculation from the large contract to the mini in 2003.

Our σ_t volatility forecasts are all out-of-sample. For this, on day t, we use all information available up to date $t-21$ trading days, estimate the model with OLS, and then forecast volatility using day t information only. This is done daily in an expanding window fashion. We start the sample in 1988, in order to have at least two years as burn-in for the first forecast on Jan 02, 1990. We note that our model forecasts volatility very well with an out-of-sample $R^2_{OOS} = 60.4\%$.

References

Aït-Sahalia, Y. and A. W. Lo (2000): "Nonparametric risk management and implied risk aversion," *Journal of Econometrics*, 94, 9–51.

Aretz, K., M.-T. Lin, and S.-h. Poon (2022): "Moneyness, Volatility, and the Cross-Section of Option Returns," *Review of Finance, forthcoming*.

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

Barone-Adesi, G., R. F. Engle, and L. Mancini (2008): "A GARCH option pricing model with filtered historical simulation," *Review of Financial Studies*, 21, 1223–1258.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," 121, 823–866.

Beason, T. and D. Schreindorfer (2022): "Dissecting the Equity Premium," *Journal of Political Economy*, 130, 2203–2222.

Berger, D., I. Dew-Becker, and S. Giglio (2020): "Uncertainty shocks as second-moment news shocks," *The Review of Economic Studies*, 87, 40–76.

Black, F. (1976): "Studies of stock price volatility changes," *Proceedings of the 1976 Meetings of the American Statistical Association*, 171–181.

Black, F. and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Bliss, R. R. and N. Panigirtzoglou (2004): "Option-implied risk aversion estimates," *Journal of Finance*, 59, 407–446.

Breeden, D. T. and R. H. Litzenberger (1978): "Prices of state-contingent claims implicit in option prices," *Journal of Business*, 51, 621–651.

Campbell, J. Y. and J. H. Cochrane (1999): "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 107, 205–251.

Chapman, D. A. (1997): "Approximating the asset pricing kernel," *Journal of Finance*, 52, 1383–1410.

Christoffersen, P., S. Heston, and K. Jacobs (2013): "Capturing option anomalies with a variance-dependent pricing kernel," *Review of Financial Studies*, 26, 1963–2006.

Cochrane, J. H. (2005): *Asset pricing: Revised edition*, Princeton university press.

Cohn, A., J. Englemann, E. Fehr, and M. A. Maréchal (2015): "Evidence for Countercyclical Risk Aversion: An Experiment with Financial Professionals," *American Economic Review*, 105, 860–885.

Constantinides, G. M. and D. Duffie (1996): "Asset Pricing with Heterogeneous Consumers," 104, 219–240.

Constantinides, G. M. and A. Ghosh (2017): "Asset Pricing with Countercyclical Household Consumption Risk," *Journal of Finance*, 72, 415–460.

Corsi, F. (2009): "A simple approximate long-memory model of realized volatility," *Journal of Financial Econometrics*, 7, 174–196.

Dew-Becker, I. and S. Giglio (2022): "Risk preferences implied by synthetic options," *Working Paper*.

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

Dittmar, R. F. (2002): "Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns," *Journal of Finance*, 57, 369–403.

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

Driessen, J., J. Koëter, and O. Wilms (2020): "Behavioral in the Short-run and Rational in the Long-run? Evidence from S&P 500 Options," *Working Paper, Tilburg University*.

Gabaix, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance," 127, 645–700.

Gatheral, J. (2004): "A parsimonious arbitrage-free implied volatility parametrization with application to the valuation of volatility derivatives," Talk at Global Derivatives.

——— (2006): *The Volatility Surface: A Practitioner's Guide*, Wiley Finance.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993): "On the relation between the expected value and the volatility of the nominal excess return on stocks," *Journal of Finance*, 48, 1779–1801.

Gormsen, N. and C. Jensen (2022): "Higher-Moment Risk," *Working paper*.

Guiso, L., P. Sapienza, and L. Zingales (2018): "Time varying risk aversion," *Journal of Financial Economics*, 128, 403–421.

Hu, G. and K. Jacobs (2020): "Volatility and Expected Option Returns," *Journal of Financial and Quantitative Analysis*, 55, 1025–1060.

Jackwerth, J. C. (2000): "Recovering risk aversion from option prices and realized returns," *Review of Financial Studies*, 13, 433–451.

Jones, C. S. (2006): "A nonlinear factor analysis of S&P 500 index option returns," *Journal of Finance*, 61, 2325–2363.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105, 1177–1216.

Kim, H. J. (2022): "Characterizing the Conditional Pricing Kernel: A New Approach," *Working Paper, University of Houston*.

Linn, M., S. Shive, and T. Shumway (2018): "Pricing kernel monotonicity and conditional information," *Review of Financial Studies*, 31, 493–531.

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

Newey, W. K. and K. D. West (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55, 703–708.

Rietz, T. A. (1988): "The equity risk premium: a solution," 22, 117–131.

Rosenberg, J. V. and R. F. Engle (2002): "Empirical pricing kernels," *Journal of Financial Economics*, 64, 341–372.

Ross, S. (2015): "The recovery theorem," *Journal of Finance*, 70, 615–648.

Rubinstein, M. (1994): "Implied Binomial Trees," 771–818.

Schreindorfer, D. (2023): "By Force of Habit and Cyclical Leverage," *Working Paper*.

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