The Implied Equity Premium*

Paul C. Tetlock† Jack McCoy ‡ Neel Shah§

November 11, 2024

Abstract

We propose and test a simple model of the equity premium implied by stock index options, assuming frictionless and complete markets for the index and index options. The model's forecasts are more accurate than forecasting benchmarks, especially when arbitrage costs are low. They closely match real market behavior, with an average equity premium of 7.9% and an implied Sharpe ratio of 0.36. Our model offers fresh economic insights into why the premium varies, including its sustained increase after the 2008 crisis, and provides a unified explanation of risk premiums for variance and higher-order moments of market returns.

*Comments welcome. First draft: January 2023. We thank Matt Massicotte for outstanding research assistance. We appreciate helpful feedback from Geert Bekaert, John Campbell, Fousseni Chabi-Yo, Kent Daniel, Ralph Koijen, Lars Lochstoer, Johnathan Loudis, Ian Martin, Stijn van Nieuwerburgh, Tano Santos, and Jeremy Stein, as well as seminar participants at the NBER, Columbia, Harvard, Pomona, Arrowstreet, and the Q Group. All errors are our own.

†Columbia Business School, paul.tetlock@columbia.edu.

‡Columbia Business School, JMccoy26@gsb.columbia.edu

§Columbia Business School, ns3481@columbia.edu

Conflict-of-interest disclosure statement

Paul C. Tetlock
I have nothing to disclose.

Jack McCoy
I have nothing to disclose.

Neel Shah
I have nothing to disclose.

The equity risk premium, the expected return of the stock market over a risk-free bond, is a key determinant of financial wealth with significant implications for the real economy. Accurate and timely estimates of the equity premium are crucial for understanding the drivers of stock price movements (Cochrane 2011) and have far-reaching effects on firm investment (Baker et al. (2003)), personal consumption (DiMaggio et al. (2020)), and monetary policy (Cieslak and Vissing-Jorgensen (2021)). Despite extensive research on the equity premium, economists' estimates of its value differ widely,[1] hindering intellectual and economic progress.

This paper proposes a new method for estimating the equity premium that is based on robust economic and econometric principles and makes accurate predictions of stock returns. The main assumptions are that markets for the stock index and its options are frictionless and complete,[2] reasonable approximations for S&P 500 index and option markets. These assumptions are necessary to ensure a coherent benchmark for the equity premium. If there are arbitrage opportunities, large trading costs, or incomplete markets, multiple values of the equity premium could be consistent with observed prices.

The model extracts maximal information from the prices of stocks and options. These prices depend not only on the equity premium, but also on the risk premiums for stock return variance, skewness, and higher-order moments. Just as investors can buy stocks to earn the equity premium, they can construct portfolios of derivatives with payoffs based on return variance and higher-order moments from portfolios of options to earn the variance and higher-order premiums. Since the prices of these derivatives depend only on option prices, they reveal market or option-implied values of stock return variance, skewness, and higher-order moments. In complete and frictionless markets, one can infer the risk premiums on these assets from allocations to these securities in a "growth-optimal" portfolio that maximizes an investor's expected long-run wealth.[3]

[1] In December 2021, the median one-year equity premium forecast is 6.0%, whereas the cross-sectional standard deviation of forecasts is 6.2% (Livingston Survey of Professional Forecasters).

[2] The market is complete over all possible realizations of market returns if options with a continuum of strike prices are available (Ross (1976)). Frictionless markets have no arbitrage, trading costs, or leverage and short sale constraints.

[3] In frictionless markets, the unique pricing kernel for any set of securities is the reciprocal of the gross return of the portfolio with growth-optimal weights on these securities (Long (1990)). This growth-optimal portfolio maximizes expected long-run wealth, which is equivalent to maximizing expected log utility of wealth. The first-order conditions of this maximization for each security show that the growth-optimal portfolio return is the reciprocal of the pricing kernel.

Our central theoretical finding is that each of these risk premiums is a weighted sum of option-implied moments of returns, where weights are positions in the growth-optimal portfolio. For instance, if the growth-optimal portfolio has positions of 130% in the stock market and -50% in variance-based derivatives, the equity premium is 1.3 times option-implied variance minus 0.5 times option-implied skewness. However, since the true weights depend on unobservable rational expectations of market returns, computing risk premiums requires additional assumptions or data. To address this issue, a seminal paper by Martin (2017) imposes plausible constraints on investor preferences and wealth, equivalent to assuming growth-optimal portfolio weights of at least 100% on stocks and 0% on variance and all higher-order derivatives, to obtain a useful lower bound on the equity premium.[4]

Instead of restricting growth-optimal weights, we estimate time-varying weights using real-time data. These weights determine the equity premium as well as the variance premium, measurable as option-implied variance minus expected market return variance. High-frequency stock prices allow precise estimation of expected variance (Andersen et al. (2001)). Applying the main theoretical result, the expected variance premium is a linear function of observable option-implied moments of returns. We estimate this relationship to recover growth-optimal portfolio weights and thus market risk premiums using real-time data on stock returns and option prices. These regressions include a finite number of option-implied moments, up to four, to approximate the growth-optimal portfolio and risk premiums.

Importantly, this new framework enables point estimates, not bounds, of option-implied risk premiums with minimal assumptions about investor preferences and wealth and minimal data requirements. These implied risk premiums are firmly grounded in asset pricing theory and forecast empirical equity and variance risk premiums well, outperforming those from Martin's lower bound at horizons ranging from monthly to annual.[5] The improvement in realized equity premium predictability is substantial with a quarterly out-of-sample R^2 of 3.17% for the implied equity premium as compared to 1.15% for Martin's lower bound. The model fits the expected variance risk premium even better, attaining an R^2 of 71.8%

[4]Chabi-Yo and Loudis (2020) provide tighter bounds based on stricter assumptions on investor preferences and wealth, incorporating information from higher order moments such as skewness and kurtosis, but use static preference weights derived from the full sample. Bakshi et al. (2019) derive expected returns from option prices alone, but assume a specific functional form where the log of the projected SDF is linear in excess returns. We briefly compare them to our estimates in Section 3.

[5]We compute risk premium forecasts for the lower bound from the log utility pricing kernel implied by the exact version of Martin's lower bound, $1/R_m$, where R_m is the gross market return.

compared to 5.4% from forecasts of the quarterly premium based on the lower bound's key assumption.

Stock return predictability from the implied equity premium is economically large, too, with monthly (annual) forecasts enabling an increase of 15% (11%) in a log utility investor's annualized expected returns. The economic magnitudes are statistically imprecise due to the short sample period (1997-2021) and the high volatility of stock returns. However, consistent with theory, excluding a 61-day period in 2008 in which the model's assumption of costless arbitrage is violated significantly strengthens return predictability from the implied equity premium in statistical and economic terms.

The average implied equity premium is 7.9% per year with 4.4% volatility, whereas Martin's lower bound averages 4.2% per year with 2.2% volatility. Our model also produces a reasonable average Sharpe Ratio of 0.37, in line with the realized Sharpe ratio of the market. Consistent with Martin's preference and wealth restrictions, the implied equity premium exceeds the lower bound on 99% of days and is greater than 0.5% below the lower bound on all days.[6] The largest deviations between the lower bound and implied equity premium forecasts occur in the post-2008 period, when the average deviation increases from 2.1% to 5.0% per year, which is larger than the average lower bound. The implied equity premium consistently falls within a narrow margin of Chabi-Yo and Loudis' (2020) conditional bounds: it is within 0.5% of the lower bound (averaging 6%) and the upper bound (averaging 10%) on 92% days. The premium tends to be closer to the lower bound before 2008, and it becomes closer to the upper bound after 2008.[7] Our estimates exhibit an 81% correlation with Bakshi et al. (2019)'s expected excess returns, which average 7.7% during our sample period.[8]

One can interpret the growth-optimal portfolio and implied equity premium as arising from a heterogeneous agent model like the classic behavioral models of Shiller (1984) and Campbell and Kyle (1993). The estimated growth-optimal portfolio weights are the stock and option positions chosen by an unconstrained rational log utility investor. The equity

[6]These minor bound violations could arise from estimation error, approximation error, or illiquidity in the prices of long-horizon options.

[7]For the full sample, the difference between the implied equity premium and Martin's lower bound averages 3.7%, and the difference averages 2.0% from the conditional upper and lower bounds of Chabi-Yo and Loudis.

[8]We replicate the conditional bounds in Chabi-Yo and Loudis (2020) and Bakshi et al. (2019) using our estimates of risk-neutral moments to make these comparisons and to calculate bounds for each date in the sample.

premium reflects the compensation required by this investor for bearing stock and derivative market risk that other investors, hereafter "behavioral" investors, do not hold in equilibrium. Behavioral investors include any individuals or institutions with different preferences from log utility of wealth, biases in beliefs, or constraints on investments.

The portfolio weights of behavioral investors can be inferred from market clearing conditions in the stock and option markets. For example, if the growth-optimal weight on the stock market exceeds 100%, as it usually does empirically, behavioral investors must hold less than 100% weight on stocks, indicating higher risk aversion than log utility. If the growth-optimal weight on variance derivatives is negative, as it is empirically, behavioral investors must hold a positive weight in these securities, consistent with a desire to hedge variance risk. Interestingly, empirical estimates of the growth-optimal weight on skewness securities suggest behavioral investors have exposure to market crashes, much like many popular hedge fund strategies with negative skewness (Mitchell and Pulvino (2001) and Malliaris and Yan (2021)).

This heterogeneous agent perspective could explain why the empirical equity premium is so high on average and why it varies over time. The high average premium arises because behavioral investors are averse to stock market as well as variance risk, relative to the growth-optimal investor, who bears a disproportionate share of these risks. The premium varies over time as behavioral investors' beliefs, preferences, and constraints vary. If they become pessimistic and demand fewer stocks, the premium increases to entice the growth-optimal investor to buy stocks. Empirical estimates of the growth-optimal weight on stocks suggest that this mechanism can explain the persistent post-2008 increase in the equity premium. Although the model estimates these time-varying portfolio weights solely from variance premium regressions on option-implied moments, they are remarkably consistent with survey data on individual investors' beliefs and stock market participation (Greenwood and Shleifer (2014) and Gallup (2022)).

The implied risk premium model fundamentally links the stock market risk premium to the variance and higher-order risk premiums in the option market. Consequently, these risk premiums exhibit large common variation over time, which aligns with evidence on market return predictability from the variance premium and tail risk (Bollerslev, Tauchen, and Zhou (2009) and Bollerslev, Todorov, and Xu (2015)). The model also accounts for the average magnitude of the expected variance premium and most of its variation over time with a

median R^2 of 71% across monthly to annual horizons.[9]

Furthermore, the model correctly predicts the existence and sign of the equity risk premium's dependence on higher-order moments of stock returns. The reason is that the pricing kernel is the reciprocal of the growth-optimal portfolio return, implying that it depends nonlinearly on stock returns with alternating signs on higher-order powers of returns, which we show using a Taylor expansion in Section 1. As a result, expected stock returns depend on higher-order return moments with alternating signs, consistent with the empirical pricing of volatility, coskewness, and cokurtosis risks identified by Ang et al. (2006), Harvey and Siddique (2000), and Christoffersen et al. (2021), respectively. Because of this nonlinearity in the pricing kernel, the model also correctly predicts that the Capital Asset Pricing Model (hereafter CAPM; Sharpe (1964) and Lintner (1965)) will fail to account for empirical risk premiums since the CAPM pricing kernel is linear in market returns.

This paper's main contribution is the implied risk premium model and estimation methodology. The model enables new estimates of all market-related risk premiums, including the equity and variance premiums, as it yields the empirical projection of the pricing kernel on to market returns. This paper differs from prior studies by exploiting the relationship between the variance risk premium and risk-neutral moments of excess market returns to obtain explicit estimates of the market pricing kernel. This kernel provides point estimates of the equity premium and insights into the pricing of market risks throughout the economy.

The implied risk premium model relates to recent studies that provide bounds on the equity premium: Martin (2017), Schneider and Trojani (2019), Chabi-Yo and Loudis (2020), and Kadan and Tang (2020). Our approach is most similar to the models by Martin, Chabi-Yo and Loudis, and Bakshi et al. (2019), but it uniquely exploits the variance risk premium for nearly exact recovery without restrictive assumptions on investor preferences and wealth or the functional form of the SDF.[10]

Prior studies examine whether risk-neutral moments, like skewness and kurtosis contribute to stock risk premiums. Bollerslev and Todorov (2011) relate tail risks in options to the

[9]For evidence that the expected variance premium is positive, see Coval and Shumway (2001), Bakshi and Kapadia (2003), Carr and Wu (2009), and Christoffersen, Heston, and Jacobs (2013).

[10]Back, Crotty and Kazempour (2022) test the risk premium bounds obtained from option prices and find them to be valid but quite loose, reducing their effectiveness for market return forecasts. Our study also is related to studies by Schneider (2019) and Beason and Schreindorfer (2022) that use options to decompose the equity premium and realized market returns.

equity premium. Kraus and Litzenberger (1978) and Harvey and Siddique (2000) show that coskewness carries a negative price of risk in the cross-section of stocks. Neuberger (2012) and Kozhan et al. (2013) examine risk premiums related to skewness. The implied risk premium framework provides a unified explanation for many of these findings, predicting a strong common component among these risk premiums, consistent with evidence of a near-perfect correlation between variance and skewness risk premiums in Kozhan et al. (2013).

1 Theory

We derive the equilibrium pricing kernel or stochastic discount factor with minimal assumptions. In Section 1.2, we interpret this pricing kernel as arising from the interaction between rational and behavioral investors.

1.1 Pricing Securities Based on the Stock Market

The model features risky securities those returns depend on the stock market's return. In each period t, investors can invest in an asset with return $R_{f,t}$, known at time t, or a risky market portfolio, with an uncertain return $R_{m,T}$ where $T > t$. The stock market's excess return is then $\tilde{R}_{m,T} = R_{m,T} - R_{f,t}$. The model also includes markets for $K-1 > 0$ derivative securities offering excess returns of $\tilde{R}^k_{m,T} - c_k$, where $k = 2, 3, ..., K$ and c_k is a constant ensuring that the cost of derivative securities is zero as discussed below. We set $c_1 = 0$ so that the market's excess return matches the form of the derivative excess returns for $k = 1$.

These securities resemble the market ($k = 1$) and swaps based on market variance, skewness, and kurtosis ($k = 2, 3, 4$) and so on. The spanning result of Bakshi and Madan (2000), as specialized in Carr and Madan (2001), shows that combinations of options with a continuum of strike prices, maturing at time T, can achieve any payoff based on the market return, $R_{m,T}$.[11] The availability of options with a continuum of strike prices effectively completes the market, equivalent to assuming $K \to \infty$ securities where all higher-order moments of market excess returns are tradable. To simplify the theory, we initially assume that $K \to \infty$ to approximate the many strike prices of traded options on the stock market. For the empirical implementation, we consider approximations of this economy with finite $K \leq 4$.

[11] Martin (2017) applies this result to replicate the squared market return, $R^2_{m,T}$.

We assume that there are no risk-free arbitrage opportunities and no trading frictions, such as transaction costs or restrictions on leverage or short sales. We also assume that gross market returns are finite and strictly positive, meaning at least one stock in the market retains positive value. Under these conditions, Long (1990) shows that there exists a strictly positive pricing kernel given by the reciprocal of the return on the portfolio combining the tradable securities with the maximum expected log return:

$$M_T = \left[R_{f,t} + \sum_{k=1}^{\infty} w_{k,t} \left(\tilde{R}^k_{m,T} - c_k \right) \right]^{-1}, \tag{1}$$

where $w_{k,t}$ is the weight of security k in this portfolio.[12] This portfolio is growth-optimal, hereafter GO, since it maximizes the expected long-run growth of investor wealth. Under these minimal assumptions, the GO portfolio prices all market-related tradable securities:

$$\mathbb{E}_t \left[M_T R_T \right] = 1, \tag{2}$$

where R_T is the return of any of the K market-related securities or the risk-free rate. Applying this pricing kernel equation to risk-free rate, yields the familiar identity: $\mathbb{E}_t M_T = R^{-1}_{f,t}$.[13]

One can express market prices in terms of "risk-neutral" expectations denoted by $\mathbb{E}^*_t$ and defined by the change of measure $M_T / \mathbb{E}_t M_T > 0$. The risk-neutral expectation of any tradable return is $\mathbb{E}^*_t R_T = R_{f,t} \mathbb{E}_t \left[M_T R_T \right] = R_{f,t}$. By subtracting $R_{f,t}$ from both sides, the risk-neutral expectation of any tradable excess return becomes zero:

$$\mathbb{E}^*_t \tilde{R}_{k,T} = 0, \forall k, \tag{3}$$

where $\tilde{R}_{k,T}$ denotes the excess return of any market-related security. Solve for c_k to obtain:

$$c_k = \mathbb{E}^*_t \tilde{R}^k_{m,T}, \forall k. \tag{4}$$

[12]The robustness of our model allows for some of these conditions to be relaxed without undermining the core theoretical framework. For example, the market may still function effectively with a sufficiently dense but finite set of options and the assumption of no trading frictions can be moderated to consider low but non-zero transaction costs or minor short selling and leverage restrictions reflecting more realistic market conditions. This adaptability makes the model both theoretically rigorous and practically applicable in various empirical settings where ideal conditions may not hold.

[13]We omit brackets in expectations that apply to a single term—i.e, $\mathbb{E}^*_t \left[\tilde{R}_{k,T} \right] = \mathbb{E}^*_t \tilde{R}_{k,T}$.

For the market, $c_1 = 0$ because $\mathbb{E}_t^* \tilde{R}_{m,T} = 0$. For the variance-based security, $c_2 = \mathbb{E}_t^* \tilde{R}_{m,T}^2$. These constants are known at time t because they are based on observable market prices—i.e., risk-neutral expectations. Empirically one can identify the constants from option prices, as we will show in equation (20).

Now we relate rational expectations of returns to risk-neutral expectations based on stock and option prices. The expected excess return of any market-related security is:

$$\begin{aligned}
\mathbb{E}_t \tilde{R}_T &= \mathbb{E}_t \left[\frac{M_T}{\mathbb{E}_t M_T} \frac{\mathbb{E}_t M_T}{M_T} \tilde{R}_T \right] \\
&= R_{f,t}^{-1} \mathbb{E}_t^* \left[M_T^{-1} \tilde{R}_T \right] \\
&= R_{f,t}^{-1} \mathbb{E}_t^* \left[\left[R_{f,t} + \sum_{k=1}^{\infty} w_{k,t} \left(\tilde{R}_{m,T}^k - \mathbb{E}_t^* \tilde{R}_{m,T}^k \right) \right] \tilde{R}_T \right], \qquad (5)
\end{aligned}$$

where the last equality substitutes for M_T^{-1} and c_k using equations (1) and (4), respectively.

We apply equation (5) to the excess returns of the stock market and traded derivatives by setting $\tilde{R}_T = \tilde{R}_{m,T}^n$ for $n = 1, ..., K$ and simplifying:

$$\mathbb{E}_t \tilde{R}_{m,T}^n - \mathbb{E}_t^* \tilde{R}_{m,T}^n = R_{f,t}^{-1} \sum_{k=1}^{\infty} w_{k,t} \left(\mathbb{E}_t^* \tilde{R}_{m,T}^{k+n} - \mathbb{E}_t^* \tilde{R}_{m,T}^k \mathbb{E}_t^* \tilde{R}_{m,T}^n \right), \forall n. \qquad (6)$$

The risk premium of any market-related security is linear in risk-neutral moments, where the term in parentheses is the excess risk-neutral moment. The only unknowns are the GO portfolio weights: $w_{k,t}$. However, since these weights vary over time and market returns are highly unpredictable, one cannot simply regress market securities' excess returns on risk-neutral moments to obtain accurate estimates of the weights.

The implied risk premium in (6) is simplest for the equity premium, $n = 1$:

$$\mathbb{E}_t \tilde{R}_{m,T} = R_{f,t}^{-1} \sum_{k=1}^{\infty} w_{k,t} \mathbb{E}_t^* \tilde{R}_{m,T}^{k+1}. \qquad (7)$$

Martin's (2017) expected equity premium formula corresponds to assuming that the stock market is the GO portfolio—i.e., setting $w_1 = 1$ and $w_k = 0$ for $k \geq 2$. In this special case,

one obtains the exact version of Martin's (2017) equation:

$$\mathbb{E}_t R_{\widetilde{m},T} \quad = \quad R_{f,t}^{-1} \mathbb{E}_t^* \tilde{R}_{m,T}^2, \tag{8}$$

where the right-hand side is the stock market's risk-neutral variance discounted by the risk-free rate. This equity premium formula applies only if a log utility investor chooses to hold a weight of 100% on stocks and 0% on all market derivatives, an unlikely special case.

The reasoning behind equation (5) offers a less restrictive alternative to assuming that the stock market is the GO portfolio. If the GO investor exploits equity premium variation by timing the market, the market weight would not be constant or equal to 1. To explore this further, we analyze a K^{th}-order approximation of the GO portfolio in which $w_{k,t} = 0$ for $k > K$, which implies that the inverse pricing kernel is a K^{th}-order polynomial:

$$M_T \approx \left[R_{f,t} + \sum_{k=1}^{K} w_{k,t} \left(\tilde{R}_{m,T}^k - \mathbb{E}_t^* \tilde{R}_{m,T}^k \right) \right]^{-1}, \tag{9}$$

where K is finite. The time-varying weights $w_{k,t}$ enable optimal market timing in K market-related securities—e.g., stocks and variance-, skewness-, and kurtosis-based derivatives when $K = 4$. With this approximation, the expected equity premium is:

$$\mathbb{E}_t \tilde{R}_{m,T} \quad = \quad R_{f,t}^{-1} \sum_{k=1}^{K} w_{k,t} \mathbb{E}_t^* \tilde{R}_{m,T}^{k+1}. \tag{10}$$

Hereafter we refer to the model of the pricing kernel in equation (9) as the implied risk premium and its application to the stock market in equation (10) as the implied equity premium (IEP). The IEP generalization of Martin (2017) shows that the predictive power of risk-neutral variance will depend on the weight of the market, $w_{1,t}$, and any nonzero weights on market-related derivative securities in the GO portfolio. For example, a nonzero quadratic term in the pricing kernel implies that the GO portfolio has a nonzero weight in variance derivatives—i.e., $w_{2,t} \neq 0$. If all additional GO weights are zero, the expected equity premium would be linear in risk-neutral variance and skewness:

$$\mathbb{E}_t \tilde{R}_{m,T} \quad = \quad R_{f,t}^{-1} \left[w_{1,t} \mathbb{E}_t^* \tilde{R}_{m,T}^2 + w_{2,t} \mathbb{E}_t^* \tilde{R}_{m,T}^3 \right].$$

If instead one allows for nonzero weights on higher-order derivatives, such as bets on market skewness, the expected equity premium would be linear in additional risk-neutral moments, such as kurtosis. Since trading strategies based on skewness or even kurtosis in market returns are feasible using deep-out-of-the-money options, this study examines approximation degrees up to $K = 4$.

The remaining challenge is estimating up to $K = 4$ unknown GO weights in the implied risk premium.[14] To tackle this problem, we exploit the pricing equation for the variance risk premium, which enables estimates of the unknown weights in terms of observable option prices and precisely estimated expected physical variance. Applying the risk premium equation (6) to market variance, $n = 2$, and setting a finite K value to approximate the GO portfolio, we obtain the following expression for the variance premium implied by option prices:

$$\mathbb{E}_t^* \tilde{R}_{m,T}^2 - \mathbb{E}_t \tilde{R}_{m,T}^2 \quad = \quad -R_{f,t}^{-1} \sum_{k=1}^{K} w_{k,t} \left(\mathbb{E}_t^* \tilde{R}_{m,T}^{k+2} - \mathbb{E}_t^* \tilde{R}_{m,T}^k \mathbb{E}_t^* \tilde{R}_{m,T}^2 \right), \tag{11}$$

Equation (11) shows that the implied variance premium equals a linear combination of higher-order risk-neutral moments weighted by the GO portfolio weights. As discussed in Section 2, we obtain precise estimates of the left-hand side variables using realized high-frequency moments for physical variance and option prices for risk-neutral variance. We also use option prices to measure the risk-neutral moments on the right-hand side. Since equation (11) holds at all times t, we use time-series regressions of the variance premium on higher-order risk-neutral moments to estimate the coefficients, $w_{k,t}$, in the GO portfolio.[15]

To provide intuition for the IEP, consider the linear, $K = 1$, approximation of the GO pricing kernel. Setting $K = 1$ in equation (11), the GO portfolio's weight on the market

[14] One could estimate predictive regressions of excess market returns on risk-neutral variance, skewness, kurtosis, and fifth moments. But the four regression coefficients, $w_{k,t}$, would be poorly estimated since realized market returns are weakly correlated with expected returns.

[15] Instead one could use a purely cross-sectional approach by applying equation (5) to derivatives based on market skewness, kurtosis, and higher-order moments, $R_T = \tilde{R}_{m,T}^k$, with $k \geq 3$. The empirical challenges are obtaining accurate estimates of higher-order physical moments of market returns and dealing with illiquidity in the options needed for very high-order risk-neutral moments.

equals the variance risk premium divided by negative risk-neutral skewness:

$$w_{1,t} = \frac{\mathbb{E}_t^* \tilde{R}_{m,T}^2 - \mathbb{E}_t \tilde{R}_{m,T}^2}{-R_{f,t}^{-1} \mathbb{E}_t^* \tilde{R}_{m,T}^3}.$$

Since empirical estimates of the variance risk premium and negative risk-neutral skewness are both consistently positive, the estimate of the GO market weight is always positive. Using this optimal market weight, we obtain a closed-form expression for the equity premium in terms of the variance premium, risk-neutral variance, and risk-neutral skewness:

$$\mathbb{E}_t \tilde{R}_{m,T} \quad = \quad \frac{\mathbb{E}_t^* \tilde{R}_{m,T}^2 - \mathbb{E}_t \tilde{R}_{m,T}^2}{-\mathbb{E}_t^* \tilde{R}_{m,T}^3} \times \mathbb{E}_t^* \tilde{R}_{m,T}^2.$$

The IEP will be positive whenever the GO market weight is positive, which it is in general. In fact, as noted in the introduction, the empirical GO market weight almost always exceeds one, which is consistent with the Martin lower bound in equation (8).

For readers seeking more intuition, Appendix 5 provides a simple two-state example that illustrates how empirical estimates of the variance premium and risk-neutral skewness enable recovery of the true equity premium and pricing kernel. The following subsection provides interpretation for the general case with any distribution of market returns.

1.2 A Heterogeneous Agent Model of Market-Related Securities

Here we offer one interpretation of the implied risk premium framework that rationalizes the proposed GO pricing kernel and offers insights into the meaning of the parameters. This interpretation exploits the result that a GO kernel prices a set of securities, e.g., the stock market and options on the market, if and only if it satisfies the first-order condition (FOC) derived from the portfolio optimization of a log utility investor who invests in these securities. Thus, we analyze a partial equilibrium model featuring the hypothetical portfolio choices of log utility or growth-optimal (GO) investors in frictionless, complete markets for market-related securities.

There are two types of investors: type G, growth-optimal, who seek to maximize expected log utility of (long-run) wealth in period T and have rational expectations of returns;

and type B who select portfolios based on potentially biased expectations of returns, non-standard preferences, and unspecified constraints. Because maximizing expected log utility is equivalent to maximizing expected long-run wealth, type G investors choose to hold the GO portfolio.[16] Although this specification of sophisticated investors has strong normative appeal (e.g., Markowitz (1976)) and support from evolutionary arguments, the existence of such investors is not necessary for the implied risk premium's validity. In period t, the endowed wealth of type G is $e_{G,t}$ and that of type B is $e_{B,t}$. We normalize the supply of stock market shares to 1 and the supply of each derivative's shares to zero.

This interpretation of the implied risk premium is notable for what it does not assume. There are no restrictions on type B investors' beliefs, preferences, or constraints, and no distributional assumptions about returns, which need not even be stationary. The tractability of frictionless, complete markets combined with the myopia of log utility investors enables this generality.

Type G (GO) investors solve the following portfolio optimization:

$$\max_{w_{G,k,t}} \quad \mathbb{E}_t \left[\ln \left(e_{G,t} R_{G,p,T} \right) \right], \forall k = 1, ...K, \tag{12}$$

$$\text{subject to} \quad R_{G,p,T} = R_{f,t} + \sum_{k=1}^{K} w_{G,k,t} \left(\tilde{R}^k_{m,T} - \mathbb{E}^*_t \tilde{R}^k_{m,T} \right), \tag{13}$$

where $R_{G,p,T}$ is the uncertain return on the type G (GO) portfolio. The FOC for optimal weights shows that the reciprocal of the GO portfolio return prices all assets:

$$\mathbb{E}_t \left[R^{-1}_{G,p,T} \left(\tilde{R}^k_{m,T} - \mathbb{E}^*_t \tilde{R}^k_{m,T} \right) \right] = 0, \forall k = 1, ...K, \tag{14}$$

where $k = 1$ is the FOC for the market. These FOCs imply that the risk-free rate satisfies:

$$\mathbb{E}_t R^{-1}_{G,p,T} = R^{-1}_{f,t}. \tag{15}$$

[16] Allowing for intermediate consumption has no impact on the results.

From these FOCs, the GO pricing kernel, $M_T = R_{G,p,T}^{-1}$, prices all assets because:

$$\begin{aligned} \mathbb{E}_t\left[M_T\left(\tilde{R}_{m,T}^k - \mathbb{E}_t^*\tilde{R}_{m,T}^k\right)\right] &= 0 \\ R_{f,t}^{-1}\mathbb{E}_t^*\tilde{R}_{m,T}^k - R_{f,t}^{-1}\mathbb{E}_t^*\tilde{R}_{m,T}^k &= 0, \forall k = 1, ...K, \end{aligned} \tag{16}$$

where $R_{f,t} = (\mathbb{E}_t M_T)^{-1}$.

Since the asset pricing equations ((3) and (16)) and pricing kernel equations ((1) and (13)) are the same, the solution to the log utility investor's portfolio choice problem results in the same GO pricing kernel as the no-arbitrage argument from Section 1. So, the same equity and variance risk premium formulas apply to the partial equilibrium considered here.

Market clearing equations in this heterogeneous agent model suggest type G (GO) investors must hold the stocks and derivatives that type B investors do not hold at equilibrium prices. These equations relate type G investors' portfolio weights, which are the unknowns in the GO pricing kernel, to interesting economic variables. Stock market clearing implies:

$$\begin{aligned} w_{G,1,t}e_{G,t} + w_{B,1,t}e_{B,t} &= P_t \\ w_{G,1,t} &= e_{G,t}^{-1}\left(P_t - w_{B,1,t}e_{B,t}\right), \end{aligned} \tag{17}$$

where P_t is the price and total capitalization of the stock market. Market clearing in each derivative security implies:

$$w_{G,k,t} = -e_{G,t}^{-1}w_{B,1,t}e_{B,t}, \forall k = 2, ...K. \tag{18}$$

Market clearing links GO weights to the portfolios of type B investors. For example, if type B investors hold only stocks and have no net holdings of any derivatives, then $w_{G,k,t} = 0$ for $k = 2, ..., K$. In this case, the equity premium is proportional to risk-neutral variance, $\mathbb{E}_t\tilde{R}_{m,T} = R_{f,t}^{-1}w_{G,1,t}\mathbb{E}_t^*\tilde{R}_{m,T}^2$, and the GO portfolio return becomes linear in excess market returns, $R_{G,p,T} = R_{f,t} + w_{G,1,t}\tilde{R}_{m,T}$. If type B investors hold no stocks or the same stock weight as type G investors, then $w_{G,1,t} = 1$ and the equity premium equals Martin's lower bound.

As another example, if type B investors hold both stocks and variance derivatives, then type G investors are exposed to variance derivatives and stocks not held by type B in-

vestors. The GO portfolio return is then quadratic in excess market return, $R_{G,p,T} = R_{f,t} + w_{G,1,t}\tilde{R}_{m,T} + w_{G,2,t}\left(\tilde{R}^2_{m,T} - \mathbb{E}^*_t\tilde{R}^2_{m,T}\right)$, implying that the second-order approximation in the no-arbitrage model will exactly account for the equity premium as in equation (10). Thus, the complexity of type B investor portfolios determines the maximum degree of risk-neutral moments needed to account for the equity premium and the polynomial degree of the GO portfolio return.

1.3 Implications of the Model

The asset pricing implications of the model depend on the GO pricing kernel, which is the reciprocal of GO portfolio return. When behavioral (type B) investors do not hold derivatives derivatives, i.e., $w_{k,t} = 0$ for $k = 2, ..., K$, this model resembles the conditional CAPM.[17] To see how the nonlinear transformation of the CAPM pricing kernel affects equilibrium pricing, consider a Taylor expansion of this one-parameter GO pricing kernel around $\tilde{R}_{m,T} = 0$:

$$M_T = R^{-1}_{f,t}\left[1 - w_{1,t}\left(\tilde{R}_{m,T}/R_{f,t}\right) + w^2_{1,t}\left(\tilde{R}_{m,T}/R_{f,t}\right)^2 - w^3_{1,t}\left(\tilde{R}_{m,T}/R_{f,t}\right)^3 + ...\right], \quad (19)$$

where the ellipsis summarizes higher-order terms that follow the same sign-flipping pattern as above. The term that is linear in the market return is the same as that in the conditional CAPM. As in the conditional CAPM, the equilibrium risk-return trade-off varies over time. In this model, variation in the weight of the market in the GO portfolio, $w_{1,t}$, determines the risk-return trade-off, including time-variation in the equity premium and the slope of the cross-sectional relation between securities' risk premiums and their market betas.

Interestingly, the same parameter, $w_{1,t}$, affects how higher-order moments are priced. The model's distinct feature is the risk premium's dependence on higher-order moments. The quadratic term mimics the pricing kernel proposed by Harvey and Siddique (2000), who find that there is a significant price of coskewness risk coming from this term. The cubic term gives rise to a cokurtosis premium of the opposite sign, consistent with the findings in Christoffersen et al. (2021). Allowing for nonzero GO pricing kernel weights on market derivatives based on variance, skewness, and kurtosis, as in Section 2, introduces additional

[17] That is, the GO pricing kernel is a monotonic function of the conditional CAPM pricing kernel.

higher-order terms in the pricing kernel. Yet even the simplest version of the GO pricing kernel has the potential to explain many empirical shortcomings of the conditional CAPM.

2 Empirical Methodology

We now return to the minimalist IEP framework in Section 1.1 and apply it to provide daily estimates of the equity and variance premiums at horizons ranging from monthly (30 days) to annual (360 days). There are three key steps in the estimation procedure:

1. estimating risk-neutral moments, including option-implied market variance
2. estimating expected (physical) market variance based on realized variance
3. regressing the variance premium on risk-neutral moments.

The main data are daily physical and risk-neutral moments of the stock market. Our proxy for the market is the Standard and Poor's (S&P) 500 index. As in Martin (2017), we focus on the 1996 to 2021 period in which OptionMetrics data are available for options on S&P 500. To measure realized market variance, we use intraday data on market returns from 1994 to 2021 for the SPDR S&P 500 Exchange-Traded Fund (ETF) Trust, ticker SPY. As a historical equity premium estimate, we use the market minus risk-free (MktRf) factor from Ken French's website.

To measure risk-neutral moments of market returns, we use the prices of actively traded S&P 500 options from OptionMetrics. We use only cash-settled European options with a.m. settlements that do not expire at quarter ends, following Martin (2017) and others. Following the options literature (e.g., Bakshi, Kapadia, and Madan (2003), and Chang et al. (2013)), we employ thorough data cleaning procedures to reduce the impact of illiquidity on estimates of risk-neutral moments. We describe procedures for cleaning data and approximating integrals, including alternative methods for robustness, in the Internet Appendix.

We measure physical moments of market returns using SPY ETF trades in the Trades and Quotes (TAQ) database. We use daily realized variance (RV) based on SPY ETF returns over 78 intraday intervals spaced equally in business (i.e., transaction) time throughout regular trading hours: 9:30am to 4:00pm Eastern time. We employ thorough data cleaning procedures, such as averaging subsample estimates, to reduce the impact of illiquidity on

estimated physical return variance, following the microstructure literature (e.g., Barndorff-Nielsen et al. (2008, 2009), Patton and Sheppard (2015)). We provide details in the Internet Appendix.

For the risk-free rate, we use constant-maturity market yields of Treasury bills of 1, 3, 6, and 12 months from the Federal Reserve Economic Database (FRED) to match equity premium horizons.[18] Option-implied volatility calculations use risk-free rates, dividend yields, and index values from OptionMetrics. We interpolate rates between option maturities and use dividend yields and closing prices for the index, ticker SPX.

2.1 Variance Premium Estimation

We estimate the IEP and GO portfolio weights using equation (11), which requires estimates of the expected variance premium. Here we compute this premium as the difference between estimates of the risk-neutral and physical variances of market excess returns.

2.1.1 Risk-Neutral Moments

We estimate risk-neutral moments of excess market returns, $\tilde{R}_{m,T}$, from the prices of call and put options on the market using the general formula of Bakshi and Madan (2000), as specialized in Carr and Madan (2001). This formula shows that risk-neutral expected excess market returns raised to the jth power are:

$$R_{f,t}^{-1}\mathbb{E}_t^*\tilde{R}_{m,T}^j = \frac{j!}{S_t^j}\left[\int_{F_{t,T}}^{\infty}(K-F_{t,T})^{j-2}\,C\,(K)\,dK + \int_0^{F_{t,T}}(K-F_{t,T})^{j-2}\,P\,(K)\,dK\right], \quad (20)$$

where $j = 2, 3, 4, 5, 6$ for relevant moments, S_t is the market index value, K, the strike price of calls and puts priced at $C(K)$ and $P(K)$, respectively, and $F_{t,T} = R_{f,t}S_t$ is the futures price of the market for maturity T.[19] Denoting the risk-neutral moment of order j for maturity T at time t by $\mathrm{Mj}_{t,T}$—e.g., $\mathrm{M2}_{t,30}$ is risk-neutral market variance at the 30-day horizon at t—we estimate $\mathrm{Mj}_{t,T}$ using the right-hand side of equation (20) multiplied by $R_{f,t}$. We refer the right-hand side of equation (20) as "discounted" risk-neutral moment, denoted by $\tilde{\mathrm{M}}\mathrm{j}_{t,T}$.

[18] We use the average of one and three-month rates as the 60-day rate and the one-month risk-free rate from Ken French's website before July 31, 2001, when FRED data on the one-month yield become available.

[19] This formula for moments of simple returns is analogous to the formula for the moments of log returns in Bakshi, Kapadia, and Madan (2003), except that we do not scale higher-order moments by variance.

2.1.2 High-Frequency Identification of Expected Variance

Our forecasts of market variance at horizons up to yearly are primarily based on a one-parameter model of daily realized variance (rv_t) as a fractionally integrated stochastic process. The key parameter is the order of fractional differencing, $0 \leq d \leq 0.5$. This model captures the well-established long memory property of return variance (e.g., Andersen et al. (2001)) where autocorrelations decay at a hyperbolic rate, much slower than the exponential decay in classic autoregressive (AR) models.

We assume the rv_t process satisfies:

$$(1-L)^d rv_t = \epsilon_t, \tag{21}$$

$$rv_t = (1-L)^{-d}\epsilon_t, \tag{22}$$

where ϵ_t is white noise and L is the lag operator defined by $Lx_t = x_{t-1}$ for any process x_t. Expanding equation (22) shows the hyperbolic decay in impulse response weights:

$$rv_t = \epsilon_t + d\epsilon_{t-1} + \frac{d(d+1)}{2!}\epsilon_{t-2} + \frac{d(d+1)(d+2)}{3!}\epsilon_{t-3} + ..., \tag{23}$$

We recursively estimate d using maximum likelihood (ML), expanding the estimation window annually. We demean rv_t on a recursive basis before estimating d_t.

Applying these real-time d_t estimates to the rv_t series enables estimation of ϵ_t via equation (21) and thus real-time forecasts of rv_{t+h} at any horizon $h \geq 1$ via equation (22) or its equivalent equation (23). The recursive ML estimates of d_t are within 0.40 ± 0.03 in 24 of 26 years. These estimates are precise in all years with standard errors ranging from 0.01 to 0.03, with the latter applying only to the first few sample years.

We use forecasts of rv_{t+h} from this fractionally integrated model at horizons ranging from monthly ($\text{RV}_{30} = \sum_{h=1}^{30} rv_{t+h}$) to annual ($\text{RV}_{360} = \sum_{h=1}^{360} rv_{t+h}$) as the basis for computing the variance premium at these horizons. We adjust these raw fractionally integrated forecasts because they are based on intraday realized variance during the trading day, whereas the expected market variance in the variance premium is based on the variance of long-run market returns. The adjustment accounts for the overnight return period and illiquidity from short-run reversals in intraday, daily, and weekly returns that do not affect long-run variance. Although these effects partially offset, it is unlikely that they exactly cancel.

We address both issues by applying a simple regression-based multiplier, κ_t, to fractionally integrated forecasts (e.g., $\mathbb{E}_t \mathrm{RV}_{30}$) to convert them into predictions of long-run return variance (e.g., $\mathrm{PVar}_{t,30} = \kappa_t \mathbb{E}_t \mathrm{RV}_{30}$). The multiplier, κ_t, is the coefficient from a regression of squared 30-day market returns on 30-day predicted realized variance based on a recursive expanding window. We omit the constant term so that the multiplier reflects the ratio of the average squared market return to the average cumulative realized variance. We use the lagged squared 30-day market return to maximize the correlation with the predicted 30-day realized variance, which also depends on (intraday) squared market returns over the same window as well as additional returns. Consistent with offsetting effects, the average coefficient estimate is 1.02 with a volatility of 0.20.[20]

Table 1, Panel A summarizes these estimates of risk-neutral and predicted market variance (M2 and PVar), and the expected variance risk premium at the monthly and annual horizons. It also reports realized monthly and annual equity premiums for comparison purposes. We annualize all quantities and multiply them by 100 to convert to annual percentages, except for the monthly equity premium which is a monthly percentage.

[Insert Table 1 here]

The bottom rows in Panel A show that the average estimated variance premiums are positive at the monthly and annual horizons at 1.52% and 1.55%, respectively. The middle rows confirm that average risk-neutral monthly (annual) variance of 4.28% (4.30%) exceeds the corresponding predicted market variance of 2.76% (2.75%). The top row shows that the monthly equity premium averaged 0.74% with a volatility of 4.74% and an annualized variance of 2.70% , almost identical to the average predicted monthly variance, PVar_{30}, of 2.76% in row four, providing evidence that the fractionally integrated model is well-calibrated. The variance premium estimates exhibit a very narrow range—the 5th to 95th percentile range is -0.03% to 5.35% (0.35% to 3.79%) for the monthly (annual) variance premium—showing that the model's variance prediction closely tracks market pricing of variance. Table 1, Panel B confirms that the correlations between model-predicted and risk-neutral variance are extremely high at 0.93 monthly and 0.91 annually, even though the fractionally integrated model does not use option prices.

[20]The coefficient estimate averages 0.90 in the pre-2008 period and 1.10 in the post-2008 period, indicating high-frequency return reversals are more pronounced in the first half of the sample.

We also find that unadjusted rv_{t+h} forecasts from the fractionally integrated (FI) model perform well relative to two challenging benchmarks: forecasts from the heterogeneous autoregressive (HAR) model proposed by Corsi (2009) and those from risk-neutral variance (M2). Table 2, Panel A compares the ability of the HAR, M2, and FI models to predict realized variance RV_T, as measured by out-of-sample (OOS) R^2, at horizons ranging from monthly ($T = 30$ days) to annual ($T = 360$ days). The null model for computing OOS R^2 is average in-sample realized variance, which is 0.0260 annualized (16.1% volatility). The M2 model uses coefficients from a recursive linear regression of RV_T to make out-of-sample predictions. The fractionally integrated model beats the HAR model at all horizons, especially those longer beyond 30 days, even though it uses fewer parameters. It also beats the M2 model at all horizons, except monthly, despite one fewer parameter. The last row shows that all three models have negative point estimates of OOS R^2 at the annual horizon, driven by two unanticipated crisis years (2008 and 2020).

[Insert Table 2 here]

Table 2, Panel B shows the in-sample R^2 values resulting from adjusting each model's forecasts using linear transformations that best fit realized variance at each horizon. This panel shows that the HAR model could perform better if its coefficients were better calibrated, e.g., using shrinkage, but it would still underperform the fractionally integrated model at long horizons, which is striking given the fewer degrees of freedom in the fractionally integrated model. The fractionally integrated model exhibits an in-sample R^2 value of 10.2% at the annual horizon, which contrasts with the negative R^2 OOS values in Panel A.

Table 2, Panel C, which reports pairwise correlations between the model forecasts, provides arguably the strongest evidence for the fractionally integrated model. The simplest ex ante forecast of realized variance is a linear transformation of the market's pricing of variance (M2). At all horizons, the fractionally integrated model's prediction of unadjusted realized variance exhibits a correlation $\sim 85\%$ with M2. As noted earlier, the adjusted RV forecasts are even more highly correlated with risk-neutral moments than the unadjusted forecasts, suggesting that the multiplier parameter, κ_t, serves its intended purpose.

Hereafter our main estimate of the expected variance premium for maturity T is:

$$\mathrm{VP}_{t,T} = \mathrm{M2}_{t,T} - \kappa_t \mathbb{E}_t \mathrm{RV}_{t,T}, \tag{24}$$

where $\mathbb{E}_t \mathrm{RV}_{t,T}$ comes from the fractionally integrated model of expected market variance. Figure 1 shows that the predicted variance from the fractionally integrated model closely tracks risk-neutral variance. Panel A (Panel B) illustrates this finding for variance at the 30-day (360-day) horizon. Panel C displays the resulting 30-day and 360-day estimates of the expected variance premium, as computed in equation (24).

[Insert Figure 1 here]

The fractionally integrated model of expected market variance provides a "natural explanation" to the "puzzle" posed by Bollerslev et al. (2009), Bekaert and Hoerova (2014), and Cheng (2019), who find that the variance premium seems to be negative in periods of sudden market turmoil, such as September 2008, which would be inconsistent with investor aversion to variance risk. The fractionally integrated model does not have this puzzling feature and forecasts return variance just as well, if not better, than the models used in prior work, as shown in Table 2. Although the limited and noisy data on crises do not permit clear statistical inferences about which model is most accurate, the fractionally integrated model has strong theoretical appeal and is maximally parsimonious with its single parameter.

2.2 Equity Premium Estimation

We estimate the equity premium using the key equation (10), substituting empirical estimates of the expected variance premium from equation (24) and risk-neutral moments from equation (20). We estimate the unknown GO portfolio weights in this equation using the analogous variance premium equation (11) with these same GO portfolio weights. We estimate up to four GO weights as coefficients in a linear regression of the expected variance premium on up to four excess discounted risk-neutral moments—M3, M4, M5, M6 multiplied by $R_{f,T}^{-1}$.

The implementation hurdles are time-varying GO portfolio weights, heteroskedasticity in variance premiums, multicollinearity in risk-neutral moments, persistence in these measures, and approximation error in the theoretical variance premium equation (11). To estimate the time-varying GO weights, we use recursive regressions with exponentially declining weights on distant data. The half-life of the exponentially declining weights is 1250 days, which matches the five year rolling windows frequently used in the realized variance literature.

Next, to enhance efficiency and robustness to heteroskedasticity, we apply inverse variance (i.e., precision-based) weights to the errors in the least squares objective function, as in weighted least squares. We use $\text{PVar}_{t,30}^{-1}$ as the precision weights.

We address multicollinearity in risk-neutral moments by imposing theory-driven assumptions to reduce dimensionality. As shown in Table 3, Panel A risk-neutral moments of different orders are strongly correlated within a given horizon. These strong correlations pose a challenge to accurate estimates of the GO portfolio weights, as multicollinearity among the regressors inflates standard errors. Table 3, Panel A shows correlations between M2, M3, M4, M5, and M6 at monthly and annual horizons.[21] Risk-neutral moments also are strongly correlated across horizons. Table 3, Panel B reports the correlations between discounted third- and fourth-order risk-neutral moments, M3 and M4, at monthly through annual horizons ($T = 30, 60, 90, 180, 360$ days). The correlations between the same-order moments with monthly to quarterly horizons range from 0.92 to 0.98. The correlations between the same-order moments with semiannual and annual horizons are also very high at 0.86 (0.97) for third-order (fourth-order) moments.

[Insert Table 3 here]

Motivated by theory, we impose a strong but reasonable assumption that the GO portfolio weights on each market-related security are equal across all horizons up to one year:

$$w_{k,t,T_1} = w_{k,t,T_2}, \forall T_1 = 30, ..., 360, T_2 = 30, ..., 360 \tag{25}$$

For example, the GO weights on the stock market are equal at the monthly and annual horizons; and the weights on variance-based derivatives are equal at these horizons. These cross-horizon restrictions facilitate identification of the GO weights because cross-moment correlation is much weaker across different horizons, as shown in Table 3, Panel A. Asset pricing theory suggests GO weights on a given market-related security are likely to be similar across horizons. For log utility investors, who maximize long-term wealth, the GO portfolio

[21] The table includes M2 for reference even though it is not a regressor in the variance premium regressions.

maintains its optimality irrespective of the horizon.[22] In the model from Section 1.2, behavioral investor fear of stocks drives time-series variation in $w_{1,t,T}$ for all maturities T; and investor desire to hedge variance drives variation in $w_{2,t,T}$ for all T.[23]

Relaxing the cross-section restriction could provide insights into how investors treat risks related to various moments across horizons. Short-term investors may focus on minimizing variance, while long-term investors might prioritize skewness and kurtosis. However, precise estimation of GO weights for individual horizons is challenging due to high correlations among risk-neutral moments. One could try to jointly estimate weights using variance and skewness premium equations, the latter derived from a second-order Taylor expansion of the pricing kernel equation (9). Difficulties in reliably estimating long-horizon skewness (Neuberger (2012), and Aretz & Arisoy (2023)) coupled with the observation that variance and skewness premium tend to be highly correlated, make the cross-horizon assumption a practical necessity.

We adopt two methods to address approximation error from the truncation (finite K) of the expected variance premium equation (11) arising from the pricing kernel approximation in equation (9). We allow an intercept in each variance premium regression to account for the average value of omitted higher-order moments. Our intercept varies across different horizons, enabling us to capture fixed effects specific to each horizon. We also evaluate how variance premium estimates improve as K increases, using $K = 1, 2, 3, 4$.

We jointly estimate variance premium regressions in equation (11) at all horizons, monthly through annual, to account for the strong correlation in the error terms. We impose the cross-horizon restrictions in equation (25) using the feasible generalized least squares (FGLS) estimator, using weights inversely proportional to variance and declining exponentially with time. Because it allows for correlation across equations, this FGLS estimator is a weighted seemingly unrelated regression (SUR), as defined in Zellner (1962). We require a minimum of one year of data for these recursive regressions, meaning that the first estimates become available in January 1997. We use the estimated time-varying GO weights from these regressions, along with risk-neutral moments, to estimate the time-varying variance and equity

[22] The pricing kernel from Martin's (2017) lower bound on the equity premium trivially satisfies the cross-horizon restriction in equation (25) as $w_{1,t,T} = 1 \; \forall \; T$ and $w_{k,t,T} = 0 \; \forall \; T$ when $k > 1$.

[23] Although the model applies to a each maturity T, maturities are linked through the market clearing condition when there is common time-series variation in $w_{B,k,t,T}$ across T for any value of k, such as $k = 1$ (stock holdings) or $k = 2$ (variance-based derivative holdings).

premiums at each horizon as in equations (11) and (10), respectively.

3 Risk Premium Estimates

Here we report the model's estimates of the equity and variance premiums and evaluate its performance against benchmarks.

3.1 Variance Premium Fit

We measure the fit of the variance premium regression, R^2, across all horizons for approximation degrees $K = 1, 2, 3, 4$. As a benchmark, we consider the representative log utility model from Martin's (2017) lower bound. In this log utility model, the variance premium is simply the negative of discounted risk-neutral skewness—i.e., equation (11) with $w_{1,t} = 1$ and $w_{k,t} = 0 \ \forall k > 1$. This analysis shows whether variance premiums implied by various models can explain the empirical expected variance premium and enables tests of the validity of the cross-horizon restriction on GO weights, which necessarily impairs model performance at some horizons.

We examine the ability of models to explain the expected variance premium in equation (24), not the realized premium. Using the realized premium would reduce the power of the test, as shown in Table 2, Panel A. The precise estimate of the expected variance premium depends on the model of physical variance but alternative models provide similar estimates, as shown in Table 2, Panel C.

We report percentage R^2 values based on recursive rolling regressions for the expected variance premium as described in Section 2.2. Table 4 shows these R^2 values for monthly to annual horizons (rows) for the five models: log utility and the $K = 1, 2, 3, 4$ models, which include K regressors of risk-neutral moments, orders $3, ..., K + 2$. All four implied risk premium models nest the log utility model. The null model for computing R^2 values is the historical mean of the expected variance premium, estimated using the same recursive procedure with declining exponential weights as in the four models. The recursive historical mean of the quarterly (90-day) variance premium ranges from 1.09% to 2.44%, averaging 1.71% per year in 1997-2021, closely matching the average variance premium of 1.85% generated by our model.

[Insert Table 4 here]

The first column of Table 4 shows that the log utility model has modest predictive power for the empirical variance premium at monthly to quarterly horizons ($R^2 \approx 5\%$ to 8%) but has no predictive ability at the semiannual and annual horizons ($R^2 < 0$). In contrast, the implied variance premium from all four models explains most of the variation in the empirical variance premium, as shown by median R^2 statistics that exceed 50%. The lowest explanatory power of any model at any horizon is 41%. The median R^2 across horizons increases with the approximation degree of the model, K, suggesting that using a high-order approximation could be necessary to capture relevant variation in the GO pricing kernel. The $K = 4$ implied variance premium, which uses risk-neutral moments M3, M4, M5, and M6 to explain the variance premium, has a median R^2 value of 71% compared to the median R^2 of 57% for the $K = 1$ model.

Table 4 provides evidence that the cross-horizon restriction, equation (25), imposed on GO weights reduces the model's ability to explain the variance premium at extreme horizons. The R^2 is lowest at the shortest (30-day) and longest (360-day) horizons, comparing the first and last rows to the middle rows. Even so, the $K = 4$ model explains 64% of monthly and 55% of annual variation in the variance premium, indicating that the cross-horizon restriction is plausible.

3.2 Growth-Optimal Portfolio Weights and Pricing Kernel

Recursive variance premium regressions provide estimates of GO portfolio weights on each day. Figure 2, Panel A shows the GO portfolio weights on the stock market, $w_{1,t}$ from all $K = 1, 2, 3, 4$ models. All models include the first ($k = 1$) term in the pricing kernel. This term represents investor aversion to stock risk, making the equity premium depend on the second-order ($k = 2$) risk-neutral moment (variance) and the variance premium on the third-order ($k = 3$) risk-neutral moment (skewness). In the interpretation in Section 1.2, the GO portfolio weight on the stock market reflects behavioral investors' reluctance to hold stocks, producing risk premiums that induce the GO investor to hold more stocks.

Estimates of the GO weight on the stock market are positive in all $K = 1, 2, 3, 4$ models. They are stable except in the first sample year, 1997, due to limited historical data. The market's GO weight in the $K = 1$ model is easy to interpret because the GO investor holds

no other market-related positions in this model. The market's GO weight in this model always exceeds one, suggesting that the typical investor is more risk averse than the log utility benchmark as discussed in Section 1.2.

[Insert Figure 2 here]

We interpret the market's GO weight in the $K > 1$ models along with the GO investor's positions in other market-related securities, such as variance-based derivatives correlated with stocks. Figure 2, Panel B shows the GO weights in the four market-related securities from the $K = 4$ model. This figure omits the first sample year, when estimates are noisy.

The GO weights on the stock market and the $k = 4$ security (swap on market kurtosis) are always positive, whereas the weights on the variance- and skewness-based market securities are consistently negative. Most weights fluctuate widely in the first few years of the sample, stabilizing shortly after 2008, showing the impacts of estimation noise and uncertainty. The four weights exhibit strong time-series correlations partly because of correlations among the underlying risk-neutral moments, which exhibit alternating signs as shown in Table 3, Panel A. As a result, weights with an odd degree $k = 1, 3$ are negatively (positively) correlated with weights of an even degree $k = 2, 4$ if these weights are of the same (opposite) signs. For example, the stock market ($k = 1$) weight is positively correlated with the variance derivative ($k = 2$) weight because these weights have opposite signs—i.e., stock weight is positive and the variance weight is negative.

We calculate standard errors of GO weights using a moving block bootstrap with 10,000 iterations and a block length of 252 days. Figure 2 shows GO weights with their 95% confidence bounds.[24] Using a heteroskedasticity-robust Wald test, we reject the null that the GO weights are as predicted by representative log utility ($w_1 = 1, w_2 = w_3 = w_4 = 0$). We also reject the null that market weights are unity for the full sample and the post-crisis subsample (p-value $< 10^{-4}$), but not in the pre-crisis period (p-value $= 0.078$).[25]

Figures 3 and 4 show how GO portfolio weights and standard errors change with shorter half-lives of exponentially declining weights in GO regressions. While shorter half-lives align

[24] The moving block bootstrap (Künsh (1989)) resamples overlapping blocks of consecutive observations to preserve any time-series dependence in the data. We split periods of length n into $n - b + 1$ overlapping blocks of length b, draw n/b blocks with replacement, and concatenate them.

[25] Our standard error estimates are increasing in block length. Methods for optimizing the block length (Politis and White (2004)) could improve these estimates.

portfolios more closely with recent market conditions, they become increasingly sensitive to short-term shocks. Standard errors for all asset categories rise as the half-life shortens. The effects are pronounced for weights tied to high-order moments, where estimation challenges are compounded by the leverage inherent in these positions. While modest reductions in half-life (e.g., from 1250 to 1000 or 750 days) do not significantly compromise weight stability, aggressive reductions produce large estimation errors and extreme portfolio rebalancing.

[Insert Figure 3 here]
[Insert Figure 4 here]

By placing a positive weight on stocks, the GO investor benefits from the unconditional equity premium. With a negative weight on variance, the GO investor takes advantage of the variance premium. The GO weight on the stock market is persistently higher after the 2008 crisis in the $K \geq 3$ models. We estimate the time-varying GO weight on stocks solely from rolling regressions of the variance premium on risk-neutral skewness, controlling for other moments. The mechanical reason for the increasing market weight is the rising partial correlation between the variance premium and risk-neutral skewness. Controlling for higher-order moments is important as the relationship between higher-order risk-neutral moments and the variance premium in the early sample absorbs much of the variation in the variance premium.

To ensure that the main properties of the GO weights are robust to different weighting approaches, we test two alternatives: one based on a Kalman filter and the other inspired by dynamic conditional beta (Engle, 2016). The weights generated by these methods share key features with our original findings, such as the signs and magnitudes of GO market and variance weights and the steady rise in market weights after the 2008 crisis. As a result, these methods produce equity premiums that are comparable to those obtained using FGLS. Modest differences arise, however, from methodological limitations and computational challenges.[26] We present these results in the Internet Appendix.

To understand the economic implications of these GO weights, we analyze estimates of the GO pricing kernel on select dates. Figure 5, Panel A shows GO pricing kernels based on different approximation degrees, $K = 1, 2, 3, 4$, using GO weight estimates on the day

[26]The Kalman filter is slow to react to shocks, resulting in lower market weights and a lower equity premium from 2008 to 2011, even though it closely matches our main results in other periods.

when the risk-neutral variance is at its median level. The excess market returns plotted on the x-axis span the 5th-to-95th percentile range of annual returns in the sample. The values of the pricing kernel range from 0.5 to 3.7, with positive values being consistent with no arbitrage. All pricing kernels are monotonically decreasing over this range, consistent with investor risk aversion. The $K = 3$ pricing kernel takes on the most extreme values, from 0.53 when the excess market return is 34.3% to 3.71 when the excess return is -26.2%, whereas the $K = 2$ kernel is the flattest.[27]

[Insert Figure 5 here]

Figure 5, Panel B shows pricing kernels on a date with expected monthly returns exactly at the 95th percentile: November 12, 2020 in the aftermath of the Covid-19 Pandemic. As expected, all pricing kernels are steeper on this date, especially those that include cubic and quartic terms. The $K = 3$ ($K = 4$) pricing kernel peaks at 4.19 (2.93) and has a low of 0.52 (0.58) in this time of turmoil. Outside the typical range of market returns, the pricing kernel approximation is not necessarily accurate, and non-monotonic kernels and negative values can arise in such ranges.[28] Fortunately, the closed-form expressions for implied risk premiums do not require kernel estimates in such extreme scenarios.[29]

3.3 Equity Premium Forecasts

Table 5, Panel A summarizes equity premium estimates from the four IEP models and the log utility model. The average annual (360-day) IEP estimates range from 6.84% to 9.47% in the IEP models; and the monthly (30-day) IEP averages range from 0.57% to 0.84% per month. These estimates are significantly higher than the log utility estimates of 4.22% per year and 0.35% per month. The 95th percentile of the IEP is 15.1% in the $K = 4$ model as compared to just 7.6% in the log utility model, reflecting the higher skewness in the IEP model. The main reason for the higher average, volatility, and skewness of the IEP relative to the log utility model is that the estimated GO weight on the stock market consistently

[27]The $K = 3$ kernel also has the highest curvature as measured by the ratio of slopes at extreme market returns, changing from a slope of -40.8 in bad times to a slope of just -0.57 in good times.

[28]The estimated pricing kernels are positive and monotonic mainly because the positive GO stock market weight dominates other weights in the range of realized returns.

[29]For applications that require pricing kernel values in rare disaster scenarios, linear extrapolation of the GO kernel from the realized range of returns could be sufficient and would simplify practical risk assessment.

exceeds one. Because risk-neutral variance is positive, volatile, and positively skewed, a higher GO weight on the market increases the average, volatility, and skewness of the IEP.

[Insert Table 5 here]

The average 360-day Sharpe ratios range from 0.32 to 0.44 in the IEP models; and the (annualized) 30-day Sharpe ratios range from 0.34 to 0.48. These estimates are nearly twice the average log utility estimate of 0.22. As seen in Figure 8 the implied Sharpe ratio is well-behaved, reaching a high of 1.0 at the annual horizon during the 2008 financial crisis.

[Insert Figure 8 here]

We analyze the forecasting performance of the IEP with the caveat that stock returns are highly unpredictable and the sample contains just 24 non-overlapping yearly periods. Table 5, Panel B shows out-of-sample R^2 values for the IEP and the log utility models at horizons up to one year. The null model for R^2 computation is a constant expected return equal to the sample average.[30] The log utility model has a positive R^2, replicating the finding in Martin (2017). The first- and second-order IEP models (IEP1 and IEP2) have lower R^2 values than the log utility model (RLUEP).

However, the equity premium predictions from the third- and especially fourth-order models (IEP3 and IEP4) generally outperform the log utility predictions (RLUEP). The IEP model predictions improve as the approximation degree of the GO pricing kernel increases. At all horizons, the IEP4 R^2 values are high as judged against the low benchmarks in the return predictability literature (Goyal and Welch (2008, 2021)). The R^2 is 1.23% monthly, 3.17% quarterly, and 8.09% annually.

In economic terms, these IEP4 R^2 values translate into large improvements in the GO investor's expected return. We quantify the value of forecasting from the perspective of a GO (log utility) investor holding the market and risk-free asset using the one-period framework of Campbell and Thompson (2008). The ability to forecast the equity premium increases this investor's expected excess portfolio return by $\frac{R^2}{1-R^2}(1+S^2)$, where S^2 is the squared unconditional Sharpe ratio on the market and R^2 is the model's ability to predict market returns. Using the S^2 and R^2 values from Tables 1 and 5, the investor increases annualized

[30] Using a recursively estimated historical mean would result in slightly worse performance for the null model because returns are modestly negatively autocorrelated in the sample.

returns by 15% (10.6%) with the monthly (annual) IEP4 forecasts. We interpret these large estimates of investment gains with caution given the short sample.

We now assess the calibration and incremental explanatory power of IEP forecasts using regressions to predict stock market returns. In theory, the IEP would predict the realized equity premium with a coefficient of unity, and no other variable, such as the variance premium, would forecast returns. Table 7 tests this hypothesis by regressing excess market returns on IEP4, the Martin (2017) lower bound (RLUEP), and the variance premium. The table shows results for horizons ranging from monthly to annual (30 to 360 days) in Panels A to E, where the independent and dependent variables in each regression have matching horizons. There are $h-1$ overlapping days in each daily predictability regression with a horizon of h days. Standard errors of regression coefficients are based on the truncated Hansen and Hodrick (1980) kernel with a bandwidth of $h-1$ days.

[Insert Table 7 here]

At all horizons, equity premium predictions from the main model (IEP4) compare well to those in Martin (2017) and those based on the variance premium. The univariate coefficients on IEP4 are statistically indistinguishable from unity[31] with point estimates ranging from 0.69 to 1.08 at monthly to annual horizons in the full-sample regressions in columns (1) to (3). Interestingly, when controlling for IEP4, all of the point estimates on risk-neutral variance, $RLUEP$, and the variance premium, VP, are negative, which contrasts with the (unreported) uniformly positive univariate coefficients on these predictors. Standard errors are high because of the short and volatile sample period and the strong correlations among the predictors: IEP4, RLUEP, and VP.

Since our model assumes no arbitrage, the IEP should forecast market returns more accurately when this assumption holds. Columns (4) to (6) in Table 7 show predictability regressions based on a sample that excludes 61 trading days in 2008, from September 19 to December 15, when arbitrage between stock and option markets is limited. We refer to the period excluding these 61 days as the no-arbitrage sample. The United States Securities

[31] We cannot reject $H_0 : \beta = 1$ with two-sided p-values ranging from 0.58 to 0.86 across various horizons. We also cannot reject $H_0 : \alpha = 0$ (p-values ranging from 0.54 to 0.78). These results are robust to including controls for $RLUEP$ and VP.

and Exchange Commission began placing restrictions on short sales on September 17, 2008, which "dramatically increased bid-ask spreads for options" (Battalio and Schultz (2011)).[32]

Stock return predictability from IEP4 roughly triples when excluding the 1% of days (61 of 6,293) with severely limited arbitrage—i.e., R^2_{adj} in column (4) is three times its value in column (1)—in the specifications with monthly to quarterly horizons. The statistical significance of the coefficients on IEP4 is much stronger, too. The incremental R^2_{adj} from the other predictors remains low or even negative in the no-arbitrage period. These findings suggest that the IEP model performs better when arbitrage costs are low.

Figure 6 shows that the one-year (360-day) IEP from $K = 4$ is strongly correlated (89%) with Martin's (2017) lower bound on the equity premium and exhibits equally strong correlation with Chabi-Yo and Loudis' (2020) conditional bounds (90%). It shows that IEP4 is greater than Martin's lower bound from 1997 to 2021 on 99% of days for the annual horizon and always greater than the lower bound for other horizons. The model deviation is, however, substantial in the post-2008 period, averaging 5.0% per year with 2.5% volatility. While IEP4 generally lies between Chabi-Yo and Loudis' conditional bounds—averaging 5.9% (annual) and 6.1% (monthly) for the lower bound and 10% (annual) and 10.6% (monthly) for the upper bound—its position relative to these bounds shifts over time. In the pre-2008 period, IEP4 is quite close to the conditional lower bound, deviating by only 0.5%. However, after 2008, IEP4 moves closer to the conditional upper bound, with the deviation from the lower bound increasing to over 3% and the deviation from the upper bound declining to an average of 1%. IEP4 is also 1.9% lower than the expected excess return from Bakshi et al. (2019) before the crisis, but it exceeds their post-crisis estimate. These two measures are highly correlated: 82% (73%) at the annual (monthly) horizon.

[Insert Figure 6 here]

Many properties of IEP4 arise from the properties of the growth-optimal weight on stocks, as shown in Figure 2. The GO weight on stocks almost always exceeds 100%, which produces a high required equity premium in equilibrium. This high GO weight on stocks is consistent

[32]We define the 61-day exclusion based on the first doubling of the weekly average spread, September 19, and the first halving of the monthly average spread, December 15, after the initial short sale ban, September 17, 2008. For each horizon, h, we measure the spread on a synthetic stock market index, created from S&P 500 options and risk-free Treasury bills maturing on day h, as a percentage of the index value. We compute daily average synthetic spreads across all strike prices weighting options by dollar volume.

with the idea that most other investors, who hold a mix of stocks and bonds, are more risk averse than log utility. The post-2008 increase in the GO weight on stocks drives the higher post-crisis equity premium, which is consistent with evidence on realized returns.[33] One interpretation is that GO stock holdings increase in response to a buying opportunity driven by behavioral investors' fear of holding stocks after the crisis. Indeed survey evidence in Greenwood and Shleifer (2014) shows that individual investors' beliefs about stock returns were pessimistic after 2008.

Analogously, the consistently negative weight on skewness in the GO portfolio could suggest that some behavioral investors exhibit a preference for negative skewness. This preference might stem from biases such as loss aversion or strategic concerns; notably, hedge funds may favor strategies with negative skewness due to reputation concerns, as argued in Malliaris and Yan (2021), and their assets under management have dramatically increased since 2008. Although the model does not explain why some investors are willing to speculate on stock market crashes, it can rule out rational unconstrained maximization of log utility.

Figure 7 shows the IEP for models with approximation degrees $K = 1, 2, 3, 4$. Panel A (Panel B) shows the one-year (monthly) IEPs. Reassuringly, the correlations between the IEP estimates from different models are very high, exceeding 93% for all pairwise correlations between the $K = 2, 3, 4$ models at the monthly and annual horizons. There is also evidence for a common factor in expected returns across horizons. The cross-horizon correlations range from 0.85 to 0.88 between IEPs at the monthly and annual horizons for all models.

[Insert Figure 7 here]

Both panels indicate that the IEP rose and fell abruptly during the economic crises of 2008 and 2020. Recall that the IEP is a linear combination of risk-neutral moments weighted by the GO portfolio allocations. These temporary spikes in the IEP arise mainly from changes in the risk-neutral variance, as the GO weight on the market did not change dramatically during these crises as shown in Figure 2.

Figure 9 decomposes the IEP into contributions from each market exposure in the GO portfolio, as represented by each term in the weighted sum of risk-neutral moments in equation (10). The primary contributor to the equity premium is the $k = 1$ term representing

[33]The pre-crisis IEP4 averages 6.2% with a Sharpe ratio of 0.30, whereas the post-crisis IEP4 is 9.4% with a Sharpe ratio 0.42.

the GO investor's required compensation for bearing stock market risk, which depends on risk-neutral variance (M2). This term averages 8% per year with 5.4% volatility, accounting for the entire mean of IEP and more than 100% of IEP variance. It peaks at roughly 40% per year during the 2008 and 2020 crises. The $k = 3$ term representing the GO investor's exposure to crash risk partially offsets the $k = 1$ term during crises, as the GO investor reduces exposure to market risk by hedging crash risk, $\tilde{R}_m^3$, as shown in Figure 2, Panel B.

[Insert Figure 9 here]

Figure 10 shows that the term structure of the equity premium is time-varying: flat or mildly upward sloping under normal conditions, but sharply downward sloping during crises. This aligns with the predictions of equilibrium asset pricing models that incorporate mean-reverting expected returns—see van Binsbergen and Koijen (2017). Figure 11 shows an upward-sloping term structure of the implied variance premium, aligning with existing literature (Aït-Sahalia, Karaman, and Mancini (2020) and Dew-Becker et al. (2017)).[34] Combining these findings, unconditional Sharpe ratios decline modestly with an investor's horizon, as shown in Panel C of Table 5. The level and slope of the IEP term structure are strongly negatively correlated with the slope having a correlation of -60% with IEP4_{360} and -94% with IEP4_{30}.[35]

[Insert Figure 10 here]
[Insert Figure 11 here]

4 The IEP in a Macroeconomic Model

Here we test how the IEP methodology performs in a simulated environment by applying it to an equilibrium pricing model. This exercise allows us to compare IEP estimates to the true equity premium in the model. We identify this value by simulating the state variables of the model and computing implied moments.

[34] We define the variance premium as the difference between risk-neutral and physical measures, whereas some studies use the opposite convention.

[35] The level is IEP4_{360} in equation (9) and the slope is the difference between the annualized one-year and one-month IEPs divided by the difference in these maturities.

We focus on the extended "bad environment-good environment" (BEGE) model of Bekaert, Engstrom, and Ermolov (BEE, 2023). This model offers several advantages in testing the IEP methodology. First, it is designed to fit the variance premium, the dependent variable in our regressions for estimating the pricing kernel. Second, it has a quasi-closed form solution, facilitating computation of high-order risk-neutral moments, the key independent variables in these regressions. Third, shocks in the model are gamma distributed with a time-varying shape parameter, and non-normal shocks are necessary to generate high-order risk-neutral moments in discrete time (BEE). Lastly, this model includes two state variables in excess returns, and more than one state variable is necessary to avoid a singular covariance matrix when estimating IEP regressions. We describe key components of the model and calculation of the moments in the Internet Appendix. For consistency with the IEP model, which is based on simple returns, we differ slightly from BEE by calculating conditional moments of simple returns $R_{t+1} = e^{r_{t+1}}$ instead of log returns.

4.1 Simulated IEP in a BEGE Model

We conduct 10,000 simulations of the BEGE model at a monthly frequency with $T = 324$ months in each iteration. We use the extended pure preference shock model calibration in BEE (their Table IV, third column). We obtain the true expected return, physical variance forecast, and the $k = 2, 3, \ldots, 6$ risk-neutral moments from the simulated state variables, as described in the Internet Appendix. Table 8, Panel A displays the levels and standard deviations of key moments in the model, Panel B displays the correlations across moments, and Panel C displays the correlations between risk-neutral moments.

[Insert Table 8 here]

The calibrated model has several realistic features, including the average values of the equity and variance risk premiums and stock return variance, and the rich patterns in correlations among the risk-neutral moments. However, implied correlations of the realized equity premium with other price variables are roughly ten times lower than in the data, which will reduce stock return predictability relative to empirical estimates.

Using the implied risk-neutral moments and variance premium, we estimate the IEP regression as in equation (11). This exercise does not exactly replicate our IEP methodology in practice for several reasons. First, we obtain risk-neutral moments directly in the model,

as opposed to using option prices, which would require numerical simulation. Second, for tractability reasons, we use only the single monthly horizon. Third, the model is simulated at the monthly, not daily, frequency. Lastly, in our baseline specifications, we use the entire sample for our parameter estimates. Table 9 shows the full-sample estimated coefficients for the IEP with $K = 1, \ldots, 4$ models.

[Insert Table 9 here]
[Insert Table 10 here]

We compare models' out-of-sample R^2 values to the truth, which is the maximum R^2, and report these values in Table 10. We consider models with and without a constant, as well as one with time-varying weights, which we estimate using weighted least squares with a five-year half life. We do this separately for each simulation, requiring a minimum of 12 observations (one simulated year) for initial parameter estimates.

In general, there is relatively low predictability within the model. However, the true equity premium calculated from rational expectations provides an upper bound on the OOS R^2 of 13 basis points. Relative to this, the IEP for $K = 3, 4$ with no constants performs fairly well, generating OOS R^2 values of just over 10 basis points. Notably, the $K = 1$ model with time-varying weights generates a similarly high R^2 of just over 11 basis points, suggesting a trade-off between flexibility and approximation degree. We interpret these results with caution given the low upper bound on predictability. But this exercise shows the potential of the IEP methodology in approximating the equity premium, as compared to the RLU model and in absolute terms.

5 Conclusion

The economic insight from the IEP model is that the equilibrium equity premium depends on the amount of stock market risk borne by a GO investor, who is exposed to the market through stocks and market-related derivatives. One can interpret this GO investor's exposures as market positions that other investors are unwilling to hold. Variation in these residual positions drives variation in the market risk premiums, including the equity, variance, and skewness risk premiums.

The model's econometric innovation is using high-frequency data to identify the stock market's physical return variance, which facilitates recovery of market risk premiums from option prices. Although future research could improve the precision of moment and parameter estimates, this paper's implementation of the IEP model provides useful new estimates of market risk premiums and the empirical pricing kernel. These estimates can help test whether market-related risks are priced in other securities, including individual stocks and bonds. Lastly, since the modeling approach here applies equally well to the set of individual securities with traded options, one can use the IEP model to estimate expected returns of individual stocks with options, even without the preference assumptions in Martin and Wagner (2019) or Kadan and Tang (2020). We are pursuing this idea in ongoing research.

References

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43–76.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Aretz, Kevin and Y. Eser Arisoy, 2023, The pricing of skewness over different return horizons, *Journal of Banking & Finance* 148, 1–19.

Aït-Sahalia, Yacine, Mustafa Karaman, and Loriano Mancini, 2020, The term structure of equity and variance risk premia, *Journal of Econometrics* 219, 204–230.

Back, Kerry, Kevin Crotty, and Seyed Mohammad Kazempour, 2022, Validity, tightness, and forecasting power of risk premium bounds, *Journal of Financial Economics*, 144, 732–760.

Baker, Malcolm, Jeremy C. Stein, and Jeffrey Wurgler, 2003, When does the market matter? Stock prices and the investment of equity-dependent firms, *Quarterly Journal of Economics* 118, 969–1005.

Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies* 16, 527—566.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics skew laws and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bakshi, Gurdip, and Dilip Madan, 2000, Spanning and derivative-security evaluation, *Journal of Financial Economics* 55, 205–238.

Bakshi, Gurdip et al., 2019, A New Formula for the Expected Excess Return of the Market. *Fox School of Business Research Paper*.

Barndorff-Nielsen, Ole E., Peter Reinhard Hansen, Asger Lunde, and Neil Shephard, 2008, Designing realized kernels to measure the ex post variation of equity prices in the presence of noise, *Econometrica* 76, 1481–1536.

Barndorff-Nielsen, Ole E., Peter Reinhard Hansen, Asger Lunde, and Neil Shephard, 2009, Realized kernels to practice: Trades and quotes, *Econometrics Journal* 12, 1–32.

Battalio, Robert, and Paul Schultz, 2011, Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets, *Journal of Finance* 66, 2013–2053.

Bekaert, Geert, and Marie Hoerova, 2014, The VIX, the variance premium, and stock market volatility, *Journal of Econometrics* 183, 181–192.

Bekaert, Geert, Eric Engstrom, and Andrey Ermolov, 2023, The variance risk premium in equilibrium models, *Review of Finance* 27, 1977-2014.

Beason, Tyler, and David Schreindorfer, 2022, Dissecting the equity premium, *Journal of Political Economy* 130, 2203–2222.

van Binsbergen, Jules H., and Ralph S.J. Koijen, 2017, The term structure of returns: Facts and theory. *Journal of Financial Economics* 124, 1–21

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, Tim, and Viktor Todorov, 2011, Tails, fears, and risk premia, *Journal of Finance* 66, 2165–2211.

Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail risk premia and return predictability, *Journal of Financial Economics* 118, 113—134.

Campbell, John Y., and Albert S. Kyle, 1993, Smart money, noise trading and stock price behaviour, *Review of Economic Studies* 60, 1–34.

Campbell, John Y. and Samuel B. Thompson, 2008, Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21, 1509–1531.

Carr, Peter, and Dilip Madan, 2001, Optimal positioning in derivative securities, *Quantitative Finance* 1, 19–37.

Carr, Peter, and Liuren Wu, 2009, Variance risk premiums, *Review of Financial Studies* 22, 1311–1341.

Chabi-Yo, Fousseni, and Johnathan Loudis, 2020, The conditional expected market return *Journal of Financial Economics* 137, 752–786.

Chang, Bo-Young, Peter Christoffersen, and Kris Jacobs, 2013, Market skewness risk and the cross section of stock returns, *Journal of Financial Economics* 107, 46–68.

Cheng, Ing-Haw, 2019, The VIX premium, *Review of Financial Studies* 32, 180–227.

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2013, Capturing option anomalies with a variance-dependent pricing kernel, *Review of Financial Studies* 1962–2006.

Christoffersen, Peter, Mathieu Fournier, Kris Jacobs, and Mehdi Karoui, 2021, Option-based estimation of the price of coskewness and cokurtosis risk, *Journal of Financial and Quantitative Analysis* 56, 65–91.

Cieslak, Anna, and Annette Vissing-Jorgensen, 2021, The economics of the Fed put, *Review of Financial Studies* 34, 4045—4089.

Cochrane, John, 2011, Discount rates, *Journal of Finance* 66, 1047–1108.

Corsi, Fulvio, 2009, A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics* 7, 174–196.

Coval, Joshua D., and Tyler Shumway, 2001, Expected option returns, *Journal of Finance* 56, 983–1009.

Dew-Becker, Ian, Stefano Giglio, Ahn Le, and Marius Rodriguez, 2017, The price of variance risk, *Journal of Financial Economics* 123, 225-–250.

Di Maggio, Marco, Amir Kermani, and Kaveh Majlesi, 2020, Stock market returns and consumption, *Journal of Finance* 75, 2851–3372.

Engle, Robert F., 2016, Dynamic conditional beta, *Journal of Financial Econometrics* 14, 643–667.

Gallup Organization, 2022, Washington, D.C., website retrieved March 22, 2023, available at https://news.gallup.com/poll/266807/percentage-americans-owns-stock.aspx

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Goyal, Amit, Ivo Welch, and Athanasse Zafirov, 2021, A comprehensive look at the empirical performance of equity premium prediction II, Swiss Finance Research Institute working paper

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

Hansen, Lars P., and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829–853.

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263–1295.

Kadan, Ohad, and Xiaoxiao Tang, 2020, A bound on expected stock returns, *Review of Financial Studies* 33, 1565—1617

Kozhan, Roman, Anthony Neuberger, and Paul Schneider, 2013, The skew risk premium in the equity index market, *Review of Financial Studies* 26, 2174—2203.

Kraus, Alan, and Robert H. Litzenberger, 1978, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085–1100.

Künsch, Hans R., 1989, The jackknife and the bootstrap for general stationary observations, *The Annals of Statistics*, 17, 1217–1241.

Lintner, John, 1965, Security prices, risk and maximal gains from diversification, *Journal of Finance* 20, 587—615.

Long, John B., 1990, The numeraire portfolio, *Journal of Financial Economics* 26, 29–69.

Malliaris, Steven, and Hongjun Yan, 2021, Reputation concerns and slow-moving capital, *Review of Asset Pricing Studies* 11, 580–609.

Markowitz, Harry M., 1976, Investment for the long run: New evidence for an old rule, *Journal of Finance* 31, 1273–1286.

Martin, Ian, 2017, What is the expected return on the market? *Quarterly Journal of Economics* 132, 367—433.

Martin, Ian, and Christian Wagner, 2019, What is the expected return on a stock? *Journal of Finance* 74, 1887—1929.

Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of risk and return in risk arbitrage, *Journal of Finance* 56, 2135–2175.

Neuberger, Anthony, 2012, Realized skewness, *Review of Financial Studies* 25, 3423–3455.

Patton, Andrew J., and Kevin Sheppard, 2015, Good volatility, bad volatility: Signed jumps and the persistence of volatility, *Review of Economics and Statistics* 97, 683–697.

Politis, Dimitris N., & Hal White, 2004, Automatic block-length selection for the dependent bootstrap, *Econometric Reviews* 23, 53—70

Ross, Stephen A., 1976, Options and efficiency, *Quarterly Journal of Economics* 90, 75–89.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Schneider, Paul, 2019, Anatomy of the market return, *Journal of Financial Economics* 132, 325–350.

Schneider, Paul, and Fabio Trojani, 2019, (Almost) model-free recovery, *Journal of Finance* 74, 323–370.

Shiller, Robert J., 1984, Stock prices and social dynamics *Brookings Papers on Economic Activity* 2, 457–498.

Zellner, Arnold, 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association* 57, 348–368.

Table 1: Summary Statistics

Panel A summarizes estimates of risk-neutral and predicted market variance ($M2$ and $PVar$), variance risk premiums ($VP = M2 - PVar$), and equity premiums (EP) at the monthly and annual horizons (30 and 360 days). We annualize all quantities and multiply them by 100 to convert to annual percentages, except for the monthly EP which is in monthly percentage terms. The statistics in columns are the mean, standard deviation (StdDev), percentiles (P5, P25, P50, P75, P95), and number of days in the 1996 to 2021 sample.

	Mean	StdDev	P5	P25	P50	P75	P95	Days
EP_{30}	0.74	4.74	-7.20	-1.58	1.26	3.48	7.16	6545
EP_{360}	8.06	17.75	-25.37	1.14	10.60	18.57	33.61	6545
$M2_{30}$	4.28	4.37	1.17	1.87	3.14	5.03	10.97	6545
$PVar_{30}$	2.76	2.53	1.03	1.45	2.07	3.33	5.80	6545
$M2_{360}$	4.30	2.23	2.05	2.68	3.84	5.33	7.87	6545
$PVar_{360}$	2.75	1.38	1.43	1.86	2.38	3.23	4.84	6545
VP_{30}	1.52	2.23	-0.03	0.31	0.85	1.83	5.35	6545
VP_{360}	1.55	1.13	0.35	0.76	1.22	2.07	3.79	6545

Panel B reports correlations among estimates of risk-neutral and predicted market variance ($M2$ and $PVar$), and variance risk premiums ($VP = M2 - PVar$) at the monthly and annual horizons (30 and 360 days).

	EP_{30}	EP_{360}	$M2_{30}$	$PVar_{30}$	$M2_{360}$	$PVar_{360}$	VP_{30}	VP_{360}
EP_{30}	1.00	0.33	0.11	0.10	0.10	0.11	0.10	0.06
EP_{360}	0.33	1.00	0.22	0.27	0.18	0.28	0.13	-0.00
$M2_{30}$	0.11	0.22	1.00	0.93	0.85	0.82	0.91	0.68
$PVar_{30}$	0.10	0.27	0.93	1.00	0.85	0.91	0.68	0.57
$M2_{360}$	0.10	0.18	0.85	0.85	1.00	0.91	0.69	0.86
$PVar_{360}$	0.11	0.28	0.82	0.91	0.91	1.00	0.56	0.57
VP_{30}	0.10	0.13	0.91	0.68	0.69	0.56	1.00	0.68
VP_{360}	0.06	-0.00	0.68	0.57	0.86	0.57	0.68	1.00

Table 2: Predicting Realized Variance

This table compares realized variance forecasts from the heterogeneous autoregressive (HAR), risk-neutral variance (M2), and fractionally integrated (FI) models. The left columns of Panel A show out-of-sample (OOS) R^2 for the three models at horizons ranging from monthly ($T = 30$ days) to annual ($T = 360$ days). The benchmark for OOS R^2 is average in-sample realized variance, which is 0.0260 annualized (16.1% volatility). The next three columns show R^2 values resulting from adjusting each model's forecasts using linear transformations that best fit realized variance at each horizon. The last column shows the number of non-overlapping observations (NNObs) at each horizon in the 1996 through 2021 sample.

	OOS R^2			Adjusted OOS R^2			
Horizon	HAR	M2	FI	HAR	M2	FI	NNObs
30	0.380	0.399	0.397	0.422	0.427	0.409	304
60	0.212	0.224	0.275	0.303	0.271	0.294	152
90	0.113	0.151	0.200	0.232	0.205	0.233	101
180	0.005	0.039	0.078	0.134	0.129	0.153	51
360	-0.086	-0.112	-0.070	0.039	0.080	0.086	25

Panel B reports pairwise correlations between the models' forecasts at each horizon.

Horizon	HAR-M2	HAR-FI	FI-M2	NNObs
30	0.751	0.820	0.847	304
60	0.677	0.747	0.841	152
90	0.636	0.693	0.850	101
180	0.541	0.589	0.852	51
360	0.410	0.485	0.844	25

Table 3: Correlations among Risk-Neutral Moments

Panel A shows correlations between risk-neutral moments of different orders—M2 (variance), M3 (skewness), M4 (kurtosis), M5, and M6—at monthly (30-day) and annual (360-day) horizons.

	$M2_{30}$	$M3_{30}$	$M4_{30}$	$M5_{30}$	$M6_{30}$	$M2_{360}$	$M3_{360}$	$M4_{360}$	$M5_{360}$	$M6_{360}$
$M2_{30}$	1.00	-0.93	0.92	-0.85	0.81	0.85	-0.44	0.83	-0.46	0.76
$M3_{30}$	-0.93	1.00	-0.92	0.93	-0.86	-0.78	0.53	-0.78	0.55	-0.71
$M4_{30}$	0.92	-0.92	1.00	-0.95	0.97	0.70	-0.38	0.78	-0.44	0.78
$M5_{30}$	-0.85	0.93	-0.95	1.00	-0.96	-0.65	0.43	-0.72	0.48	-0.71
$M6_{30}$	0.81	-0.86	0.97	-0.96	1.00	0.60	-0.34	0.70	-0.42	0.73
$M2_{360}$	0.85	-0.78	0.70	-0.65	0.60	1.00	-0.67	0.95	-0.68	0.84
$M3_{360}$	-0.44	0.53	-0.38	0.43	-0.34	-0.67	1.00	-0.65	0.96	-0.55
$M4_{360}$	0.83	-0.78	0.78	-0.72	0.70	0.95	-0.65	1.00	-0.70	0.96
$M5_{360}$	-0.46	0.55	-0.44	0.48	-0.42	-0.68	0.96	-0.70	1.00	-0.64
$M6_{360}$	0.76	-0.71	0.78	-0.71	0.73	0.84	-0.55	0.96	-0.64	1.00

Panel B reports the correlations between third- and fourth-order risk-neutral moments, M3 (skewness) and M4 (kurtosis), at monthly through annual horizons (30, 60, 90, 180, 360 days).

	$M3_{30}$	$M3_{60}$	$M3_{90}$	$M3_{180}$	$M3_{360}$	$M4_{30}$	$M4_{60}$	$M4_{90}$	$M4_{180}$	$M4_{360}$
$M3_{30}$	1.00	0.96	0.92	0.81	0.53	-0.92	-0.91	-0.89	-0.86	-0.78
$M3_{60}$	0.96	1.00	0.98	0.89	0.63	-0.87	-0.92	-0.92	-0.91	-0.85
$M3_{90}$	0.92	0.98	1.00	0.94	0.71	-0.82	-0.88	-0.91	-0.92	-0.87
$M3_{180}$	0.81	0.89	0.94	1.00	0.86	-0.67	-0.75	-0.79	-0.85	-0.82
$M3_{360}$	0.53	0.63	0.71	0.86	1.00	-0.38	-0.44	-0.50	-0.60	-0.65
$M4_{30}$	-0.92	-0.87	-0.82	-0.67	-0.38	1.00	0.97	0.94	0.87	0.78
$M4_{60}$	-0.91	-0.92	-0.88	-0.75	-0.44	0.97	1.00	0.98	0.94	0.86
$M4_{90}$	-0.89	-0.92	-0.91	-0.79	-0.50	0.94	0.98	1.00	0.98	0.91
$M4_{180}$	-0.86	-0.91	-0.92	-0.85	-0.60	0.87	0.94	0.98	1.00	0.97
$M4_{360}$	-0.78	-0.85	-0.87	-0.82	-0.65	0.78	0.86	0.91	0.97	1.00

Table 4: Predicting the Expected Variance Premium

This table compares the fit of five models in predicting the expected variance premiums in equation (24). The measure of fit is model R^2 based on recursive rolling regressions. The table shows percentage R^2 values for monthly to annual horizons (rows) for the five models (columns): representative log utility (RLUVP) and the $K = 1, 2, 3, 4$ implied variance premium (IVP) models, which include K regressors consisting of risk-neutral moments of orders $3, ..., K+2$. The null model for computing R^2 values is the historical mean of the variance premium, using the same recency weights as in the IVP models. The last column reports the number of days in the model evaluation period spanning 1997 to 2021.

	RLUVP	IVP1	IVP2	IVP3	IVP4	Days
30	7.8	48.7	46.1	63.3	63.5	6293
60	8.8	58.2	55.9	69.9	71.4	6293
90	5.4	58.6	57.3	69.4	71.8	6293
180	-4.6	57.2	61.2	63.9	71.0	6293
360	-27.7	41.2	64.1	44.5	55.4	6293

Table 5: Summary of Equity Premium Predictions

Panel A summarizes equity premium estimates from four implied equity premium (IEP) models and from the representative log utility model (RLUEP) at monthly (30-day) and annual (360-day) horizons. The IEP estimate that uses K terms to approximate the pricing kernel at horizon h days is $IEPK_h$. The RLUEP estimate at horizon h days is $RLUEP_h$. Equity premiums are not annualized. The columns show the mean (Mean), standard deviation (StdDev), and percentiles (P5, P25, P50, P75, P95) of equity premiums for all 6,293 days from 1997 to 2021.

	Mean	StdDev	P5	P25	P50	P75	P95	Days
RLUEP_{30}	0.35	0.36	0.09	0.15	0.26	0.41	0.91	6293
IEP1_{30}	0.70	0.85	0.18	0.29	0.45	0.81	1.82	6293
IEP2_{30}	0.57	0.73	0.16	0.23	0.37	0.65	1.46	6293
IEP3_{30}	0.84	1.04	0.23	0.37	0.56	0.92	2.14	6293
IEP4_{30}	0.69	0.85	0.17	0.30	0.46	0.77	1.81	6293
RLUEP_{360}	4.22	2.16	1.95	2.69	3.83	5.12	7.63	6293
IEP1_{360}	8.22	5.02	4.01	4.98	6.79	10.18	15.24	6293
IEP2_{360}	6.84	4.29	3.41	4.14	5.41	8.40	12.99	6293
IEP3_{360}	9.47	5.47	4.48	6.23	7.88	10.87	18.51	6293
IEP4_{360}	7.87	4.35	3.40	5.21	6.75	9.39	15.10	6293

Panel B reports out-of-sample R^2 values of equity premium predictions from the IEP and RLUEP models at horizons ranging from monthly (30 days) to annual (360 days). The last column shows the number of non-overlapping observations (NNObs) at each horizon in the 1997 through 2021 sample.

	RLUEP	IEP1	IEP2	IEP3	IEP4	NNObs
30	0.61	-0.88	-0.18	0.13	1.23	304
60	1.09	-1.18	0.01	0.72	2.36	152
90	1.15	-2.05	-0.46	1.01	3.17	101
180	2.58	0.08	2.05	6.62	8.10	51
360	1.91	-1.39	1.23	7.70	8.09	25

Panel C shows the average annualized Sharpe ratios implied by the four implied equity premium (IEP) models and representative log utility (RLU) model from 1997 to 2021 for all horizons.

	30	60	90	180	360
RLU	0.222	0.225	0.227	0.226	0.223
IEP1	0.413	0.411	0.410	0.404	0.398
IEP2	0.340	0.340	0.339	0.334	0.324
IEP3	0.483	0.479	0.476	0.462	0.437
IEP4	0.396	0.396	0.395	0.386	0.365

Table 7: Market Return Predictability Regressions

This table shows regressions of the realized market equity premium, EP_h, on the implied equity premium, $IEP4_h$, the representative log utility equity premium, $RLUEP_h$, and the variance premium, VP_h, where h is the horizon in days and the regressors are lagged by h days for horizons $h = 30, 60, 90, 180, 360$ days. Hansen-Hodrick (1980) standard errors based on a bandwidth of $h - 1$ days appear in parentheses.

Panel A contains results for the full sample of days from 1997 through 2021, whereas Panel B shows results for the "No-Arbitrage" sample that excludes 61 days from September 19, 2008 to December 15, 2008.

Panel A: Full Sample (6293 Observations)

	30-Day Horizon			60-Day Horizon			90-Day Horizon			180-Day Horizon			360-Day Horizon		
	EP_{30}	EP_{30}	EP_{30}	EP_{60}	EP_{60}	EP_{60}	EP_{90}	EP_{90}	EP_{90}	EP_{180}	EP_{180}	EP_{180}	EP_{360}	EP_{360}	EP_{360}
Const	0.002	0.003	0.002	0.003	0.004	0.003	0.004	0.005	0.004	-0.007	-0.003	-0.005	-0.014	-0.011	-0.008
	(0.004)	(0.004)	(0.004)	(0.009)	(0.008)	(0.008)	(0.013)	(0.012)	(0.013)	(0.021)	(0.017)	(0.02)	(0.051)	(0.042)	(0.043)
IEP4	0.694	1.036	0.709	0.736	1.003	0.861	0.762	1.144	0.902	1.083**	1.502	1.656***	1.112***	1.306	2.112***
	(0.55)	(1.163)	(0.843)	(0.657)	(1.179)	(0.998)	(0.688)	(1.237)	(0.944)	(0.455)	(1.566)	(0.508)	(0.393)	(1.814)	(0.737)
RLUEP		-0.845			-0.658			-0.937			-0.997			-0.436	
		(2.727)			(2.124)			(2.255)			(3.079)			(3.425)	
VP			-0.007			-0.113			-0.191			-1.554			-5.402
			(0.231)			(0.455)			(0.633)			(1.316)			(3.299)
Adj. R^2	0.015	0.015	0.015	0.027	0.027	0.027	0.035	0.035	0.035	0.080	0.080	0.094	0.075	0.076	0.137

Panel B: No-Arbitrage Sample (6232 Observations)

	30-Day Horizon			60-Day Horizon			90-Day Horizon			180-Day Horizon			360-Day Horizon		
	EP_{30}	EP_{30}	EP_{30}	EP_{60}	EP_{60}	EP_{60}	EP_{90}	EP_{90}	EP_{90}	EP_{180}	EP_{180}	EP_{180}	EP_{360}	EP_{360}	EP_{360}
Const	-0.001	-0.002	-0.001	-0.004	-0.004	-0.004	-0.01	-0.009	-0.01	-0.021	-0.019	-0.02	-0.019	-0.015	-0.013
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)	(0.01)	(0.01)	(0.01)	(0.024)	(0.022)	(0.022)	(0.054)	(0.045)	(0.045)
IEP4	1.393***	1.153	1.44	1.371***	1.417	1.633**	1.525***	1.692	1.933***	1.49***	1.69	2.222***	1.183***	1.346	2.248***
	(0.507)	(1.307)	(0.998)	(0.448)	(1.209)	(0.737)	(0.429)	(1.188)	(0.592)	(0.412)	(1.522)	(0.527)	(0.446)	(1.818)	(0.79)
RLUEP		0.605			-0.114			-0.417			-0.487			-0.383	
		(2.585)			(2.4)			(2.412)			(3.14)			(3.536)	
VP			-0.02			-0.226			-0.527			-1.912			-5.662
			(0.254)			(0.377)			(0.536)			(1.473)			(3.442)
Adj. R^2	0.040	0.040	0.040	0.065	0.065	0.067	0.103	0.103	0.106	0.118	0.118	0.136	0.072	0.072	0.138

Table 8: Moments in the BEGE Model (BEE, 2023)

Panel A reports the levels of monthly variables in the "bad environment-good environment" (BEGE) model of Bekaert, Engstrom, and Ermolov (BEE, 2023) averaged across 10,000 simulations of the model for 324 months.

	Mean	StdDev	P5	P25	P50	P75	P95
EP_{30}	0.41	3.81	-4.94	-1.81	0.05	2.38	6.97
$M2_{30}$	2.39	0.42	1.90	2.08	2.30	2.59	3.18
$PVar_{30}$	1.75	0.33	1.36	1.51	1.68	1.92	2.39
VP_{30}	0.64	0.13	0.54	0.57	0.61	0.67	0.79

Panel B reports the monthly correlations of simulated monthly variables in the BEGE model calculated across 10,000 simulations of the model for 324 months.

	EP_{30}	$M2_{30}$	$PVar_{30}$	VP_{30}
EP_{30}	1.00	0.03	0.04	0.02
$M2_{30}$	0.03	1.00	0.97	0.79
$PVar_{30}$	0.04	0.97	1.00	0.61
VP_{30}	0.02	0.79	0.61	1.00

Panel C reports the monthly correlations of the monthly risk-neutral moments calculated across 10,000 simulations of the BEE model for 324 months.

	$M2_{30}$	$M3_{30}$	$M4_{30}$	$M5_{30}$	$M6_{30}$
$M2_{30}$	1.00	-0.54	0.44	-0.31	0.30
$M3_{30}$	-0.54	1.00	-0.99	0.97	-0.96
$M4_{30}$	0.44	-0.99	1.00	-0.99	0.99
$M5_{30}$	-0.31	0.97	-0.99	1.00	-1.00
$M6_{30}$	0.30	-0.96	0.99	-1.00	1.00

Table 9: IEP Coefficients in the BEGE Model

The table shows IEP coefficients estimated from pooled regressions across all simulations of the BEGE model of Bekaert, Engstrom, and Ermolov (BEE, 2023). Hansen-Hodrick (1980) standard errors based on a bandwidth of 12 periods appear in parentheses.

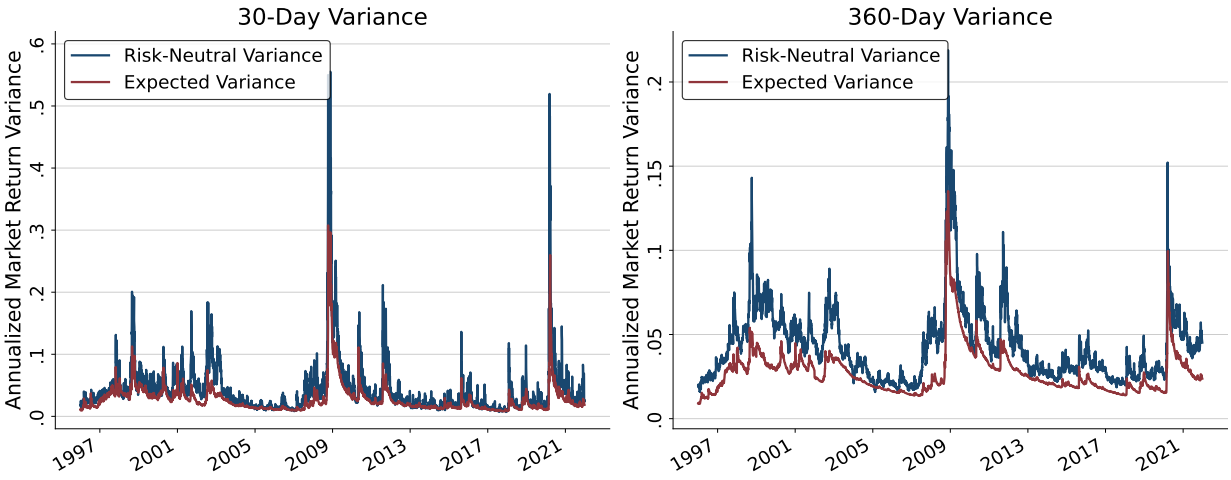
	Constant				**No Constant**			
	K=1	K=2	K=3	K=4	K=1	K=2	K=3	K=4
Const	0.000	0.000	0.000	0.000				
	(0.000)	(0.000)	(0.000)	(0.000)				
1	1.981	5.957	3.451	3.468	2.194	3.413	2.182	2.470
	(0.003)	(0.004)	(0.010)	(0.010)	(0.001)	(0.030)	(0.001)	(0.011)
2		7.239	-4.825	-4.561		1.861	-10.827	-10.327
		(0.007)	(0.051)	(0.068)		(0.045)	(0.011)	(0.025)
3			-11.930	-9.261			-17.784	-28.554
			(0.051)	(0.375)			(0.019)	(0.432)
4				3.336				-15.576
				(0.441)				(0.632)
Adj. R^2	0.897	0.998	1.000	1.000	0.995	0.996	1.000	1.000
Observations	3240000	3240000	3240000	3240000	3240000	3240000	3240000	3240000

Table 10: Restricted R^2 for Equity Premium Estimates in BEGE Model

The table shows out-of-sample R^2 values for alternative IEP estimates in simulations of the BEGE model of Bekaert, Engstrom, and Ermolov (2023). The R^2 values come from monthly (one-period) comparisons of the IEP to the unconditional sample mean equity premium across all simulations. The second column shows the R^2 as a fraction of the R^2 from the model's true equity premium, which is an upper bound.

Estimate	R^2 (%)	As Fraction of Truth
Truth	0.129	1.000
IEP4, No Const, TV Weights	0.045	0.351
IEP3, No Const, TV Weights	0.050	0.388
IEP2, No Const, TV Weights	-2.463	-19.130
IEP1, No Const, TV Weights	0.112	0.867
IEP4, No Const	0.107	0.828
IEP3, No Const	0.105	0.816
IEP2, No Const	-0.229	-1.778
IEP1, No Const	0.095	0.740
IEP4, Const	-0.174	-1.350
IEP3, Const	-0.151	-1.174
IEP2, Const	-2.400	-18.645
IEP1, Const	0.093	0.722
RLU	-0.248	-1.923

Figure 1: The Estimated Variance Premium



Figures show predicted (physical) variance from the fractionally integrated model and risk-neutral variance from option prices at the monthly (30-day) and annual (360-day) horizons.

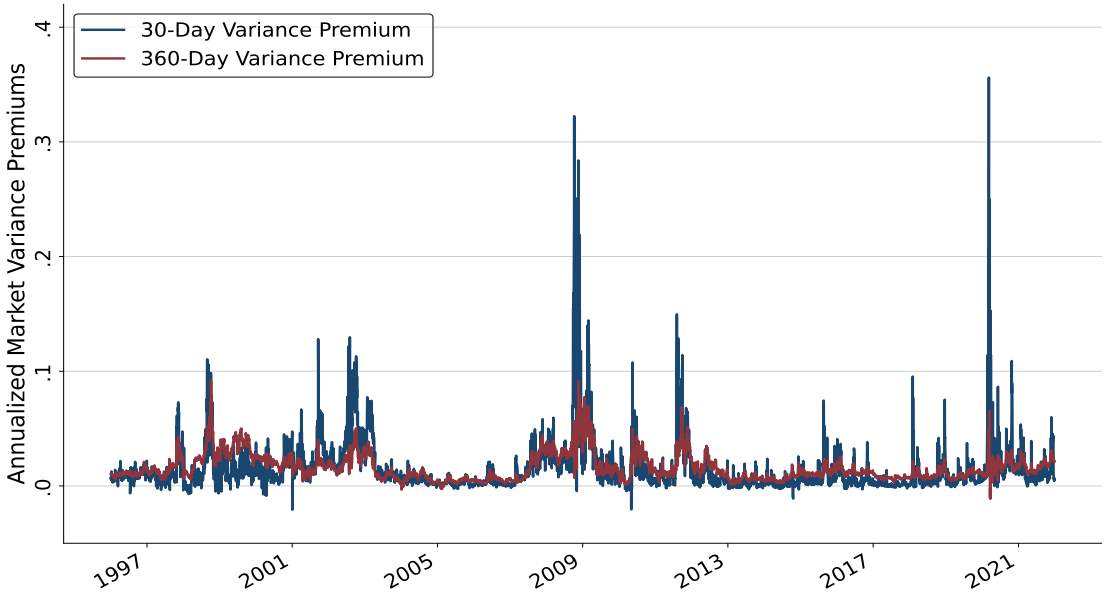
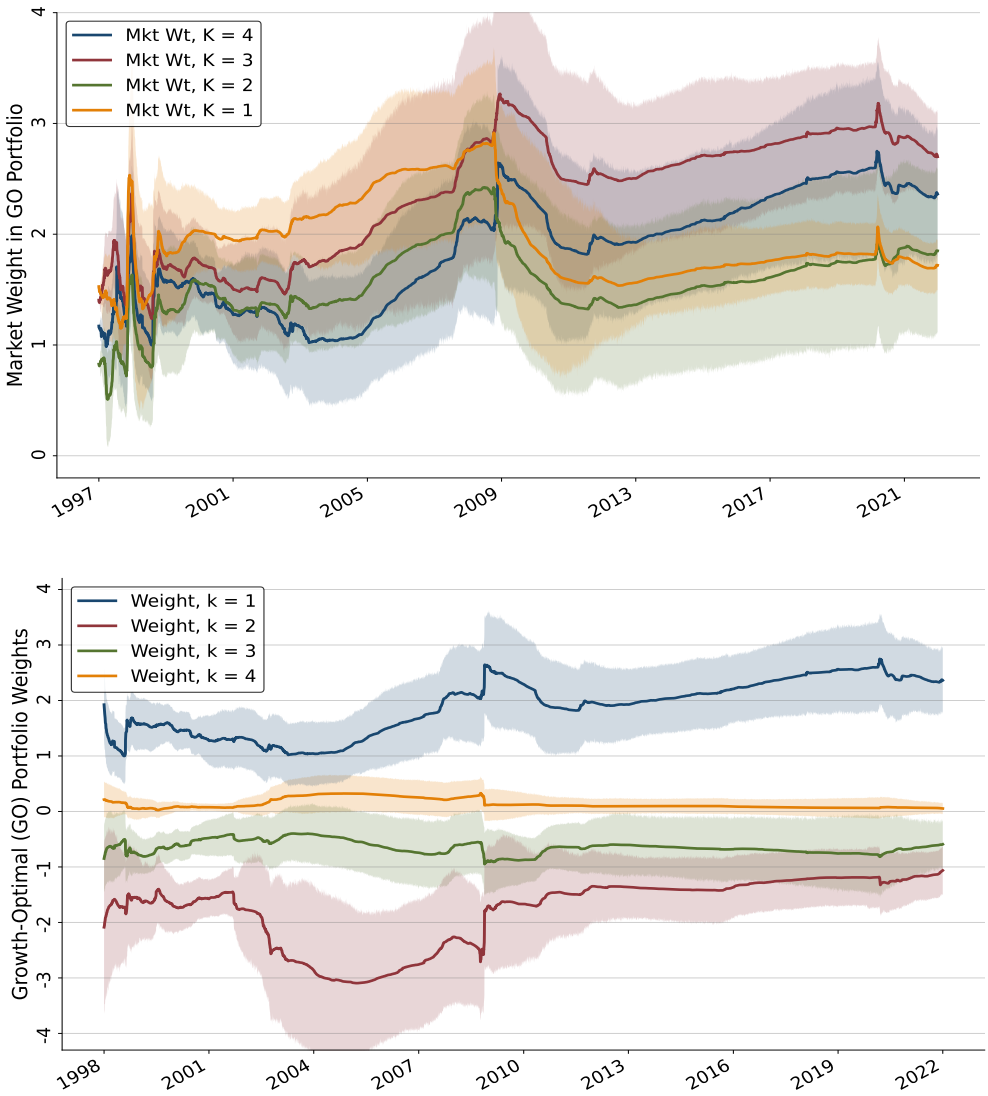


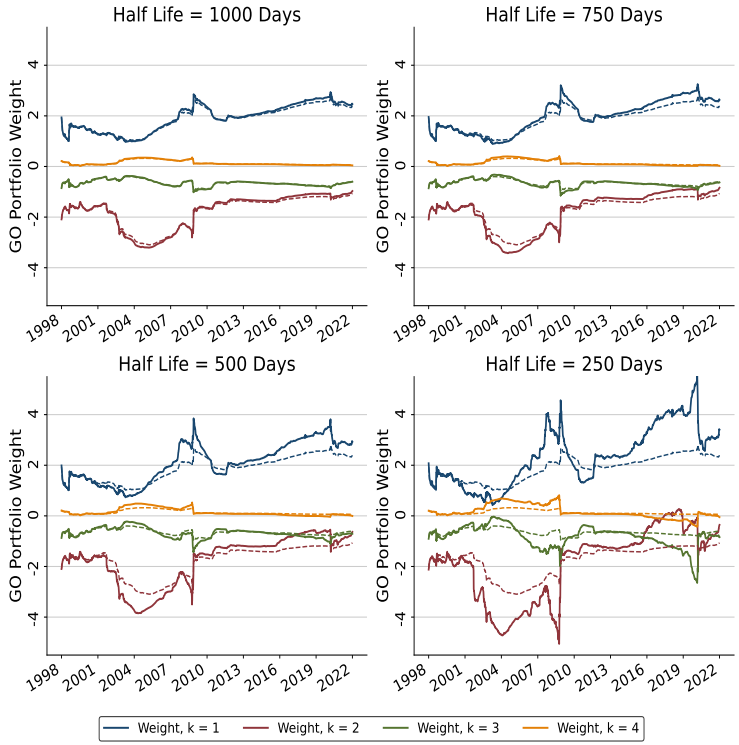
Figure shows annualized variance premiums at the monthly (30-day) and annual (360-day) horizons. Each variance premium is the difference between risk-neutral variance and predicted variance from the fractionally integrated model at each horizon.

Figure 2: Growth-Optimal Portfolio Weights



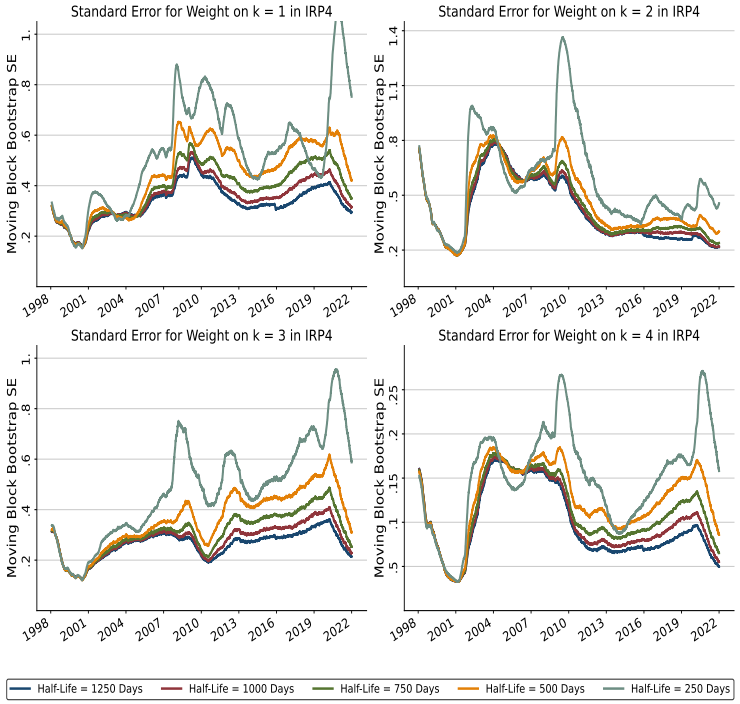
Figures show estimated growth-optimal (GO) portfolio weights in models of implied risk premiums. The top panel shows the weight on the stock market in the GO portfolio for models that include the risk-free asset, the market, and $K-1$ market-related securities for $K = 1, 2, 3, 4$. The bottom panel shows weights on all ($K = 4$) risky assets in the growth-optimal portfolio based on the risk-free asset, the market ($k = 1$), and three market-related securities ($k = 2, 3, 4$). Weight estimates come from recursive regressions of the variance premium on K risk-neutral moments, as in equation (11). Shaded areas represent 95% confidence intervals based on standard errors from a moving block bootstrap with a block length of 252 days.

Figure 3: Growth-Optimal Portfolio Weights Alternative Recency Weights



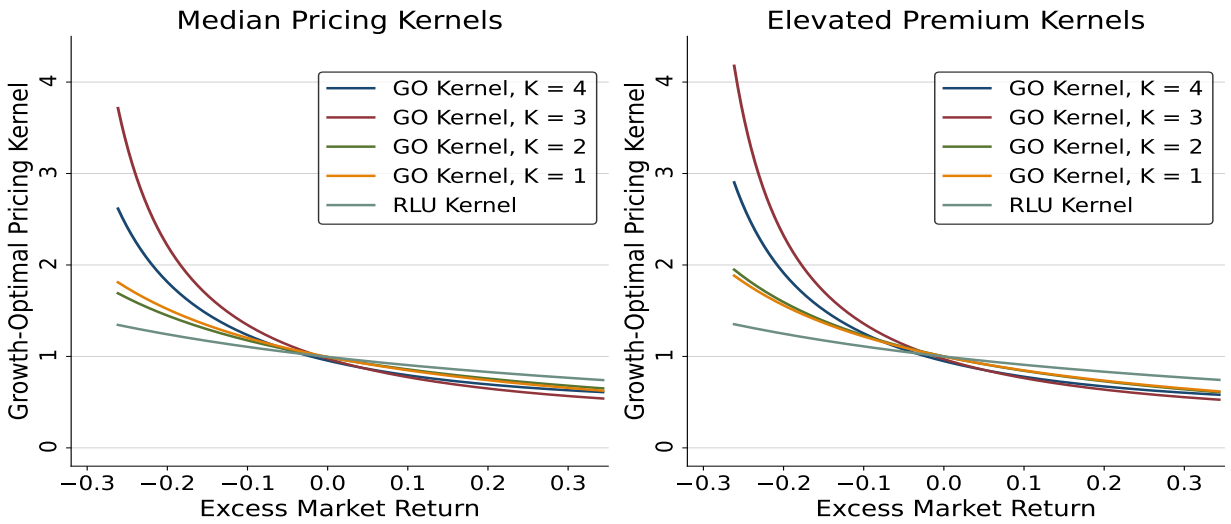
Figures show the weights on all (K = 4) risky assets in the growth-optimal portfolio based on the IEP4 model using exponentially declining weights with half lives ranging from 250 to 1000 days. Dotted lines represent weights obtained using a half life of 1250 days as in figure 2.

Figure 4: Standard Errors of GO Weights Alternative Recency Weights



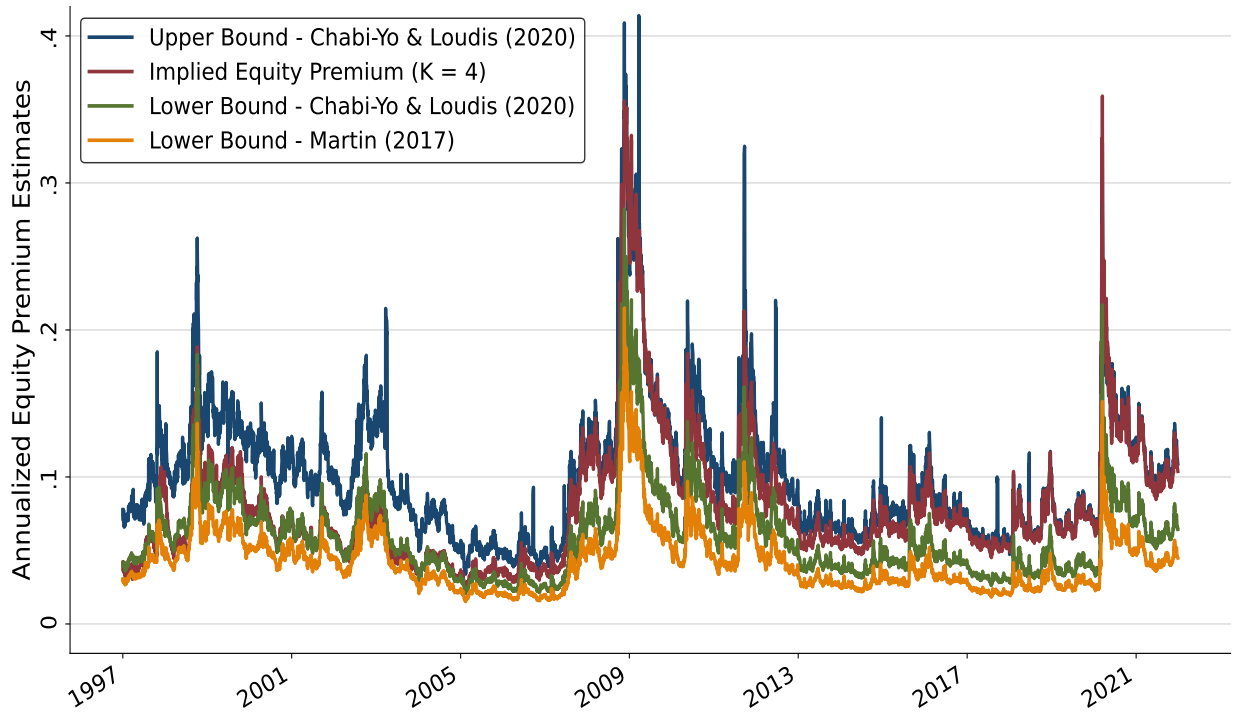
Figures show standard errors of weights on all (K = 4) risky assets in the growth-optimal portfolio based on the implied risk premium model using exponentially declining weights with half lives of 250 to 1000 days. Standard errors are from a moving block bootstrap with a block length of 252 days.

Figure 5: Growth-Optimal Pricing Kernel



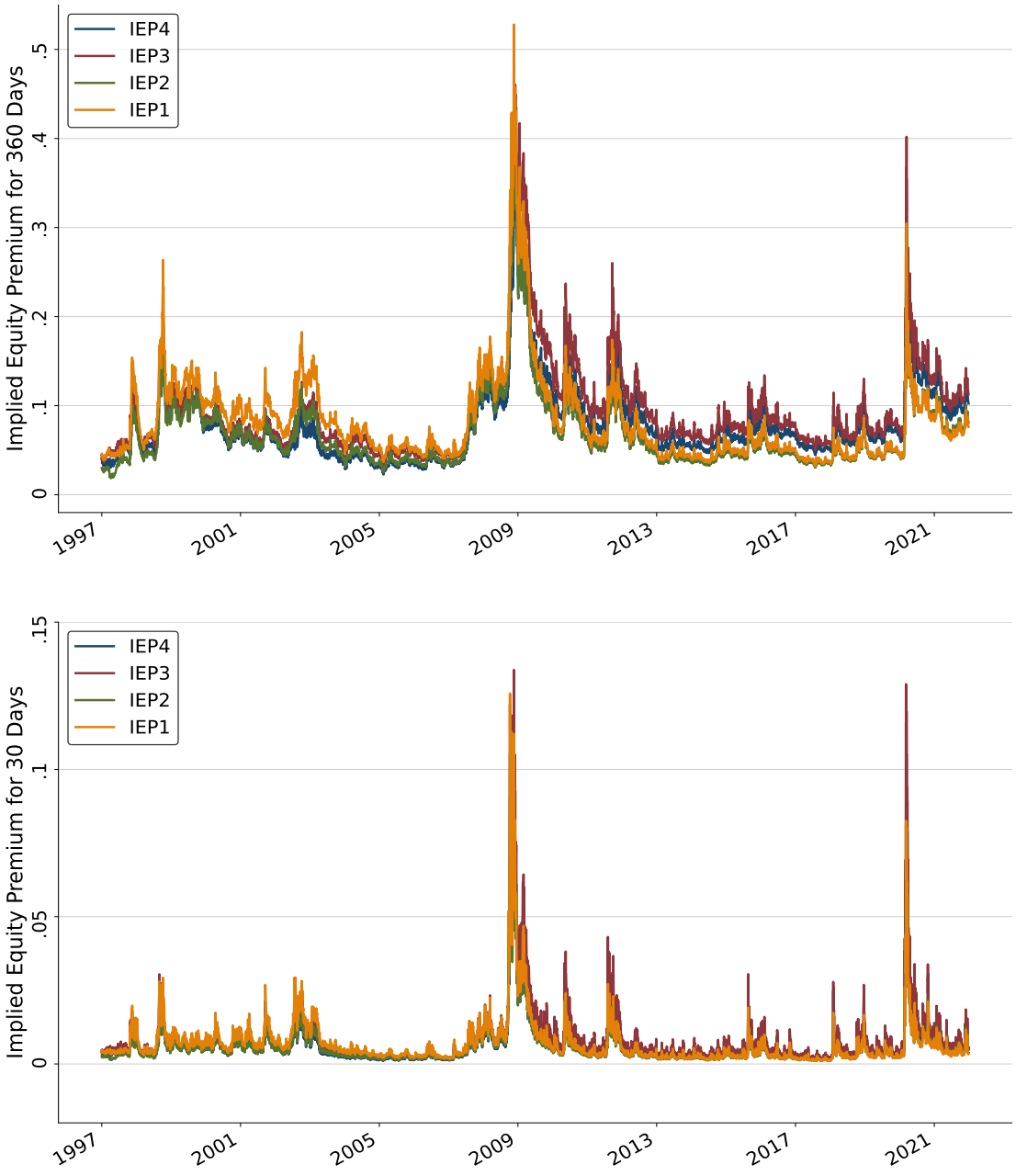
Figures show growth-optimal (GO) pricing kernels that include weights on the risk-free asset, stock market, and $K-1$ market-related securities. The market-related securities have excess returns equal to the market excess return raised to the k^{th} power, where $k = 2, ..., K$. The left panel shows the GO pricing kernel for implied risk premium models with $K = 1, 2, 3, 4$ risky assets on June 13, 2016, which is the day on which the risk-neutral variance is at its median level. The right panel plots the GO pricing kernel on November 12, 2020, when the expected equity premium was at the 95th percentile for the sample.

Figure 6: The Implied Equity Premium versus Alternatives



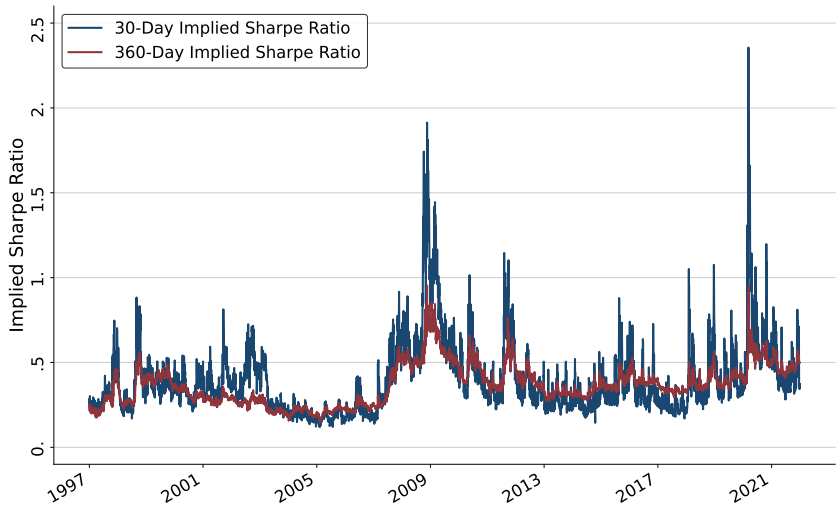
This figure compares the one-year (360-day) implied equity premium (IEP) to the lower bound from Martin (2017) and conditional bounds from Chabi-Yo and Loudis (2020) from 1997 to 2021. The IEP (IEP4_{360}) is based on the fourth-order ($K = 4$) approximation of the growth-optimal pricing kernel in equation (9). The lower bound is the representative log utility equity premium, Lower Bound_{360}, which is a special case of equation (9) with $w_{1,t} = 1$ and $w_{k,t} = 0 \; \forall \; k > 1$. Chabi-Yo and Loudis bounds are based on a replication using our moments. The data are daily from 1997 through 2021.

Figure 7: Implied Equity Premiums



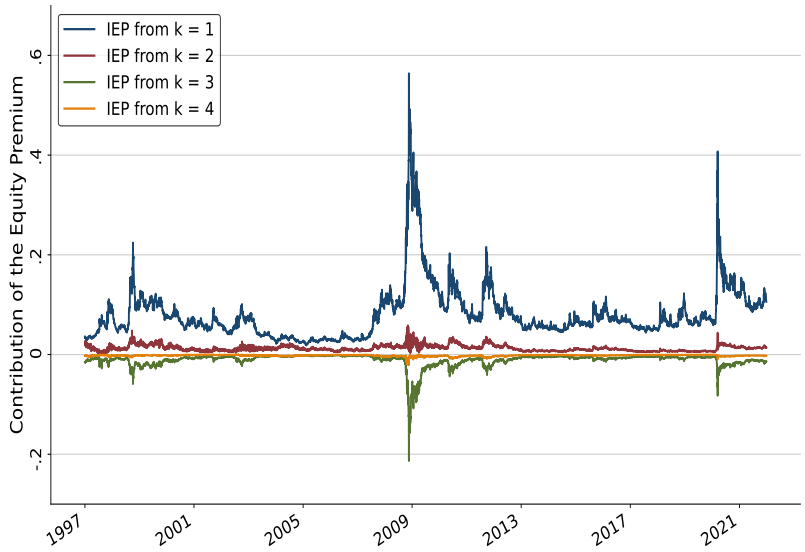
Figures show one-year (360-day) and one-month (30-day) implied equity premiums (IEP1, ..., IEP4) for models that approximate the growth-optimal pricing kernel with degrees $K = 1, 2, 3, 4$. The data are daily from 1997 through 2021.

Figure 8: Implied Sharpe Ratio



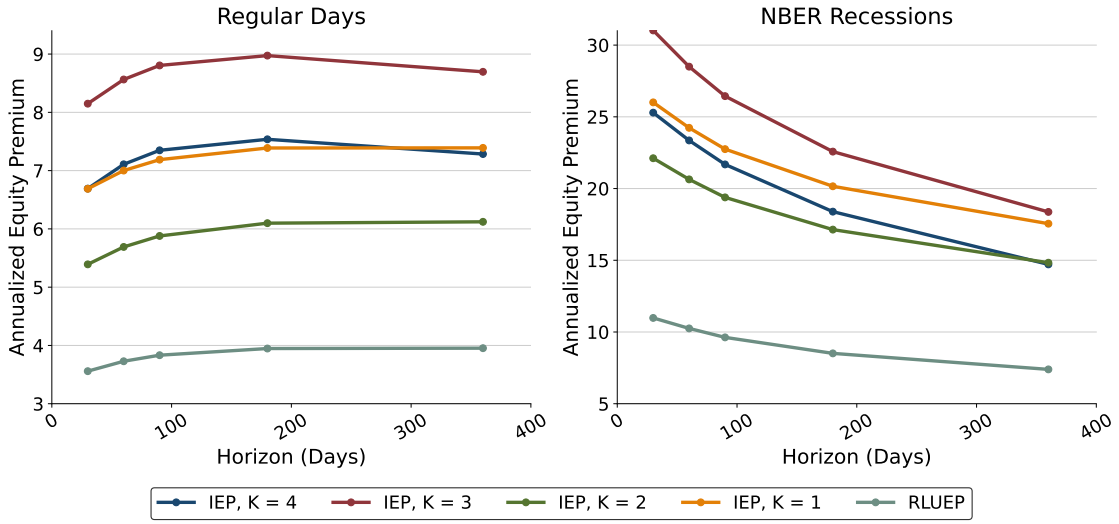
This figure shows 30-day and 360-day implied Sharpe ratios for the IEP4 model. The data are daily from 1997 through 2021.

Figure 9: Decomposing the Implied Equity Premium



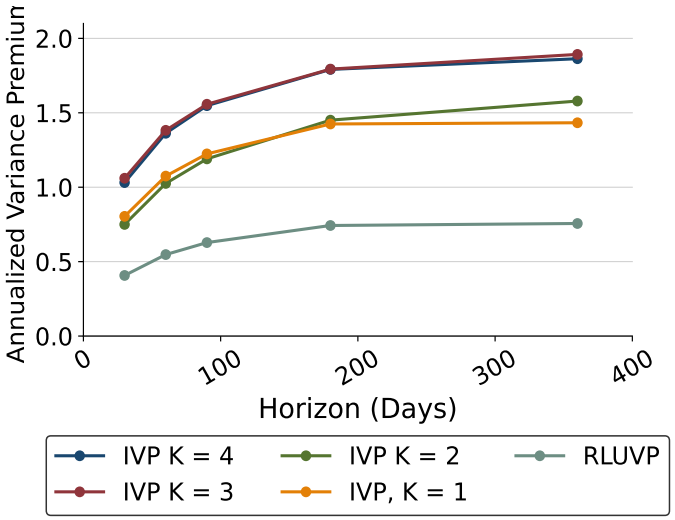
This figure shows components of the implied equity premium (IEP) at the 360-day horizon. IEP is based on the fourth-order ($K = 4$) approximation of the growth-optimal pricing kernel, IEP4_{360}. IEP components are the four terms ($k = 1, 2, 3, 4$) in equation (9), where k is the exponent applied to the market excess return. Data are daily from 1997 through 2021.

Figure 10: Term Structure of the Implied Equity Premium



This figure shows the term structure of implied equity premiums (IEPs) for models with degrees $K = 1, 2, 3, 4$ and for representative log utility (RLU). The term structures in the right (left) panel are averages during NBER recessions (normal times).

Figure 11: Term Structure of the Implied Variance Premium

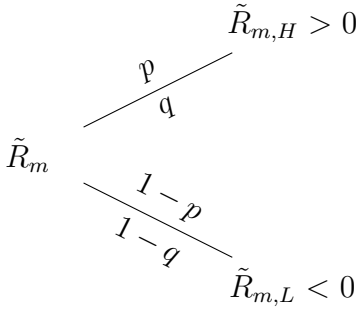


This figure shows the term structure of average implied variance premiums (IVPs in %) for models with degrees $K = 1, 2, 3, 4$ and for representative log utility (RLU).

Appendix : Recovery in a Two-State Model

In a simple setting, we show how to use empirical estimates of the variance premium and risk-neutral skewness to find the true equity premium and pricing kernel. As in the main model, available assets include a risk-free security with a gross rate of R_f, a stock market with an excess return of $\tilde{R}_m$, and options on the market with a continuum of strike prices. We assume that there are no arbitrage opportunities.

The model has two states: state H has high market excess returns, $\tilde{R}_{m,H} > 0$, and state L has low market excess returns, $\tilde{R}_{m,L} < 0$. State H occurs with probability $0 < p < 1$. These restrictions on return realizations and probabilities are necessary to ensure the absence of arbitrage. We parameterize return realizations so that the true equity premium is μ and return volatility is σ. This implies that $\tilde{R}_{m,H} = \mu + p^{-0.5}(1-p)^{0.5}\sigma$ and $\tilde{R}_{m,L} = \mu - p^{0.5}(1-p)^{-0.5}\sigma$. Negative excess returns in the low state implies an upper bound on the market Sharpe ratio: $\mu/\sigma < p^{0.5}(1-p)^{-0.5}$.



The econometrician does not observe the equity premium, μ, but has useful information about it. She observes the second moment of market excess returns, $\mathbb{E}\tilde{R}_m^2$, which is like empirical market variance. She effectively observes all risk-neutral moments of excess market returns, $\mathbb{E}^*\tilde{R}_m^j$ for $j = 1, ..., \infty$, because they are weighted sums of prices of market options. In this simple model, observation of risk-neutral variance and skewness, $\mathbb{E}^*\tilde{R}_m^2$ and $\mathbb{E}^*\tilde{R}_m^3$, will be sufficient to recover the physical return distribution.

By the no-arbitrage assumption, the growth-optimal portfolio consisting of the risk-free asset and stock market is the reciprocal of the pricing kernel, $m = \left(R_f + w\tilde{R}_m\right)^{-1}$, that prices these two assets. The two parameters are R_f, which is known, and w, which is the unknown weight on stocks. In this two-state model with no arbitrage, any pricing kernel

that prices the risk-free rate and the stock market must also price all options based on the market because the former assets span the latter.

One can explicitly compute risk-neutral probabilities and pricing kernel values for the two states in terms of the unknown equity premium and market variance. Because the risk-neutral excess return of the market is zero, the risk-neutral probabilities of the high and low states must be $q = -\frac{\tilde{R}_{m,L}}{\tilde{R}_{m,H}+\tilde{R}_{m,L}} > 0$ and $1-q = \frac{\tilde{R}_{H,L}}{\tilde{R}_{m,H}+\tilde{R}_{m,L}} > 0$, respectively. Because the pricing kernel is proportional to the ratio of risk-neutral to physical probabilities, its values in the high and low states must be $R_f\left(R_f + w\tilde{R}_{m,H}\right)^{-1} = q/p$ and $R_f\left(R_f + w\tilde{R}_{m,L}\right)^{-1} = (1-q)/(1-p)$, respectively. From the pricing kernel definition, this implies that the unknown weight on stocks is:

$$w = -\frac{R_f \mu}{\tilde{R}_{m,L}\tilde{R}_{m,H}}.$$

This weight on stocks is positive, $w > 0$, because $-\tilde{R}_{m,L} > 0$.

We now show how to estimate w and thus μ from option prices and physical return variance. The first step is to express observable quantities in terms of the unknown parameters of the physical return distribution. Risk-neutral variance is:

$$\begin{aligned}\mathbb{E}^*\tilde{R}_m^2 &= q\tilde{R}_{m,H}^2 + (1-q)\tilde{R}_{m,L}^2 \\ &= -\tilde{R}_{m,L}\tilde{R}_{m,H},\end{aligned}$$

which is positive because $-\tilde{R}_{m,L} > 0$. Risk-neutral skewness is:

$$\begin{aligned}\mathbb{E}^*\tilde{R}_m^3 &= q\tilde{R}_{m,H}^3 + (1-q)\tilde{R}_{m,L}^3 \\ &= -\tilde{R}_{m,L}\tilde{R}_{m,H}\left(\tilde{R}_{m,H} + \tilde{R}_{m,L}\right).\end{aligned}$$

The variance premium is:

$$\begin{aligned}\mathbb{E}^*\tilde{R}_m^2 - \mathbb{E}\tilde{R}_m^2 &= -\tilde{R}_{m,L}\tilde{R}_{m,H} - \left(\sigma^2 + \mu^2\right) \\ &= \left[p^{0.5}(1-p)^{-0.5} - p^{-0.5}(1-p)^{0.5}\right]\mu\sigma - 2\mu^2.\end{aligned}$$

The second step is applying the variance premium equation (11) in the main model to the case here with a single parameter, w, in the pricing kernel and solving for w to obtain:

$$\begin{aligned} w &= \frac{\mathbb{E}_t^* \tilde{R}_m^2 - \mathbb{E}\tilde{R}_m^2}{-R_f^{-1}\mathbb{E}_t^* \tilde{R}_m^3} \\ &= -\frac{R_f \mu}{\tilde{R}_{m,L}\tilde{R}_{m,H}}, \end{aligned}$$

which is the same as the true value for w.

The third and last step is applying the equity premium equation (10) in the main model to the linear pricing kernel with slope w to obtain:

$$\begin{aligned} \mathbb{E}\tilde{R}_m &= R_f^{-1} w \mathbb{E}^* \tilde{R}_m^2 \\ &= R_f^{-1} \frac{R_f \mu}{\tilde{R}_{m,L}\tilde{R}_{m,H}} \left(\tilde{R}_{m,L}\tilde{R}_{m,H} \right) \\ &= \mu. \end{aligned}$$

This last equation shows that the option-implied equity premium equals the true equity premium in this simple model.

Recovery of the true pricing kernel, m, and equity premium, μ, is exact for any physical distribution parameters (p, μ, σ) that satisfy the no-arbitrage assumption. In particular, the recovery procedure works regardless of the signs of the variance premium and risk-neutral skewness. The variance premium is positive whenever $p > 0.5 + 0.5 \left[\mu^2 \left(\mu^2 + \sigma^2 \right)^{-1} \right]^{0.5}$ or $p < 0.5 - 0.5 \left[\mu^2 \left(\mu^2 + \sigma^2 \right)^{-1} \right]^{0.5}$. The former condition is satisfied for reasonable equity premium and volatility values, such as $\mu = 0.05$ and $\sigma = 0.15$, because stocks outperform Treasuries in most years ($p \approx 2/3$). The conditions for negative risk-neutral skewness are the same as those for a positive variance premium. Since these two quantities always have opposite signs, the pricing kernel parameter, w, which is the negative of the ratio of these quantities, is always positive. In the empirically relevant case, the variance premium is positive, risk-neutral skewness is negative, and w is positive.

Internet Appendix for "The Implied Equity Premium"

by

Paul C. Tetlock Jack McCoy Neel Shah

Columbia Business School

Abstract

This Internet Appendix (IA) provides additional information on our data cleaning processes; robustness checks for different methods of estimating implied volatility and growth-optimal portfolio weights; and mathematical and simulation details for calculating the implied equity premium in the BEGE model. We also provide additional summary statistics and figures documenting the properties of the estimated equity premium.

1 Measuring Option Prices and Market Variance

1.1 Option Prices

We compute risk-neutral moments of market returns from SPX option prices using OptionMetrics data. We adopt data filtering methods from Chang et al. (2013) and Martin (2017) to eliminate securities with low liquidity and unreliable prices. We augment these methods using a customized no-arbitrage filter for options with extreme strike prices, which are important for higher-order risk-neutral moments.

As in Martin (2017), we remove SPX options with:

- maturities less than 7 days or greater than 549 days
- duplicated data
- p.m. settlements
- non-positive closing bids
- quarter-end expiration dates
- non-null expiration indicators
- arbitrage violations relative to the SPX index
- the higher midpoint price, comparing the put and call for each date, maturity, and strike price.

The minimum price filter restricts the sample to out-of-the-money options. For each maturity, we set the SPX futures price as the lowest strike price such that an SPX put exceeds the value of an SPX call by one penny.[1] We also drop options with open interest less than 1000 contracts to ensure sufficient liquidity. We drop options with zero or missing implied volatility, which violate arbitrage bounds.

[1]This definition is almost perfectly correlated with the futures price implied by futures-spot parity.

We drop a few options with glaring data errors to reduce outliers, though these omissions have little substantive impact on the results. Based on sudden changes in data coverage and extreme illiquidity, we drop seven sets of options with the following date-maturity combinations and nine contracts with the following date-maturity-strike-type combinations:

Date	Maturity	Strike	Type
1998-03-24	1999-03-20	All	Both
1998-09-21	1999-09-18	All	Both
1999-09-20	2000-09-16	All	Both
1999-09-21	2000-09-16	All	Both
1999-09-22	2000-09-16	All	Both
2007-06-19	2008-06-21	All	Both
2001-09-17	2002-06-22	1900	Call
2001-09-18	2002-06-22	1900	Call
2001-09-19	2002-06-22	1900	Call
2001-09-20	2002-06-22	1600	Call
2001-09-20	2002-06-22	750	Put
2001-09-21	2002-06-22	800	Put
2006-01-19	2006-12-16	1900	Call
2008-10-10	2009-09-19	1525	Call
2011-06-06	2012-06-16	3000	Call

We design a set of filters to ensure sufficient information for the computation of risk-neutral moments at each horizon. For each date-maturity pair, we require at least:

- a moneyness (i.e., strike/index) range from 95% to 105%
- two calls
- two puts
- five options of any type.

The minimum moneyness range drops 0.8% of date-maturity pairs and the last three requirements drop only 0.1% of pairs.

Whereas Chang et al. (2013) drop options with bid-ask midpoints of less than \$0.375 (3/8), we drop options with extreme moneyness until options with adjacent strike prices differ by at least a penny. This filter eliminates most remaining arbitrage violations and illiquid options.

1.2 Risk-Neutral Moments

As in Chang et al. (2013) and Martin (2017), we compute risk-neutral moments based on observed option prices and strike prices for each date and maturity. We use a discrete approximation of the integral in equation (20) of the paper based on the range of available strike prices supplemented using extrapolated option prices, following Chang et al. (2013). We assume options with strikes outside the available range of moneyness have implied volatility equal to a trend-line extrapolation of the volatility skew based on the three options with the closest strike prices. We estimate the trend-line parameters by minimizing the sum of absolute deviations from the three closest strikes. We constrain the absolute value of the slope to be no greater than 1.0, which affects almost no data. We only extrapolate insofar as the range of available moneyness is within two standard deviations of the futures price. We also winsorize the implied volatility range at 5% and 100% and the moneyness range at 1% and 300% of the futures price, though these limits rarely bind. This extrapolation procedure creates an additional 1.8% of options beyond the original cleaned data.

Linear extrapolation of the implied volatility (IV) surface might raise concerns, such as underestimating tail risk and violating no-arbitrage by introducing a kink in the IV surface. We address these issues in Internet Appendix (IA) Section 6 by using a procedure that allows rich tail features that capture investor preferences during uncertainty. There we implement the stochastic volatility inspired (SVI) model in Gatheral and Jacquier (2014) to obtain arbitrage-free IV surfaces.

To infer the values of risk-neutral moments at consistent horizons of monthly, bimonthly, quarterly and annual, we interpolate between moment values at the nearest available maturities, as in Martin (2017) and related studies. Our interpolation procedure ensures data availability and reduces the likelihood of errors. First, we set the raw odd (even) risk-neutral moments to missing for date-maturity pairs with non-negative (non-positive) values, which would be inconsistent with theory and prior evidence. This filter only affects 43 third-order moments (0.1%) and 179 fifth-order moments (0.5%) out of 51,013 date-maturity pairs; and it affects no even moments. We also set a handful of third- and fifth-order moments that exceed a very small negative (annualized) value, -0.0001, to missing.

Second, for each moment, date, and irregular maturity, we fill in missing daily moments using adjusted lagged weekly moments. To this end, we partition all moments into six maturity groups with cutoffs defined by the uniform maturities of 30, 60, 90, 180, and 360 days, creating groupings of [7,30], [31, 60], [61, 90], [91, 180], [181, 360], and [361, 549] days.

For each group, date, and moment, we estimate the lagged weekly moment using a rolling median of the moment's non-missing values on days -7 to -1. We fill in missing weekly medians for the longest horizons, e.g., [361, 549], using values from the next longest horizon, e.g., [181, 360]. We then apply this procedure to the shortest horizons with missing weekly data using the next shortest horizon. These missing data procedures apply to fewer than 0.7% of observations for all weekly moments.

Third, to obtain values for standardized maturities, we interpolate between adjacent horizon groups based on the local slope of the term structure of the risk-neutral moment. For example, we estimate the third risk-neutral moment at a 30-day maturity using a [1/3, 2/3] weighted average of the third risk-neutral moments with 10-day and 40-day maturities. We apply this interpolation to the raw daily moments and to the lagged weekly moments, filling in any remaining missing values of weekly moments with most recently available values. Lastly, we fill in any missing daily moments using adjusted weekly moments. The adjusted weekly moment is the unadjusted weekly moment plus the median difference between the daily and weekly moments across all horizons for which both are nonmissing. We apply this same interpolation procedure to OptionMetrics' risk-free rates with irregular maturities to obtain risk-free rates with standardized maturities.

1.3 ETF Variance

We estimate the physical variance of market returns using the realized variance (RV) of the ETF with ticker SPY. The RV estimator is the sum of squared log SPY returns over 78 intraday intervals, where 78 is the number of 5-minute periods in regular exchange trading hours. The final RV estimator is the average of 10 sub-sampled RV estimators based on 10 staggered sets of 78 non-overlapping intervals. Each set of 78 return intervals is based on 79 trade prices that are equally spaced in transaction time—i.e., equal numbers of trades—throughout regular trading hours. Thus, the 10 RV estimators use a total of 790 prices equally spaced in business time. The first RV estimator uses prices 1, 11, 21, ..., 781, the second uses prices 2, 12, 22, ..., 782, and the tenth uses prices 10, 20, 30, ..., 790.

For these estimators, we use only trades that satisfy the following standard conditions, applied in the order below:

- price is positive
- quantity is positive

- time must be between 9:30 a.m. and 4:00 p.m. Eastern Time
- correction code is "00"
- condition code is not 1, 4, 7, 8, 9, A through D, G, H, K, L, N, P, R, S, U through W, Y, or Z
- exchange is the most active for SPY on that day
- price must be within the range of the ETF's daily low and high.

We aggregate all trades by transaction time and use the median trade price for each time as the price. We drop trades with extreme reversals as defined an absolute return exceeding five times the 50-observation rolling average. The number of unique transaction times is the basis for the partition into 790 prices equally spaced in business time.

We fix two glaring data errors in trade prices. We obtain the daily low and high prices from the Center for Research in Securities Prices (CRSP) database. On March 31, 1997, we use high and low prices from Yahoo! Finance because the CRSP high and low prices are incorrect. On May 12, 1999, we set SPY prices in TAQ that are below the daily CRSP low to the index low for trades before 12 noon and to the previous valid trade price for trades after noon. These changes reduce outliers but have little substantive impact on the results.

2 Sharpe Ratios for All Models

In Table IA.I, we present the Sharpe ratios implied by various equity premium models (for $K = 1, 2, 3, 4$) at the monthly (30-day) and annual (360-day) horizons. The implied equity premium (IEP) model exhibits a Sharpe ratio of 0.36 (for K=4) at the annual horizon, which is much higher than the RLU Sharpe of 0.22. These Sharpe ratios vary substantially over time, with standard deviations for annual Sharpe ratios between 0.11 and 0.15.

3 Bootstrap Standard Errors of GO Weights

As discussed in Section 3.2, we calculate the standard errors of GO weights using a moving block bootstrap. We chose a block length of 252 days (one year), which corresponds to the maximum horizon for our equity premium analysis. Figure IA.1 shows that our standard error estimates are generally increasing in block length. Standard errors nearly double when

extending the block length from 63 to 252 days; however further increases in block length only modestly increase the standard errors.

4 Term Structure of the Implied Equity Premium

Figure IA.2 shows the term structure of the implied equity premium (IEP). The level is defined as the annualized IEP at the annual horizon. The slope is defined as the difference between the annualized one-year and one-month IEPs, normalized by the maturity difference: $(365/360 \times \text{IEP4}_{360} - 365/30 \times \text{IEP4}_{30}) / ((360 - 30)/365)$. The slope of the IEP term structure exhibits a strong negative correlation with the level: -0.94 (-0.60) at the annual (monthly) horizon. While the IEP slope is generally close to zero, it turns sharply negative during crises, consistent with Figure IA.2.

5 Comparing Monthly Equity Premium Estimates

Whereas Figure 6 in the main paper compares the one-year (360-day) implied equity premium (IEP) to the lower bound from Martin (2017) and conditional bounds from Chabi-Yo and Loudis (2020), Figure IA.3 compares the monthly (30-day) IEP measures to the corresponding monthly benchmarks.

6 Robustness to Arbitrage-Free IV Surfaces

6.1 Carr and Madan (2005) Arbitrage Filter Algorithm

Here we analyze how estimates of risk-neutral moments depend on static arbitrage opportunities by applying the algorithm of Carr and Madan (2005). They show that applying three filters on the costs of option trading strategies ensures that there is no static arbitrage. The filter is straightforward to apply and model-free, fitting with our goal of minimizing the dependency of the moments and IEP on any one particular model.

Following Carr and Madan's notation, let $i = 0, \ldots, N$ index strikes K_i, let $j = 1, \ldots, M$ index available maturities T_j, and let $C_{i,j}$ denote the corresponding call option price ($P_{i,j}$ the put option price). The first filter is for the price of a vertical spread which is long

$1/(K_i - K_{i-1})$ at strike K_{i-1} and short $1/(K_i - K_{i-1})$ of strike K_i:

$$Q_{i,j} = \frac{C_{i-1,j} - C_{i,t}}{K_i - K_{i-1}}, \qquad Q_{0,j} = 0$$

the payoff of which is bounded between zero and one (for put options, we reverse the sign in the numerator). Therefore, we filter for

$$0 \leq Q_{i,j} \leq 1, \ \forall i, j.$$

The second filter is for the cost of a butterfly spread

$$BSpr_{i,j} = C_{i-1,j} - \frac{K_{i+1} - K_{i-1}}{K_{i+1} - K_i} C_{i,j} + \frac{K_i - K_{i-1}}{K_{i+1} - K_i} C_{i+1,j}$$

which is long one share of the call at K_{i-1}, short $(K_{i+1} - Ki - 1)/(K_{i+1} - K_i)$ shares of the call at K_i, and long $(K_i - K_{i-1})/(K_{i+1} - K_i)$ shares of the call at K_{i+1}. The payoff is always greater than zero, and the formula is the same for puts. We therefore keep only strikes i, j such that

$$BSpr_{i,j} \geq 0.$$

The last filter imposes that the calendar spread between maturities j and $j + 1$ has a nonnegative cost:

$$C_{i,j+1} - C_{i,j} \geq 0.$$

We apply these filters iteratively, as arbitrages may still remain between the two strikes K_{i-1} and K_{i+1} after removing K_i. We apply the algorithm until the percent of new observations removed in iteration n relative to the number of observations remaining in iteration $n - 1$ is less than a specified tolerance level, which we set at 0.1%. We do this before extrapolating the tails with a mean-absolute deviation line, with the intent of preserving the same plus-or-minus two standard deviation moneyness range, but it does have the downside of potentially allowing for arbitrages in the extrapolated tails. However, this occurs in less than 2.5% of the extrapolated observations.

Table IA.II shows the means and standard deviations of the baseline risk neutral moments in the paper compared to those where we apply Carr and Madan's filter. Notably, the levels and volatility of the moments are extremely close, suggesting that our baseline filters already eliminate the most significant arbitrages and bad price quotes. Table IA.II also shows that

the pairwise correlations between the baseline and arbitrage-free moments are also nearly 100% across all moments and horizons.

6.2 SVI Option IV Surface

Estimating risk-neutral moments of excess market returns, $\tilde{R}_{m,T}$, is challenging when available options have a limited range of strike prices. This finite range introduces a truncation error that varies over time, influenced by shifts in investor preferences, risk aversion, and market liquidity. Additionally, the discretization of strike prices poses challenges, particularly in the earlier parts of the sample, when the grid of available strike prices may not be sufficiently dense.

Our primary strategy for addressing truncation error is to generate synthetic options using a minimum absolute deviation (MAD) extrapolation scheme. Specifically, we assume that the implied volatility (IV) function remains constant beyond the range of available liquid options up to ± 2 standard deviations. However, this approach creates two potential problems: it could violate no-arbitrage conditions by creating discontinuities in the IV surface and distorting the shape, thus affecting the estimation of risk-neutral moments.

To address these issues, we calibrate Gatheral and Jacquier's (2014) stochastic-volatility-inspired (SVI) model. The SVI model, which is widely regarded for its practical relevance and interpretive value, expresses implied variance as

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

where $x = \log\left(K/F_{t,\tau}\right)$ is the option's log moneyness, $F_{b,\tau}$ is the futures price for maturity τ, and a, b, ρ, m, and σ are parameters. For each date, we calibrate the SVI model by minimizing the RMSE between observed IVs and SVI-implied IVs for all horizons separately and use the parameters to generate an arbitrage-free surface. The fit has an average R^2 of 98.3% and a median R^2 of 99.7% across all date-maturities. Finally, we extrapolate out to a minimum of two standard deviations OTM for each horizon and create a fine grid with an interval of 0.01 moneyness units.

6.3 Robustness to Additional IV Surface Methods

Table IA.II shows that the moments obtained from the Carr and Madan filtering algorithm and the SVI-implied surface are highly correlated with those from our base model and have

similar means. Moments obtained from the Carr and Madan filtering algorithm are effectively indistinguishable from the moments obtained from the IV surface without those filters. However, the SVI exerts a more substantial impact by more carefully modeling the tails, particularly the right-hand side of the IV smile in the earlier part of our sample and produces moments that are slightly different for the higher order moments, especially for the odd (third and the fifth) moments.

Figure IA.4 shows the effect of these alternative approaches on our GO weight estimation for the $K = 4$ IEP model. Using the moments from SVI, the GO weights for k = 1 and k = 3 are larger, particularly before the 2008 crisis, because of the higher partial correlation between the variance premium and the odd risk neutral moments.

Figure IA.5 shows the resulting IEP estimates for the monthly and annual horizons, applying the re-estimated weights to the risk-neutral moments corresponding to each IV surface cleaning and extrapolation method. Because they result in similar GO weights, these alternative methods have little impact on our baseline estimates. Table IA.III shows that the equity premium predictions are very similar across all models. The SVI surface produces slightly higher equity premium estimates (by 0.84% for the 360-day horizon and by 0.18% for the 30-day horizon) because of higher GO weights in the pre-crisis period. Thus, our main results are not overly sensitive to how one estimates risk-neutral moments in the IV surface.

7 Weighting Methods

Our IEP equation (10) suggests that one can recover the GO portfolio weights by estimating a linear regression of the variance premium on the discounted risk neutral moments. In Section 2.2, we use feasible generalized least squares (FGLS) regressions with cross-sectional restrictions and inverse variance precision weights to address multicollinearity, heteroskedasticity, and autocorrelation. To demonstrate robustness, we consider alternate specifications.

In the first alternative, we consider a state-space representation in which time-varying GO weights result from a learning process. We model this learning using the Kalman filter in which GO weights are latent variables that we estimate recursively through prediction and update steps as in Adrian and Franzoni (2009). We obtain priors from a pooled regression with horizon-specific intercepts, using the first three years of data for reasonably precise priors. The state transition uses an identity matrix (random walk). We jointly update system uncertainty and Kalman gain using data from all horizons while imposing the cross-sectional restriction for weights, following the spirit of multi-sensor fusion Kalman filter (Sun and Deng, 2004).[2]

We use 30-day expected realized variance ($\text{PVar}_{t,30}$)) as a measure of process noise and derive measurement noise recursively from the target variable with a learning rate of 1%. A higher process noise increases the filter's sensitivity to data. In our approach, process noise typically outweighs measurement error, allowing the algorithm to prioritize data. This approach is sensitive to the choice of prior and noise processes, which often involve subjective decisions.

As another alternative, we incorporate insights from the dynamic conditional correlation (DCC) literature (Engle (2002, 2016)). The dynamic conditional beta (DCB) model uses the theory of conditional heteroskedastic factor models to model the dynamics of betas, capturing potential non-linearities and asymmetries in their evolution over time. The coefficients (GO weights) are calculated as:

$$\hat{\beta}_t^{DCB} = \mathbf{\Sigma}_{XX,t}^{-1} \cdot \mathbf{\Sigma}_{XY,t} \tag{1}$$

where the conditional covariance matrix $\mathbf{\Sigma}_t$ is specified as a DCC-GARCH(1,1) model, meaning $\mathbf{\Sigma}_t = \text{diag}(h_t)^{1/2}\ R_t\ \text{diag}(h_t)^{1/2}$, h_t consists of univariate GARCH(1,1) variances for each

[2]Alternatively, we can update the weights for each horizon separately and compute a precision-weighted average. This approach produces GO weights that are both qualitatively and quantitatively comparable to those derived from joint estimation.

excess risk neutral moment, and the conditional correlation matrix R_t comes from a multivariate GARCH (MGARCH). Specifically, we model the conditional means of the regressors as a first-order VAR process and the conditional covariances as a DCC-MGARCH process in which the variance of each error term follows GARCH(1,1). We jointly estimate conditional correlation matrices for all moments and horizons every week using expanding windows with the same recency weights as in Section 2.2.[3] The joint estimation of conditional correlation allows us to capture cross-correlation effects. Finally, we average weights over a one-year rolling window to smooth noisy daily estimates of the conditional variance.

Like our FGLS weights in Section 2.2, these weights are available in real time, but start from 1999, after observing three years of data. Figure IA.6 compares Kalman filter and DCB weights. They share characteristics with FGLS weights: a positive market weight (k=1) typically exceeding one that rises after the 2008 crisis, and a large negative weight on the variance asset (k=2). However, higher measurement noise during the 2008 crisis causes the Kalman filter to react less to new information, producing a lower equity premium as seen in Figure IA.7. The DCB approach, on the other hand produces higher weights around the same period, driven by an increase in both the conditional correlations and the conditional variances. Another notable departure is the flipped signs of the DCB weights on higher order (k = 3, 4) assets, driven by the conditional correlation between the variance premium and the fifth (sixth) excess moment being negative (positive).

The IEP derived from these weights correlates strongly with that from FGLS weights (correlations of 83% to 97% across horizons) and exhibits similar means and out-of-sample return predictability, especially in the no-arbitrage period. This evidence shows that, despite variations in GO weight estimation from high multicollinearity, these different methods product equity premium estimates with similar properties.

8 BEGE Model

This section describes the solution for the extended "preference shock" version of Bekaert, Engstrom, and Ermolov (BEE, 2023)'s model. Their original appendix includes only the solution for the baseline version of the model. While the solution for the preference shock version is a straightforward application of their algorithm, we include it here for completeness.

[3]We estimate this weekly for ease of computation.

8.1 Model Set Up

Utility has a habit form:

$$V = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{(C_t - H_t)^{1-\gamma} - 1}{1-\gamma}.$$

One-period consumption evolves according to

$$\begin{aligned} g_{t+1} &= \bar{g} + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\ \omega_{p,t+1} &\sim \Gamma(p_t, 1) - p_t \\ \omega_{n,t+1} &\sim \Gamma(n_t, 1) - n_t \end{aligned}$$

The shape parameters of the shocks follow a mean-reverting AR(1) process with the same shocks as consumption:

$$\begin{aligned} p_{t+1} &= \bar{p} + \rho_p(p_t - \bar{p}) + \sigma_{pp}\omega_{p,t+1} \\ n_{t+1} &= \bar{n} + \rho_n(n_t - \bar{n}) + \sigma_{nn}\omega_{n,t+1} \end{aligned}$$

Log inverse surplus $q_t = \log(Q_t) = \log(C_t/(C_t - H_t))$ is modeled directly as

$$\begin{aligned} q_{t+1} &= \bar{q} + \rho_q(q_t - \bar{q}) + \sigma_{qp}\omega_{p,t+1} + \sigma_{qn}\omega_{n,t+1} + \sigma_{qq}\omega_{q,t+1}, \\ \omega_{q,t+1} &\sim \Gamma(s_t, 1) - s_t, \\ s_{t+1} &= \bar{s} + \rho_s(s_t - \bar{s}) + \sigma_{sq}\omega_{q,t+1}. \end{aligned}$$

Lastly, dividends are a more volatile claim to the consumption shocks:

$$d_{t+1} = \bar{g} + \gamma_g(\sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1}).$$

All prices and returns here refer to the dividend-paying asset.

8.2 Pricing Kernel and Risk-Free Rate

The one period pricing kernel is

$$\begin{aligned} M_{t,t+1} &:= M_{t+1} \\ &= \frac{\partial V/\partial C_{t+1}}{\partial V/\partial C_t} \\ &= \frac{\beta^{t+1}(C_{t+1}-H_{t+1})^{-\gamma}}{\beta^{t}(C_{t+1}-H_{t+1})^{-\gamma}} \\ &= \beta^1 \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{C_{t+1}-H_{t+1}}{C_{t+1}}\right)^{-\gamma} \left(\frac{C_t}{C_t-H_t}\right)^{-\gamma} \\ &= \beta^h e^{\gamma(q_{t+1}-q_t-g_t)}. \end{aligned}$$

This implies a log pricing kernel that is linear in the state variables and shocks:

$$m_{t+1} = m_0 + m_q q_t + m_{\omega p}\omega_{p,t+1} + m_{\omega n}\omega_{n,t+1} + m_{\omega q}\omega_{q,t+1} \tag{2}$$

where

$$\begin{aligned} m_0 &= \log\beta - \gamma\bar{g} + \gamma(1-\rho_q)\bar{q} \\ m_q &= -\gamma(1-\rho_q) \\ m_{\omega,p} &= \gamma(\sigma_{qp} - \sigma_{cp}) \\ m_{\omega,n} &= \gamma(\sigma_{qn} + \sigma_{cn}) \\ m_{\omega,q} &= \gamma\sigma_{qq}. \end{aligned}$$

This pricing kernel is similar to the kernel in the baseline model, but with the additional preference shock term $m_{\omega,q}\omega_{q,t+1}$.

All moment calculations here depend on the fact that, as shown by BEE, for $X \sim \Gamma(k,\theta) - k\theta$

$$\mathbb{E}\, e^X = e^{-g(\theta)k}, \qquad g(\theta) = \theta + \log(1-\theta).$$

The one-period log risk-free rate is given by

$$r_{f,t} = -\log \mathbb{E}_t\, e^{m_{t+1}} = f_0 + f_q q_t + f_p p_t + f_n n_t + f_s s_t \tag{3}$$

where

$$
\begin{aligned}
f_0 &= -m_0 \\
f_q &= -m_q \\
f_p &= g(m_{\omega p}) \\
f_n &= g(m_{\omega n}) \\
f_s &= g(m_{\omega q})
\end{aligned}
$$

8.3 Equity Pricing

In this model, the price-dividend ratio of the equity asset has an exact solution, which we log-linearize around the expected values of the state variables to enable feasible computations of expected returns and higher order moments. We follow BEE's algorithm here.

By the definition of P/D with no bubbles:

$$\frac{P_t}{D_t} = \mathbb{E}_t \sum_{i=1}^{\infty} e^{\sum_{j=1}^{i} m_{t+h} + d_{t+j}}.$$

The solution method is to use recursive coefficients and iterated expectations to obtain the form

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty} e^{A_i + B_i p_t + C_i n_t + D_i q_t + F_i s_t} \tag{4}$$

The base case in the inductive algorithm is

$$
\begin{aligned}
\mathbb{E}_t\, e^{m_{t+1} + d_{t+1}} &= \mathbb{E}_t\, e^{m_0 + m_q q_t + m_{\omega p}\omega_{p,t+1} + m_{\omega n}\omega_{n,t+1} + m_{\omega q}\omega_{q,t+1} + \bar{g} + \gamma_g(\sigma_{cq}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1})} \\
&= e^{m_0 + \bar{g} - g(m_{\omega p} + \gamma_g \sigma_{cp}) p_t - g(m_{\omega n} - \gamma_g \sigma_{cn}) n_t + m_q q_t - g(m_{\omega q}) s_t}
\end{aligned}
$$

which yields the initial coefficients

$$
\begin{aligned}
A_1 &= m_0 + \bar{g} \\
B_1 &= -g(m_{\omega p} + \gamma_g \sigma_{cp}) \\
C_1 &= -g(m_{\omega n} - \gamma_g \sigma_{cn}) \\
D_1 &= m_q \\
F_1 &= -g(m_{\omega q}).
\end{aligned}
$$

Assuming that the coefficients hold for term $i-1$, the algorithm applies iterated expectations and coefficient matching to term i:

$$\begin{aligned} e^{A_{i-1}+B_{i-1}p_t+C_{i-1}n_t+D_{i-1}q_t+F_{i-1}s_t} &= \mathbb{E}_t\, e^{\sum_{i=1}^{j} m_{t+i}+d_{t+i}} \\ &= \mathbb{E}_t\left[e^{m_{t+1}+d_{t+1}}\,\mathbb{E}_{t+1}\, e^{\sum_{j=2}^{i} m_{t+j}+d_{t+j}}\right] \\ &= \mathbb{E}_t\, e^{m_{t+1}+d_{t+1}+A_{i-1}+B_{i-1}p_{t+1}+C_{i-1}n_{t+1}+D_{i-1}q_{t+1}+F_{i-1}s_{t+1}} \end{aligned}$$

where the third line applies the inductive hypothesis. Expanding the pricing kernel, dividends, and state variables using their definitions shows that the term in the exponent on the right-hand side is

$$\begin{aligned} & m_0 + \bar{g} + B_{i-1}\bar{p}(1-\rho_p) + C_{i-1}\bar{n}(1-\rho_n) + D_{i-1}\bar{q}(1-\rho_q) + F_{i-1}\bar{s}(1-\rho_s) \\ & \quad + (m_q + D_{i-1}\rho_q)q_t + B_{i-1}\rho_p p_t + C_{i-1}\rho_n n_t + F_{i-1}\rho_s s_t \\ & \quad + (m_{\omega p} + \gamma_g\sigma_{cp} + B_{i-1}\sigma_{pp} + D_{i-1}\sigma_{qp})\omega_{p,t+1} \\ & \quad + (m_{\omega n} - \gamma_g\sigma_{cn} + C_{i-1}\sigma_{nn} + D_{i-1}\sigma_{qn})\omega_{n,t+1} \\ & \quad + (m_{\omega q} + D_{i-1}\sigma_{qq} + F_{i-1}\sigma_{sq})\omega_{q,t+1}. \end{aligned}$$

Evaluating the expectation and matching the coefficients then yields

$$A_i = A_1 + A_{i-1} + B_{i-1}\bar{p}(1-\rho_p) + C_{i-1}\bar{n}(1-\rho_n) + D_{i-1}\bar{q}(1-\rho_q) + F_{i-1}\bar{s}(1-\rho_s) \tag{5}$$

$$B_i = B_{i-1}\rho_p - g(m_{\omega p} + \gamma_g\sigma_{cp} + B_{i-1}\sigma_{pp} + D_{i-1}\sigma_{qp}) \tag{6}$$

$$C_i = C_{i-1}\rho_n - g(m_{\omega n} - \gamma_g\sigma_{cn} + C_{i-1}\sigma_{nn} + D_{i-1}\sigma_{qn}) \tag{7}$$

$$D_i = D_1 + D_{i-1}\rho_q \tag{8}$$

$$F_i = F_{i-1}\rho_s - g(m_{\omega q} + D_{i-1}\sigma_{qq} + F_{i-1}\sigma_{sq}) \tag{9}$$

which is the conjectured form.

8.3.1 Log-linearization

Using the standard decomposition of log returns

$$r_{t+1} = \log R_{t+1} = \log\left(1 + \frac{P_{t+1}}{D_{t+1}}\right) - pd_t + d_{t+1}$$

we log-linearize the terms $pd_t = \log(P_t/D_t)$ and $\log(1 + P_{t+1}/D_{t+1})$ around the expected values of the state variables $(\bar{p}, \bar{n}, \bar{q}, \bar{s})$. For notational convenience, define

$$\overline{\Psi}_i := e^{A_i + B_i\bar{p} + C_i\bar{n} + D_i\bar{q} + F_i\bar{s}}.$$

Then, the first order approximation of pd_t is

$$pd_t \approx K_0^1 + K_p^1 p_t + K_n^1 n_t + K_q^1 q_t + K_s^1 s_t$$

where

$$\begin{aligned}
K_p^1 &= \frac{\sum_{i=1}^{\infty} B_i \overline{\Psi}_i}{\sum_{i=1}^{n} \overline{\Psi}_i} \\
K_n^1 &= \frac{\sum_{i=1}^{\infty} C_i \overline{\Psi}_i}{\sum_{i=1}^{n} \overline{\Psi}_i} \\
K_q^1 &= \frac{\sum_{i=1}^{\infty} D_i \overline{\Psi}_i}{\sum_{i=1}^{n} \overline{\Psi}_i} \\
K_s^1 &= \frac{\sum_{i=1}^{\infty} F_i \overline{\Psi}_i}{\sum_{i=1}^{n} \overline{\Psi}_i} \\
K_0^1 &= \log\left(\sum_{i=1}^{\infty} \overline{\Psi}_i\right) - K_p^1\bar{p} - K_n^1\bar{n} - K_q^1\bar{q} - K_s^1\bar{s}.
\end{aligned}$$

Similarly, the first order Taylor expansion of the nonlinear term is

$$\log\left(1 + \frac{P_{t+1}}{D_{t+1}}\right) \approx K_0^2 + K_p^2 p_{t+1} + K_n^2 n_{t+1} + K_q^2 q_{t+1} + K_s^2 s_{t+1}$$

where

$$
\begin{aligned}
K_p^2 &= \frac{\sum_{i=1}^{\infty} B_i \overline{\Psi}_i}{1 + \sum_{i=1}^{n} \overline{\Psi}_i} \\
K_n^2 &= \frac{\sum_{i=1}^{\infty} C_i \overline{\Psi}_i}{1 + \sum_{i=1}^{n} \overline{\Psi}_i} \\
K_q^2 &= \frac{\sum_{i=1}^{\infty} D_i \overline{\Psi}_i}{1 + \sum_{i=1}^{n} \overline{\Psi}_i} \\
K_s^2 &= \frac{\sum_{i=1}^{\infty} F_i \overline{\Psi}_i}{1 + \sum_{i=1}^{n} \overline{\Psi}_i} \\
K_0^2 &= \log\left(1 + \sum_{i=1}^{\infty} \overline{\Psi}_i\right) - K_p^2 \bar{p} - K_n^2 \bar{n} - K_q^2 \bar{q} - K_s^2 \bar{s}.
\end{aligned}
$$

Substituting these terms into the return decomposition and expanding the $t+1$ terms in terms of time t state variables and $t+1$ shocks, we obtain that one-period log returns are approximately

$$
r_{t+1} \approx r_0 + r_p p_t + r_n n_t + r_q q_t + r_s s_t + r_{\omega p} \omega_{p,t+1} + r_{\omega n} \omega_{n,t+1} + r_{\omega q} \omega_{q,t+1} \tag{10}
$$

where

$$
\begin{aligned}
r_0 &= \bar{g} + K_0^2 - K_0^1 + K_p^2 \bar{p}(1-\rho_p) + K_n^2 \bar{n}(1-\rho_n) + K_q^2 \bar{q}(1-\rho_q) + K_s^2 \bar{s}(1-\rho_s) \\
r_p &= K_p^1 + K_p^2 \rho_p \\
r_n &= K_n^1 + K_n^2 \rho_n \\
r_q &= K_q^1 + K_q^2 \rho_q \\
r_s &= K_s^1 + K_s^2 \rho_s \\
r_{\omega p} &= K_p^2 \sigma_{pp} + K_q^2 \sigma_{qp} + \gamma_g \sigma_{cp} \\
r_{\omega n} &= K_n^2 \sigma_{nn} + K_q^2 \sigma_{qn} - \gamma_g \sigma_{cn} \\
r_{\omega p} &= K_s^2 \sigma_{sq} + K_q^2 \sigma_{qq}.
\end{aligned}
$$

8.4 Physical Moments

To calculate physical moments of simple returns $R_{t+1} = e^{r_{t+1}}$, we apply BEE's result for the moment-generating function of a demeaned gamma distribution. For any expectation

of R^k_{t+1}, where $k \in \mathbb{N}_+$, we can obtain quasi-closed form expressions using BEE's log-linear approximation:

$$\begin{aligned}\mathbb{E}_t\, R^k_{t+1} &\approx \mathbb{E}_t\, e^{k(r_0+r_p p_t+r_n n_t+r_q q_t+r_s s_t+r_{\omega n}\omega_{n,t+1}+r_{\omega p}\omega_{p,t+1}+r_{\omega q}\omega_{q,t+1})} \\ &= e^{kr_0+kr_q q_t+(kr_p-g(kr_{\omega p}))p_t+(kr_n-g(kr_{\omega n}))n_t+(kr_s-g(kr_{\omega q}))s_t}\end{aligned}$$

This expression allows us to compute the physical variance in the model as

$$\mathrm{Var}_t(\tilde{R}_{t+1}) \approx \mathbb{E}_t\, e^{2r_{t+1}} - \left(\mathbb{E}_t\, e^{r_{t+1}}\right)^2.$$

8.5 Risk-Neutral Moments

We calculate higher-order risk-neutral moments of excess returns by expanding the polynomial

$$\mathbb{E}^*_t\, \tilde{R}^k_{t+1} = e^{r_{ft}}\, \mathbb{E}_t\left[e^{m_{t+1}}\left(R_{t+1}-R_{ft}\right)^k\right] = \sum_{j=0}^{k}\binom{k}{j}(-1)^j e^{(j+1)r_{ft}}\, \mathbb{E}_t\, e^{m_{t+1}+(k-j)r_{t+1}}. \tag{11}$$

Then, the expectation term is

$$\begin{aligned}&\mathbb{E}_t\, e^{m_{t+1}+ur_{t+1}} \\ &= e^{m_0+m_q q_t+u(r_0+r_q q_t+r_p p_t+r_n n_t+r_s s_t)} \times \mathbb{E}_t\, e^{(m_{\omega p}+ur_{\omega p})\omega_{p,t+1}+(m_{\omega n}+ur_{\omega n})\omega_{n,t+1}+(m_{\omega q}+ur_{\omega q})\omega_{q,t+1}} \\ &= e^{m_0+ur_0+(m_q+ur_q)q_t+(ur_p-g(m_{\omega p}+ur_{\omega p}))p_t+(ur_n-g(m_{\omega n}+ur_{\omega n}))n_t+(ur_s-g(m_{\omega q}+ur_{\omega q}))s_t}.\end{aligned}$$

8.6 Model Simulations

8.6.1 Errors in Approximations

Figure IA.8 shows the approximation errors in log returns that result from linearizing. Approximation errors are negligible for returns at or above -16% monthly. Returns are below -16% in 0.16% of periods in our 10,000 simulated time series of 324 observations. Even for exact returns of -30%, the approximate return understates the exact return by only 2 percentage points.

8.6.2 Distribution of Rolling Weights

Figure IA.9 shows the distribution of the IEP coefficients in the BEGE model. We estimate coefficients with separate rolling regressions in each of the 10,000 simulated time series. We use least squares with an expanding window and minimum of 12 observations to estimate the coefficients with no intercept term.

8.7 Implied SDFs in the Model

We compare implied SDFs from the IEP methodology to approximations of the true SDFs in the model. Because the model has multiple state variables and shocks, there is no closed form expression for the SDF M_{t+1} in terms of the excess return $\tilde{R}_{t+1}$. Therefore, we focus on comparing IEP estimates to the quantity:

$$\mathbb{E}\left[M_{t+1}|\mathcal{I}_t, \tilde{R}_{t+1}\right] \tag{12}$$

where the information set $\mathcal{I}_t$ includes the four state variables in the model.

We calculate the expectation of the SDF in equation (12) by considering the model's state on a given day t in which risk-neutral variance, $M2_t$, equals its median value and simulating the model's three shocks forward for one period. We truncate the sample at the 0.5 and 99.5 percentiles of returns. We then obtain the conditional expectation of the SDF as a function of returns using both a locally-weighted regression and a tenth-order polynomial expansion with restrictions corresponding to those in equation (9). Both methods produce similar shapes for the pricing kernel. We compare this kernel to SDF estimates from the IEP1 and IEP4 models with constant weights in Figure IA.10. The true SDF exhibits a pronounced kink below zero which is difficult for polynomials to capture. Nonetheless, the SDF from IEP4 reasonably approximates the curve of the true SDF, while the SDF from IEP1 is too flat, emphasizing the advantages of including higher-order risk-neutral moments in estimating the equity premium.[4]

[4]The SDF from IEP4 does curve slightly upward at higher levels of excess returns, contrasting with the monotonic pricing kernel in the empirical analysis of 5, which is likely due to errors in the estimated weights from multicollinearity.

References (Internet Appendix)

Adrian, Tobias and Franzoni, Francesco, (2009), Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM, *Journal of Empirical Finance* 16, 537–556.

Bekaert, Geert, Eric Engstrom, and Andrey Ermolov, 2023, The variance risk premium in equilibrium models, *Review of Finance* 27, 1977–2014.

Carr, Peter, and Dilip Madan, 2005, A note on sufficient conditions for no arbitrage, *Finance Research Letters* 2, 125–130.

Chang, Bo-Young, Peter Christoffersen, and Kris Jacobs, 2013, Market skewness risk and the cross section of stock returns, *Journal of Financial Economics* 107, 46–68.

Engle, Robert F., 2002, Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics*, 20, 339–350.

Engle, Robert F., 2016, Dynamic conditional beta, *Journal of Financial Econometrics* 14, 643–667.

Gatheral, J., & Jacquier, A., 2013, Arbitrage-free SVI volatility surfaces, *Quantitative Finance*, 14(1), 59–71.

Martin, Ian, 2017, What is the expected return on the market? *Quarterly Journal of Economics* 132, 367–433.

Sun, Shu-Li and Deng, Zi-Li, 2004 Multi-sensor optimal information fusion Kalman filter, *Automatica* 40, 1017–1023.

Table: IA.I: Sharpe Ratios of IEP Measures

This table summarizes the Sharpe ratio implied by the four implied equity premium models (Implied Sharpe K, where $K = 1, 2, 3, 4$) and from the representative log utility model (RLU Sharpe) at monthly (30-day) and annual (360-day) horizons. The implied Sharpe ratio that uses K terms to approximate the pricing kernel at horizon h days is Implied Sharpe K_h. The RLU Sharpe at horizon h days is RLU Sharpe$_h$. Sharpe ratios are annualized. The columns show the mean (Mean), standard deviation (StdDev), and percentiles (P5, P25, P50, P75, P95) of Sharpe ratios for all 6,293 days from 1997 to 2021.

	Mean	StdDev	P5	P25	P50	P75	P95	Days
RLU Sharpe $_{30}$	0.22	0.11	0.11	0.14	0.20	0.26	0.44	6293
Implied Sharpe 1_{30}	0.41	0.23	0.19	0.26	0.35	0.49	0.86	6293
Implied Sharpe 2_{30}	0.34	0.19	0.16	0.21	0.29	0.40	0.71	6293
Implied Sharpe 3_{30}	0.48	0.25	0.25	0.32	0.40	0.57	0.92	6293
Implied Sharpe 4_{30}	0.40	0.21	0.18	0.26	0.34	0.47	0.76	6293
RLU Sharpe$_{360}$	0.22	0.06	0.14	0.17	0.21	0.26	0.34	6293
Implied Sharpe 1_{360}	0.40	0.15	0.25	0.30	0.35	0.46	0.70	6293
Implied Sharpe 2_{360}	0.32	0.11	0.21	0.25	0.29	0.37	0.58	6293
Implied Sharpe 3_{360}	0.44	0.14	0.28	0.33	0.41	0.50	0.68	6293
Implied Sharpe 4_{360}	0.36	0.12	0.20	0.27	0.35	0.43	0.56	6293

Table IA.II: Risk-Neutral Moments from Alternative Methods

This table compares the second through sixth risk-neutral moments–M2 through M6–as computed in the main paper with those computed by applying the arbitrage filtering algorithm of Carr and Madan (2005) and those from the stochastic-volatility-inspired (SVI) algorithm (Gatheral and Jacquier (2013)). We present results for 30-day, 90-day and 360-day horizons, as shown by the subscripts on each moment. Mean, StdDev and Median columns show the average, standard deviation and the median values of risk-neutral moments from each procedure. The Corr columns display correlations between moments from our baseline process of linear extrapolating the IV surface and moments from alternative procedures.

	MAD Extrapolation			Carr & Madan Filters				Stochastic Vol. Inspired			
	Mean	StdDev	Median	Mean	StdDev	Median	Corr	Mean	StdDev	Median	Corr
$M2_{30}$	4.28	4.37	3.14	4.26	4.34	3.12	1.000	4.75	4.72	3.48	0.996
$M2_{90}$	4.40	3.40	3.53	4.39	3.39	3.52	1.000	4.50	3.49	3.58	0.998
$M2_{360}$	4.30	2.23	3.84	4.30	2.23	3.84	1.000	4.29	2.27	3.79	0.997
$M3_{30}$	-0.41	0.55	-0.24	-0.40	0.54	-0.23	0.999	-0.47	0.56	-0.28	0.967
$M3_{90}$	-0.63	0.59	-0.44	-0.62	0.59	-0.44	1.000	-0.64	0.59	-0.45	0.992
$M3_{360}$	-0.76	0.41	-0.63	-0.75	0.41	-0.63	0.999	-0.73	0.41	-0.61	0.967
$M4_{30}$	0.34	0.82	0.13	0.33	0.80	0.13	1.000	0.38	0.80	0.15	0.976
$M4_{90}$	0.71	1.19	0.38	0.70	1.18	0.37	1.000	0.73	1.22	0.40	0.996
$M4_{360}$	1.57	1.57	1.11	1.56	1.56	1.10	1.000	1.61	1.72	1.11	0.991
$M5_{30}$	-0.30	0.82	-0.09	-0.29	0.79	-0.08	0.998	-0.30	0.65	-0.10	0.906
$M5_{90}$	-0.77	1.34	-0.35	-0.75	1.33	-0.34	0.999	-0.74	1.28	-0.35	0.979
$M5_{360}$	-1.68	1.42	-1.10	-1.66	1.41	-1.09	0.999	-1.60	1.39	-1.06	0.976
$M6_{30}$	0.61	2.28	0.11	0.57	2.16	0.10	0.998	0.54	1.64	0.11	0.910
$M6_{90}$	1.89	4.68	0.61	1.83	4.60	0.57	0.999	1.74	4.46	0.58	0.975
$M6_{360}$	6.36	10.08	3.30	6.25	9.88	3.26	0.999	6.81	13.46	3.47	0.972

Table IA.III: Equity Premiums from Alternative Risk-Neutral Moments

The table summarizes implied equity premiums (IEPs) based on risk-neutral moments from alternative arbitrage-free surfaces. The first seven columns present the mean, standard deviation, and percentiles of each IEP4 estimate obtained from different IV surfaces. The eighth and ninth columns list the mean IEP of the baseline model–based on linear extrapolation of the IV surface–and correlations with IEPs from alternative IV surface estimates. The last two columns show out-of-sample R^2 for the full-sample from 1996 to 2021 and the "No Arbitrage" sample, as defined in the paper. We label rows by the name of each model, where "CM" refers to the IV surface with the additional Carr and Madan (2005) filters applied; "SVI" refers to the estimates from risk-neutral moments calculated from a fitted SVI model (Gatheral and Jacquier (2013)); and the subscripts refer to horizons in days.

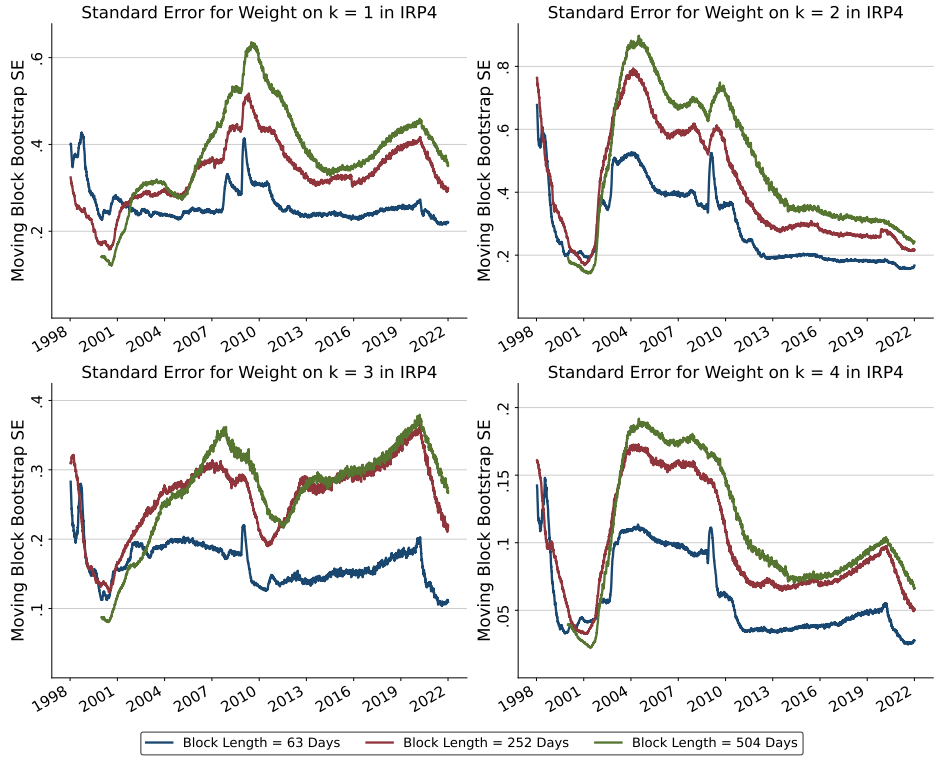
	Mean	StdDev	P5	P25	P50	P75	P95	Mean (MAD)	Corr	OOS R^2	OOS R^2 (No Arb)
CM_{30}	0.68	0.84	0.17	0.29	0.45	0.76	1.79	0.69	1.0	1.20	3.62
CM_{60}	1.4	1.44	0.40	0.68	1.0	1.61	3.33	1.42	1.0	2.28	5.96
CM_{90}	2.13	1.92	0.68	1.11	1.59	2.48	4.86	2.14	1.0	3.03	8.93
CM_{180}	4.19	3.00	1.58	2.44	3.34	4.95	8.92	4.22	1.0	7.83	10.42
CM_{360}	7.82	4.29	3.42	5.18	6.72	9.3	14.92	7.87	1.0	7.84	7.51
SVI_{30}	0.87	0.99	0.24	0.38	0.59	0.96	2.29	0.69	0.984	-0.31	3.38
SVI_{60}	1.66	1.62	0.54	0.82	1.21	1.91	3.88	1.42	0.989	0.40	5.27
SVI_{90}	2.47	2.12	0.90	1.31	1.88	2.88	5.55	2.14	0.988	0.80	8.12
SVI_{180}	4.77	3.23	2.08	2.86	3.85	5.61	9.94	4.22	0.986	6.09	9.44
SVI_{360}	8.71	4.50	4.31	5.88	7.49	10.2	16.43	7.87	0.980	6.52	6.32

Table IA.IV: Implied Equity Premium with Alternative Weights

This table summarizes estimates of the implied equity premium (IEP) with $K = 4$ at all horizons for various weighting methods as described in IA Section 7 and Section 2.2. The columns display the mean (Mean), standard deviation (StdDev), and percentiles (P5, P25, P50, P75, P95) of equity premiums for all 5,788 days from 1999 to 2021. Equity premiums are not annualized. The table also shows correlations (Corr) between the baseline IEP and IEPs from other GO-weighting methods as well as out-of-sample R^2 values for predictions from IEP models for all dates and for the No Arbitrage subsample, which excludes 61 trading days from September 19 to December 15, 2008.

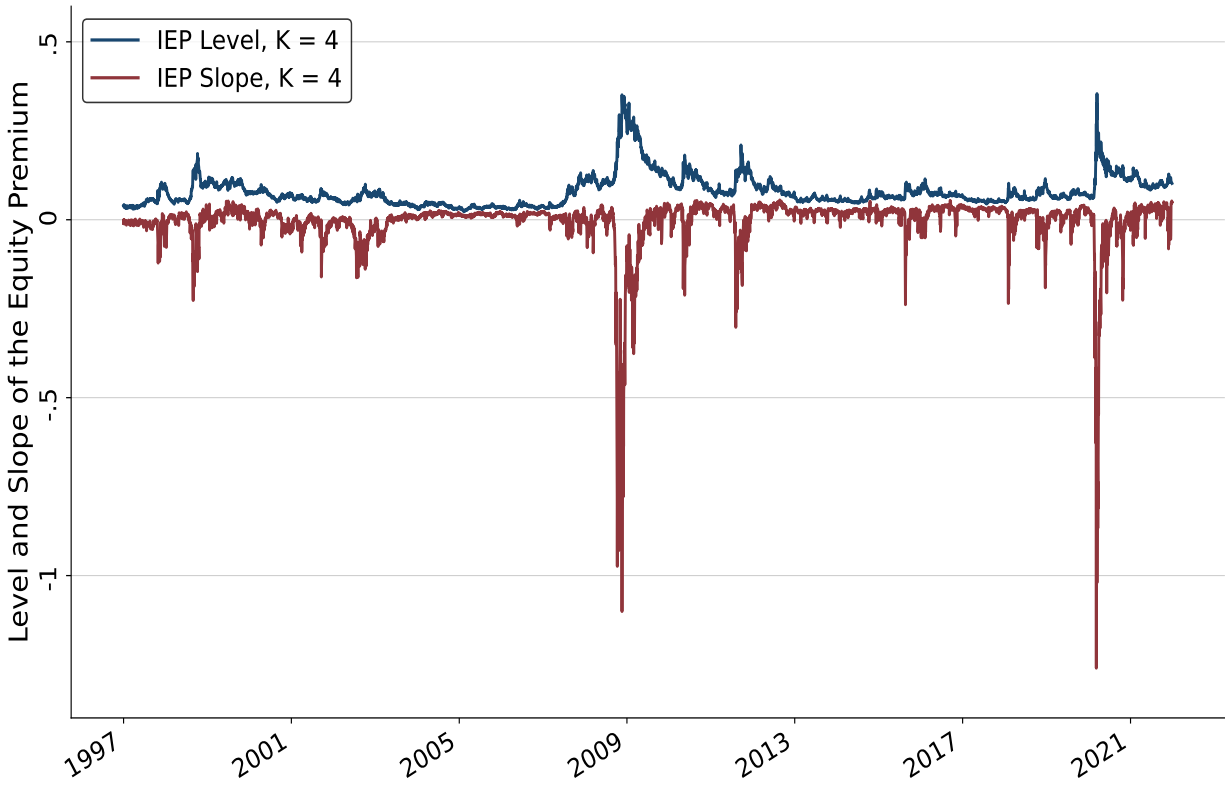
	Mean	StdDev	P5	P25	P50	P75	P95	Corr	OOS R^2	OOS R^2 (No Arb)
Kalman Filter$_{30}$	0.69	0.77	0.16	0.27	0.47	0.85	1.88	0.94	0.88	3.09
Dynamic Conditional Beta$_{30}$	0.62	1.02	0.12	0.20	0.34	0.59	1.96	0.97	0.32	3.76
Feasible Generalized LS$_{30}$	0.70	0.88	0.17	0.29	0.46	0.77	1.80	-	1.02	3.76
Kalman Filter$_{360}$	8.07	3.76	3.11	5.41	7.37	9.89	15.20	0.83	3.59	3.16
Dynamic Conditional Beta$_{360}$	8.43	7.40	3.13	4.76	6.06	9.32	20.61	0.91	11.21	12.53
Feasible Generalized LSs$_{360}$	7.98	4.43	3.39	5.29	6.85	9.46	15.45	-	9.00	8.60

Figure IA.1: Block-Bootstrap Standard Errors of GO Weights.



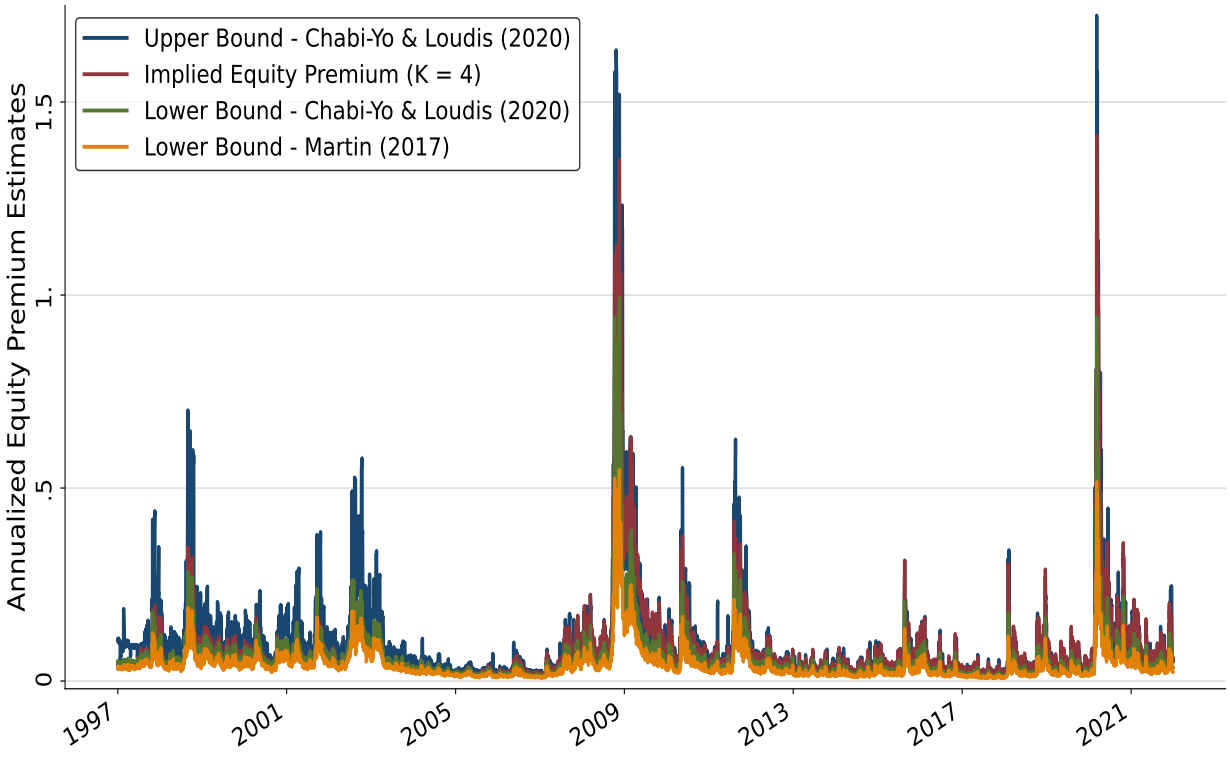
These figures show standard errors (SEs) of weights on all risky assets ($k = 1, 2, 3, 4$) in the growth-optimal portfolio based on the IEP4 model. The three SE lines in each panel come from moving block bootstraps with block lengths of 1 quarter (63 days), 1 year (252 days), and 2 years (504 days).

Figure: IA.2: Term Structure of the Implied Equity Premium



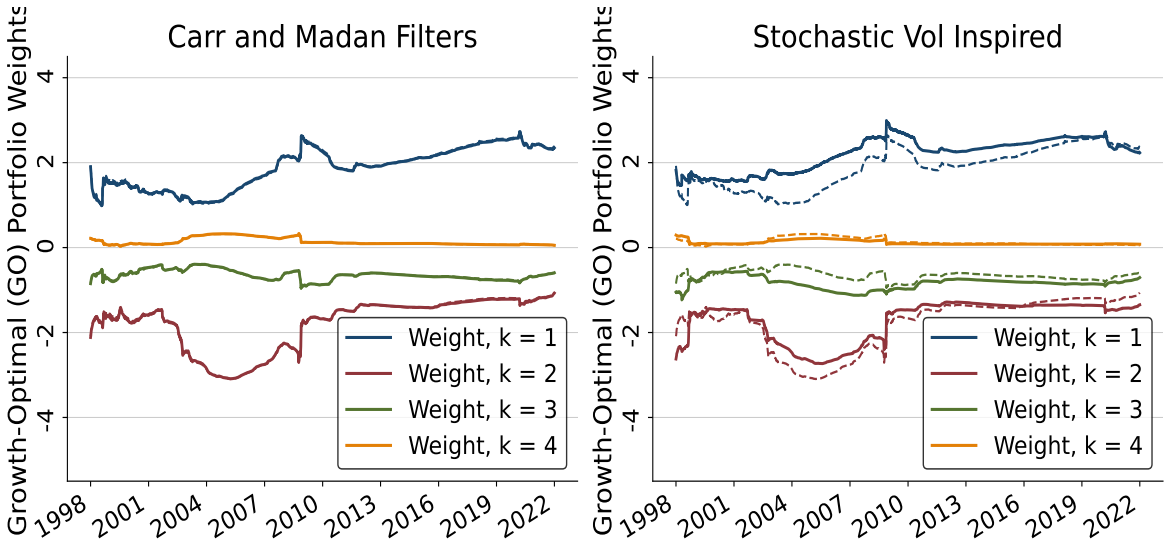
This figure shows the level and slope of the term structure of the implied equity premium (IEP). The level is IEP4_{360} in equation (9) in the paper based on the fourth-order ($K = 4$) approximation of the growth-optimal pricing kernel. The slope is the difference between the annualized one-year and one-month IEPs divided by the difference in these maturities: $(365/360 \times \text{IEP4}_{360} - 365/30 \times \text{IEP4}_{30}) / ((360 - 30)/365)$. The data are daily from 1997 through 2021.

Figure IA.3: The Implied Equity Premium and Alternative Bound Estimates



This figure shows four estimates of the annualized equity premium at the 30-day horizon: the implied equity premium (IEP), Martin's (2017) lower bound, and Chabi-Yo and Loudis (2020) upper and lower bounds on the equity premium. The IEP is based on the fourth-order ($K = 4$) approximation of the growth-optimal pricing kernel, IEP4_{30}, in equation (9). The lower bound is the representative log utility equity premium, Lower Bound_{30}, which is a special case of equation (9) with $w_{1,t} = 1$ and $w_{k,t} = 0 \; \forall \, k > 1$. Chabi-Yo and Loudis bounds are based on a replication using our moments. The data are daily from 1997 through 2021.

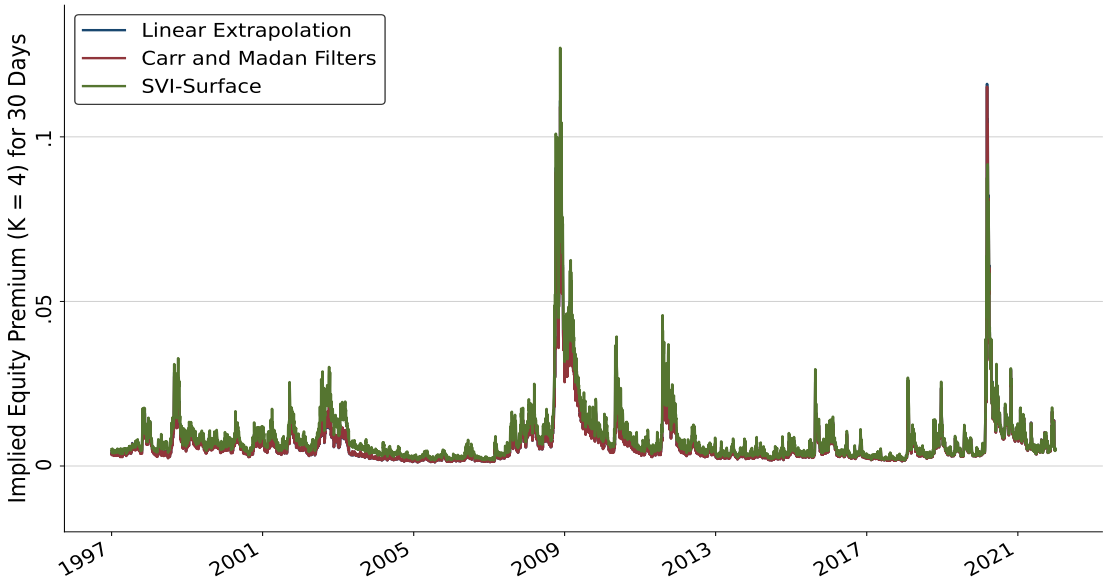
Figure IA.4: Growth-Optimal Portfolio Weights with SVI-based IV Surface



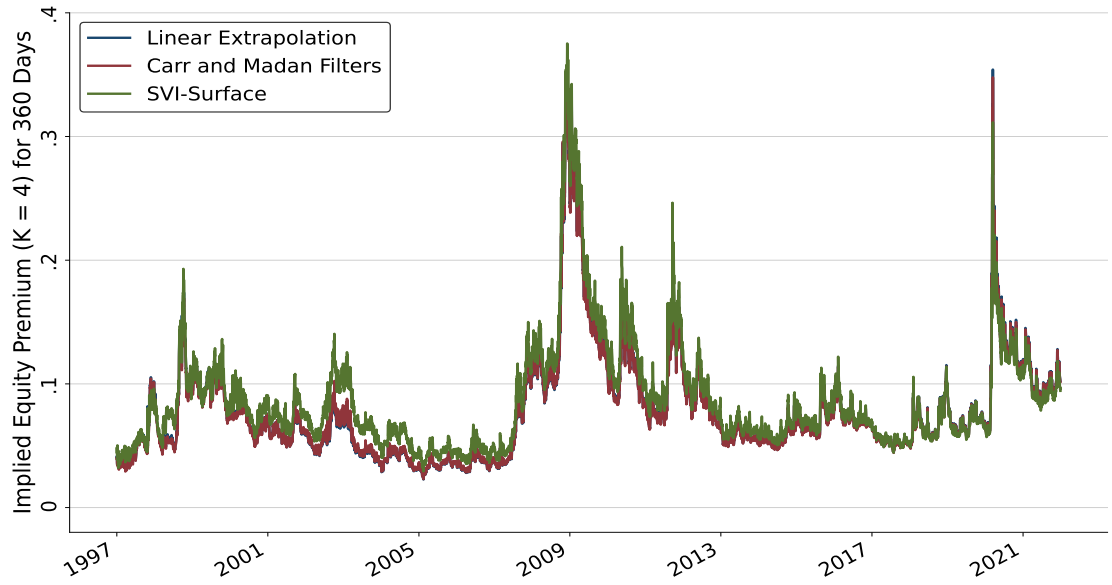
This figure shows the weights on all ($k = 1, 2, 3, 4$) risky assets in the GO portfolio based on the IEP4 model. The left panel shows the weights obtained using risk-neutral moments based on options that meet the Carr and Madan filters, while the right panel shows the weights obtained using risk-neutral moments with the implied volatility (IV) surface extrapolated by fitting an SVI model. Dotted lines in both figures correspond to weights obtained using our baseline procedure without the extra filters and with linear extrapolation of the IV surface.

Figure IA.5: IEP4 with Alternative IV Surfaces

Panel A: 30-day horizon

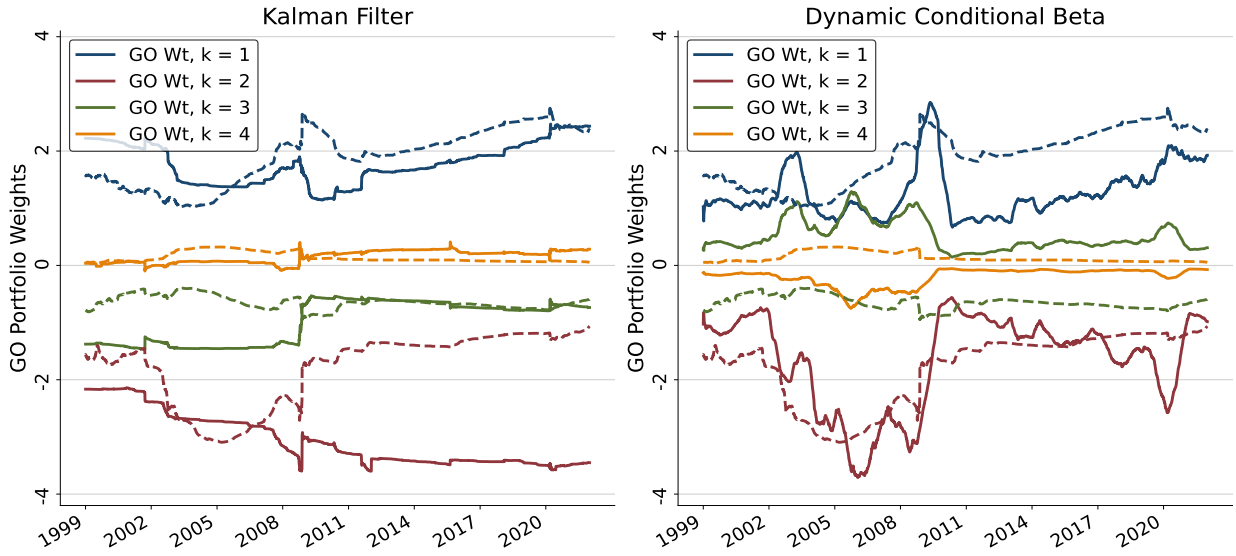


Panel B: 360-day horizon



This figure compares the baseline implied equity premium (IEP4) from linear extrapolation to IEPs from two alternative estimations of the implied volatility surface: Carr and Madan filters and SVI. Panel A shows IEP4 estimates for the 30-day horizon, and Panel B shows those for the 360-day horizon.

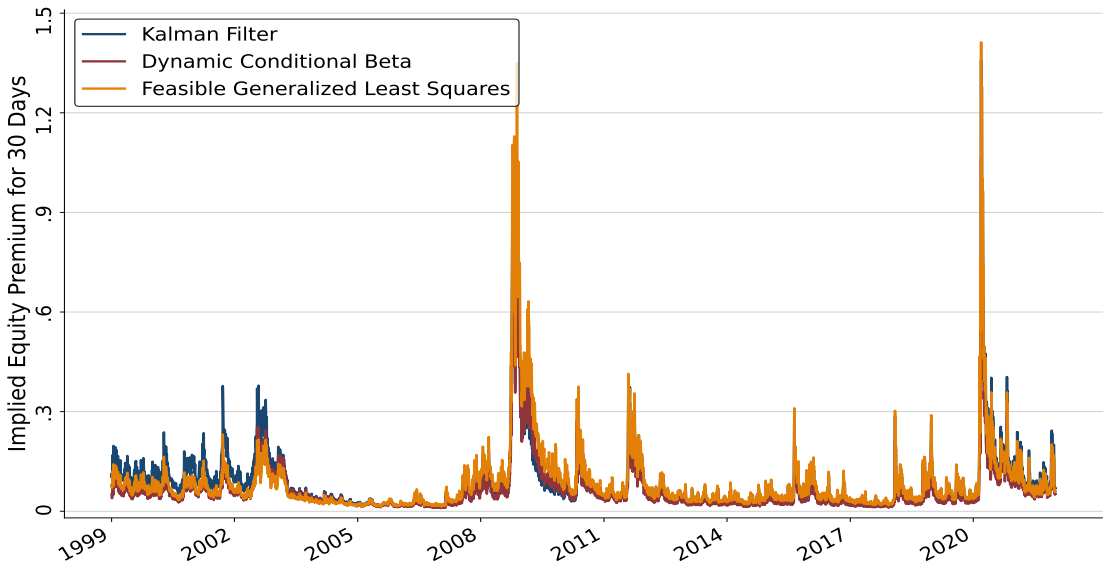
Figure IA.6: Growth-Optimal Portfolio Weights (Alternative Weighting Methods)



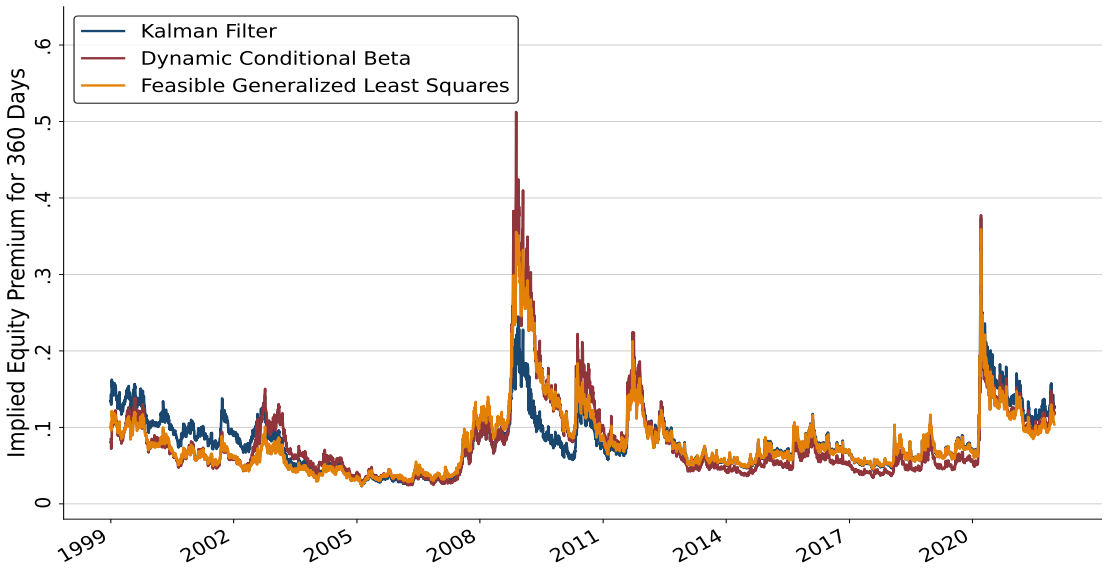
The figure shows the weights on all risky assets ($k = 1, 2, 3, 4$) in the GO portfolio based on the IEP4 model using the Kalman filter and dynamic conditional betas as described in IA Section 7. Dotted lines in figures represent baseline GO weights obtained using the FGLS regression as in Figure 2 in the main paper.

Figure IA.7: 30-Day Implied Equity Premium (Alternative Weighting Methods)

Panel A: 30-day horizon

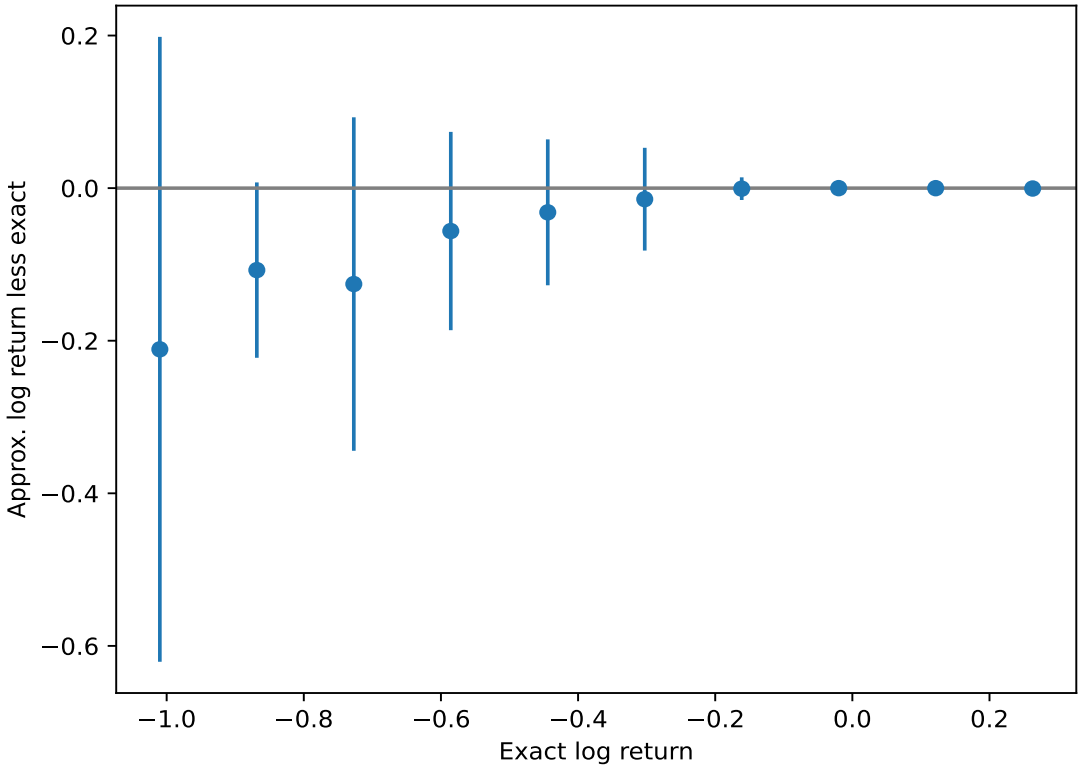


Panel B: 360-day horizon



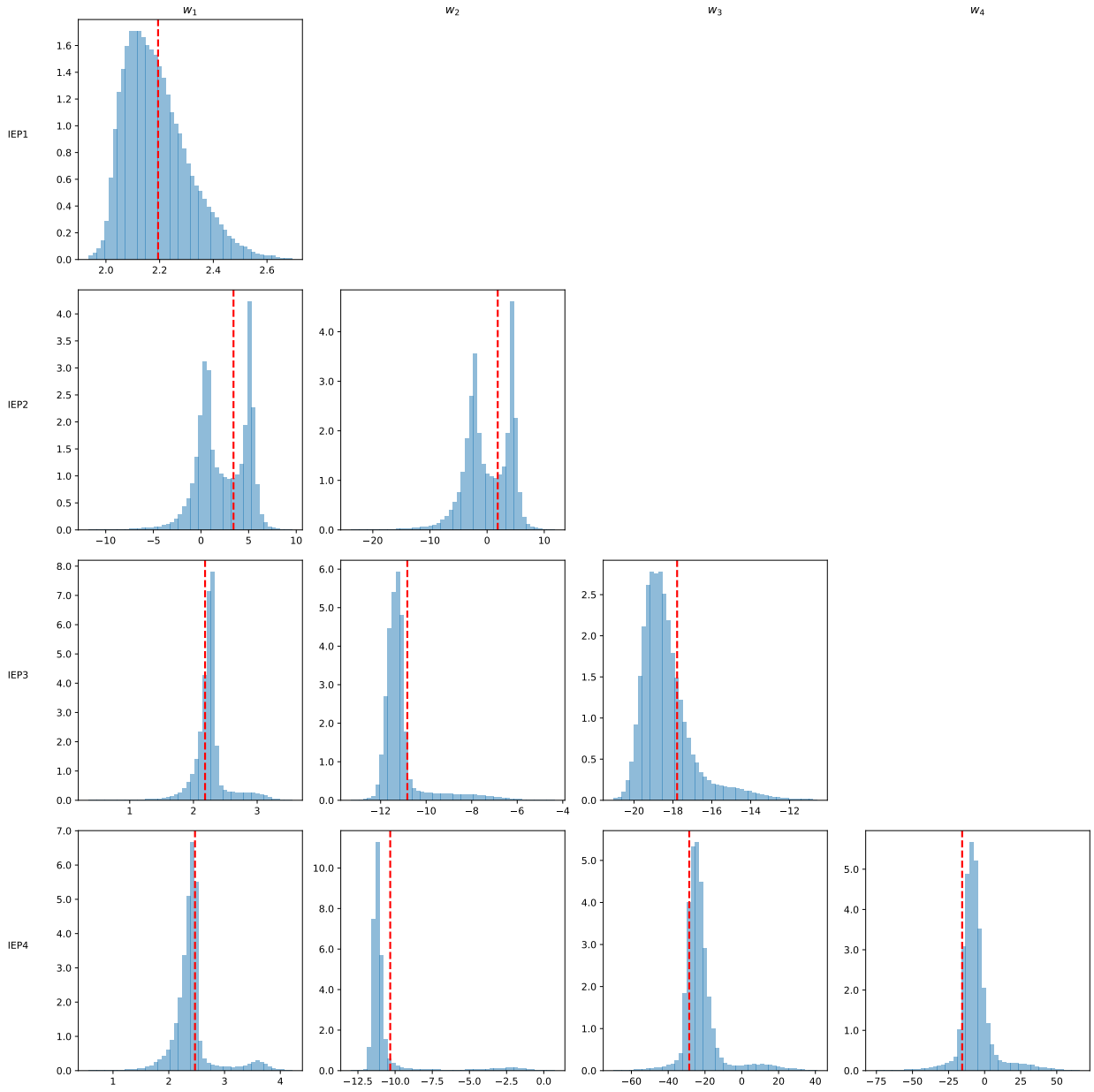
Each panel in this figure shows the annualized implied equity premium (IEP) obtained from three methods. The IEP comes from the fourth-order ($K = 4$) approximation of the growth-optimal pricing kernel, IEP4, in equation (9). The three methods estimate this equation using the Kalman filter, dynamic conditional betas, and the feasible generalized least squares to obtain GO (i.e., regression) weights as described in IA Section 7 and Section 2.2. Panel A shows IEP4 estimates for the 30-day horizon, and Panel B shows IEP4 for the 360-day horizon Data is daily from 1999 through 2021.

Figure IA.8: Approximation Errors in the BEGE Model Return Linearization



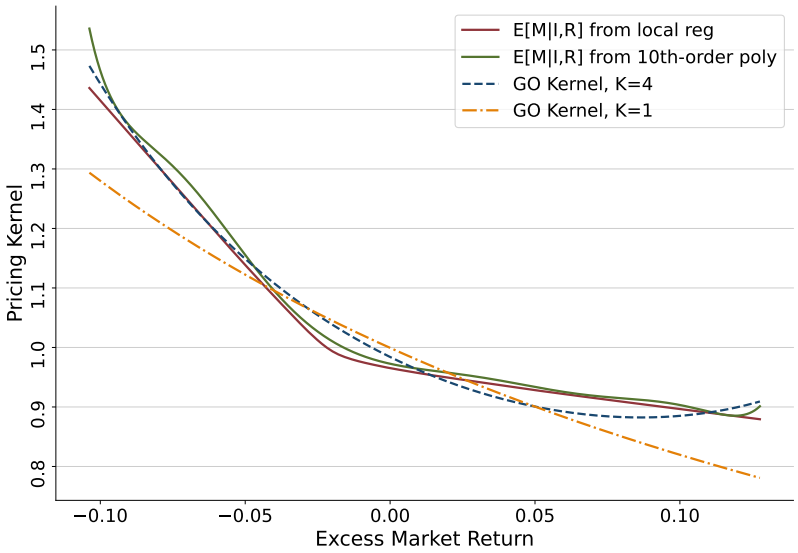
The figure shows the approximation error in log returns for the BEGE log-linearization as a function of the exact returns using a binned scatterplot. The y axis shows approximate log returns minus exact log returns, and the x axis shows exact log returns. Each point corresponds to a binned mean of the errors: the y value is the mean error for each bin, and the x value is the midpoint of the upper and lower bin boundaries. The whiskers around each point show two-standard-deviation ranges for each bin. Because bins have equal width, the leftmost bins have far fewer data points than the rightmost bins.

Figure IA.9: Distribution of Portfolio Weights in the Model



The figure shows the distribution of estimated growth-optimal (GO) portfolio weights in the implied equity premium (IEP) in the simulated BEGE model. Each row corresponds to a given IEP model ($K = 1, 2, 3, 4$), and each column corresponds to a given weight parameter (w_k, where $k = 1, 2, 3, 4$). The y axis values represent counts in the hundreds of thousands. The dashed red line shows the full-sample GO weight estimate across all 10,000 simulated time series.

Figure IA.10: SDF in a BEGE Model



This figure compares the true pricing kernel in a simulated BEGE model to IEP pricing kernels with approximation degrees $K = 1, 4$. We approximate the true kernel using the expected pricing kernel given returns for the day with median risk-neutral variance in the BEGE simulations. The solid lines show the approximate true function using both locally weighted regressions and a tenth-order polynomial, while the dashed lines show the implied growth-optimal kernel for the IEP models with constant weights.