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 30, 2022

Abstract

We use options and return data to show that negative stock market returns are significantly more painful to investors when they occur in periods of low volatility. In contrast, popular asset pricing theories imply that the pricing of stock market risk does not vary with volatility, or that it moves in the opposite direction. Our finding suggests that stock market volatility evolves largely independently from the pricing kernel. We embed this assumption into a consumption-based model with a disappointment averse investor. The model captures the dynamics of the pricing kernel and resolves four recent puzzles about stock market risk premia.

JEL Classification: G12, G13, G33

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

^{*}Contact: david.schreindorfer@asu.edu and tobias.sichert@hhs.se. We thank Caio Almeida, Tyler Beason, Oliver Boguth, Ian Dew-Becker (discussant), Kris Jacobs (discussant), Christian Skov Jensen (discussant), Christian Julliard, Lars-Alexander Kuehn, Jonathan Payne, Paola Pederzoli (discussant), 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, and ITAM, the Virtual Derivatives Workshop, the JEF seminar, the BI-SHOF Conference 2022, the CICF 2022, the SoFiE Conference 2022, and the EFA 2022 for helpful comments and suggestions. We thank Ian Dew-Becker for sharing data. Tobias Sichert acknowledges support from the Jan Wallanders and Tom Hedelius Foundation and the Tore Browaldhs Foundation, grant no. Fh21-0026.

Research on asset pricing centers on the idea of a pricing kernel, i.e., a stochastic process that is guaranteed to exist in the absence of arbitrage opportunities and prices all payoffs. The pricing kernel has to be very *volatile* to explain the large average excess return on the aggregate stock market (Hansen and Jagannathan 1991; Alvarez and Jermann 2005). We document that the pricing kernel displays very little *comovement* with return volatility. This finding implies a surprising amount of time-variation in the pricing of stock market risks and it stands in sharp contrast to the economic mechanisms of prominent asset pricing theories.

For our purposes, it is essential to distinguish between the pricing kernel, M_{t+1} , which is a function of many different shocks, and its projection onto stock market returns, $E_t[M_{t+1}|R_{t+1}]$, which is a function of only returns. The "projected pricing kernel" reveals how marginal utility is expected to vary with returns, conditional on investors' information set at time t, 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 of the projected pricing kernel for the 10th and 90th percentile of volatility. The steeper curve in periods of low volatility implies that negative returns are considerably *more* painful to investors in calm markets. We show that the difference is statistically significant and robust to alternative estimation methods.

To interpret the empirical evidence, we derive the projected pricing kernel 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 changes in the conditional distribution of the pricing kernel, such as an increase in its conditional volatility. Counter to first intuition, the finding in



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.

Figure I therefore suggests that return volatility evolves mostly independently from the pricing kernel and that investors' risk aversion does not vary with volatility.

Our findings provide a unified view of four facts about stock market risk premia that have been highlighted as puzzling in prior work. First, it is challenging to detect a risk-return trade-off in the time series of stock market returns (Glosten et al. 1993) and simple volatility-timing strategies earn positive CAPM alphas as a result (Moreira and Muir 2017). Figure I shows that this is the case because risk prices for negative returns are lower when such returns are more likely to occur. Second, whereas shocks to realized stock market variance carry a large negative risk premium (Carr and Wu 2009), shocks to future variance do not (Dew-Becker et al. 2017). Since a large realized variance frequently coincides with large negative returns, the variance risk premium can be explained by the steep slope of the projected pricing kernel. In turn, investors' indifference towards shocks to expected variance is consistent with our result that return volatility evolves largely independently from the pricing kernel. Third, the risk-neutral return distribution implied by equity index option prices is much more non-normal than the physical distribution, which is reflected in a large difference between the VIX index and a related simple VIX, or SVIX, index (Martin 2017). This sizable risk-adjustment reflects the steep slope of the projected pricing kernel, which increases the amount of risk-neutral tail risk relative to its physical counterpart. Fourth, the majority of the equity premium reflects shocks that coincide with monthly stock market returns between -30% and -10% (Beason and Schreindorfer 2022). This is the case because such returns tend to coincide with very high levels of marginal utility, as shown in Figure I. Hence, all four puzzles can be traced back to properties of the projected pricing kernel. While the prior literature has shown that popular consumption- and intermediary-based asset pricing models are inconsistent with these facts, it has not recognized the connection between them.

Contrary to Figure I, we show that 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. This counterfactual pricing kernel behavior is an inherent feature of the models' economic mechanisms, which rely on a tight connection between return volatility and the pricing kernel to rationalize the level and predictability of stock returns. We argue that, without this tight connection, the habit model is not able to explain the countercyclical volatility and long-horizon predictability of stock returns, whereas the long-run risks model loses its ability to explain return predictability and most of the equity premium.

We consider a number of additional models with habits and recursive utility, most of which were explicitly designed to capture features of stock market volatility and option markets. Specifically, we show that the models of Drechsler and Yaron (2011), Wachter (2013), Constantinides and Ghosh (2017), and Bekaert and Engstrom (2017) closely resemble the original habit and long-run risks models in terms of their implications for the projected pricing kernel.

In the last section, we propose a consumption-based asset pricing model to rationalize our empirical finding. The model assumes that consumption growth is IID, whereas dividend growth is heteroscedastic. We do not provide an explicit microfoundation, but argue that this assumption is consistent with countercyclical leverage (Halling et al. 2016). IID consumption growth implies that the pricing kernel's conditional distribution is time-invariant. Heteroscedastic dividend growth implies that the volatility of returns is time-varying, which makes the pricing kernel's projection onto returns flatter in volatile times. This feature is implied by our analytical results and it reflects the intuition that negative returns are more indicative of deteriorating macroeconomic conditions if they occur in calm markets, rather than the middle of a recession. We assume additionally that the representative agent is disappointment averse (Gul 1991; Routledge and Zin 2010). This assumption captures the steep slope of the projected pricing kernel, as previously shown by Schreindorfer (2020). A calibration of the model is quantitatively consistent with Figure I and it resolves the four aforementioned puzzles.

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 one exception 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

 $^{^{1}}$ 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.

the projected pricing kernel to macro-finance puzzles.

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.

We also relate to Bliss and Panigirtzoglou (2004), who use option prices to estimate the representative investor's relative risk aversion. Their study focuses on the level of risk aversion, but shows 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 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 projection. Kim shows how the projected pricing kernel varies with many different macroeconomic covariates and analyzes implications for conditional risk premia. In contrast, we focus on covariation of the *projected* pricing kernel with volatility and link it to properties of the (*multivariate*) pricing kernel and the economic mechanisms of equilibrium models.²

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.

²There are also numerous implementation differences. Our estimation approach has the advantage of being likelihood-, rather than moment-based, and it guarantees that the implied density of stock market returns integrates to one on every day of the sample, which is not the case for Kim's estimator. Additionally, our parameterization of the projected pricing kernel implies considerably more time-variation (as a function of volatility) and results in a better fit to the data.

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 equals³

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

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

³The fact that the pricing kernel equals the ratio of risk-neutral to physical probabilities (scaled by R^{f}) is a well-know 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

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

$$LL(\theta) = \sum_{t=1}^{T} \ln f_t(R_{t+1}; \theta).$$
(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 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,⁴

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

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

⁴Prior 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.^{5,6} 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),

$$M(R_{t+1}, \sigma_t; \theta) = \exp\left\{\delta_t + \sum_{i=0}^N c_i \times \left(\frac{\ln R_{t+1}}{\sigma_t}\right)^i\right\}.$$
(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 $H_0: b = 0$.

⁵The 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.

⁶Instead of computing δ_t based on the theoretical restriction $\int f = 1$, one could add a timevarying 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.⁷

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.

⁷Volatility 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.

Ν	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).⁸ 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.

⁸There 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 (riskneutral) 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 = 2and 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.⁹ 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-

⁹The linear case (K = 1) implies too little time-variation in c_{it} 's and therefore too little timevariation 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 = 1case 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 6 for details.

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

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

Finally, time-variation in the projected pricing kernel implies that expected option returns vary systematically with volatility. In the online appendix, we use this observation to provide non-parametric support for our parametrization of E[M|R]. Specifically, Figure IA.VI shows that average put option returns in our sample are significantly more negative in periods of low volatility than in periods of high volatility. These returns are computed based on observed option quotes and realized returns, i.e., without any parametric assumptions, and we show that they line up closely with the expected put returns that are implied by our parametric estimate of E[M|R]. Average option returns therefore provide additional support for the assumed functional form of E[M|R] and the amount of time-variation in its shape.



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.

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 $\sigma_t^2 - (\frac{VIX_t}{100})^2$ has a bootstrapped *p*-value of 0.00% under the sampling distribution of the N = 1 estimator. Hence, a log-linear projected pricing kernel is inconsistent with risk premia on higher moments.

Whereas the log-linear estimator 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 = 2.¹⁰ 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 logquadratic specification captures it well.

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

This section derives the determinants of time-variation in the projected pricing kernel. Based on these findings, we illustrate that prominent consumption-based models are inconsistent with the dynamics of E[M|R] because they assume a counterfactually tight connection between return volatility and the pricing kernel.

A. Determinants of Time-variation in E[M|R]

The following result relates the slope of the *projected* pricing kernel to the second moments of returns and the pricing kernel.

¹⁰The 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 $\left(\int_0^{0.9} (R-1)[f(R) - f^*(R)]dR\right) / \left(\int_0^{\infty} (R-1)[f(R) - f^*(R)]dR\right)$.

PROPOSITION 1 (E[M|R] under lognormality). If the log pricing kernel and log returns are jointly normal with variance $[\sigma_m^2, \sigma_r^2]$ and correlation ρ_{mr} , the slope of the projected pricing kernel equals

$$\frac{\partial E[M|R]}{\partial R} = \rho_{mr} \frac{\sigma_m}{\sigma_r} \times \frac{E[M|R]}{R}.$$
(10)

Proof. Normality implies that the log pricing kernel, conditional on log returns, is distributed as $\ln M | \ln R \sim N \left(\mu_m + \rho_{mr} \frac{\sigma_m}{\sigma_r} \left[\ln R - \mu_r \right], \sigma_m^2 (1 - \rho_{mr}^2) \right)$, where μ_m is the mean of $\ln M$ and μ_r the mean of $\ln R$. Using the moment generating function of a normal random variable, the conditional expectation of the pricing kernel equals

$$\mathbb{E}[M|R] = \exp\left\{\mu_m + \rho_{mr} \frac{\sigma_m}{\sigma_r} \left[\ln R - \mu_r\right] + \sigma_m^2 (1 - \rho_{mr}^2)/2\right\}.$$

Differentiating with respect to R yields (10).

Note that a positive equity premium requires $\rho_{mr} < 0$, so that $\frac{\partial E[M|R]}{\partial R}$ is negative. Proposition 1 generates two useful insights. First, all else equal, an increase in σ_r makes E[M|R] flatter, similar to what we saw empirically in Figure I. The intuition for this result is that a decrease in σ_r increases the informativeness of returns about the macroeconomy. For example, a -10% 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. Second, if an increase in σ_r leaves E[M|R] unchanged or makes it steeper, it must be accompanied by at least a proportional increase in $|\rho_{mr} \times \sigma_m|$. If σ_m is proportional to σ_r , for example, a -10% drop in the market will be equally informative about macroeconomic fundamentals during low and high volatility times.

Figure V shows that the second scenario describes the models of Campbell and Cochrane (1999) and Bansal and Yaron (2004). In particular, the habit model implies that E[M|R] becomes *steeper* when volatility is high, whereas long-run risks model implies that the shape of E[M|R] does not vary with volatility. Since both models are conditionally lognormal, we can rely on Proposition 1 to connect this (lack of) time-variation in E[M|R] to their economic mechanisms.



Figure V: E[M|R] in log-normal asset pricing models. This figure represents the theoretical equivalent to Figure I. We plot the projected pricing kernel, $E_t[M_{t+1}|R_{t+1}]$, for periods in which the conditional volatility of stock market returns is at its 10th or 90th percentile.

B. Habits

The Campbell and Cochrane (1999) model provides a simple economic mechanism for the long-horizon predictability and countercyclical volatility of stock market returns. When consumption falls relative to the agent's habit level, risk aversion rises. Higher risk aversion leads to higher expected returns and a lower price-dividend ratio, because the distribution of future dividend growth rates is time-invariant. Hence, the price-dividend ratio predicts future returns. At the same time, a drop in consumption makes the habit more sensitive to additional consumption shocks, which increases the conditional volatility of the pricing kernel, and thereby the conditional volatility of the price-dividend ratio and returns. Hence, stock market volatility moves countercyclically.

Proposition 1 makes it straightforward to understand time-variation in E[M|R]in the habit model. First, the fact that the habit controls the conditional volatilities of both returns and the pricing kernel implies that σ_r and σ_m are positively related to one another. We illustrate this effect in the left panel of Figure VI (dotted black



Figure VI: Sources of E[M|R]'s time-variation in log-normal models. This figure shows how the conditional volatility of the log pricing kernel (σ_m , left y-axis, dotted line) and the conditional correlation between log returns and the log pricing kernel (ρ_{mr} , right y-axis, solid line) move in relation to the conditional volatility of log returns (σ_r , x-axis). For the long-run risks model, where different combinations of states can lead to the same σ_r but different σ_m and ρ_{mr} , we plot medians and 95% confidence bounds of σ_m and ρ_{mr} . Proposition 1 shows how the shape of E[M|R] relates to σ_r , σ_m , and ρ_{mr} .

line; left y-axis). Second, the time-varying sensitivity of the habit process implies that both the pricing kernel and the price-dividend ratio become more exposed to consumption shocks when risk aversion (and volatility) is high. As a result, returns and the pricing kernel become more negatively correlated to one another in volatile times (solid blue line; right y-axis). Jointly, these effects imply that an increase in σ_r is accompanied by more than a proportional increase in $|\rho_{mr} \times \sigma_m|$. Proposition 1 shows that, as a result, E[M|R] becomes steeper when return volatility rises.

The model's counterfactual prediction about the variation in E[M|R] results directly from the habit mechanism. In particular, if the habit did not become more sensitive to consumption shocks when risk aversion is high, return volatility would not move countercyclically. Additionally, as explained in Section II.C of Campbell and Cochrane (1999), the habit's countercyclical sensitivity is required to ensure that the habit level moves nonnegatively with consumption, i.e., that it behaves in an economically sensible manner. The model's prediction about the projected pricing kernel can therefore not be changed without altering its main economic mechanism.

C. Recursive Utility

The mechanism of Bansal and Yaron (2004) rationalizes the level and predictability of the equity premium. Their model assumes persistent variation in the conditional mean (x_t) and volatility (σ_t) of consumption growth, along with an Epstein and Zin (1989) agent who strongly dislikes such "long run risks". Negative shocks to x_t and positive shocks to σ_t are therefore associated with high levels of marginal utility. Dividends are subject to the same variation in conditional moments as consumption, so that equity becomes less attractive when marginal utility is high. Hence, investors require a premium for holding stocks. Additionally, higher levels of σ_t are associated with riskier dividends and hence higher discount rates, which lowers the price-dividend ratio. As a result, the price-dividend ratio predicts returns.

Based on Proposition 1, it is straightforward to show that the volatility channel is responsible for the model's counterfactual implications for E[M|R]. In particular, in order for volatility risk to generate an equity premium, it is essential that σ_t controls the volatility of both dividends and consumption, as it would otherwise not induce covariation between returns and the pricing kernel. This assumption implies, however, that σ_t also controls the volatility of returns and the pricing kernel, which are therefore positively related to one another. We show in the right panel of Figure VI that σ_m is roughly proportional to σ_r , whereas ρ_{mr} is approximately constant. As Proposition 1 shows, these features imply that the slope of E[M|R] is time-invariant.

To eliminate this undesirable implication of the model, one could generate the equity premium entirely from x-risk. Unfortunately, doing so leads to two undesirable consequences. First, the model looses its ability to rationalize the predictability of stock returns, which results from variation in σ_t . Second, Beeler and Campbell (2012) show that relying on x_t as a source of risk premia leads to counterfactual implications about the variance ratios of consumption growth. In response to this critique, Bansal et al. (2012) proposed a new calibration of the model that attributes most of the equity premium to σ -risk, i.e., to the source of the model's inconsistency

with E[M|R]. It therefore appears infeasible to align the model with the dynamics of the pricing kernel without substantial changes to its basic structure.

D. Other Models

In Section V of the online appendix, we consider a number of additional recursive utility models with non-normal shocks. Like the original long-run risks model, Drechsler and Yaron (2011), Wachter (2013), and Constantinides and Ghosh (2017) rationalize the equity premium by interacting recursive utility with a persistent state variable that controls the volatility of returns. We show that, as in Bansal and Yaron (2004), σ_m is approximately proportional to σ_r in these models and E[M|R] is approximately time-invariant as a result. We also consider the model of Bekaert and Engstrom (2017), which combines the utility function of Campbell and Cochrane (1999) with exogenous variation in the volatility of consumption and dividends in the spirit of Bansal and Yaron (2004). In this model, σ_m increases approximately linearly in σ_r , ρ_{mr} decreases approximately linearly in σ_r , and E[M|R]is approximately time-invariant as a result. With regard to the pricing kernel, the Bekaert-Engstrom model therefore behaves more like a long-run risks than an external habit model. These results show that Proposition 1 is helpful for understanding the time-variation of E[M|R] in models with non-normal shocks, despite the fact that it was derived under the assumption of log-normality.

In sum, other models with habits and recursive utility have similarly counterfactual implications about the projected pricing kernel as the original Campbell and Cochrane (1999) and Bansal and Yaron (2004) models. In addition to failing to capture the time-variation of E[M|R], the appendix shows that all of the aforementioned models fail to capture its steep slope. As a result, the models are inconsistent with the four facts about stock market risk premia that we highlighted in the introduction, as shown in the prior literature.

III. Model

We propose a consumption-based asset pricing model to explain the shape and timevariation of the projected pricing kernel. It is an extension of the disappointment aversion model in Schreindorfer (2020), which can explain the steep slope of the projection but implies time-invariance due to the model's IID environment. We augment the model with time-varying volatility in dividend growth rates. This simple extension accounts for the observed variation in the projected pricing kernel and allows the model to explain the four puzzling facts about risk premia that we discussed in the introduction.

A. Economy

The environment is a representative-agent pure-exchange economy with a single nonstorable consumption good. The agent trades a risk-free bond, which is in zero net supply, and equity, which is a claim to the dividends in all future periods. Log growth rates of aggregate consumption and dividends equal

$$\Delta c_{t+1} = g + \sigma^c \varepsilon_{t+1}^c,$$

$$\Delta d_{t+1} = g + \sigma_t^d \varepsilon_{t+1}^d.$$
(11)

Innovations have a mean of zero, a standard deviation of one, and follow

$$\begin{aligned} \varepsilon_{t+1}^{c} &= \sqrt{1 - \omega^2} \eta_{t+1}^{c} + \omega (\eta_{t+1}^{e} - 1), \\ \varepsilon_{t+1}^{d} &= \sqrt{1 - \omega^2} \eta_{t+1}^{d} + \omega (\eta_{t+1}^{e} - 1), \end{aligned}$$
(12)

where η^c and η^d are standard normal and η^e is exponentially distributed with a unit rate parameter. The three shocks (η^c, η^d, η^e) are mutually independent and IID over time. The joint distribution of ε^c and ε^d is therefore characterized by a single free parameter, ω , which controls their correlation. For $\omega < 0$, as assumed in our calibration, consumption and dividend growth rates are left-skewed and positively correlated. Schreindorfer (2020) provides empirical support for the correlation structure between consumption and dividends that is implied by (12). Whereas consumption growth is IID, the volatility of dividend growth evolves stochastically as

$$\ln \sigma_{t+1}^d = \nu + \varrho(\ln \sigma_t^d - \nu) + \sigma^s \eta_{t+1}^s, \tag{13}$$

where η^s is IID standard normal and independent of all other shocks. A possible micro foundation for dividend heterogeneity is time-varying leverage. Specifically, when firms' sales decline, costs that are fixed in the short-term, such as rent and wages (operating leverage) and coupon payments to debt holders (financial leverage) increase relative to profits. As a result, dividends fall and become more exposed to future shocks, i.e., their conditional volatility increases.¹¹ Because the aforementioned costs are not dead weight, however, they do not affect the volatility of aggregate consumption. Our setting is broadly consistent with the empirical evidence in Jurado et al. (2015), who show that the volatility of stock market returns is only weakly correlated with macroeconomic uncertainty. To keep things simple, we assume that macroeconomic uncertainty (the conditional volatility of consumption) is time-invariant, rather than weakly correlated with the volatility of cash flows and returns. III.G discusses an extension that would make the model more realistic along this dimension.

B. Preferences

The representative agent is generalized disappointment averse (GDA), as in Routledge and Zin (2010). GDA belongs to the class of Epstein and Zin (1989) recursive utility functions. As such, the pricing kernel involves a term for consumption risk, a term for long-run risks, and a term for disappointments. Because our setting assumes IID consumption growth, however, the long-run risks term collapses to a time-invariant constant and the pricing kernel simplifies to

$$M_{t+1} = \begin{cases} \beta_v \times e^{-\gamma \Delta c_{t+1}} & \Delta c_{t+1} > x_v \\ \beta_v \times e^{-\gamma \Delta c_{t+1}} \times (1+\theta) & \Delta c_{t+1} \le x_v. \end{cases}$$
(14)

¹¹Realistically, the shock to dividends (η^d) should therefore be negatively correlated with the shock to conditional volatility (η^s) . We abstract from this feature for tractability.

Here, β_v and x_v are composite parameters that depend on both endowment and preference parameters.¹² We refer interested readers to Routledge and Zin (2010) for additional details on the GDA utility function and to Schreindorfer (2020) for the special case with IID consumption growth.

Because consumption growth is IID, the pricing kernel's conditional distribution is time-invariant by construction. It is therefore clear that any time variation in the projected pricing kernel must result from the heteroscedasticity of dividends. It is also worth noting that, for $\theta = 0$, the pricing kernel simplifies to that of a constant relative risk aversion (CRRA) utility function. Relative to this case, GDA overweights left-tail outcomes in consumption (below the disappointment threshold x_{v}) by a factor of $(1 + \theta)$. Disappointment averse preferences were originally developed to address puzzling experimental results, such as the Allais (1979) paradox, but their generalization can also be interpreted as capturing the value-at-risk constraints of financial institutions. To illustrate the effect of disappointment aversion in the clearest possible way, our benchmark calibration eliminates aversion against regular consumption risk by setting $\gamma = 0$. The resulting pricing kernel is highly stylized. It equals a step function with a value of $\beta_v(1+\theta)$ when consumption growth falls below the disappointment threshold x_v and value of β_v otherwise. Nevertheless, as we show below, the model makes quantitatively realistic predictions about the shape of the projected pricing kernel and it matches a multitude of moments related to returns, option prices, and stock market volatility.

C. Model Solution

We solve for asset prices based on the analytical formulas in Schreindorfer (2020), with a slight modification to account for the heteroscedasticity of dividends. Specifically, we discretize the volatility process based on the method in Rouwenhorst (1995) with 31 gridpoints. Next, we evaluate asset prices conditional on the current and

¹²The constants are given by $\beta_v \equiv \beta \frac{(v/m)^{1/\psi-\gamma}}{1+\theta\delta^{1-\gamma}\mathbb{E}[1\{\Delta c_{t+1}\leq x_v\}]}$ and $x_v = \ln\left(\frac{\delta m}{v}\right)$, respectively, where v and m are the equilibrium utility-to-consumption and certainty equivalent-to-consumption ratios of the underlying Epstein and Zin (1989) utility function.

future volatility state based on the formulas in Schreindorfer (2020). Lastly, for each current volatility state, we integrate over future states based on the transition matrix of the discretized process.

D. Calibration

We calibrate the model at a monthly frequency. Table II shows parameter values in Panel A and targeted moments in Panel B. Because consumption data is noisy, especially at higher frequencies, we calibrate the consumption process based on the longest available annual dataset, which spans 1929-2019. To that end, we timeaggregate simulated monthly consumption data in the model to an annual frequency. All other parameters are calibrated based on moments of monthly returns in our 1990-2019 sample.

We measure consumption by the sum of real nondurables and services consumption per capita from the Bureau of Economic Analysis (BEA) and impute nominal dividends from cum- and ex-dividend market returns from the Center for Research in Security Prices (CRSP). Dividends are converted to real terms using the personal consumption expenditure deflator from the BEA. We set g = 0.018/12 and $\sigma^c = 0.026/\sqrt{12}$ to match the mean and standard deviation of annual log consumption growth and $\omega = -0.75$ to match its correlation with annual log dividend growth. As noted above, ω also controls the skewness of consumption and dividend growth rates and, as a result, the amount of tail risk in returns. Table II shows that our calibration implies that monthly returns have an average conditional skewness of -0.62 and kurtosis of 4.2, close to the empirical values of -0.61 and 4.4, despite the fact that the model was not calibrated to match these moments.

We set ν , ρ , and σ^s to match the mean, standard deviation, and first-order autocorrelation of the conditional return volatility. This is a simplification, as the heteroscedasticity of returns likely results from both cash flow and discount rate volatility in reality.

In most models with recursive utility, the agent's elasticity of intertemporal sub-

Table II: Calibration and moments

Panel A reports parameters for two alternative monthly calibrations and indicates which moments were used to calibrate them. The calibrations differ in the assumed risk preferences, which equal either constant relative risk aversion ("CRRA") or generalized disappointment aversion ("GDA"). Panel B shows moments in the data and the two calibrations. Δc denotes log growth rate of annual consumption, Δd the log growth rate of annual dividends, and R the monthly log ex-dividend return. Quantity moments are based on annual data over 1929-2019. The Sharpe ratios (SR) of variance swaps are from Dew-Becker et al. (2017) and the fraction of the equity premium due to returns below x% from Beason and Schreindorfer (2022), abbreviated by EP(x). All other moments are computed based on monthly returns and the corresponding density estimates f_t and f_t^* in our 1990-2019 sample.

Parameter	Description	CRRA	GDA	Targeted Moment(s)
g	Average growth rate	0.018/12	0.018/12	$E[\Delta c_t]$
σ^c	Consumption volatility	$0.026/\sqrt{12}$	$0.026/\sqrt{12}$	$\sigma[\Delta c_t]$
ω	Mixture parameter	-0.75	-0.75	$\rho[\Delta c_t, \Delta d_t]$
ν	Mean of $\ln \sigma_t^d$	-3.31	-3.31	$E\left[\sigma_t^2(R_{t+1})\right]$
σ^s	Conditional vol. of $\ln \sigma_t^d$	0.225	0.225	$\sigma \left[\sigma_t^2(R_{t+1}) \right]$
Q	Persistence of $\ln \sigma_t^d$	0.845	0.845	$AC1\left[\sigma_t^2(R_{t+1})\right]$
γ	Risk aversion	10	—	literature
θ	Disappointment magnitude	_	20	ERP/VRP
δ	Disappointment threshold	_	0.97	ERP/VRP
ψ	EIS	1.5	1.5	literature
β	Time discount rate	$0.984^{1/12}$	$0.984^{1/12}$	R^f

Panel B: Moments

Moment	Data	CRRA	GDA	Description
Targeted:				
$\overline{E[\Delta c_t]}$	1.8	1.8	1.8	avg. consumption growth $(\%/\text{year})$
$\sigma[\Delta c_t]$	2.1	2.1	2.1	std. consumption growth $(\%/\text{year})$
$\rho[\Delta c_t, \Delta d_t]$	0.52	0.52	0.52	corr. consumption/dividend growth
$E\left[\sigma_t(R_{t+1})\right]$	13.8	13.6	13.8	avg. conditional vol $(\%/\text{year})$
$\sigma\left[\sigma_t(R_{t+1})\right]$	5.9	5.9	5.9	std. conditional vol (%/year)
$AC1\left[\sigma_t(R_{t+1})\right]$	0.83	0.83	0.83	autocorr. conditional vol
$E[E_t(R_{t+1}) - E_t^*(R_{t+1})]$	8.1	2.1	8.3	equity premium $(\%/\text{year})$
$E\left[\sigma_t^2(R_{t+1}) - \sigma_t^{2*}(R_{t+1})\right]$	-12.4	-1.1	-11.6	variance premium $(\%^2/\text{month})$
$R^f - 1$		2.3	1.0	risk-free rate (%/year)
Not targeted:				
$1 - Util/Util^{no-risk}$		27.7	25.1	welfare costs (%)
$E\left[skew_t(R_{t+1})\right]$	-0.61	-0.66	-0.62	avg. conditional skewness
$E\left[kurt_t(R_{t+1})\right]$	4.4	4.3	4.2	avg. conditional kurtosis
Volatility-managed alpha	3.5	2.2	3.2	Moreira and Muir (2017) ($\%$ /year)
SR, 1m variance swap	-1.3	-0.1	-1.3	Dew-Becker et al. (2017) (annualized)
SR, 3m forward variance swap	0.1	0.0	0.0	Dew-Becker et al. (2017) (annualized)
SR, 6m forward variance swap	0.1	0.0	0.0	Dew-Becker et al. (2017) (annualized)
$E[VIX_t - SVIX_t]$	0.64	0.20	0.67	Martin (2017)
$\sigma[VIX_t - SVIX_t]$	0.56	0.18	0.56	Martin (2017)
EP(-10%)	0.80	0.32	0.84	Beason and Schreindorfer (2022)
EP(-30%)	0.13	0.01	0.09	Beason and Schreindorfer (2022)

stitution (EIS) plays a crucial role for risk premia via its interaction with persistent state variables. That is not the case in our model because consumption growth is IID, so that the EIS gets absorbed into the time-invariant constant β_v (see Footnote 12). We set the EIS to $\psi = 1.5$, the value used in Bansal and Yaron (2004), and use a time discount rate of $\beta = 0.984^{1/2}$ to generate a risk-free rate of 1% per year. We eliminate curvature in the utility function by setting $\gamma = 0$. The disappointment magnitude $\theta = 20$ and disappointment threshold $\delta = 0.97$ (which enters the composite parameter x_v ; see Footnote 12) are set to jointly match the equity premium and the variance premium. The fact that preferences can be calibrated to match both risk premia jointly is a unique feature of GDA. Specifically, when the disappointment threshold δ is lowered so that fewer and more extreme consumption events are disappointing, the disappointment magnitude θ can be increased to keep the equity premium constant. This change increases the variance premium because tail events play a disproportionately big role for the variance, compared to the mean. A similar calibration is not feasible based on standard utility functions because they rely on a single parameter to control agents' aversion against all risks.

To illustrate the model's mechanism, we also show results for a second calibration that changes risk preferences but assumes otherwise identical parameters. Specifically, the "CRRA" columns in Table II eliminate disappointment aversion by setting $\theta = 0$ and instead assume a relative risk aversion coefficient of $\gamma = 10$.

An important question is whether our GDA calibration implies a "reasonable" degree of risk aversion. To provide an assessment, we compute the welfare costs of risk for both calibrations, i.e., the fraction of consumption that the agent would give up (today and at every future date and state) in order to exchange her endowment for an alternative endowment with the same mean consumption growth rate but no risk. As shown in Table II, we find costs of 27.7% for the CRRA agent and 25.1% for the GDA agent. Since both calibrations assume an identical amount of endowment risk, this comparison shows that our GDA calibration implies less risk aversion than a CRRA calibration with a relative risk aversion coefficient of $\gamma = 10$. The model



Figure VII: E[M|R] in the CRRA and GDA models. We plot the projected pricing kernel for the 10th and 90th percentile of conditional stock market volatility. Model calibrations are shown in Table II.

therefore resolves the equity premium puzzle of Mehra and Prescott (1985).

E. The Projected Pricing Kernel

Figure VII shows the projected pricing kernel for both model calibrations. As in the data, the projection becomes steeper in times of low volatility. Since the pric-



Figure VIII: Sources of E[M|R]'s time-variation in the CRRA and GDA models. This figure is the analogue to Figure V for our model. It shows how the conditional volatility of the log pricing kernel (σ_m , left y-axis, dotted line) and the conditional correlation between log returns and the log pricing kernel (ρ_{mr} , right y-axis, solid line) move in relation to the conditional volatility of log returns (σ_r , x-axis). Model calibrations are shown in Table II.

ing kernel's distribution is time-invariant by constriction, it is clear that the timevariation in its projection onto returns results exclusively from the heteroscedasticity of dividends. However, the CRRA calibration in the bottom-left panel shows that heteroscedasticity alone can only qualitatively account for the variation in E[M|R], because the CRRA pricing kernel is considerably too flat. The bottom-right panel shows that this shortcoming can be addressed by adding GDA risk preferences to the model. Because GDA implies very high aversion against tail events in consumption, which tend to coincide with tail events in cash flows, it generates a projected pricing kernel that is very steep in the left tail. Furthermore, despite the fact that the pricing kernel itself is a discontinuous function of consumption growth (see equation 14), the model implies that its projection onto returns is a smooth function of returns. The reason is that consumption and dividend growth rates are imperfectly correlated. In particular, the more negative the realized dividend growth rate (and return), the more likely it is to coincide with a disappointing consumption event. As a result, the projected pricing kernel rises continuously when moving further into the left tail of the return distribution.

While the model is not conditionally log-normal, the intuition we derived from Proposition 1 applies, as Figure VIII shows. When dividends become more volatile, returns become more volatile as well ($\uparrow \sigma_r$). However, the pricing kernel's conditional distribution, including its volatility (σ_m) and correlation with returns (ρ_{mr}), does not change because dividend volatility evolves independently from consumption and the pricing kernel. As a result, a return of, e.g. -10%, is more likely to coincide with a bad consumption draw when volatility is high than when it is low.

F. Four Related Puzzles

The model's consistency with the projected pricing kernel allows it to explain four additional facts about stock market risk premia that have been highlighted as puzzling in prior work.

First, it is well-known that it is difficult to detect a linear risk-return trade-off in the time series of stock market returns (Glosten et al. 1993). Moreira and Muir (2017) show that, as a result, a "volatility-managed portfolio" that is invested in the market and a risk-free asset, with a market weight that is inversely proportional to the market's conditional variance, earns a significantly positive CAPM alpha. The authors show that, in contrast, existing asset pricing models generate alphas that are insignificantly different from zero because they imply an unrealistically strong risk-return trade-off. We use the volatility-managed alpha to quantify the risk-return trade-off in our model. For consistency with our other results, we use the cum-dividend return of the S&P 500 (rather than the CRSP market index) and the conditional volatility estimate implied by f_t to measure the volatility-managed alpha in our 1990-2019 sample.¹³ Our empirical estimate of 3.5% (annualized) is slightly lower than Moreira and Muir's estimate of 4.9% over 1926-2015, but it is comparable to the model-implied value of 3.2%. Based on this metric, our model is therefore consistent with the weak risk-return trade-off in the data. This prediction

¹³We otherwise follow the approach of Moreira and Muir (2017). Specifically, let $X = R_{t+1}^{\text{cum-div}} - R_t^f$, let $Y = X/\sigma_t^2$, and $c = \sigma[X]/\sigma[Y]$. The volatility-managed alpha is the intercept of a linear regression of $c \times Y$ on X.

originates in the assumption that volatility involves independently from the pricing kernel, so that the model implies little variation in expected stock market returns.

Second, Dew-Becker et al. (2017) show that variance swaps – claims to realized stock market variance between now and sometime in the future – earn very negative returns and Sharpe ratios. In contrast, forward variance swaps – claims to realized stock market variance in a future month – have holding period returns and Sharpe ratios that are insignificantly different from zero. This finding implies that investors are not concerned about shocks to future volatility in aggregate stock returns. Because our model assumes that volatility evolves independent from the pricing kernel, it naturally matches this fact. Specifically, the model implies that forward variance swaps of any maturity earn expected excess returns and Sharpe ratios of zero, as shown in Table II.

Third, Martin (2017) uses the difference between the VIX index and a related simple VIX (SVIX) to quantify deviations from lognormality in the risk-neutral distribution. He shows that existing consumption-based models are far removed from capturing the properties of VIX–SVIX, because they either imply too little or way too much (in the case of rare disaster models) tail risk in the option-implied distribution. We compute VIX and SVIX based on the estimates of f_t^* in our 1990-2019 sample and find that their difference has an average of 0.64 and standard deviation of 0.56, close to Martin's estimates of 0.77 and 0.75 over 1996-2012.¹⁴ The model provides an excellent match for these metrics with an average of 0.67 and a standard deviation of 0.56. In contrast, the CRRA calibration implies a much smaller average of 0.20. Because the two calibrations imply very similar return distributions under the physical measure, as indicated by the conditional moments in Table II, the value of VIX–SVIX in the GDA calibration predominantly reflects a risk adjustment. This adjustment reflects investors' large aversion against tail risk. Martin's metric therefore highlights the importance of capturing the steep slope of

¹⁴The squared VIX equals $VIX_t^2 = 2 (\ln E_t^*[R_{t+1}] - E_t^*[\ln R_{t+1}])$, whereas the squared SVIX equals $SVIX_t^2 = \frac{1}{(R_t^f)^2} var_t^*[R_{t+1}]$ – see Martin (2017). We evaluate these moments based on our daily estimates of $f_t^*(R_{t+1})$ and convert both indices to the usual annualized percentage units.

the projected pricing kernel.

Fourth, Beason and Schreindorfer (2022) use option prices and realized returns to quantify the importance of tail risk for the equity premium. They find that monthly returns below -10% account for 80/100 of the equity premium in the data, whereas returns below -30% account for only 13/100. In contrast, existing consumption- and intermediary-based models either attribute almost none of the equity premium to returns below -10% or they counterfactually attribute it mostly to returns below -30% (in the case of rare disaster models). Table II shows that our model attributes 84/100 of the equity premium to monthly returns below -10% and 9/100 to returns below -30%, both close to Beason and Schreindorfer's estimates over 1990-2019. This implication results directly from the steep projected pricing kernel: Because returns below -10% tend to coincide with very high levels of marginal utility, they play a disproportionately large role for the equity premium. Returns below -30%coincide with even higher values of the pricing kernel, but the fact that they occur so rarely makes them negligible for the equity premium. In the CRRA calibration of our model, returns below -10% occur with roughly the same frequency as in the GDA calibration. Because the CRRA pricing kernel is substantially flatter, however, they only contribute 33/100 to the equity premium. Hence, a realistic account of the equity premium relies crucially on capturing the steep slope of the projected pricing kernel.

G. A Model Shortcoming

Our model is kept deliberately parsimonious to illustrate the role of two channels – a lack of comovement between return volatility and the pricing kernel, and agents' attitudes towards tail risks – driving properties of the projected pricing kernel. A limitation of this simple setting is that it fails to explain the long-horizon predictability and excess volatility of stock market returns. In particular, because expected dividend growth is time-invariant, apart from a small Jensen's effect¹⁵, expected

¹⁵Expected dividend growth equals $E[e^{\Delta d}] = \frac{\exp\{g - \omega \sigma_t^d + (1 - \omega^2)(\sigma_t^d)^2/2\}}{1 - \omega \sigma_t^d}$.

returns and the dividend yield display only small amounts of variation we well.

One way to overcome this issue would be to add a second, more persistent and less volatile volatility component to the model that affects both consumption and dividends. As long as a less persistent and more volatile volatility component continues to affect only dividends, as in our baseline specification, most variation in the volatility of returns would continue to be independent from the pricing kernel over short horizons, such as a month. As a result, the model's implications about timevariation in E[M|R] would likely not change substantially. An attractive feature of this alternative setting is that the conditional volatility of consumption growth ("macroeconomic uncertainty") is more persistent and only weakly correlated with the conditional volatility of cash flows and returns, which is in line with the empirical evidence in Jurado et al. (2015). In addition, GDA preferences endogenously generate countercyclical variation in agents' effective risk aversion when paired with a persistent state variable, such as the volatility of consumption growth, because disappointments events are more likely to occur when volatility is high (Routledge and Zin 2010). This feature allows models with GDA preferences to generate more return predictability than models with standard Epstein and Zin (1989) utility based on the same amount of variation in consumption risk. We did not add a second volatility component to our model in order to highlight its mechanism for time-variation in E[M|R] in the clearest possible way.

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. This evidence provides a useful diagnostic test for asset pricing models, which routinely assume difficultto-measure dynamics in preferences and fundamentals to rationalize asset prices. We show that many popular models require counterfactual dynamics of the pricing kernel in order to explain the mean and predictability of stock market returns.

Our Proposition 1 shows that the observed variation in the projected pricing kernel is consistent with return volatility evolving close to independently from the pricing kernel. This theoretical finding is 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.

We propose a consumption-based model to explain the empirical dynamics of the pricing kernel, and show that it provides a unified explanation for four puzzles about stock market risk premia that have been documented in prior work. In retrospect, this is perhaps unsurprising, because the projected pricing kernel contains all pricing-relevant information for claims on the stock market, including options. In our view, properties of the projected pricing kernel should therefore be a primary empirical target of any theory for stock market returns.

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 riskneutral 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

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

where P is the price of a put option and K the associated strike price. The riskneutral 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.¹⁶ SVI describes implied variance (the square of IV) for a given maturity τ with the function

$$\sigma_{BSM}^2(x) = a + b \left(\rho(x - m) + \sqrt{(x - m)^2 + \sigma^2} \right),$$
 (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_{BSM}^2 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 noarbitrage bounds $P \leq K/R^f$ or $C \leq S$, (ii) have a bthbid quote of zero, (iii) have the

¹⁶SVI 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,

$$\hat{\theta}_t = \underset{\theta}{\operatorname{argmin}} \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} \left[\sigma_{BSM,t,i} - \sigma_{BSM}(\kappa_{t,i}, \tau_{t,i}; \theta) \right]^2},$$
(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 .¹⁷ 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:

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

¹⁷The 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 $R_{OOS}^2 = 60.4\%$.

References

- AÏT-SAHALIA, Y. AND A. W. LO (2000): "Nonparametric risk management and implied risk aversion," Journal of Econometrics, 94, 9–51.
- ALLAIS, M. (1979): "The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school," in *Expected Utility Hypothesis and the Paradox*, ed. by M. Allais and O. Hagon, Dordrecht, Holland: D. Reidel Publishing Co.
- ALVAREZ, F. AND U. JERMANN (2005): "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," *Econometrica*, 73, 1977–2016.
- BANSAL, R., D. KIKU, AND A. YARON (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.
- 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.
- BEASON, T. AND D. SCHREINDORFER (2022): "Dissecting the Equity Premium," Journal of Political Economy, 130, 2203–2222.
- BEELER, J. AND J. Y. CAMPBELL (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.
- BEKAERT, G. AND E. ENGSTROM (2017): "Asset return dynamics under habits and bad environment–good environment fundamentals," *Journal of Political Economy*, 125, 713–760.
- 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. 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.
- CARR, P. AND L. WU (2009): "Variance risk premiums," *Review of Financial Studies*, 22, 1311–1341.
- 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.
- 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.
- EPSTEIN, L. G. AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.
- 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.
- GUL, F. (1991): "A Theory of Disappointment Aversion," Econometrica, 59, 667-686.
- HALLING, M., J. YU, AND J. ZECHNER (2016): "Leverage dynamics over the business cycle," Journal of Financial Economics, 122, 21–41.
- HANSEN, L. P. AND R. JAGANNATHAN (1991): "Implications of SecurityMarket Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.
- 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.
- MARTIN, I. (2017): "What is the expected return on the market," *Quarterly Journal of Economics*, 132, 367–433.
- MEHRA, R. AND E. C. PRESCOTT (1985): "The equity premium: A puzzle," 15, 145-161.
- MOREIRA, A. AND T. MUIR (2017): "Volatility-managed portfolios," The Journal of Finance, 72, 1611–1644.
- 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.
- ROUTLEDGE, B. R. AND S. E. ZIN (2010): "Generalized Disappointment Aversion and Asset Prices," *Journal of Finance*, 65, 1303–1332.
- ROUWENHORST, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in Frontiers of Business Cycle Research, ed. by T. F. Cooley, Princeton, NJ: Princeton University Press, 294–330.
- SCHREINDORFER, D. (2020): "Macroeconomic tail risks and asset prices," The Review of Financial Studies, 33, 3541–3582.
- WACHTER, J. A. (2013): "Can time-varying risk of rare disasters explain aggregate stock market volatility?" Journal of Finance, 68, 987–1035.