(Almost) Model-Free Recovery

Author(s): PAUL SCHNEIDER and FABIO TROJANI

Source: *The Journal of Finance*, Vol. 74, No. 1 (February 2019), pp. 323-370
Published by: Wiley for the American Finance Association
Stable URL: https://www.jstor.org/stable/26656080
Accessed: 10-02-2025 15:01 UTC

(Almost) Model-Free Recovery

PAUL SCHNEIDER and FABIO TROJANI*

ABSTRACT

Under mild assumptions, we recover the model-free conditional minimum variance projection of the pricing kernel on various tradeable realized moments of market returns. Recovered conditional moments predict future realizations and give insight into the cyclicality of equity premia, variance risk premia, and the highest attainable Sharpe ratios under the minimum variance probability. The pricing kernel projections are often U-shaped and give rise to optimal conditional portfolio strategies with plausible market timing properties, moderate countercyclical exposures to higher realized moments, and favorable out-of-sample Sharpe ratios.

A SEMINAL FINDING IN BREEDEN and Litzenberger (1978) shows that the conditional forward-neutral distribution of an asset return in arbitrage-free option markets is uniquely identified by the second derivative of the European call price function with respect to the option's strike. No corresponding model-free result is available for the conditional distribution of returns under the physical belief.

Recent model-based recovery theorems study conditions on physical transition dynamics and pricing kernels under which the physical belief can be determined from Arrow-Debreu prices alone. Broadly speaking, this literature shows that recovery is possible whenever physical probabilities are induced by a family of path-independent pricing kernels and satisfy further stability conditions.[1] Path-independence is essential for these theorems, as, in general, only a particular probability different from the physical probability can be

*Paul Schneider is at USI Lugano and the Swiss Finance Institute. Fabio Trojani is at the University of Geneva and Swiss Finance Institute. We are indebted to Editor Ken Singleton and two anonymous referees for key comments that strongly improved our paper. Our thanks for helpful discussions and comments also go to Jaroslav Borovička, Peter Carr, Gianluca Cassese, Jerome Detemple, Damir Filipovic, Patrick Gagliardini, Stefano Giglio, Peter Gruber, Leonid Kogan, Abraham Lioui, Eric Renault, Steve Ross, Mirela Sandulescu, Olivier Scaillet, Raman Uppal, Christian Wagner, Liuren Wu, and workshop participants at Boston University, Brown University, EDHEC Business School, EPFL, IFSID Montreal, MIT, Morgan Stanley, University of Zurich, SFI Annual Research Days, and USI. We are indebted to Dacheng Xiu for providing us with parameters and state variables. Financial support from the Swiss Finance Institute (Project "Term structures and cross sections of asset risk premia,") and the Swiss National Science Foundation (Projects "Trading Asset Pricing Models," "Model-Free Asset Pricing," and "Higher Order Robust Resampling and Multiple Testing Methods") is gratefully acknowledged. We have read the *Journal of Finance* disclosure policy and have no conflicts of interest to disclose.

[1] These conditions are naturally nested by the requirement of recurrent physical state dynamics. Recurrence is the requirement that the expected number of times a stochastic process visits every

DOI: 10.1111/jofi.12737

recovered. Under further stability assumptions, this probability can be identified with the long forward probability in Hansen and Scheinkman (2009).[2] Hence, the physical belief remains unidentifiable without further information.

Acknowledging the impossibility of identifying physical probabilities without additional restrictions, we step back from the recovery approaches in the literature and ask a different question: *Which family of pricing kernels is consistent with mild sign restrictions on (i) particular covariances of market returns with the pricing kernel and (ii) the risk premiums for variance and higher order moment risks, given various pricing constraints for the tradeable second and higher realized moments of returns?*[3] We first characterize the kernel projections that exactly price tradeable second and higher realized return risks under the given sign restrictions. We next determine the minimum variance kernel projection (MVP) consistent with the most conservative lower bound on the maximal Sharpe ratio in the economy. Finally, we show that mild, economically motivated and empirically supported restrictions are sufficient to obtain an informative minimum variance probability, which corresponds to an MVP that is tradeable with delta-hedged option portfolios. The MVP naturally defines a lower bound on the maximal Sharpe ratio that is attainable in any economy satisfying the initial economic constraints. Our recovery approach therefore corresponds to an (almost) model-free kernel selection that recovers the MVP under a conservative max-min Sharpe ratio criterion.

We determine the MVP from mild hypotheses on the risk premia of certain tradeable realized moments of returns. These hypotheses can include, for instance, a negative risk premium for particular variance payoffs, a positive risk premium for payoffs that are increasing in market returns, or ranking relations between the Sharpe ratios of various higher moments. Together with the maintained no-arbitrage assumption that equivalent physical and forward conditional moments are defined on a common information set, these hypotheses imply a binding lower bound for the variance of kernel projections, which is elicited from Arrow-Debreu prices alone. In this sense, our methodology is (almost) model-free, as it depends exclusively on reasonable and testable assumptions about particular risk premia, which allow us to avoid hypotheses on the structure of the economy that might be difficult to test in practice. For instance, we make no ex-ante assumption on the state

subset of its state space is infinite. See Carr and Yu (2012), Ross (2015), Walden (2017), and Qin and Linetsky (2017), among others, for various recovery theorems based on recurrent physical dynamics.

[2] See Borovička, Hansen, and Scheinkman (2016) for a thorough discussion. This probability also yields a unique long-run factorization of the pricing kernel into transitory and permanent (martingale) components.

[3] While, theoretically, the market equity premium is often assumed positive in equilibrium, further natural assumptions can be motivated for the risk premia on higher moments. For instance, theoretical and empirical literature on market variance risk typically concludes that the risk premium for even realized moment risk is negative; see Carr and Wu (2009), Martin (2013), and Schneider and Trojani (2014), among others. Similarly, the risk premium for odd realized moment risk, such as skewness risk, tends to be positive; see Kozhan, Neuberger, and Schneider (2013) and Schneider and Trojani (2014).

space of returns, Markovianity or stationarity of the relevant state processes under the forward or physical probabilities, semimartingale properties, or path-independence of the pricing kernel. Clearly, the cost of this generality arises from our assumptions on the risk premia of particular payoffs, which are nonetheless easier to interpret theoretically and to validate empirically.

Our empirical analysis covers 25 years of monthly Standard & Poor's (S&P) 500 option data from January 1990 to August 2016. We find that our MVPs produce useful insights into the conditional structure of pricing kernels and their dependence on the investment horizon. MVPs for horizons above one quarter are usually U-shaped. In contrast, monthly short-horizon projections can be downward-sloping in states of moderate market uncertainty. More informative economic assumptions may also strengthen the U-shape of the MVP. For instance, the requirement that variance Sharpe ratios be at least as large as market Sharpe ratios (labeled as constraint "SR" below) already leads to consistently U-shaped projections at monthly horizons.

MVPs naturally embed potentially useful information about the time-varying structure of any probability belief that supports these projections. We find that the recovered moments of returns are highly time-varying and countercyclical. Annualized monthly and annual equity premia are less than 1% in periods of low and transitory volatility, but can be as large as about 22% and 10%, respectively, in turbulent markets, such as during the financial crisis of 2008. Recovered first and second moments have predictive power for future returns and second realized moments of returns. They also provide insights into the cyclicality of equity premia, variance risk premia, and highest attainable Sharpe ratios.

Tradeable minimum variance projections (MVPs) allow us to study the properties of optimal risk-return trade-offs under the minimum variance probability. Optimal delta-hedged option portfolios that short the MVP imply a countercyclical Sharpe ratio that is decreasing (increasing) with the investment horizon in periods of large (moderate) uncertainty. In phases of greater uncertainty, the optimal investment strategy is tilted to short out-of-the-money call option payoffs, in order to replicate the countercyclical U-shapedness of MVPs. We find that, overall, these optimal option portfolios have plausible market timing properties, moderate exposures to higher realized moments, and favorable out-of-sample Sharpe ratios when SR constraints are considered.

Our paper is related to different fields in the literature. The empirical properties of pricing kernels in index option markets have been considered by a large number of authors. Aït-Sahalia and Lo (1998), Jackwerth (2000), Rosenberg and Engle (2002), Aït-Sahalia and Duarte (2003), Chabi-Yo, Garcia, and Renault (2008), Chabi-Yo (2012), Christoffersen, Heston, and Jacobs (2013), and Song and Xiu (2016) among others estimate U-shaped pricing kernel projections using different parametric and nonparametric methods. A key finding of this research is that the model-implied kernel projection on the linear span of index and option returns is increasing (decreasing) in return states that correspond to positive out-of-the-money call (put) payoffs. U-shaped pricing kernels can be naturally motivated in economies with disagreeing investors. Bakshi and Madan (2008) obtain a pricing kernel that is a U-shaped

function of returns in a stylized static model where pessimistic investors are short the equity. In this setting, the increasing region of the kernel directly reflects the risk aversion of short equity investors who seek protection against upside market moves by purchasing out-of-the-money index calls. Taking a different view, Christoffersen, Heston, and Jacobs (2013) show that, in a reduced-form stochastic volatility model with a negative price of variance risk, the projection of the pricing kernel on returns is a quadratic function of returns. Our findings add new model-free and truly forward-looking evidence on the link between U-shaped pricing kernels and negative variance risk premia, which is independent of assumptions about underlying state dynamics or the conditioning information set. Our evidence shows that U-shapes are countercyclical and more likely to appear when the Sharpe ratios for selling market variance are larger than market Sharpe ratios. Our evidence also highlights the structure of optimal investment strategies supporting the U-shape, which are short a larger amount of variance in periods of heightened uncertainty.

Our approach is also related to the recent literature on pricing kernel recovery following the Ross (2015) recovery theorem. Motivated by the Hansen and Scheinkman (2009) long-term factorization of the pricing kernel, Borovička, Hansen, and Scheinkman (2016) emphasize that, in general, only a stationary ergodic belief incorporating long-run risk adjustments can be recovered when the kernel is path-dependent.[4] Jensen, Lando, and Pedersen (2017) prove a recovery theorem for parametric path-independent kernels on a finite state space with time-inhomogeneous transition probabilities. We also contribute to this literature by developing a model-free approach to forward-looking MVP recovery, which is implementable from mild and empirically verifiable economic assumptions. This approach avoids introducing other (explicit or implicit) high-level assumptions that might be difficult to test empirically, such as Markovianity, market completeness, or path-independence. In contrast to extant literature, we do not attempt to directly recover physical probabilities or to characterize the long-run structure of path-dependent pricing kernels. We instead recover a tradeable MVP that identifies the smallest maximal Sharpe ratio among all trading strategies involving nonlinear claims on the market.

Our work is also naturally related to the literature on moment bounds for asset returns. Hansen and Jagannathan (1991, HJ) bound the Sharpe ratio of any traded return using the minimal pricing kernel variance. Cochrane and Saa-Requejo (2000) consider minimum variance pricing kernels that exclude unrealistically large Sharpe ratios. Bansal and Lehmann (1997), Alvarez and Jermann (2005), and Backus, Chernov, and Zin (2014) obtain a direct bound on the expected log return of any traded asset using the entropy of the pricing

[4] Chabi-Yo, Bakshi, and Gao (2015) test the path-dependence assumption in the bond market, while Qin and Linetsky (2017) and Qin and Linetsky (2016) give extended recovery theorems to general recurrent Markovian processes and general semimartingale state dynamics, respectively. Christensen (2017) introduces a general nonparametric method for estimating the long-term factorization of the pricing kernel in dynamic Markov environments.

kernel.[5] All of these bounds depend on unobservable pricing kernel dispersion properties. Martin (2017) motivates a negative covariance condition (NCC) on returns and unobservable pricing kernels, which yields an observable lower bound for expected returns. We contribute to this literature with a broader set of NCCs that correspond to a wider and more flexible family of observable lower bounds on the moments of returns. We also contribute with a second set of observable upper bounds on the moments of returns, which follow from the economically and empirically motivated assumption of a negative risk premium for (generalized) variance risk. We show that these empirically supported bounds are informative about MVPs.

Finally, by employing tools and methods from the literature on the trading of realized variance and realized higher moments, we are able to compute the model-free upper and lower bounds for the moments of returns, which imply bounds for the variance of the MVP. We are also able to replicate the MVPs with appropriate option portfolios.[6]

The remainder of the paper is organized as follows. Section I motivates our approach starting from a simplified Black-Scholes economy, while Section II introduces the more general approach, which uses HJ pricing kernel projections. In Section III, we develop various economic hypotheses for the risk premia of tradeable realized moments of returns. We then explain the resulting upper and lower bounds for the conditional moments of returns. Section IV presents our empirical study of minimum variance pricing kernel recovery. Section V concludes. Finally, Appendix A contains proofs of the results in the main text, Appendix B reports implementation details, and Appendix C illustrates the properties of MVPs in exponentially affine jump-diffusion models.

I. Motivating (Almost) Model-Free Recovery

In this section, we develop in a simple Black-Scholes economy the intuition motivating our approach. Below, we provide extensions to general model-free settings that do not rely on particular assumptions for the distribution of asset returns under the pricing or physical probabilities.

Consider the forward price $F_{t,T}$ at time t for maturity T on the underlying S&P 500, where for simplicity maturity $T - t = 1$. In Black-Scholes's world, the log forward return follows a normal distribution under a fixed forward probability $\mathbb{Q}$,

$$r := \log(F_{T,T}/F_{t,T}) \overset{\mathbb{Q}}{\sim} N\left(-\frac{1}{2}\sigma^2, \sigma\right), \tag{1}$$

[5] Snow (1991) derives an additional set of bounds on the nonlinear moments of returns, which embed entropy bounds as limit cases and are nested in the tight minimum power divergence and arbitrage-free dispersion bounds in Almeida and Garcia (2017) and Orłowski, Sali, and Trojani (2015).

[6] The literature is too large to discuss in more detail here. Key contributions include Neuberger (1994), Carr and Madan (2001), Britten-Jones and Neuberger (2000), Carr and Lewis (2004), Lee (2008), Lee (2010), Neuberger (2012), Martin (2013), Kozhan, Neuberger, and Schneider (2013), Bondarenko (2014), and Andersen, Bondarenko, and Gonzalez-Perez (2015), among many others.

where σ is a constant volatility parameter. Broadly speaking, Ross (2015) pricing kernel recovery addresses identification of a unique physical probability $\mathbb{P}$, equivalent to $\mathbb{Q}$, by means of an underlying pricing kernel $\frac{d\mathbb{Q}}{d\mathbb{P}}$. Without additional requirements on the physical probability, it is apparent that arbitrarily many such probabilities (or pricing kernels) may exist. Indeed, any probability having the same zero-probability events as the forward probability would be one. Various recovery theorems in the literature focus on the conditions under which the physical probability is uniquely determined from the forward probability alone. They show that in general Ross (2015) recovery is possible for probabilities induced by a path-independent pricing kernel, when additional stability conditions on the physical transition dynamics are satisfied.[7] In contrast, the physical probabilities induced by path-dependent kernels cannot be recovered. Thus, additional information is generally needed to identify investors' beliefs from Arrow-Debreu prices.

The challenge in identifying physical probabilities from Arrow-Debreu prices appears in the Black-Scholes economy under a transition-independent pricing kernel parameterized by risk-aversion parameter $\gamma > 0$. Under this choice, physical probabilities are parameterized as

$$\frac{d\mathbb{P}}{d\mathbb{Q}}(r, \gamma) := \frac{e^{\gamma r}}{\mathbb{E}_t^{\mathbb{Q}}[e^{\gamma r}]} = e^{\gamma(r-(\gamma-1)\sigma^2/2)} \Leftrightarrow \frac{d\mathbb{Q}}{d\mathbb{P}}(r, \gamma) = e^{-\gamma(r-(\gamma-1)\sigma^2/2)}. \quad (2)$$

In this setting, knowledge of the physical probability is equivalent to identifying the marginal investor's risk aversion. Note that while functional form (2) is particularly convenient and intuitive, it precludes nonmonotonic kernels with U-shaped or locally concave return properties, which are both theoretically and empirically undesirable.[8] Moreover, while the physical probability induced by pricing kernel (2) is naturally supported in a continuous-time Black-Scholes model, no such probability can be identified from Arrow-Debreu price information alone; see Walden (2017).[9] In other words, even in the simplest Black-Scholes economy with transition-independent kernels, Ross (2015) recovery is impossible.

Given that it is impossible to identify physical probabilities without further information, we ask a different question: *Under relevant pricing constraints for the moments of returns, which minimum variance pricing kernel is consistent with mild economic assumptions on these moments?* Theoretically, this approach is motivated by the fact that minimum variance kernels provide a conservative lower bound on the maximally attainable Sharpe ratio of any economy satisfying the initial constraints; see again Hansen and Jagannathan (1991).

[7] As mentioned, these conditions are naturally nested by the requirement of recurrent physical state dynamics. See footnote 1 for details.

[8] For instance, Christoffersen, Heston, and Jacobs (2013) provide recent semiparametric evidence on U-shaped pricing kernel projections in the presence of priced volatility risk and negative variance risk premia.

[9] This result is an immediate consequence of the fact that, in a Black-Scholes model, no recurrent physical probability exists.

In a Black-Scholes model, the forward probability implied by (2) directly incorporates the pricing constraint on the first moment of asset returns ($\mathbb{E}_t^{\mathbb{Q}}[e^r] = 1$). A second pricing constraint on, for example, the second moment of returns can be introduced to fix volatility parameter σ. The minimum kernel variance probability then solves the optimization problem

$$\min_{\gamma \in \mathcal{R}} \mathbb{E}_t^{\gamma} \left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}(r, \gamma) \right)^2 \right], \tag{3}$$

where $\mathbb{E}_t^{\gamma}[\cdot]$ denotes physical expectations under $\mathbb{P}$ induced by kernel (2) and $\mathcal{R}$ corresponds to economic constraints on admissible probabilities.

An unconstrained optimization ($\mathcal{R} = [0, \infty)$) yields the solution $\gamma^\star = 0$, that is, the minimum variance probability equals the forward probability. This uninformative solution has obvious unappealing implications, such as a zero market equity premium and variance risk premium. Similar to the motivation underlying the good-deal bounds approach in Cochrane and Saa-Requejo (2000), we therefore introduce informative solutions using economically plausible lower bounds on the kernel variance. For instance, a lower bound $L > 1$ on expected gross returns yields an informative minimum variance kernel from the equivalence[10]

$$L \leq \mathbb{E}_t^{\gamma}[e^r] \iff \frac{\ln L}{\sigma^2} \leq \gamma. \tag{4}$$

Further plausible economic constraints can be incorporated, provided that the family of admissible kernels is flexible enough to accommodate them. For instance, one might require a negative variance risk premium from trading the log contract from Neuberger (1994). Unfortunately, no informative Black-Scholes kernel can satisfy this constraint, which highlights an important limitation of family (2) for pricing tradeable nonlinear market claims.

To overcome the limitations of the Black-Scholes and similar parametric assumptions, we extend our methodology to general model-free specifications of forward and physical probabilities. Our methodology is model-free in the sense that it initially restricts pricing and physical probabilities only by a model-free no-arbitrage condition, under which a finite set of moments of returns is exactly priced. Moreover, it ensures that minimum variance kernels are replicable using appropriate delta-hedged portfolios of European options. While being model-free, our approach naturally also applies to parametric models based on general stochastic volatility and jump specifications with path-dependent pricing kernels.[11]

[10] Under this bound, the solution is the pricing kernel (2) with risk aversion $\gamma^\star = \frac{\ln L}{\sigma^2}$.

[11] We demonstrate below how our approach performs in these settings. Testing our approach with flexible parametric models also helps us assess more precisely various constraints on risk premia used in our empirical study.

II. Minimum Variance Recovery

In this section, we introduce our approach to minimum variance pricing kernel recovery independent of model-based assumptions about the functional form of the pricing kernel or the structure of the underlying state dynamics.[12]

A. Model-Free Arbitrage-Free Markets

As in Section I, we denote by $F_{t,T}$ the time-t forward price of the S&P 500 index for delivery in $T \geq t$. The corresponding gross forward return is given by $R := F_{T,T}/F_{t,T}$. Given a set of tradeable option payoffs on R, we rely on the model-free fundamental theorem of asset pricing (FTAP) of Acciaio et al. (2016) to justify the existence of a forward probability $\mathbb{Q}$ on the positive real line, which identifies the forward price of market returns ($\mathbb{E}_t^{\mathbb{Q}}[R] = 1$) and the forward prices of R-measurable payoffs that are replicable by delta-hedged portfolios of traded options on R.[13] Relying on the model-free FTAP allows us to develop our theory without making tight assumptions on the structure of physical beliefs in the first place. We consider physical distributions $\mathbb{P}$ for returns on the positive real line, which are equivalent to $\mathbb{Q}$ on the information set generated by traded R-measurable payoffs, and focus on the problem of recovering $\mathbb{P}$ from $\mathbb{Q}$ by imposing mild economic assumptions on the structure of the normalized forward pricing kernel

$$\mathcal{M}_{\mathbb{P}}(R) := \frac{d\mathbb{Q}}{d\mathbb{P}}(R). \tag{5}$$

Denoting by $\mathbb{E}_t^{\mathbb{P}}[\cdot]$ ($\mathbb{E}_t^{\mathbb{Q}}[\cdot]$) expectations under $\mathbb{P}$ ($\mathbb{Q}$) conditional on time t information, $\mathcal{M}_{\mathbb{P}}(R)$ is such that

$$\mathbb{E}_t^{\mathbb{P}}\left[\mathcal{M}_{\mathbb{P}}(R) f(R)\right] = \mathbb{E}_t^{\mathbb{Q}}\left[f(R)\right] \tag{6}$$

for any traded payoff $f(R)$ that is replicable from a corresponding delta-hedged option portfolio.

B. Pricing Kernels Parameterized by HJ Projections

We optimally approximate admissible pricing kernels using their HJ projection on various tradeable realized moments of market returns. This approach achieves several goals at the same time. First, it allows us to approximate in a flexible way any admissible pricing kernel using information generated by returns alone. Second, it automatically builds in key information from

[12] As we avoid tight assumptions on underlying state dynamics, such as recurrence or ergodicity, we do not attempt to characterize minimum variance pricing kernels by means of a corresponding Hansen and Scheinkman (2009) long-run factorization.

[13] Under a model-free no arbitrage condition on the family of tradeable delta-hedged option payoffs, $\mathbb{Q}$ exists without additional assumptions on the physical data-generating process of returns, for example, under a complete option market or a bounded state space.

observable pricing constraints on various realized moments of returns. Third, it is completely described by a finite number of physical and forward-neutral moments. Fourth, it enables us to trade any admissible pricing kernel projection by means of delta-hedged option portfolios. Finally, it is also naturally applicable to model-based settings in which the pricing kernel may be a path-dependent function of returns and other sources of risk.

We denote the vector of the first J realized moments of returns by $\mathbf{R}_J := (R^0, R^1, \ldots, R^J)'$. The vector of forward prices of these payoffs is the vector $\mathbb{E}_t^{\mathbb{Q}}[\mathbf{R}_J]$ of forward moments of returns, which are observable through the prices of appropriate delta-hedged option portfolios. Consistent with this notation, $\mathbb{E}_t^{\mathbb{P}}[\mathbf{R}_J]$ is a vector of physical moments of returns. Different vectors of physical moments may correspond to different admissible physical probabilities and pricing kernels. We take the first J conditional forward moments as given and consider the family of HJ projections on $\mathbf{R}_J$ that exactly price power payoffs

$$\mathcal{M}_{\mathbb{P}}^J(R) := \arg\min_{\mathcal{M}_{\mathbb{P}}(R)} \mathbb{E}_t^{\mathbb{P}}[(\mathcal{M}_{\mathbb{P}}(R))^2], \tag{7}$$

subject to

$$\mathbb{E}_t^{\mathbb{P}}\left[\mathcal{M}_{\mathbb{P}}(R)\mathbf{R}_J\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J\right]. \tag{8}$$

Therefore, HJ projections are explicitly given by[14]

$$\mathcal{M}_{\mathbb{P}}^J(R) = \mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J'\right]\mathbb{E}_t^{\mathbb{P}}\left[\mathbf{R}_J\mathbf{R}_J'\right]^{-1}\mathbf{R}_J =: \mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J'\right]H_0^{-1}(\mathbb{P})\mathbf{R}_J, \tag{9}$$

with the following $(J+1)\times(J+1)$ *Hankel* matrix of moments to order $2J$

$$H_0(\mathbb{P}) := \mathbb{E}_t^{\mathbb{P}}\left[\mathbf{R}_J\mathbf{R}_J'\right] = \begin{pmatrix} 1 & \mathbb{E}_t^{\mathbb{P}}[R] & \cdots & \mathbb{E}_t^{\mathbb{P}}[R^J] \\ \mathbb{E}_t^{\mathbb{P}}[R] & \mathbb{E}_t^{\mathbb{P}}[R^2] & \cdots & \mathbb{E}_t^{\mathbb{P}}[R^{J+1}] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}_t^{\mathbb{P}}[R^J] & \mathbb{E}_t^{\mathbb{P}}[R^{J+1}] & \cdots & \mathbb{E}_t^{\mathbb{P}}[R^{2J}] \end{pmatrix}. \tag{10}$$

Similarly, the HJ projection's variance equals the conditional HJ distance

$$\mathbb{D}_t^J(\mathbb{P}) := \left(\mathbb{E}_t^{\mathbb{P}}\left[\mathbf{R}_J\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J\right]\right)' H_0^{-1}(\mathbb{P})\left(\mathbb{E}_t^{\mathbb{P}}\left[\mathbf{R}_J\right] - \mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J\right]\right). \tag{11}$$

It follows that different HJ projections correspond to different Hankel matrices of physical moments $H_0(\mathbb{P})$ and resulting HJ distances. Importantly, HJ projections in equation (9) are completely determined by the first J forward

[14] These finite-order projections are of the form

$$\mathcal{M}_{\mathbb{P}}^J(R) = a_0 + a_1R^1 + \cdots + a_JR^J.$$

Filipović, Mayerhofer, and Schneider (2013) discuss explicit conditions for the convergence of the sequence of finite-order projections to a well-defined infinite-order projection $\mathcal{M}_{\mathbb{P}}(R)$ as $J \to \infty$. A sufficient condition is existence of exponential moments, which is ensured, for instance, under a physical return distribution with bounded support.

moments of returns, which are fixed and observable from the prices of appropriate option portfolios, and by the unobservable first $2J$ physical moments of returns, which parsimoniously parameterize HJ projections in our setting. In our empirical study, we focus on HJ projections of order $J = 3$, as they are flexible enough to capture pricing kernel nonlinearities of various forms, including U-shaped or locally concave return dependencies. This flexibility is important, for example, to incorporate information from path-dependent pricing kernels with priced volatility risks.

Model-based examples of path-dependent pricing kernels with priced volatility risk are introduced in Appendix C, where we rely on Duffie, Pan, and Singleton (2000) and Song and Xiu (2016) specifications of forward probabilities in double-jump diffusion settings. Using the exponential affine pricing kernels in equation (C4), path dependence arises through the kernel dependence on the whole path of the volatility factors. Figures IA.10 and IA.11 in the Internet Appendix[15] illustrate the various shapes of HJ kernel projections produced by these models, which can be monotonic, U-shaped, or inversely U-shaped. In particular, HJ projections in Figure IA.10F can be pronouncedly U-shaped when volatility has a negative risk premium, consistent with the evidence in Christoffersen, Heston, and Jacobs (2013). This key information is lost for HJ projections that do not sufficiently incorporate the pricing characteristics of higher moments of returns, for example, linear CAPM-type projections.

C. HJ Projections and Supporting Probabilities

We make use of $(J + 1) \times (J + 1)$ Hankel matrices to parameterize admissible HJ projections in our minimum variance recovery approach. To this end, we need to ensure that these matrices can be interpreted as matrices of moments of a probability distribution on the positive real line. The question of whether a finite set of $2J$ numbers corresponds to moments of a probability distribution is called the *truncated moment problem* in the literature. The solution to this problem, which is provided by Curto and Fialko (1991, theorem 5.1) and detailed in Appendix A, relies on various conditions. Among these, a key requirement is that any admissible Hankel matrix H_0 be invertible, which is essentially the requirement of a nondegenerate HJ projection (9) implied by this matrix. Beyond existence of a solution of the truncated moment problem, a second important requirement concerns uniqueness. Indeed, without additional requirements, HJ projections can be supported by several probabilities on the positive real line. However, among these probabilities, there always exists a minimal supporting probability d_J with the smallest moment of order $2J + 1$. This probability satisfies the conditions detailed in Appendix A and is defined on a minimal discrete state space $\{x_0, x_1, \ldots, x_J\}$. In this way, we obtain a powerful parameterization of Hankel matrices that greatly improves

[15] The Internet Appendix is available in the online version of the article on the *Journal of Finance* website.

the characterization and computation of minimum variance HJ projections in our approach.[16]

PROPOSITION 1 (Vandermonde factorization of Hankel moment matrix):

(i) Let H_0 be a $(J+1)\times(J+1)$ positive definite Hankel matrix. Under the conditions given in Appendix A, matrix H_0 satisfies the Vandermonde factorization

$$H_0 = VDV' =: H_0(d_J), \tag{12}$$

where $D := diag(d_J)$ is a diagonal matrix of strictly positive probabilities d_J on strictly positive discrete support $\{x_0, x_1, \ldots, x_J\}$ and

$$V := \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_0 & x_1 & \ldots & x_J \\ \vdots & \vdots & & \vdots \\ x_0^J & x_1^J & \ldots & x_J^J \end{bmatrix} \tag{13}$$

is the transpose of a Vandermonde matrix.

(ii) Probability d_J is given explicitly by $d_J = V^{-1} H_0 e_1$, where e_1 is the first unit vector in $\mathbb{R}^{J+1}$.

(iii) There is a unique minimal probability d_J such that (12) holds and matrix H_1 defined in equation (A2) of Appendix A is singular.

Proposition 1 offers a direct model-free parameterization of HJ projections supported by physical probabilities d_J and corresponding matrices of moments $H_0(d_J)$. This is a consequence of the particular structure of the Vandermonde factorization (12), which is illustrated in the next example.

EXAMPLE 1 (Linear Projection): For $J = 1$ and $d_J = d_1 := (d_{01}, d_{11})'$, Vandermonde factorization (12) yields

$$H_0(d_1) = \begin{pmatrix} d_{01} + d_{11} & d_{01}x_0^1 + d_{11}x_1^1 \\ d_{01}x_0^1 + d_{11}x_1^1 & d_{01}x_0^2 + d_{11}x_1^2 \end{pmatrix},$$

where $d_{01}, d_{11} > 0$ and $x_0, x_1 > 0$. The first diagonal element in these matrices corresponds to the requirement that probabilities sum to 1. The other matrix components correspond to definitions of first and second moments using probabilities d_{01}, d_{11} on domain $\{x_0, x_1\}$. In summary, the discrete probability d_1 supports the moment matrix $H_0(d_1)$. Proposition 1 (iii) then characterizes a unique such supporting probability with minimal moment of order $2J+1$.

[16] Working with positive state space values $x_0, x_1, \ldots, x_J$ and probability vector d_J in factorization (12), instead of working directly with moment matrix $\mathbb{E}_t^{\mathbb{P}}[\mathbf{R}_J \mathbf{R}_J']$, ensures positivity of matrix H_0, and thereby greatly facilitates the implementation of our approach.

From Proposition 1, HJ projections (9) and HJ distances (11) supported by corresponding probabilities are naturally represented as

$$\mathcal{M}_{d_J}(R) = \mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J'\right] V^{-1'} D^{-1} V^{-1} \mathbf{R}_J, \tag{14}$$

$$\mathbb{D}(d_J) = \left(d_J - V^{-1}\mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J\right]\right)' D^{-1}\left(d_J - V^{-1}\mathbb{E}_t^{\mathbb{Q}}\left[\mathbf{R}_J\right]\right). \tag{15}$$

These parameterizations are particularly convenient for determining MVPs in the presence of various nonlinear constraints, such as the positivity constraint on matrix H_0 in equation (10).[17]

D. Minimum Variance Pricing Kernel Recovery

Given the family of HJ projections induced by the prices of a finite set of moments of returns, we recover the one with the globally smallest variance, which we call the MVP. As in the example in Section I, without additional constraints, this approach yields an uninformative solution $\mathbb{P} = \mathbb{Q}$ and an HJ distance of zero in equation (11). Such a solution is incompatible with the existence of equity and variance risk premia. To obtain an informative MVP solution, we motivate mild economic moment constraints in Section III, which restrict the admissible projections and probabilities in Proposition 1.

Further constraints are also easily incorporated. For instance, we may tighten the set of kernel projections by means of additional pricing constraints for the forward moments of order higher than J. Alternatively, we may set the Sharpe ratio of particular trading instruments, such as (short) variance swaps, to be larger than the Sharpe ratio of a market investment.

For brevity, denote by $\mathcal{R}_J$ the set of probabilities in Proposition 1 that are restricted by relevant constraints. The global MVP is then obtained from the solution of the following minimization problem over all admissible HJ projections

$$d_J^\star := \arg \inf_{d_J \in \mathcal{R}_J} \mathbb{D}(d_J). \tag{16}$$

Importantly, whenever $\mathcal{R}_J$ is characterized by observable forward moments alone, $d_J^\star$ can be computed without further assumptions on the time-series properties of returns. All moment constraints in Section III depend exclusively on the observable prices of a finite set of moments of returns.

REMARK 1: *Economic constraints restricting admissible pricing kernels are adopted earlier in prior good-deal bounds literature. Cochrane and Saa-Requejo (2000) take as given an initial physical probability and consider a family of pricing kernels that price exactly a subset of returns. They then select the forward probability that implies the cheapest price of a focus payoff, subject to an upper bound on the maximal Sharpe ratio attainable. In contrast, we*

[17] In Appendix B, we collect additional properties of Vandermonde factorizations and Vandermonde matrices. We also provide a more detailed description of the numerical implementation of our method.

take as given an initial forward probability and parameterize the family of MVPs that price a given set of realized moments of returns. Using economic constraints that induce a lower bound on the pricing kernel variance, we select the HJ projection with the lowest maximal Sharpe ratio. In doing so, we crucially exploit the properties of HJ projections on various realized moments of returns and the parameterizations offered by Proposition 1.

III. Constraints on Physical Moments

In this section, we introduce various constraints on the physical moments of returns, which are helpful to restrict admissible MVPs. These constraints follow from economically motivated and empirically verifiable assumptions about the risk premia of tradeable nonlinear payoffs.

A. Tradeable Payoffs

The set of traded nonlinear payoffs depends on the structure of the underlying option market. In complete option markets, Carr and Madan (2001) show that sufficiently smooth payoff functions f can be replicated by a delta-hedged portfolio with weight $f''(x)$ in out-of-the-money options of moneyness x,[18]

$$\begin{aligned}\mathcal{D}f(R) &:= f(R) - f(1) - f'(1)(R-1)\\ &= \int_0^1 f''(x)(x-R)^+dx + \int_1^\infty f''(x)(R-x)^+dx. \end{aligned} \tag{17}$$

It follows that $f(R)$ is replicated using a risk-free investment, a linear investment in the forward index market, and an option portfolio with payoff $\mathcal{D}f(R)$. For empirical purposes, the forward price of this payoff can therefore be computed from the forward prices of available out-of-the-money options.

Our methodology is based on convex payoff functions $f^p(R) := \frac{R^p-1}{p(p-1)}$ indexed by a single parameter $p \in \mathbb{R}\backslash\{1\}$. These functions give rise to the delta-hedged power payoffs $D^p(R) := \mathcal{D}f^p(R)$ and allow us to trade various moments of returns.[19] We apply these functions both to returns and to n^{th} powers of returns, in which case we use the notation $D^p(R^n) := \mathcal{D}f^p(R^n)$.

REMARK 2: *In complete option markets, the properties of payoffs $D^p(R^n)$ have been characterized in Schneider and Trojani (2018). An important aspect of our methodology is that it is applicable without essential modifications in incomplete option markets, when the set of option strikes is discrete and bounded by a threshold l (u) from below (above). In such settings, we make use of localized (corridor) versions $D^p_{[l,u]}(R^n)$ of a delta-hedged power payoff, which*

[18] Function f should be almost-everywhere twice continuously differentiable, with first and second derivatives f' and f''.

[19] The forward-neutral moments of returns directly follow as a simple transformation of the prices of these payoffs: $\mathbb{E}_t^{\mathbb{Q}}[R^p] = p(p-1)\mathbb{E}_t^{\mathbb{Q}}[D^p(R)] + 1$.

are replicated using the option portfolios in Schneider and Trojani (2014).[20] *Therefore, while we preserve below the unannotated version* $D^p(R^n)$ *for lighter notation, our results hold for corridor payoffs.*

B. Upper Bounds on Physical Moments

For any positive power n, the convex payoff $D^p(R^n)$ can be interpreted as a generalized variance payoff of an increasing transformation of returns. The fact that variance payoffs are widely used hedging instruments against market downturns motivates the following negative risk premium condition.[21]

DEFINITION 1 (Negative divergence premium, NDP): *The negative divergence premium condition* $NDP(p, n)$ *holds if*

$$-Cov_t^{\mathbb{P}}\left[\mathcal{M}_{\mathbb{P}}(R), D^p(R^n)\right] = \mathbb{E}_t^{\mathbb{P}}[D^p(R^n)] - \mathbb{E}_t^{\mathbb{Q}}[D^p(R^n)] \leq 0. \tag{18}$$

As emphasized in Section I, the Black-Scholes model is incompatible with a negative variance risk premium, that is, condition $NDP(p, 1)$ is always violated in that model. However, the evidence in the literature consistently supports $NDP(p, 1)$ conditions. The key implication of NDP conditions for our purposes is that they imply observable constraints on the physical moments of asset returns. This property directly follows from the convexity of payoff $D^p(R)$ and Jensen's inequality.

PROPOSITION 2 (Upper bound on physical moments): *If condition* $NDP(p, n)$ *holds, then*

$$\mathbb{E}_t^{\mathbb{Q}}[D^p(R^n)] \geq D^p\left(\mathbb{E}_t^{\mathbb{P}}[R^n]\right). \tag{19}$$

Proposition 2 gives rise to an upper bound $\overline{\mathbb{E}}_t[R^n](p)$ of the n^{th} physical moment of returns, which is simply the largest moment $\mathbb{E}_t^{\mathbb{P}}[R^n]$ satisfying inequality (19). This upper bound is a function of the forward moment $\mathbb{E}_t^{\mathbb{Q}}[D^p(R^n)]$. Figure IA.1 of the Internet Appendix illustrates the determination of upper bound $\overline{\mathbb{E}}_t[R^n](p)$ for $n = 1$. Figure IA.1A highlights the relation between parameter p and the tightness of the bound, which in our setting is typically decreasing in parameter p. Note that, for trading purposes, payoff $D^p(R^n)$ is replicable by a portfolio of payoffs $D^{pn}(R)$ and $D^p(R)$ as follows:

$$D^p(R^n) = \frac{n}{p-1}\left[(np-1)D^{pn}(R) - (n-1)D^n(R)\right]. \tag{20}$$

[20] In Appendix A, we provide details on the practical implementation of our methodology in incomplete option markets.

[21] A vast literature documents that the market variance risk premium carries a negative sign. Carr and Wu (2009) are among the first to address the subject empirically, reporting large Sharpe ratios close to one in the S&P 500 market.

From its tradeability, moment bound (19) is directly testable empirically. In Section IV.B, we show that $NDP(p, n)$ conditions are empirically strongly supported for various positive powers n. Therefore, pricing kernels with plausible empirical implications need to be supported by physical probabilities consistent with NDP conditions.

C. Lower Bounds on Physical Moments of Returns

In this section, we motivate a new family of economic lower bounds on the physical moments of returns, which we obtain from NCCs.

DEFINITION 2 (NCC): *For given $p, n \in \mathbb{R}$, the negative covariance condition $NCC(p, n)$ of order (p, n) holds if*

$$Cov_t^{\mathbb{P}}\left[\mathcal{M}_{\mathbb{P}}(R)R^p, R^n\right] = \mathbb{E}_t^{\mathbb{Q}}[R^{p+n}] - \mathbb{E}_t^{\mathbb{Q}}[R^p]\mathbb{E}_t^{\mathbb{P}}[R^n] \leq 0. \tag{21}$$

The economic content of NCC conditions is a lower bound on the physical moments of returns. As this bound depends only on forward moments, it can also be directly computed from the prices of delta-hedged option portfolios.

PROPOSITION 3 (Tradeable lower bound on physical moments): *If condition $NCC(p, n)$ holds for some $p \in [0, 1]$ such that $p + n > 1$, it follows that*

$$\mathbb{E}_t^{\mathbb{P}}[R^n] \geq \frac{\mathbb{E}_t^{\mathbb{Q}}[R^{p+n}]}{\mathbb{E}_t^{\mathbb{Q}}[R^p]} =: \underline{\mathbb{E}}_t[R^n](p). \tag{22}$$

For any $p \geq 0$ and $n > 0$, this condition yields a nontrivial lower bound for the n^{th} physical moment of returns, which can be computed from forward-neutral moments alone. For $p = 0$ and $n = 1$, the bound is equivalent to a positive equity premium, while for $p = 0$ and $n > 1$, it is equivalent to a positive higher moment risk premium. Positive moment risk premia are natural economic assumptions satisfied in benchmark economic models with risk-averse investors. The choice $p = n = 1$ yields a different lower bound on the equity premium, given by the second forward moment of returns. Martin (2017) justifies the latter bound theoretically under various economic conditions.[22] The broader family of covariance conditions in Definition 2 allows us to embed existing equity premium lower bounds and more flexibly constrain the moments of returns, in a way that can be made more consistent with both the empirical evidence and mild economic assumptions.

Lower bound (22) is always informative and increases in parameter p. For $n = 1$ and $p \to 0$, we obtain the weakest condition of a positive equity

[22] For instance, he shows that the lower bound holds in economies with (i) jointly log-normal pricing kernel and returns and market Sharpe ratios larger than market volatility, (ii) a representative agent with relative risk aversion greater than one who maximizes expected utility, and (iii) an Epstein-Zin representative agent with relative risk aversion and elasticity of intertemporal substitution both greater than one.

premium, while for $p \to 1$, we obtain the tightest equity premium lower bound based on the second forward moment. Similar features hold with respect to higher physical moments.

REMARK 3: *$NCC(0, n)$ condition for $n > 1$ implies a direct lower bound on the n^{th} moment of returns,*

$$\mathbb{E}_t^{\mathbb{P}}[R^n] \geq \underline{\mathbb{E}}_t[R^n](0) = \mathbb{E}_t^{\mathbb{Q}}[R^n]. \tag{23}$$

A useful consequence of this inequality is that the tightest first moment lower bound compatible with the $NCC(0, n)$ condition and moment monotonicity is[23]

$$\mathbb{E}_t^{\mathbb{P}}[R] \geq (\mathbb{E}_t^{\mathbb{Q}}[R^n])^{1/n}. \tag{24}$$

Bound (24) *is thus a natural one for constraining the admissible MVPs consistent with a mild $NCC(0, n)$ condition.*

Similar to Martin (2017), sufficient conditions for bound (22) in benchmark economies can be provided.[24] For instance, it is easy to see that in such economies, our extended *NCC* conditions hold for relative risk aversion parameters larger than p. Figure 1 plots the time-series dynamics of first-moment lower bounds resulting from $NCC(p, 1)$ ($p = 2/5, 1$) conditions and lower bound (24) for $n = 2$. The figure highlights large bound variability and commonality, together with strong countercyclicality. In addition, it clarifies quantitatively the trade-offs between tighter and weaker *NCC* conditions, as the $NCC(1, 1)$ lower bound is uniformly higher than all other bounds in the figure, usually by a large amount. In contrast, the $NCC(2/5, 1)$ bound is only marginally lower than bound (24).

While standard economic models may illustrate the validity of bounds in simple terms, it is important to recall that they typically rely on a monotonically decreasing pricing kernel. Therefore, in Section III.B, we provide direct empirical evidence on the validity of various lower bounds related to *NCC* conditions, using general reduced-form parametric models with flexible (possibly U-shaped) pricing kernels, as well as a fully nonparametric approach. We find that the $NCC(2/5, 1)$ condition is a robust assumption for our empirical analysis. In contrast, tighter $NCC(1, 1)$ conditions are violated quite often and more systematically.

Note that in, an arbitrage-free economy, equivalent physical and forward probabilities identify the probability and the forward price of an event conditional on a common information set. As our bounds depend only on forward moments, they are based on the same information set as the unknown conditional physical moments. However, it may be the case that the prices of certain payoffs are only weakly sensitive to certain information, or do not react to such information at all. In such cases, this information may be then

[23] For any probability $\mathbb{M}$ and $n_2 \geq n_1 > 0$, moment monotonicity corresponds to the moment inequality $(\mathbb{E}_t^{\mathbb{M}}[R^{n_2}])^{1/n_2} \geq (\mathbb{E}_t^{\mathbb{M}}[R^{n_1}])^{1/n_1}$.

[24] An example of such economies is the Black-Scholes setting in Section I.

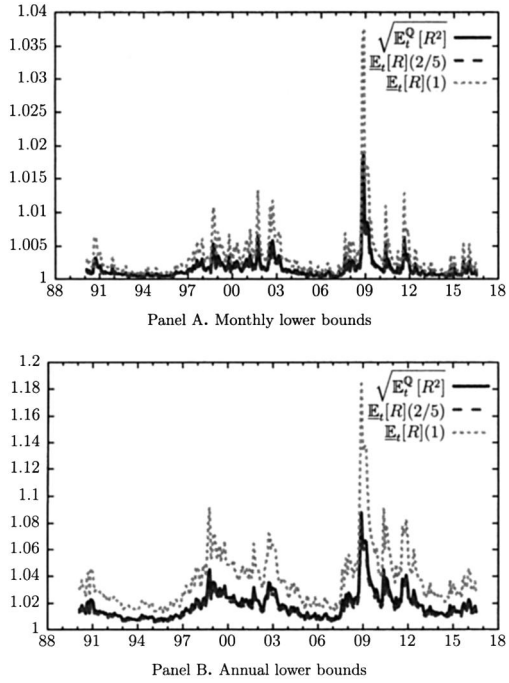


Figure 1. Various lower bounds related to $NCC(p, n)$ conditions. The figure plots the time series of various monthly (Panel A) and annual (Panel B) lower bounds for expected returns $\mathbb{E}_t[R]$, including the lower bound (22) for $NCC(2/5, 1)$ and $NCC(1, 1)$ conditions and the lower bound in equation (24) for $n = 2$. The data correspond to European options on the S&P 500 written between January 1990 and August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

revealed by the prices of higher order moments. For instance, the literature on unspanned stochastic volatility documents that state variables not present in bond yields may appear in the prices of bond options.[25] In our framework, we try to identify the information set shared by physical and forward-neutral conditional probabilities using various higher forward moments. Notably, the upper and lower bounds in Propositions 2 and 3 constrain the conditional physical moments of returns without making additional assumptions on the conditioning information set. Therefore, our setting is also compatible with situations in which some low-order conditional moments under the physical probability may be unspanned by the low-order conditional forward moments. In principle, our approach could also be extended to settings with multivariate HJ projections on the realized moments of the returns of various tradeable underlying assets, provided that option contracts on multivariate payoffs are available (for instance, quanto or spread options).

[25] We thank the Editor for highlighting this important point. Such unspanning features are well studied, for example, in the literature on unspanned volatility and risk premium features in government bond markets; see Collin-Dufresne and Goldstein (2002), Duffee (2011), and Joslin, Priebsch, and Singleton (2014), among others.

IV. Empirical Minimum Variance Recovery

We rely on a panel of S&P 500 European options, over the sample period from January 1990 to August 2016 to extract the forward moments of market returns from forward prices of delta-hedged option portfolios. The option data come from MarketDataExpress, a vendor connected to the Chicago Board Options Exchange (CBOE). The data set includes closing bid and ask quotes for each option contract, from which we compute mid prices along with the corresponding strike prices. We filter out all entries with nonstandard settlements and those that violate basic no-arbitrage conditions. As options mature on the third Friday of each month, we collect each third Friday of all months in our sample the forward prices of different option-replicating portfolios, corresponding to option maturities of 1, 2, 6, and 12 months.[26] In this way, we synthesize the observable information in the term structure of conditional moments of market forward returns. We replicate the time series of forward prices of the S&P 500 index using put call parity, consistent with the procedure in the CBOE (2009) white paper for the computation of the VIX index.

A. *(Almost) Model-Free Recovery*

For horizons of 1, 3, 6, and 12 months, we compute the MVP of order $J = 3$ from optimization problem (16) on each third Friday of the month between January 1990 and August 2016.[27] This sequence of optimizations is based on the set $\mathcal{R}_J$ first introduced in Section II.D, which constrains the physical conditional moments of returns as follows.

First, we impose a lower bound on conditional expected returns using the $NCC(2/5, 1)$ condition. Second, we impose a series of upper bounds on various conditional moments of returns using $NDP(k/n, n)$ conditions with parameters $k = 2, 3, \ldots, 6$ and $n = 1, 2, \ldots, 6$, such that $k \neq n$, so all resulting moment bounds depend on polynomial moments alone. Third, we impose $J = 3$ additional pricing constraints on HJ projections, to exactly match the forward moments of returns of orders $J + 1$ to $2J$. These last equality constraints facilitate the empirical identification of the Vandermonde factorization (12).

Our first set of findings is based on the set of moment restrictions described above, which are sufficient to obtain an informative MVP. The flexibility of our methodology allows us to further tighten the family of admissible MVPs also by means of nonlinear moment constraints that may otherwise be difficult to manage. To this end, we extend the set of constraints by additionally requiring the Sharpe ratio of (short) market variance payoffs to exceed the Sharpe ratio of market returns. This constraint provides additional insights into the properties of the optimal risk-return trade-off for trading market risk versus market

[26] To obtain a set of constant maturity options for maturities above one month, we apply standard interpolation schemes.

[27] A projection of order $J = 3$ depends on physical moments up to order 6. Projections of order $J = 4$ depend on eight physical moments, which makes computation of the inverse of Hankel matrix in equations (14) and (15) more difficult, even with the special treatment in Appendix B.

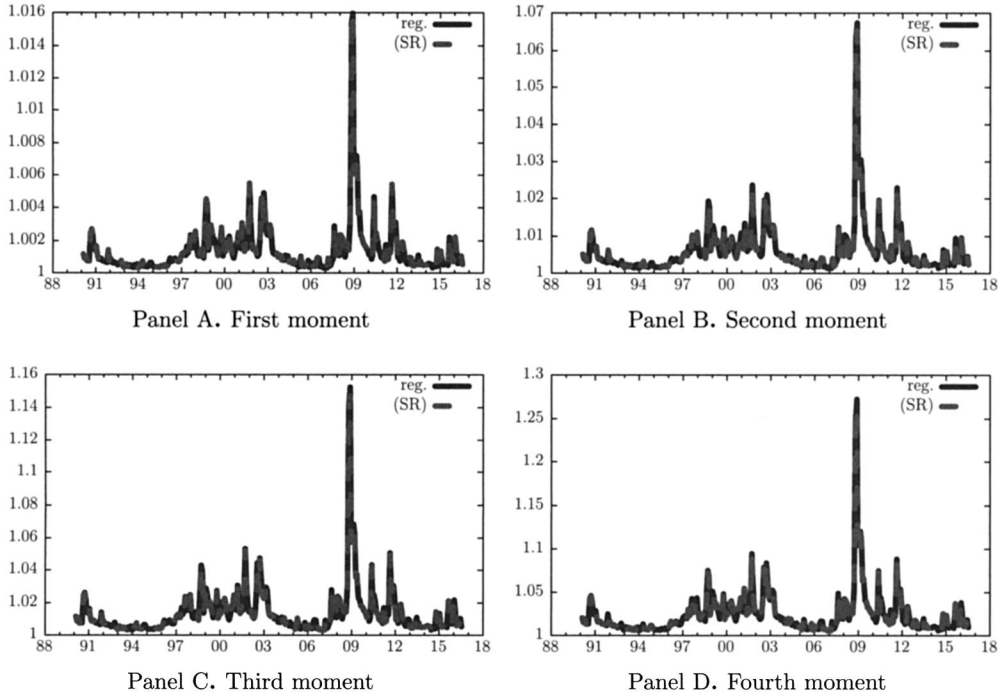


Figure 2. Recovered moments of returns (one-month maturity). The figure plots the first four monthly conditional moments of S&P 500 index gross returns, recovered from the solution of minimum variance optimization program (16) in Section II.D, using a projection of order $J = 3$ for a horizon of one month with ("SR") and without ("reg.") SR constraints. The data correspond to European options on the S&P 500 written between January 1990 and August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

variance risk. We show how this additional information affects the structure of MVPs. The results from this second empirical setting are denoted by "SR."

A.1. Moment Recovery

Figure 2 reports the time series of monthly conditional moments of market returns implied by MVPs. Recovered moments are consistently strictly below the upper bounds implied by our NDP conditions, often by a large amount, and they tend to reflect the dynamics of the lower bounds implied by condition $NCC(2/5, 1)$. While first moments are directly constrained by the lower bound, it is important to note that higher moments are as well because they are all supported by a probability distribution consistent with the statement of Proposition 1. Interestingly, the time series of first moments recovered with and without SR constraints are virtually identical. However, the higher moments recovered under the SR constraint are also similar and highly correlated with those obtained without the SR constraint; they are consistently lower,

Table I

Predictive Regressions of Realized Moments on Minimum Variance Conditional Moments

Panel A reports predictive regression coefficients and standard errors for a regression of the first two realized moments of returns onto the first two recovered minimum variance moments. Standard errors are reported in square brackets. Panel B reports similar predictive regression results using minimum variance conditional moments that also incorporate the SR constraint (the Sharpe ratio of trading variance exceeds the market Sharpe ratio). The data correspond to European options written on the S&P 500 from January 1990 to August 2016.

Panel A. Predictive Regression			Panel B. Predictive Regression (SR)		
First moments	$\hat{\alpha}[\sigma_{\hat{\alpha}}]$	$\hat{\beta}[\sigma_{\hat{\beta}}]$	First moments	$\hat{\alpha}[\sigma_{\hat{\alpha}}]$	$\hat{\beta}[\sigma_{\hat{\beta}}]$
1m	0.0047 [0.0035]	0.23 [2.98]	1m	0.0047 [0.0035]	0.23 [2.98]
3m	0.0096 [0.0085]	1.50 [1.94]	3m	0.0096 [0.0085]	1.50 [1.94]
6m	0.0032 [0.021]	3.36 [1.60]	6m	0.0032 [0.021]	3.36 [1.60]
12m	0.024 [0.058]	2.54 [2.22]	12m	0.024 [0.058]	2.54 [2.22]
Second moments			Second momente		
1m	0.0044 [0.0036]	0.19 [1.32]	1m	0.0044 [0.0037]	0.20 [1.39]
3m	0.0093 [0.0087]	0.69 [0.92]	3m	0.0091 [0.0088]	0.75 [0.98]
6m	0.0024 [0.021]	1.56 [0.81]	6m	0.0023 [0.022]	1.63 [0.87]
12m	0.024 [0.06]	1.16 [1.11]	12m	0.024 [0.061]	1.23 [1.20]

especially during periods of high uncertainty. This feature has implications for the recovered risk premia of higher order realized return payoffs.[28]

A.2. Predictability of Realized Moments

It is natural to ask whether the moments recovered by our approach have predictive power for future realized moments.[29] In Table I, we use a standard predictive regression to test whether the first and second recovered moments are unbiased predictors of future market returns and realized market variances, respectively. With the exception of the estimated intercepts for the six-month forecasting horizon, we cannot reject the null hypothesis that the regression intercept equals zero and the regression slope equals one, which is a necessary condition for the recovered conditional moments to be unbiased. As expected from the previous results, the predictive power is virtually identical for MVPs with and without SR constraints.

Given the large standard errors of the predictive regressions in Table I, we also investigate whether effective out-of-sample forecasting power results from the recovered first and second moments. Based on the Campbell and Thompson (2008) measure of predictive ability, we find in Table II, Panels A to C, that recovered moments have out-of-sample predictive power at all

[28] The lower recovered higher moments under the additional SR constraint are more apparent at longer horizons of 12 months. Figure IA.5 of the Internet Appendix plots the time series of the first four annual recovered moments of returns.

[29] Given the similarity of the moments recovered in Figure 2, we expect similar predictive power with and without SR constraints.

Table II
Predictions

Panel A reports the out-of-sample R^2s of recovered first and second moments of returns without SR (the Sharpe ratio of trading variance exceeds the market Sharpe ratio) constraints, relative to a sample mean forecast. Following Campbell and Thompson (2008), out-of-sample predictive R^2s are given by

$$R^2 := 1 - \frac{\sum_{i=2}^{n}(y_i - \hat{y}_{i-1})^2}{\sum_{i=2}^{n}(y_i - \bar{y}_{i-1})^2},$$

where y_i is the realized first or second moment of returns at time i, $\hat{y}_{i-1}$ is the recovered first (column "1") or second (column "2") conditional moment of y_i at time $i-1$, and $\bar{y}_{i-1}$ is the sample mean of the first or second realized moment of returns based on observations up to time $i-1$. We compute predictive R^2s for maturities from 1 to 12 months (rows "1m," "3m," "6m," "12m"). Panel B reports the estimated prediction bias $B = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_{i-1})$ implied by recovered first and second moments, along with 95% block-bootstrapped confidence intervals (in parentheses). Panel E reports (annualized) market sample Sharpe ratios and sample Sharpe ratios of the optimal strategy shorting the recovered MVPs. Panels C, D, and F report the same results corresponding to recovered kernels under the additional SR constraint that the Sharpe ratio of selling market variance is higher than the market Sharpe ratio. The data correspond to European options written on the S&P 500 from January 1990 to August 2016.

Panel A. Predictive R^2			Panel B. Sample Bias of Prediction		
	1	2		1	2
1m	0.0082	0.53	1m	0.65 (−4.61, 5.92)	−0.21 (−0.30, −0.098)
3m	0.031	0.27	3m	1.28 (−3.86, 6.47)	−0.10 (−0.28, 0.13)
6m	0.047	0.19	6m	1.29 (−3.57, 5.94)	−0.064 (−0.30, 0.26)
12m	0.098	0.11	12m	0.91 (−3.47, 5.04)	−0.032 (−0.32, 0.34)
Panel C. Predictive R^2 (SR)			Panel D. Sample Bias of Prediction (SR)		
	1	2		1	2
1m	0.0082	0.53	1m	0.64 (−4.64, 5.96)	−0.19 (−0.28, −0.078)
3m	0.031	0.28	3m	1.29 (−3.88, 6.38)	−0.074 (−0.26, 0.18)
6m	0.047	0.20	6m	1.29 (−3.62, 5.96)	−0.025 (−0.28, 0.31)
12m	0.098	0.14	12m	0.93 (−3.45, 5.05)	0.016 (−0.29, 0.42)

Panel E. Sample Sharpe Ratio			Panel F. Sample Sharpe Ratio (SR)		
	Kernel	Market		Kernel	Market
1m	0.39	0.39	1m	0.86	0.40
3m	0.46	0.42	3m	0.54	0.42
6m	0.42	0.39	6m	0.44	0.39
12m	0.39	0.39	12m	0.34	0.39

horizons, relative to simple forecasts implied by sample moments. Again, the forecasting power of recovered moments with and without SR constraints is similar. Consistent with intuition, given the mean-reverting properties of second conditional moments, we find that while the term structure of predictive R^2s is upward-sloping for returns, it is downward-sloping for realized variance.

A.3. Minimum Variance Sharpe Ratios

Table II, Panels E and F, report the sample market Sharpe ratios and the sample Sharpe ratios for shorting the MVP. Market Sharpe ratios are roughly 40% and largely independent of the investment horizon. For recovered MVPs not incorporating SR constraints, market Sharpe ratios are very similar to the Sharpe ratios from trading the MVP. This evidence follows from the observation that the MVPs implied by this set of economic constraints exhibit on average only relatively moderate nonlinearities, which are not sufficient to generate a substantially different unconditional mean variance trade-off when trading market returns or MVPs.[30] For recovered MVPs incorporating SR constraints, sample Sharpe ratios from optimally shorting the projection are clearly larger at horizons of one and three months, which gives rise to a downward-sloping term structure of optimal Sharpe ratios. This feature directly reflects the more pronounced nonlinearity of MVPs in this case, which derives from the fact that, under the tighter set of economic constraints, selling market variance is a more profitable trading strategy.

The term structures of conditional volatilities of MVPs, with and without SR constraints, are depicted in Figure 3 for horizons between 1 and 12 months. We find that both term structures are upward-sloping (downward sloping) during calm (turbulent) periods. All conditional volatilities are countercyclical and range on average between a minimum of about 4% in calm periods and a maximum of about 40% (32%) in periods of particular market distress whenever SR constraints are (are not) incorporated. MVP volatilities identify a tight HJ upper bound on the Sharpe ratio of any R-measurable power payoff of order $J = 3$, which is optimally tradeable by shorting the MVP itself (see Schneider (2015)). We therefore obtain countercyclical minimum variance optimal Sharpe ratios implying a term structure that is upward-sloping (downward-sloping) during calm periods (periods of distress). Minimum variance optimal Sharpe ratios also correspond to a lower bound for the optimal Sharpe ratio of any HJ projection that is consistent with the set of constraints $\mathcal{R}_J$ in optimization problem (16). In this sense, they define an economically motivated lower bound for the largest Sharpe ratio attainable in any of these economies under the corresponding economic constraints.

A.4. Equity and Variance Risk Premia

The conditional moments recovered by our approach imply time-varying market equity premia and variance risk premia at all horizons, as shown in Figure 4. Annualized equity premia range between a minimum of about 0.5% in calm periods and a maximum of about 20% (10%) at the monthly (annual) horizon during the financial crisis. The term structure of equity premia is steeply downward-sloping in phases of turbulent markets and moderately

[30] Figure IA.3 of the Internet Appendix plots average MVPs and shows that they exhibit relatively mild nonlinearities when SR constraints are not incorporated. In contrast, Figure IA.4 shows that average MVPs exhibit a pronounced U-shaped pattern when SR constraints are incorporated.

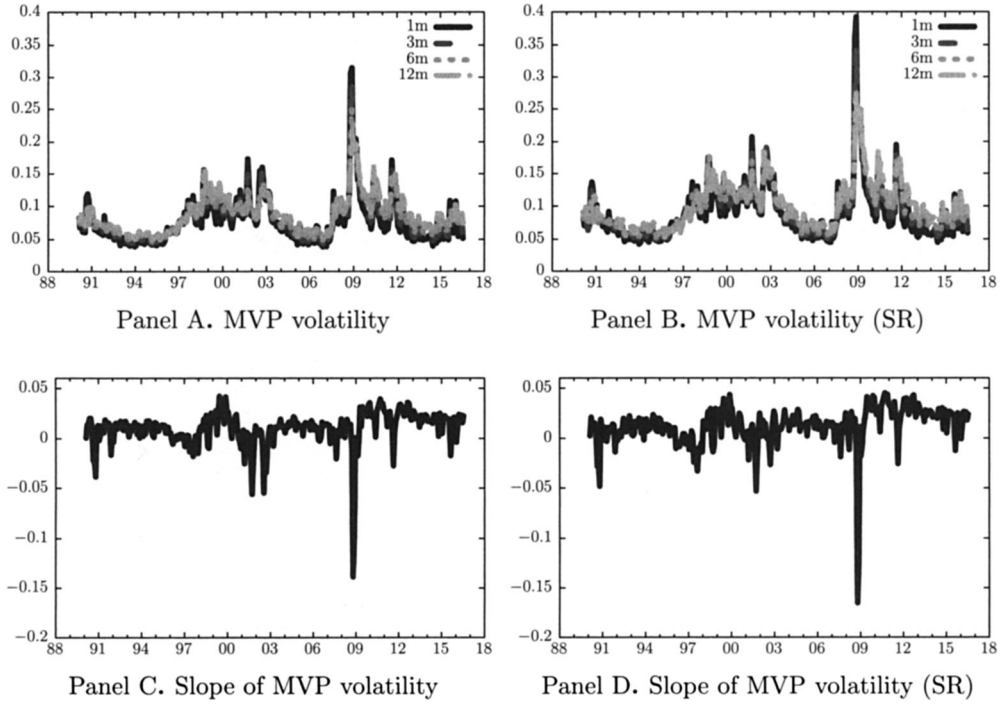


Figure 3. Minimum variance pricing kernel projection volatility. The figure plots the time series of (annualized) recovered conditional volatilities of the minimum variance pricing kernel projection on the first $J = 3$ realized moments of simply compounded S&P 500 index returns. The recovered projections solve the corresponding minimization problems (16). "Slope" is defined as the 12-month MVP volatility minus the annualized one-month minimum variance projection volatility. The data correspond to European options on the S&P 500 written between January 1990 and August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

upward-sloping in less volatile periods. Consistent with the findings above, the term structures recovered with and without SR constraints are virtually indistinguishable.

Variance risk premia are in all cases negative and highly time-varying. The comovement between variance risk premia with and without SR constraints is large. However, the SR constraint induces systematically larger variance risk premia in absolute value: variance risk premia at monthly (annual) horizons range between -5.3% (-2.6%) and -3.2% (-1.5%) when SR constraints are imposed and when they are not, respectively. These differences are consistent with the lower recovered second moments and with the more attractive (short) variance swap positions under the SR constraint. The term structures of variance risk premia are usually downward-sloping, but they can be strongly upward-sloping during infrequent periods of pronounced market distress. The inversions of the term structures of equity and variance risk premia are typically short-lived phenomena; see also Gruber, Tebaldi, and Trojani (2015) for related model-based evidence.

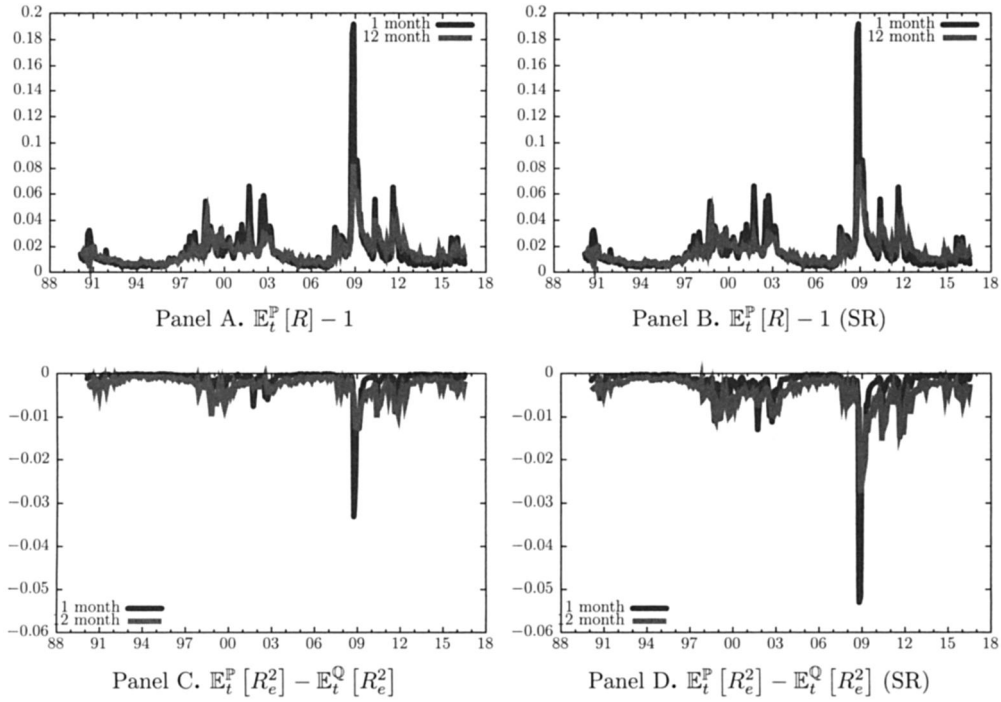


Figure 4. Minimum variance equity and variance risk premia. The figure plots the time series of equity and variance risk premia recovered from the solution of optimization program (16) in Section II.D, using a projection of order $J = 3$ for one- and two-month horizons, with ("SR") and without SR constraints (the Sharpe ratio of trading variance is higher than the market Sharpe ratio). $R_e = R - 1$ denotes for brevity the simply compounded forward return on the S&P 500 index. The data correspond to European options on the S&P 500 index written between January 1990 and August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

A.5. Supporting State Space

The time series of solutions to the minimum variance problem (16) corresponds to a time series of Hankel matrices with Vandermonde factorization $H_0 = V D V'$ in Proposition 1. This factorization directly identifies the probabilities d_J that support such solutions with representative state space $\{x_0, \ldots, x_J\}$. Figure 5 provides insights into the properties of the minimal support of these probabilities by plotting the time series of the lowest and largest elements x_0 and x_J of their state space for the MVP recovery problem without SR constraints.[31] For comparison, we also plot the time series of the minimally and maximally available option strikes in each month of our sample period.

The support of d_J is usually broader than the range of available option strikes and tends to be skewed toward negative returns, which is an indication of some degree of negative asymmetry in the supporting distribution d_J for

[31] The results for the MVP recovery problem with SR constraints are similar and depicted in Figure IA.6 of the Internet Appendix.

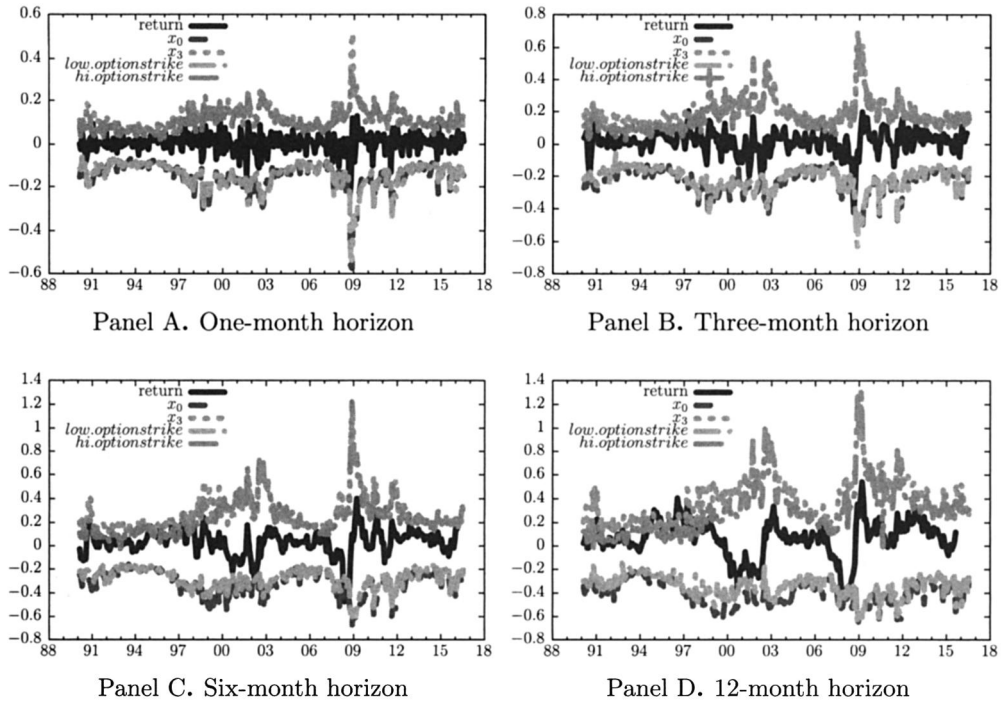


Figure 5. Ex-post realized returns and ex-ante minimal support range. The figure plots the time series of the lowest (x_0) and highest (x_J) element in the support of the recovered probability d_J in Vandermonde factorization (12), which is identified from the solution of minimization program (16) in Section II.D, for 1-, 3-, 6-, and 12-month horizons. All figures are based on a projection of order $J = 3$ and include for comparison the time series of the highest and lowest available option strikes, together with the time series of realized ex-post returns. In each graph, the ordinate is shifted by -1, to obtain a direct interpretation with simple excess returns. The data correspond to European options on the S&P 500 written between January 1990 and August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

returns.[32] We find that the support of d_J is quite informative about the range of possible future returns and more accurate than the range of available strikes. More specifically, we find that, at monthly horizons, the lowest (highest) option strike is violated ex-post by a realized return in 0 (10) out of 312 observations. In contrast, the bounds implied by the lowest and largest admissible values x_0 and x_J in the support of d_J are never violated. At longer horizons, some violations are sporadically observed, which are smaller and less frequent than those implied by the available strike range. Overall, this evidence shows that the model-free range $[x_0, x_J]$ is informative about the range of future possible values for market returns. This information is useful to gain more insight into the range of future states of returns that are empirically relevant for studying the properties of recovered MVPs.

[32] Support for future returns broader than the range of available option strikes is consistent with the positive prices of all out-of-the-money options used in our minimum variance kernel recovery.

A natural question that arises concerns how the recovered minimal range of admissible values $[x_0, x_J]$ depends on the availability of out-of-the-money options with corresponding strike prices. Figure IA.12 of the Internet Appendix plots the time series of the boundaries of the recovered state space under various assumptions on the availability of out-of-the-money option strikes, that is, when every month we drop the two and five most out-of-the-money call and put options from the sample. We find that the upper bound x_J is relatively insensitive to the choice of the most out-of-the-money options, with similar dynamics over time, even though the bounds implied by a smaller set of strikes are lower, especially in periods of high market uncertainty. The lower bound x_0 is more sensitive to the choice of the available range of option strikes than the upper bound. Indeed, while dropping the two most out-of-the-money call and put options produces similar bound dynamics as in the full sample, dropping the five most out-of-the-money options produces a time series of often implausibly high lower bounds, which are frequently violated ex-post in the data.

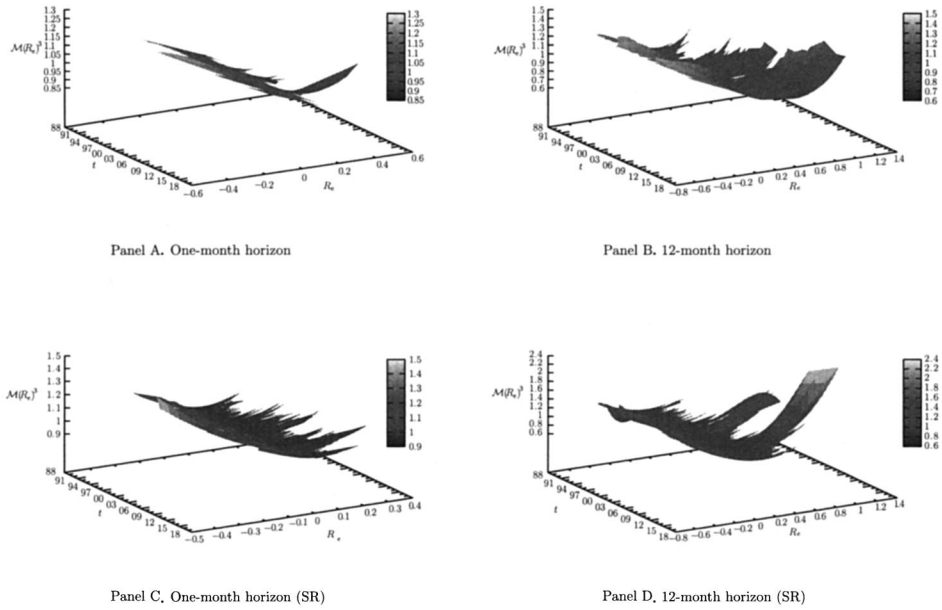
We can illustrate the information content of the minimal recovered support based on full or incomplete option strike information by counting the number of times that the support is violated ex-post by a realized return. Using all option strikes and the tighter SR constraint, we find only zero, zero, and two violations of the upper support boundary, for horizons of one, three, and six months, respectively, out of 312 data points. Similarly, we find only zero, two, and four violations of the lower support boundary. In contrast, when dropping the five most out-of-the-money call and put options, we obtain two, one, and one violations of the upper bound and 3, 9, and 11 violations of the lower bound. We therefore conclude that the information in out-of-the-money call and put options is useful to determine ex-ante an empirically relevant support for future returns using MVPs.

A.6. The Shape of Minimum Variance Projections

An important economic question concerns the shape of MVPs as a function of the market return. Since for $J \to \infty$, these projections converge to the unknown minimum variance kernel, we can gain insight into the pricing of distinct future states from the projection's functional form. Figure 6 plots the time series of MVPs on their minimal state space $[x_0, x_J]$ for horizons of 1 and 12 months.[33]

With respect to MVPs not incorporating SR constraints, we find that they are usually monotonically decreasing at monthly horizons, but also pronouncedly U-shaped in a number of economically important periods of heightened uncertainty. At longer horizons, the projections are consistently more U-shaped. We further find that SR constraints lead to a more distinctly U-shaped MVP at all horizons. The more pronounced U-shape is consistent with the incorporation of the larger Sharpe ratio for selling market volatility.

[33] Figure IA.7 of the Internet Appendix plots the time series of MVPs for horizons of three and six months.



Panel A. One-month horizon

Panel B. 12-month horizon

Panel C. One-month horizon (SR)

Panel D. 12-month horizon (SR)

Figure 6. Conditional minimum variance kernel projection (1- and 12-month horizon). The panels depict the times series of recovered minimum variance projections of order $J = 3$ for horizons 1 (Panel A) and 12 (Panel B) months, plotted as a function of simple returns $R^e = R - 1$. The bottom two panels correspond to a projection with the additional SR constraint that the Sharpe ratio on trading variance be higher than the Sharpe ratio of the forward equity premium, for horizons 1 (Panel C) and 12 (Panel D) months. Kernel projections are plotted over their recovered minimal support range $[x_0, x_J]$ in Vandermonde factorization (12). The data correspond to European options on the S&P 500 written between January 1990 and August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

It is useful to quantify and compare more systematically the degree of U-shapedness in MVPs. To this end, we measure the convexity of MVPs in regions of positive out-of-the-money call payoffs, using the convexity indices reported in Figure 7. Figure 7 shows that the projections' U-shapes are countercyclical, as they increase at all horizons when countercyclical market uncertainty, proxied by the one-month option implied volatility, also increases. In general, SR constraints produce more pronounced U-shapes at all horizons. When SR constraints are not incorporated, the convexity of MVPs at horizons of 3, 6, and 12 months is still apparent when uncertainty is not unusually low. In contrast, monthly projections are pronouncedly U-shaped only in periods of particularly high uncertainty.

In summary, our model-free evidence on the term structure of MVPs reveals a consistently countercyclical U-shape. At horizons above one month, the U-shape is a common and persistent feature of kernel projections. At shorter horizons, it appears in periods of high conditional uncertainty or if short variance positions can be assumed to imply a better mean variance trade-off than simple market investments.

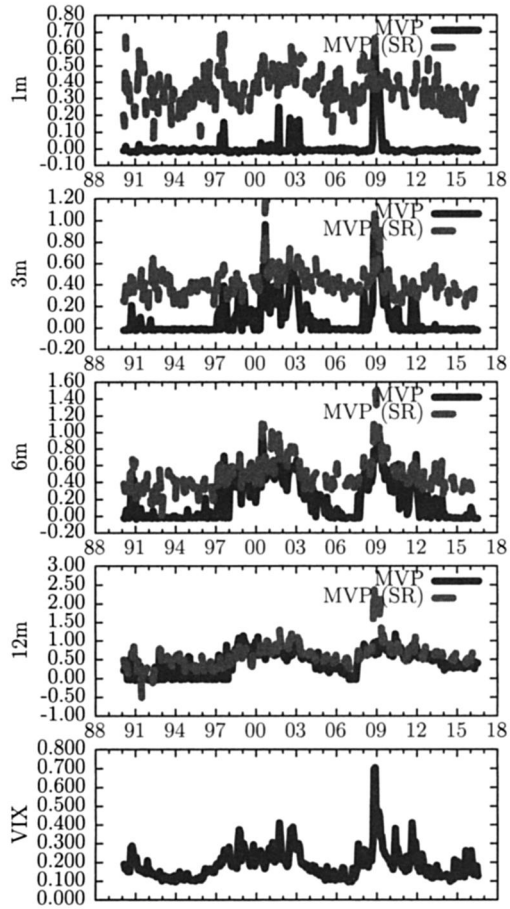


Figure 7. U-Shapedness of minimum variance pricing kernel projections. The figure plots the time series of Bregman divergences

$$D_{\mathcal{M}_{d_3^*}}(R) := \frac{\mathcal{M}_{d_3^*}(x_3) - \mathcal{M}_{d_3^*}(1) - \mathcal{M}'_{d_3^*}(1)(x_3 - 1)}{\mathcal{M}_{d_3^*}(1)(x_3 - 1)},$$

of minimum variance pricing kernel projections (MVPs) of order $J = 3$, for horizons of 1, 3, 6, and 12 months, together with the VIX-implied volatility index as a simple measure of countercyclical market uncertainty. Here, $D_{\mathcal{M}_{d_J^*}}(R) = 0$ ($D_{\mathcal{M}_{d_J^*}}(R) > 0$) when the MVP is linear (convex) over the support of positive returns. The constant $\mathcal{M}_{d_J^*}(1)(x_J - 1)$ normalizes the projection support and the level of $\mathcal{M}_{d_J^*}(R)$ in $R = 1$. $D_{\mathcal{M}_{d_J^*}}(R)$ measures the U-shapedness over support $[1, x_J]$, which corresponds to payoff regions of out-of-the-money calls. The upper value x_J in minimal support range $[x_0, x_J]$ is recovered from the Vandermonde factorization (12). In each panel, labels "MVP" correspond to solutions of optimization program (16) in Section II.D, using a projection of order $J = 3$, where the labels "MVP (SR)" correspond to solutions of the same optimization program with the additional SR constraints. Data correspond to S&P 500 index option prices written between January 1990 and January 2016. (Color figure can be viewed at wileyonlinelibrary.com)

A.7. *U*-Shape Dynamics and Countercyclical Uncertainty

To gain more insight into the dynamic relation between kernel projection features and countercyclical uncertainty, we next consider a number of important economic events in our sample. For brevity, we focus on projections not incorporating SR constraints.

Figure 8 plots the projections on the third Friday of the "crisis" months September 2002 and November 2008, together with those observed two months before and two months after these crisis months.[34] On November 2008, two months after the Lehman default and one day after a six-year low of the Dow Jones Industrial Average, projections at all horizons are U-shaped. At semiannual and annual horizons, the U-shape persists two months later and was at least partly anticipated two months earlier. At monthly horizons, the U-shape is not anticipated two months earlier and virtually disappears two months later. The term structure of pricing kernel projections is different in September 2002, when the Dow Jones Industrial Average hit a four-year low and the NASDAQ hit a six-year low, following the bursting of the dot.com bubble. While at semiannual and annual horizons, all of these projections are clearly U-shaped; monthly projections are monotonic and almost linear.[35]

Figure 9 shows that the projections on the third Friday of June 1994 and November 2006, two "conundrum" months of historically low and weakly persistent market volatility, are all monotonically downward-sloping.

A.8. *Trading the Minimum Variance Projection*

As MVPs are replicable with delta-hedged option portfolios, they are naturally interpreted as optimal payoffs of portfolios of realized moments. Here, a downward-sloping projection corresponds to an optimal directional claim that is along the market return. Analogously, a convex projection corresponds to a nonlinear payoff that is short some degree of market return dispersion. Therefore, discrepancies between optimal and market conditional Sharpe ratios are larger when MVPs are more U-shaped, that is, in periods of heightened uncertainty.

When SR constraints are not incorporated, optimal and market Sharpe ratios diverge essentially only at horizons above one month, usually by a small amount.[36] These Sharpe ratio differences correspond to optimal portfolios in Figure 10, Panel C, that usually invest less than 5% of total wealth in claims to the higher moments of returns. When SR constraints are incorporated, optimal

[34] In both periods, market volatility was historically very high.

[35] Evidence qualitatively similar to that for September 2002 is shown in Figure IA.2 of the Internet Appendix for the more recent "crisis" months of May 2010 and August 2011. Historically, these periods correspond to events that led to situations of particularly turbulent financial markets. For example, in May 2010, the Eurozone countries and the IMF agreed to a euro 110 billion loan for Greece, while on August 6, 2011, the credit rating of the United States was downgraded by S&P from AAA to AA+.

[36] Figure 10, Panel A of the Internet Appendix, provides detailed results. While annual optimal Sharpe ratios can be about 20% larger than market Sharpe ratios when uncertainty is high, monthly optimal Sharpe ratio improvements are never larger than 6%.

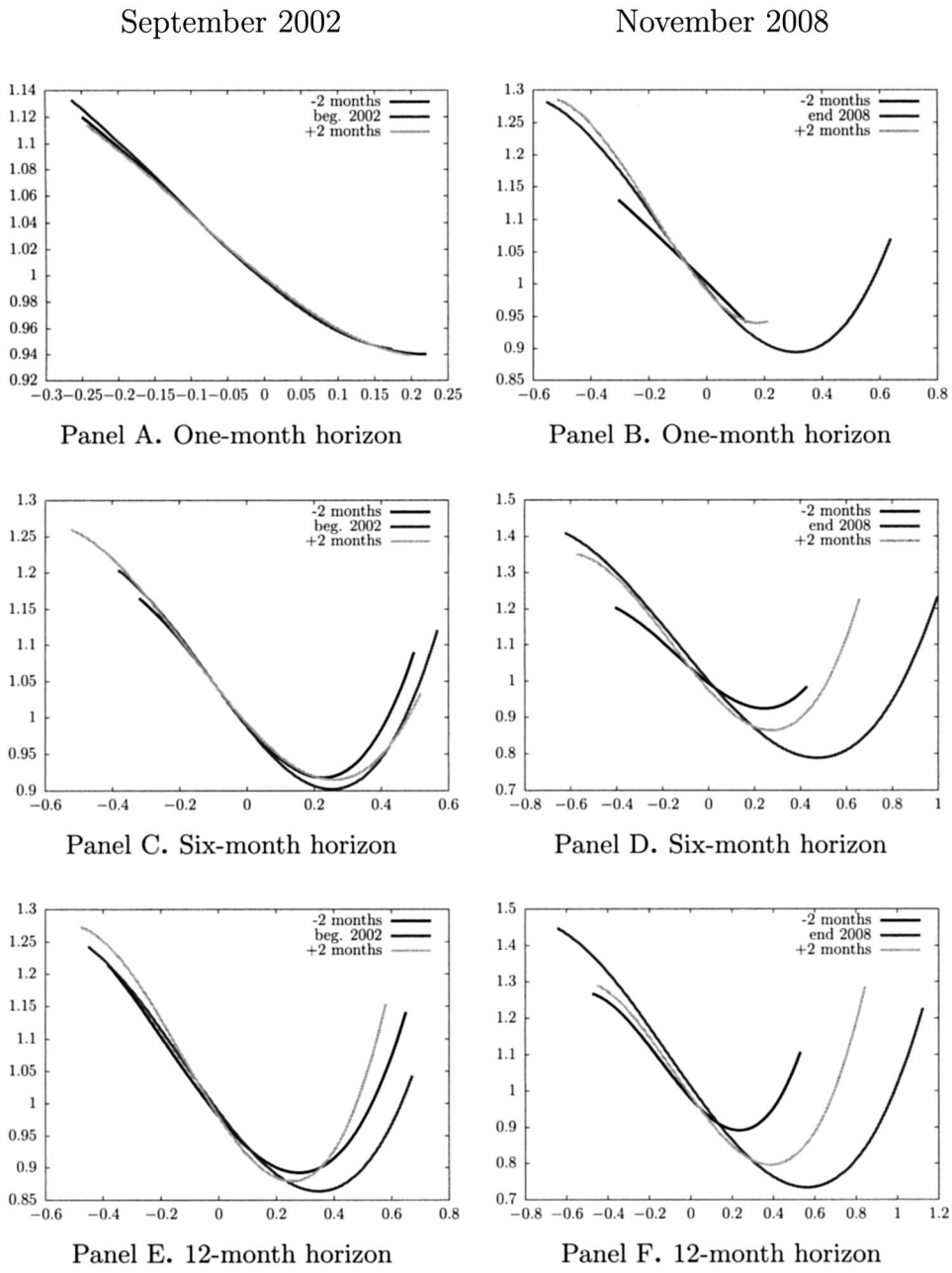


Figure 8. Shapes of minimum variance pricing kernel projections (crises). The figure plots minimum variance pricing kernel projections on the first $J = 3$ realized moments of simply compounded S&P 500 returns for horizons of 1, 6, and 12 months and in the months September 2002 and November 2008, two months in which the recovered pricing kernel projection volatility was particularly high. Kernel projections are plotted over their recovered minimal support range $[x_0, x_J]$ in Vandermonde factorization (12). The data correspond to European options on the S&P 500. (Color figure can be viewed at wileyonlinelibrary.com)

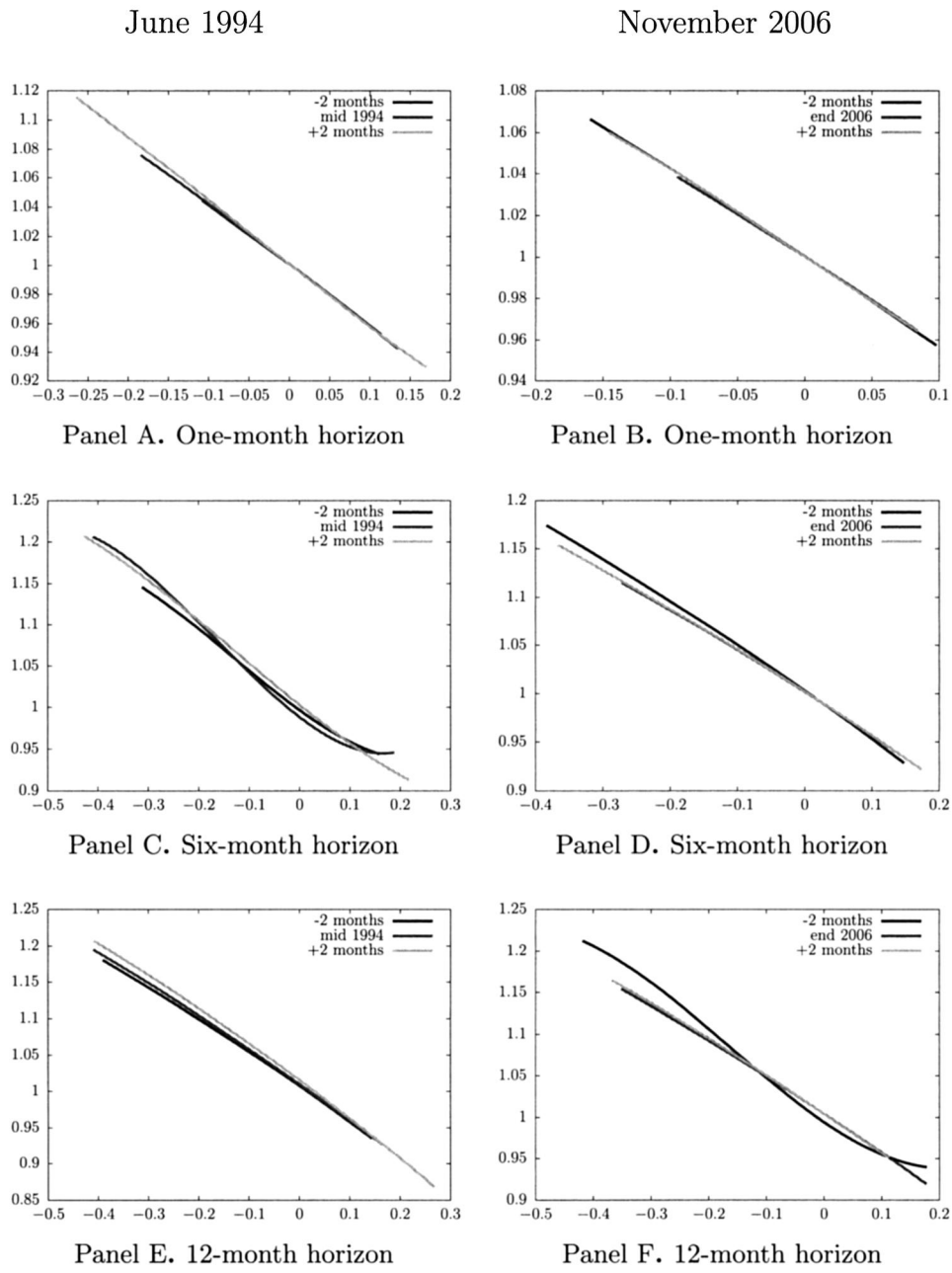


Figure 9. Shapes of minimum variance pricing kernel projections (conundrum). The figure plots minimum variance pricing kernel projections on the first $J = 3$ realized moments of simply compounded S&P 500 returns for horizons of 1, 6, and 12 months and in the months June 1994 and November 2006, two months in which the recovered pricing kernel projection volatility was particularly low. Kernel projections are plotted over their recovered minimal support range $[x_0, x_J]$ in Vandermonde factorization (12). The data correspond to European options on the S&P 500. (Color figure can be viewed at wileyonlinelibrary.com)

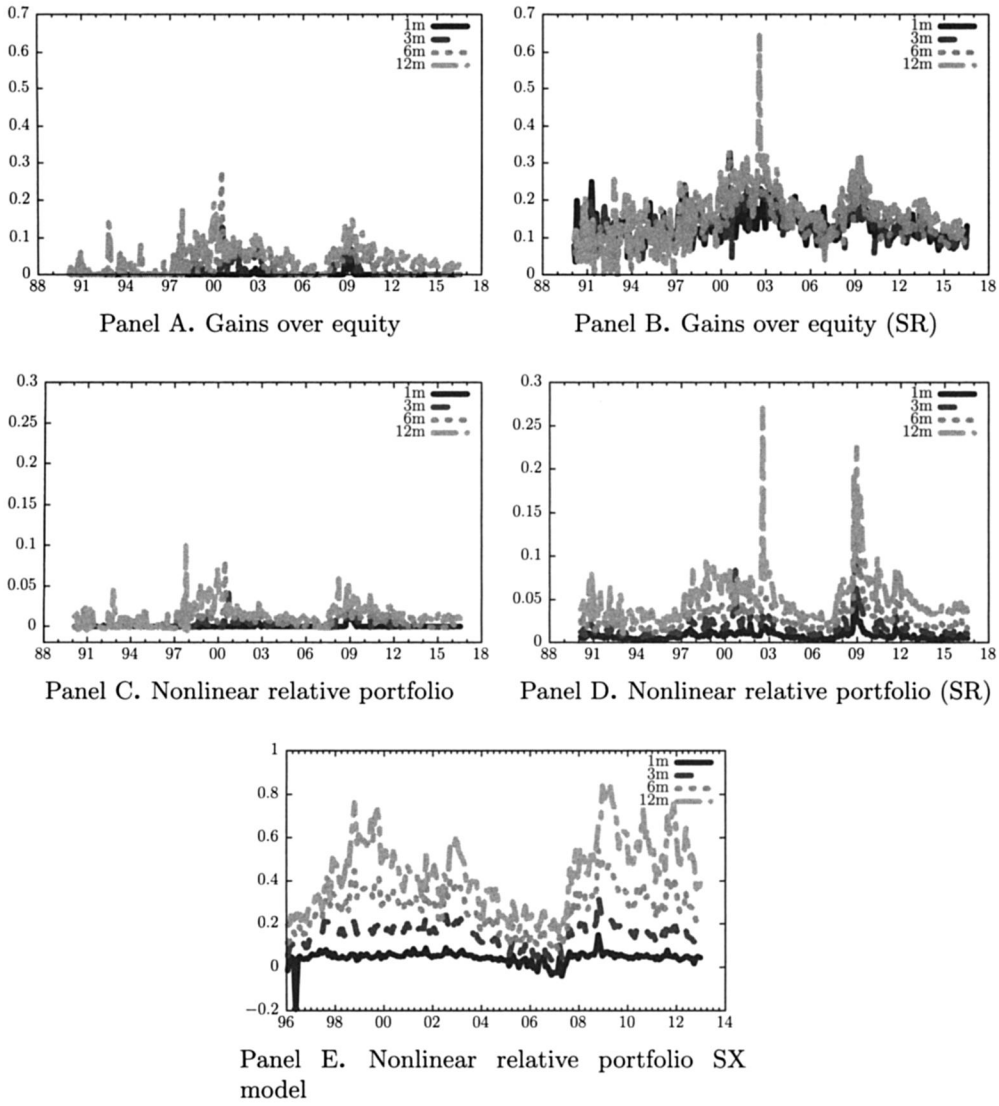


Figure 10. Sharpe ratio improvements of tradeable projections of order $J = 3$ over market Sharpe ratios. Panels A and B report for horizons of 1, 3, 6, and 12 months the time series of percentage increases of optimal recovered conditional Sharpe ratios, as a percent of recovered conditional market Sharpe ratios. Panels C to E report for the same horizons the time series of the fraction of total wealth invested by tradeable minimum variance kernel projections in nonlinear realized moments of order 2 and 3, without SR (the Sharpe ratio of trading variance exceeds the market Sharpe ratio) constraints, with SR constraints and with the model-based SR constraints in the Song and Xiu (2016) (SX) model, respectively. The data correspond to European options written on the S&P 500 from January 1990 to August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

monthly and quarterly Sharpe ratios are clearly larger than market Sharpe ratios.[37] In this case, differences are induced by optimal portfolios in Figure 10, Panel D, that invest a relatively small fraction of less than 10% of total wealth in claims to the higher realized moments of returns. At monthly horizons, optimal exposures to higher realized moments are typically less than 5%, but they produce large Sharpe ratio improvements due to the high profitability of monthly short variance investments under the assumption of SR constraints.

In summary, the countercyclical optimal risk-return trade-off of MVPs improves on conditional market Sharpe ratios in a horizon-dependent way. These improvements arise from optimal replicating portfolios that allocate a countercyclical moderate fraction of wealth to nonlinear moment exposures. Unconditionally, sample Sharpe ratios of monthly optimal projections incorporating SR constraints clearly improve on sample market Sharpe ratios.

A.9. Minimum Variance versus Model-Based Projections

It is natural to compare almost model-free MVPs with the pricing kernel projections of estimated flexible parametric models, such as the Song and Xiu (2016) two-factor double-jump stochastic volatility model. This specification incorporates conditioning information generated by two hidden volatility factors; see Appendix C for details. It is estimated from an observed panel of S&P 500 options, VIX options and S&P 500 index returns, over the sample period January 4, 1996 to December 31, 2012 for S&P 500 index options and July 1, 2007 to December 31, 2012 for VIX options. As shown in Song and Xiu (2016, Figure 10), the estimated model-based pricing kernel implies a U-shaped projection on returns consistent with the data. At the same time, it produces a more counterfactual upward-sloping projection on VIX, which may be evidence of a not easily detectable model misspecification.

In the Internet Appendix, we show that the model-based conditional HJ projections of order $J = 3$ are typically U-shaped at all horizons, usually more pronouncedly so than for the MVPs in Section A.6.[38] These projections correspond to model-based optimal portfolios that often invest large fractions of total wealth in claims to the higher realized moments of returns. These nonlinear exposures are depicted in Figure 10, Panel E, and may appear implausibly high in a number of cases, as, for instance, right after the Lehman collapse, when they imply investments in higher realized moments of about 10%, 20%, 40%, and 80% of total wealth for horizons between 1 and 12 months.

We find that these large model-based exposures to higher realized moments yield low sample Sharpe ratios ex-post. For instance, at horizons of 6 and 12 months, the optimal model-based sample Sharpe ratios of 0.36 and 0.28 are below the market Sharpe ratio of 0.39. For horizons of one and three months, the optimal model-based sample Sharpe ratios of 0.7 and 0.58 are higher than the market sample Sharpe ratios. In contrast, model-based monthly optimal

[37] Figure 10, Panel B of the Internet Appendix, presents these results.

[38] Complete results are shown in Figure IA.8 of the Internet Appendix.

Sharpe ratios are lower than both monthly optimal Sharpe ratios of model-free MVPs with SR constraints (Sharpe ratio of 0.74) and model-based MVPs with SR constraints (Sharpe ratio of 0.73).[39]

B. Testing the Conditional Moment Bounds

In Section IV.A, we report our empirical findings for MVPs of order $J = 3$. These findings follow from a combination of NDP and NCC conditions, along with different pricing constraints on the higher moments of market returns. The results show that these constraints are sufficient to imply an informative MVP. In this section, we conduct a battery of parametric and nonparametric tests of the conditional moment inequalities in Propositions 2 and 3 and show that our NCC and NDP constraints are empirically well supported.

B.1. Upper Bounds

We first check the inequalities in Proposition 2 unconditionally. Given parameters k/n and n in Section IV.A, this is a test of the moment inequality

$$\mathbb{E}^{\mathbb{P}}\left[D^{k/n}(R^n) - \mathbb{E}_t^{\mathbb{Q}}[D^{k/n}(R^n)]\right] \leq 0. \tag{25}$$

Panels B, C, and D of Figure 11 plot sample averages of monthly payoffs $D^{k/n}(R^n) - \mathbb{E}_t^{\mathbb{Q}}[D^{k/n}(R^n)]$ with 90% bootstrap confidence intervals and provide unambiguous support for moment inequality (25).

Model-free tests of conditional moment inequalities in Proposition 2 are more challenging because, even with nonparametric methods, one still has to specify the variables driving the conditioning information set. We estimate nonparametrically the time series of conditional risk premia $\mathbb{E}_t^{\mathbb{P}}[D^{k/n}(R^n)] - \mathbb{E}_t^{\mathbb{Q}}[D^{k/n}(R^n)]$ using the simple univariate specification

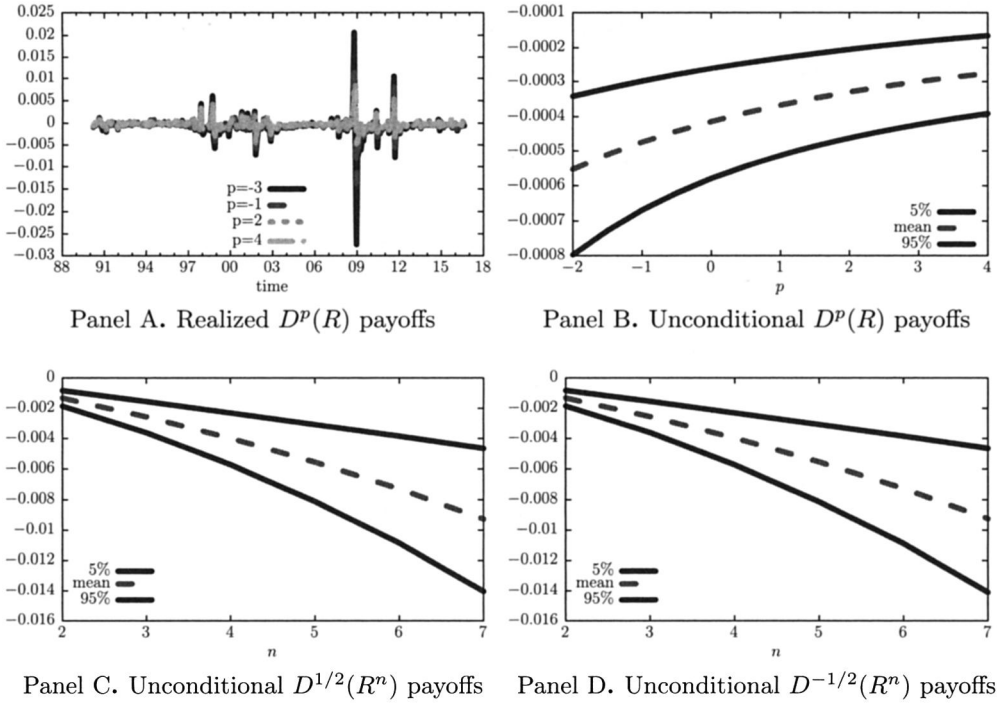
$$\mathbb{E}_t^{\mathbb{P}}[D^{k/n}(R^n)] := \mathbb{E}^{\mathbb{P}}\left[D^{k/n}(R^n) \mid \mathbb{E}_t^{\mathbb{Q}}[D^2(R^n)]\right] := f(\mathbb{E}_t^{\mathbb{Q}}[D^2(R^n)]). \tag{26}$$

This specification incorporates conditioning information generated by the forward price of simple variance swaps on R^n. Estimated nonparametric risk premia $\hat{f}(\mathbb{E}_t^{\mathbb{Q}}[D^2(R^n)]) - \mathbb{E}_t^{\mathbb{Q}}[D^{k/n}(R^n)]$ are shown in Figure 12, Panel B, and are all unambiguously negative, consistent with NDP conditions.

Finally, we test NDP conditions using the conditional transition densities of the two-factor double-jump stochastic volatility model estimated in Song and Xiu (2016).[40] Figure 13, Panels C and D, confirm that these conditions are also satisfied with respect to the model-based state dynamics.

[39] The conditional optimal portfolio allocations of model-free and model-based MVPs with SR constraints are also very similar. See Figure IA.16 of the Internet Appendix for details. These findings may be due to model misspecification, but also to the estimation approach used, which optimizes statistical fit without constraining the model-implied Sharpe ratio properties.

[40] See Example C3 in Appendix C for a detailed model description.



Panel A. Realized $D^p(R)$ payoffs

Panel B. Unconditional $D^p(R)$ payoffs

Panel C. Unconditional $D^{1/2}(R^n)$ payoffs

Panel D. Unconditional $D^{-1/2}(R^n)$ payoffs

Figure 11. Unconditional nonparametric tests of condition *NDP(p,n)* in Assumption 1. This figure plots monthly realized and average profits of trading strategies replicated by corresponding delta-hedged S&P 500 index option portfolio payoffs over the sample period from January 1990 to August 2016. Panel A plots the time series of realized divergence swap payoffs $D^p(R) - \mathbb{E}_t^{\mathbb{Q}}[D^p(R)]$ $(p = -3, -1, 2, 4)$. Panels B, C, and D plot average profits and block-bootstrapped 95% confidence intervals for trading divergence swap payoffs $D^p(R^n)$ in Section III.A for different choices of p and n. (Color figure can be viewed at wileyonlinelibrary.com)

B.2. Lower Bounds

To test the lower bound implied by the $NCC(p, 1)$ conditions in Proposition 3, we estimate conditional expected returns with a parsimonious nonparametric specification

$$\mathbb{E}_t^{\mathbb{P}}[R] := \mathbb{E}^{\mathbb{P}}\left[R \mid \mathbb{E}_t^{\mathbb{Q}}[D^2(R)]\right] := g(\mathbb{E}_t^{\mathbb{Q}}[D^2(R)]). \tag{27}$$

Panel A of Figure 12 reports estimated monthly expected returns, while the bottom panel reports the largest parameters $\hat{p} \in [0, 1]$ such that

$$\hat{\mathbb{E}}_t^{\mathbb{P}}[R] = \hat{g}(\mathbb{E}_t^{\mathbb{Q}}[D^2(R)]) \geq \underline{\mathbb{E}}_t[R](\hat{p}). \tag{28}$$

Thus, $\hat{p}$ identifies the tightest lower bound in Proposition 3 that is consistent with estimated conditional expected returns. We obtain a negative conditional expected return for only two outlying observations in the financial crisis. In all other cases, the lower bound from the $NCC(2/5, 1)$ condition is not violated. In

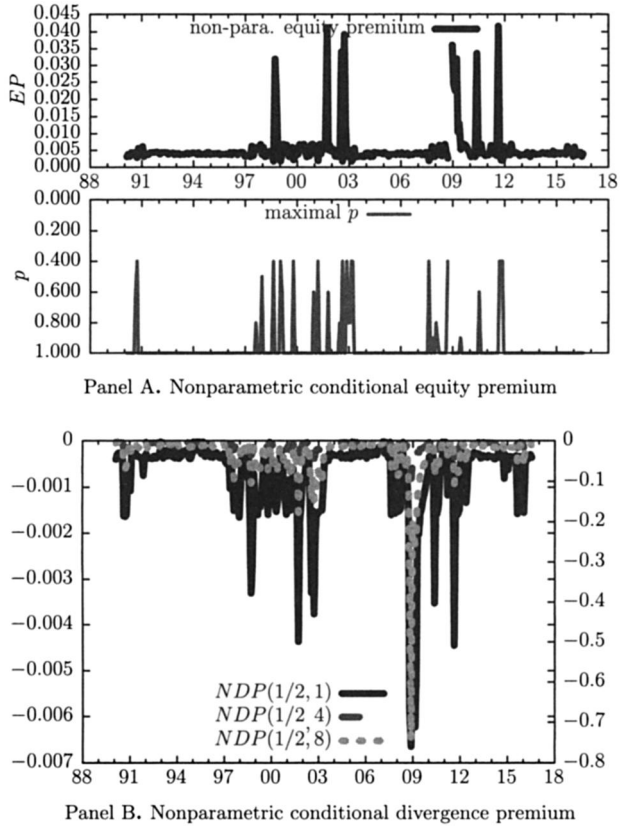


Figure 12. Nonparametric conditional tests of conditions *NDP*(*p*,*n*) and *NCC*(*p*,1) in Assumptions 1 and 2. The figure depicts time series of nonparametric regression estimates of the monthly divergence premium $\mathbb{E}_t^{\mathbb{P}}[D^p(R^n)] - \mathbb{E}_t^{\mathbb{Q}}[D^p(R^n)]$ and expected return $\mathbb{E}_t^{\mathbb{P}}[R]$ using the nonparametric specifications $\mathbb{E}_t^{\mathbb{P}}[D^p(R^n)] = f(\mathbb{E}_t^{\mathbb{Q}}[D^2(R^n)])$ and $\mathbb{E}_t^{\mathbb{P}}[R] = g(\mathbb{E}_t^{\mathbb{Q}}[D^2(R)])$. Panel A plots the time series of nonparametric estimates of $\mathbb{E}_t^{\mathbb{P}}[R]$, together with the time series of the largest power parameters $p \in [0, 1]$, such that the lower bound $\underline{\mathbb{E}}_t[R](p)$ in equation (22) holds. Panel B plots the time series of nonparametric estimates of $\mathbb{E}_t^{\mathbb{P}}[D^p(R^n)] - \mathbb{E}_t^{\mathbb{Q}}[D^p(R^n)]$, with divergence premia for $n = 1$ ($n = 4, 8$) reported on the left (right) y-axis. The figure is based on options on the S&P 500 index from January 1990 to August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

contrast, the bound resulting from tighter NCC conditions, such as $NCC(1, 1)$, is violated as much as 53.5% of the time.

Related evidence arises from the conditional expected returns estimated in Song and Xiu (2016). Model-based expected returns are always positive and smoother than nonparametric expected returns. We find that model-based violations of $NCC(2/5, 1)$ constraints are rare and concentrated in a short period during the recent financial crisis; see Figure 13, Panels A and B. They arise in 1.9% (2.5%) of observations at a monthly (annual) horizon. In contrast, violations of tighter $NCC(1, 1)$ constraints arise in 51% (69%) of observations at a monthly (annual) horizon.

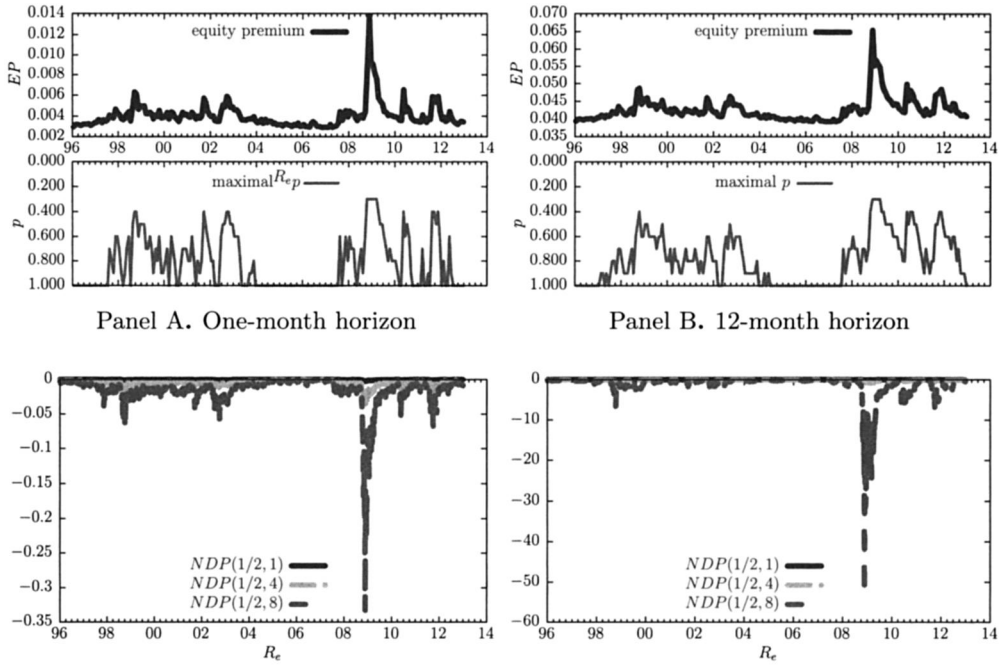


Figure 13. Model-based conditional tests of conditions *NDP*(*p*,*n*) and *NCC*(*p*,1) in Assumptions 1 and 2. The figure depicts time series of estimated model-based divergence premia and expected returns based on the Song and Xiu (2016) model. Panels A and B plot the time series of estimated monthly and annual expected returns $\mathbb{E}_t^{\mathbb{P}}[R]$, together with the time series of the largest power parameters $p \in [0, 1]$ such that the lower bound $\underline{\mathbb{E}}_t[R](p)$ in equation (22) holds. Panels C and D plot estimates of monthly and annual divergence premia $\mathbb{E}_t^{\mathbb{P}}[D^p(R^n)] - \mathbb{E}_t^{\mathbb{Q}}[D^p(R^n)]$ for $p = 1/2$ and $n = 1, 4, 8$. The figure is based on options on the S&P 500 index from January 1990 to August 2016. (Color figure can be viewed at wileyonlinelibrary.com)

V. Conclusion

In this paper, we introduce a new almost model-free approach that recovers a minimum variance pricing kernel projection (MVP) from forward-looking Arrow-Debreu prices under a conservative max-min Sharpe ratio optimization. Relying on empirically and economically motivated assumptions on the sign of the risk premia of various realized moments of returns, we recover an informative MVP that is tradeable using delta-hedged option portfolios. Besides our mild starting assumptions, our approach imposes no additional constraints on the underlying economy, such as Markovian or stationary state processes, market completeness, or path-independent pricing kernels, nor does it rely on additional time-series information about future returns.

We apply our approach to 25 years of S&P 500 option data over the period from January 1990 to August 2016. We find that the first and second recovered moments of returns provide useful information about future realized moments,

implying economically relevant out-of-sample predictability according to standard metrics. Recovered moments are highly time-varying and correspond to countercyclical term structures of equity premia, variance risk premia, and market Sharpe ratios. When plotted against future returns, MVPs are systematically U-shaped at longer horizons, consistent with the assumption of a large negative (generalized) variance risk premium. They are U-shaped at short horizons as well, when uncertainty is high, or when the conditional Sharpe ratio for selling market variance improves on the market Sharpe ratio. Optimal strategies selling the recovered MVP are profitable and improve on a linear market investment mainly by selling moderate amounts of out-of-the-money calls in states of higher countercyclical uncertainty. These optimal strategies outperform the model-based optimal kernel strategies estimated using cross-sectional and time-series information on Arrow-Debreu prices.

Initial submission: June 3, 2016; Accepted: September 13, 2017
Editors: Bruno Biais, Michael R. Roberts, and Kenneth J. Singleton

Appendix A: Proofs

A. The Moment Problem and Its Solution

Given a projection of order $J \geq 1$, the set of supporting physical probabilities $\mathcal{P}$ for projection $\mathcal{M}_{\mathbb{P}}(R)$ consists of all probability distributions $\mathbb{P}$ with the properties below (Curto and Fialko (1991), Theorem 5.1):

(1) The Hankel matrix

$$H_0(\mathbb{P}) := \mathbb{E}_t^{\mathbb{P}}\left[R^{i+j}\right]_{0\leq i,j\leq J} = \mathbb{E}_t^{\mathbb{P}}\left[\mathbf{R}_J\mathbf{R}_J'\right] \tag{A1}$$

is positive definite.

(2) The Hankel matrix

$$H_1(\mathbb{P}) := \mathbb{E}_t^{\mathbb{P}}\left[R^{1+i+j}\right]_{0\leq i,j\leq J} \tag{A2}$$

is positive semidefinite.

B. Proof of Proposition 1

Factorization (12) follows from the Vandermonde factorization of positive definite Hankel matrices (see, for example, Drmač (2015)). Constraint $\mathbf{1}_{J+1}'H_0(\mathbb{P})\mathbf{1}_{J+1} = 1$, where $\mathbf{1}_{J+1}$ is a $(J+1)$-dimensional vector of ones, then implies that D is a diagonal matrix of strictly positive probabilities. In Vandermonde factorization (12), the distinct entries $x_0, x_1, \ldots, x_J$ are defined by the roots of a uniquely determined polynomial of order $J+1$, with coefficients

$-\varphi_0, -\varphi_1, \ldots, -\varphi_J, 1$ such that

$$H_0 \begin{pmatrix} \varphi_0 \\ \varphi_1 \\ \vdots \\ \varphi_J \end{pmatrix} = H_1 e_{J+1}, \tag{A3}$$

where e_{J+1} is the $(J+1)$th unit vector in $\mathbb{R}^{J+1}$ and H_1 is defined in equation (A2) above. Let e_1 be the first unit vector in $\mathbb{R}^{J+1}$. As H_1 is positive semidefinite under the assumptions of Appendix A for the truncated moment problem, Proposition 3.3 and Theorem 5.1 in Curto and Fialko (1991) imply that

$$d_J = V^{-1} H_0 e_1 \tag{A4}$$

is a strictly positive $(J+1)$-atomic probability measure on support $\{x_0, x_1, \ldots, x_J\} \subset (0, \infty)$, such that $H_0(d_J) = H_0(\mathbb{P})$. Theorem 5.2 in Curto and Fialko (1991) finally implies that among any such representing measures, there is one such that H_1 is singular.

Appendix B: Implementation

A. *Tradeable Contracts in Detail*

To account for incomplete option markets with bounded moneyness domain $[a, b]$, we follow Schneider and Trojani (2014) and consider approximate payoff functions that are linearized outside of a corridor $[l, u]$ containing $[a, b]$. For a general payoff function f, the corridor payoff is

$$f_{[l,u]}(R) := \begin{cases} f(l) + f'(l)(R - l), & R < l, \\ f(R), & l \le R \le u, \\ f(u) + f'(u)(R - u), & R > u. \end{cases} \tag{B1}$$

Corridor payoffs are tradeable using an extension of the Carr and Madan (2001) spanning result,[41]

$$\mathcal{D} f_{[l,u]}(R) = \int_l^1 f''(K)(K - R)^+ dK + \int_1^u f''(K)(R - K)^+ dK. \tag{B2}$$

[41] We consider corridors wider than the available moneyness range $[a, b]$, in order to be consistent with the nonzero prices of deep out-of-the-money put and call options for strike prices a and b, respectively. Like the divergence operator $\mathcal{D}$, $f_{[l,u]}$ preserves the curvature of f. Moreover, it is such that $f \ge f_{[l,u]}$ ($f \le f_{[l,u]}$) whenever f is convex (concave). Equation (B2) directly implies $\mathcal{D} f \ge \mathcal{D} f_{[l,u]}$ ($\mathcal{D} f \le \mathcal{D} f_{[l,u]}$) whenever f is convex (concave).

In practice, this option portfolio is evaluated from a finite set of options with strikes $l = K_1 < \cdots < K_n = 1 < K_{n+1} < \cdots < K_N = u$ using the discrete approximation

$$\sum_{i=1}^{n-1} f''(K_i)(K_i - R)^+ \Delta K_i + \sum_{i=n}^{N} f''(K_i)(R - K_i)^+ \Delta K_i,$$

where $\Delta K_1 = K_2 - K_1$, $\Delta K_N = K_N - K_{N-1}$, and $\Delta K_i = (K_{i+1} - K_{i-1})/2$ for $1 < i < N$. This approximation is accurate for the strike range and density typically available in S&P 500 option markets.

Our approach to recovery can be parameterized completely using a finite number of physical and forward-neutral moments of returns. We therefore focus on corridor approximations of the form

$$f(R) \approx f(1) + f'(1)(R - 1) + \mathcal{D} f_{[l,u]}(R) \tag{B3}$$

for payoff functions generated by realized moments (powers) of returns. Since $\mathbb{E}_t^{\mathbb{Q}}[R] = 1$, corridor divergence $D^p_{[l,u]}(R)$ always has a finite forward price. In this case, the $NDP(p, 1)$ condition directly yields $\mathbb{E}_t^{\mathbb{P}}[D^p_{[l,u]}(R)] < \infty$. More generally, the $NDP(p, n)$ condition always implies $\mathbb{E}_t^{\mathbb{P}}[D^p_{[l,u]}(R^n)] < \infty$ if the n^{th} forward-neutral moment is finite.

B. *Vandermonde Factorization*

The Vandermonde factorization $H_0 = V D V'$ in equation (12), where diagonal matrix D has strictly positive entries and matrix V is based on distinct parameters $x_0, x_1, \ldots, x_J \subset (0, \infty)$, ensures that H_0 is positive definite. This factorization comes with the additional benefit of delivering stable closed-form expressions for the determinant and the inverse of H_0.[42] From factorization (12), it follows that $\det(H_0) = \det(D)\det^2(V)$, where $\det(D) = \prod_{i=0}^{J} d_{iJ}$ and $\det(V) = \prod_{1 \le i < j \le J}(x_i - x_j)$. Inverses of Vandermonde matrices are computed efficiently with a closed-form and numerically stable LU decomposition using the closed-form expression for the determinant. An algorithm specifically designed for this purpose is available in Turner (1966). To compute the global MVP in our recovery approach, we minimize the HJ distance $\mathbb{D}(d_3)$ with a differential evolution algorithm. Determinants of Hankel matrices H_0 for projections of order $J = 3$ are on the order of 1e-12.

Appendix C: Affine Models and Pricing Kernels

In this appendix, we illustrate the properties of our minimum variance kernel recovery approach in the context of well-known arbitrage-free jump diffusion models, based on affine state dynamics with exponentially

[42] Determinants and inverses of Hankel matrices are notoriously difficult to compute and standard inversion methods generally fail (see, for instance, Drmač (2015)).

affine pricing kernels. For this purpose, denote by X a process on probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$ with values in $D \subset \mathbb{R}^d$ and with semimartingale characteristics

$$A_t = \int_0^t a(X_{s-})ds,$$
$$B_t = \int_0^t b(X_{s-})ds,$$
$$\nu(\omega, dt, d\xi) = K(X_{t-}(\omega), d\xi)dt,$$

where $a(x)$, $b(x)$, and $K(x, d\xi)$ are affine functions of state x

$$a(x) = a + x_1\alpha_1 + \cdots + x_d\alpha_d,$$
$$b(x) = a + x_1\beta_1 + \cdots + x_d\beta_d,$$
$$K(x, d\xi) = m(d\xi) + x_1\mu_1(d\xi) + \cdots + x_d\mu_d(d\xi).$$

Define for $i \in \{1, \ldots, d\}$

$$F(y) = \frac{1}{2}\langle y, ay\rangle + \langle b, y\rangle + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{\langle \xi, y\rangle} - 1 - \langle h(\xi), y\rangle\right) m(d\xi), \tag{C1}$$

$$R_i(y) = \frac{1}{2}\langle y, \alpha_i y\rangle + \langle \beta_i, y\rangle + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{\langle \xi, y\rangle} - 1 - \langle h(\xi), y\rangle\right) \mu_i(d\xi), \tag{C2}$$

with $h(\xi) = 11_{\{|\xi|\leq 1\}}\xi$. Furthermore, let

$$\mathcal{D} := \bigcap_{x\in D} \left\{ y \in \mathbb{R}^d : \int_{|\xi|\geq 1} e^{\langle y, \xi\rangle} K(x, d\xi) < \infty \right\},$$

be the domain in which K has finite exponential moments. From Keller-Ressel and Mayerhofer (2015), the moment-generating function of X takes the form

$$\mathbb{E}^{\mathbb{Q}}\left[e^{\langle y, X_t\rangle} \mid X_0 = x\right] = e^{p(t,y) + \langle q(t,y), x\rangle}, \tag{C3}$$

where p and q solve the ordinary differential equations

$$\frac{\partial}{\partial t}p(t, y) = F(q(t, y)),\ p(0, y) = 0,$$
$$\frac{\partial}{\partial t}q(t, y) = R(q(t, y)),\ q(0, y) = y.$$

A. *Parametric Minimum Variance Kernels*

There are various possibilities for defining pricing kernels in affine models. We follow an exponentially affine approach that extends the pricing kernel specification in the introductory example of Section I. For any parameter $\theta \in \mathcal{D}$

and $t \geq 0$, define the process (as in Keller-Ressel (2008) and Dai and Singleton (2000))

$$\bar{\mathcal{M}}_t^{\theta} := \frac{d\mathbb{P}^{\theta}}{d\mathbb{Q}}\Big|_{\mathcal{F}_t} = \exp\left(\langle \theta, X_t - X_0\rangle - tF(\theta) - \left\langle R(\theta), \int_0^t X_s\, ds\right\rangle\right). \tag{C4}$$

This process is a $\mathbb{Q}$ martingale. Importantly, under probability $\mathbb{P}^{\theta}$, process X is again affine with characteristics $F^{\theta}(y) = F(y+\theta) - F(\theta)$ and $R^{\theta}(y) = R(y+\theta) - R(\theta)$. Furthermore, if the traded coordinates of X are $\mathbb{Q}$ martingales, then

$$\mathcal{M}_t^{\theta} := \frac{1}{\bar{\mathcal{M}}_t^{\theta}} = \frac{d\mathbb{Q}}{d\mathbb{P}^{\theta}}\Big|_{\mathcal{F}_t} \tag{C5}$$

is a valid forward pricing kernel for horizons $t > 0$.

EXAMPLE C1 (Black-Scholes power utility kernel): In a Black-Scholes model, $F_{t,T} := e^{X_t}$ for any $0 \leq t \leq T$ and the univariate log forward process X solves the SDE

$$dX_s = -\frac{1}{2}\sigma^2 ds + \sigma dW_s^{\mathbb{Q}}\ ;\ X_0 = x. \tag{C6}$$

The corresponding characteristics are

$$F_{BS}(y) = \frac{\sigma^2 y^2}{2} - \frac{\sigma^2 y}{2},$$
$$R_{BS}(y) = 0.$$

For parameter $\theta = \gamma$, the physical characteristics become

$$F_{BS}^{\theta}(y) = \frac{\sigma^2 y^2}{2} + \gamma\sigma^2 y - \frac{\sigma^2 y}{2},$$
$$R_{BS}^{\theta}(y) = 0,$$

along with the physical SDE

$$dX_s = \left(\gamma - \frac{1}{2}\right)\sigma^2 ds + \sigma dW_s^{\mathbb{P}^{\theta}}; X_0 = x. \tag{C7}$$

In this model, we recognize the $\mathbb{P}^{\theta}$ expected drift rate as $\gamma\sigma^2$. Given $T > 0$ and $r := X_T - X_0$, the Black-Scholes pricing kernel for horizon T is

$$\mathcal{M}_{\gamma,T}^{BS} = e^{-\gamma \cdot r + \frac{\gamma\sigma^2}{2}(\gamma-1)T}, \tag{C8}$$

the typical normalized power utility kernel from the introduction in Section I. Since X is the only state variable in this economy and kernel (C8) is r-measurable, this kernel agrees with its projection on R,

$$\mathcal{M}_{\gamma,T}^{BS} = \mathbb{E}^{\mathbb{P}}[\mathcal{M}_{\gamma,T}^{BS} | R]. \tag{C9}$$

EXAMPLE C2 (Duffie, Pan, and Singleton (2000) double-jump diffusion model): For a more flexible model, consider a specification that solves the bivariate SDE

$$d\begin{pmatrix} V_s \\ X_s \end{pmatrix} = \begin{pmatrix} \kappa_v(\bar{v} - V_s) \\ -\bar{\lambda}\bar{\mu} - \frac{1}{2}V_s \end{pmatrix} ds + \sqrt{V_s}\begin{pmatrix} \sigma_v & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} dW_s^{\mathbb{Q}} + dZ_t^{\mathbb{Q}}, \quad \text{(C10)}$$

where $F_{t,T} = e^{X_t}$ as before, $W^{\mathbb{Q}}$ is a bivariate $\mathbb{Q}$-Brownian motion and $dZ_t^{\mathbb{Q}}$ is a compound Poisson counter in $\mathbb{R}^2$ with constant mean arrival rate $\bar{\lambda}$ and a jump distribution that allows both cojumps and separate jumps in X and V. For equations (C1) and (C2), we specify[43]

$$\int_{\mathbb{R}^2\setminus\{0\}} \left(e^{\langle \xi, y\rangle} - 1\right) m(d\xi) = \bar{\lambda}(\Theta(c_1, c_2) - 1),$$

$$\int_{\mathbb{R}^2\setminus\{0\}} \left(e^{\langle \xi, y\rangle} - 1\right) \mu_i(d\xi) = 0 \; ; \; i = 1, 2,$$

where[44]

$$\Theta(c_1, c_2) := \frac{\lambda_v \Theta_v(y_1) + \lambda_y \Theta_x(y_2) + \lambda_c \Theta_c(y_1, y_2)}{\bar{\lambda}}, \quad \text{(C11)}$$

$$\Theta_v(c) := \frac{1}{1 - \mu_v c}, \quad \text{(C12)}$$

$$\Theta_x(c) := \exp\left(\mu_x c + \frac{1}{2}\sigma_x^2 c^2\right), \quad \text{(C13)}$$

$$\Theta_c(c_1, c_2) := \frac{\exp\left(\mu_x c_2 + \frac{1}{2}\sigma_x^2 c_2^2\right)}{1 - \mu_v c_1 - \rho_J \mu_v c_2}. \quad \text{(C14)}$$

Note that $\bar{\mu} = \Theta(0, 1) - 1$ for process e^{X_t} to be a martingale under $\mathbb{Q}$. With this specification, we obtain

$$F_{DPS}(y_1, y_2) = y_1 \bar{v}\kappa_v - y_2\bar{\lambda}\bar{\mu} + \bar{\lambda}(\Theta(y_1, y_2) - 1),$$

$$R^v_{DPS}(y_1, y_2) = -y_1\kappa_v + \frac{1}{2}\left(2\rho\sigma y_2 y_1 + \sigma^2 y_1^2 + y_2^2\right) - \frac{y_2}{2},$$

$$R^x_{DPS}(y_1, y_2) = 0.$$

Using the parameterization $\theta = (\eta, \gamma)$, we then have

$$F^{\theta}_{DPS}(y_1, y_2) = y_1 \bar{v}\kappa_v - y_2\bar{\lambda}\bar{\mu} + \Theta(\eta, \gamma)\bar{\lambda}\left(\frac{\Theta(y_1 + \eta, y_2 + \gamma)}{\Theta(\eta, \gamma)} - 1\right),$$

[43] We use a finite-variation specification for which the truncation function h drops out.

[44] The joint jump-size distribution in equation (C14) arises from the joint distribution of $(Y, Z + \rho_J Y)$, where Y is exponentially distributed ($Y \sim \mathcal{E}(1/\mu_v)$) and Z is normally distributed ($Z \sim N(\mu_x, \sigma_x)$).

$$R^{v,\theta}_{DPS}(y_1, y_2) = y_1(\sigma(\gamma\rho + \eta\sigma) - \kappa_v) + \frac{1}{2}\left(2\rho\sigma y_2 y_1 + \sigma^2 y_1^2 + y_2^2\right)$$
$$+ y_2\left(\gamma + \eta\rho\sigma - \frac{1}{2}\right), R^{x,\theta}_{DPS}(y_1, y_2) = 0.$$

The physical state dynamics resulting from forward dynamics (C10) is given by the SDE

$$d\begin{pmatrix} V_s \\ X_s \end{pmatrix} = \begin{pmatrix} \bar{v}\kappa_v + (\sigma(\gamma\rho + \eta\sigma) - \kappa_v)V_s \\ -\bar{\lambda}\bar{\mu} + \left(\gamma + \eta\rho\sigma - \frac{1}{2}\right)V_s \end{pmatrix} ds + \sqrt{V_s}\begin{pmatrix} \sigma_v & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} dW^{\mathbb{P}^\theta}_s + dZ^{\mathbb{P}^\theta}_s.$$

The physical average jump intensity is $\Theta(\eta, \gamma)\bar{\lambda}$ and the physical jump-size distribution has moment-generating function $\frac{\Theta(c_1+\eta, c_2+\gamma)}{\Theta(\eta,\gamma)}$.[45]

Again, using the notation $r := X_T - X_0$ and parameter vector $\theta = (0, \gamma)$, the pricing kernel for horizon $T > 0$ is given by

$$\mathcal{M}^{DPS}_{\gamma,T} = \exp\left(-\gamma \cdot r + T F_{DPS}(0, \gamma) + R^v_{DPS}(0, \gamma)\int_0^T V_s ds\right). \quad \text{(C15)}$$

This pricing kernel depends on both the forward log return and the entire path of the variance process. Therefore, it is path-dependent and not r-measurable. Its finite-order projections on various realized moments of gross returns can be computed using the framework in this paper. More specifically, $\mathbb{P}$ and $\mathbb{Q}$ moments follow from the exponentially affine moment-generating function, which allows us to compute the model-based HJ projections consistent with equation (9).

Figure IA.11 reports various MVP of order $J = 3$ for different conditioning values of the instantaneous variance and different pricing kernel parameters γ. We find that the shapes of these pricing kernel projections are relatively insensitive to the conditioning value of the instantaneous variance, despite the wide range of values for implied volatilities. In contrast, they are quite sensitive to the choice of the price of return risk parameterized by parameter γ.

EXAMPLE C3 (Song and Xiu (2016) model): This model is based on the following two-factor specification of forward dynamics

$$\begin{aligned} dV_s &= \left(\eta + \kappa(\xi_s - V_s)\right) ds + \sigma\sqrt{V_s} dB^{\mathbb{Q}}_s + J^{\mathbb{Q}}_V dN_s, \\ d\xi_s &= \alpha(\theta - \xi_s) ds + \gamma\sqrt{\xi} dM^{\mathbb{Q}}_s, \\ dX_s &= -\left(\mu\lambda_s + \frac{1}{2}V_s\right) ds + \sqrt{V_s} dW^{\mathbb{Q}} + J^{\mathbb{Q}}_X dN_s, \end{aligned} \quad \text{(C16)}$$

where $W^{\mathbb{Q}}$ and $B^{\mathbb{Q}}$ are standard Brownian motions with correlation ρ, $M^{\mathbb{Q}}$ is another independent Brownian motion, and $J^{\mathbb{Q}}_X$ and $J^{\mathbb{Q}}_V$ are random jump sizes

[45] Under $\mathbb{P}^\theta$, this jump-size distribution corresponds to the joint moment-generating function of $(Y, Z + \rho_J Y)$, where $Y \sim \mathcal{E}(1/\mu_v - \eta - \gamma\rho_J)$ and $Z \sim N(\mu_x + \gamma\sigma_x^2, \sigma_x)$.

with exponential and, respectively, double exponential distributions

$$J_V^{\mathbb{Q}} \sim \mathcal{E}(\beta_V)\,;\; J_X^{\mathbb{Q}} \sim \begin{cases} \mathcal{E}(\beta_+) & \text{with probability } q, \\ \mathcal{E}(\beta_-) & \text{with probability } 1-q. \end{cases}$$

The jump intensity $\lambda_t = \lambda_0 + \lambda_1 V_t$ is stochastic and $\mu = \mathbb{E}^{\mathbb{Q}}[e^{J_X^{\mathbb{Q}}} - 1]$. The $\mathbb{P}$ state dynamics in the model read

$$\begin{aligned} dV_s &= \left(\eta + \kappa\xi_s - \kappa^{\mathbb{P}} V_s\right) ds + \sigma\sqrt{V_s} dB_s^{\mathbb{P}} + J_V^{\mathbb{P}} dN_s, \\ d\xi_s &= \left(\alpha\theta - \alpha^{\mathbb{P}}\xi_s\right) ds + \gamma\sqrt{\xi} dM_s^{\mathbb{P}}, \\ dX_s &= (\mu_0 + \mu_1 V_s) ds + \sqrt{V_s} dW^{\mathbb{P}} + J_X^{\mathbb{P}} dN_s, \end{aligned} \tag{C17}$$

where $W^{\mathbb{P}}$ and $B^{\mathbb{P}}$ are standard Brownian motions with correlation ρ, $M^{\mathbb{P}}$ is another independent Brownian motion, and $J_X^{\mathbb{P}}$ and $J_V^{\mathbb{P}}$ are random jump sizes with exponential and, respectively, double exponential distributions

$$J_V^{\mathbb{P}} \sim \mathcal{E}(\beta_V^{\mathbb{P}})\,;\; J_X^{\mathbb{P}} \sim \begin{cases} \mathcal{E}(\beta_+^{\mathbb{P}}) & \text{with probability } q, \\ \mathcal{E}(\beta_-^{\mathbb{P}}) & \text{with probability } 1-q. \end{cases}$$

Also, for this specification, the $\mathbb{P}$ and $\mathbb{Q}$ moments follow from the exponential affine moment generating function, which is sufficient to compute the model-based MVP consistent with equation (9).

Figure IA.10 plots various model-based MVP of order $J = 3$ for different pricing kernel parameter vectors $\theta = (\theta_v, 0, \theta_s)$. The shapes of these pricing kernel projections are relatively flexible, with monotonic, U-shaped, or inverse U-shaped projections that arise for various choices of vector θ. U-shaped projections typically obtain when volatility shocks are priced ($\theta_v > 0$) and when return shocks have a relatively small price of risk (θ_s is small).

B. Nonparametric Minimum Variance Kernels

Given a forward probability $\mathbb{Q}$, the previous section introduces a family of exponentially affine pricing kernels (C5) and corresponding physical moment specifications that allow an explicit computation of a model-based finite-order MVP. More generally, given a finite set of forward moments, we can parameterize all admissible MVP of order J using Vandermonde representation (12). If the forward moments are model-based, this approach allows us to determine MVPs consistent with model-based forward moments without imposing additional structure on physical moments through a parametric pricing kernel specification.

Clearly, any set of model-based physical moments is always supported by Vandermonde factorization (12). Therefore, model-based MVPs generally induce a higher pricing kernel variance than model-free MVPs under our approach because physical moments need to satisfy the model-implied parametric moment structure. Model-based constraints on the moment structure can be

difficult to elicit and interpret, especially when several moments are simultaneously constrained, and they can have unexpected effects on the shape of MVPs. For instance, the bottom panel of Figure IA.9 shows that, even in the model of Duffie, Pan, and Singleton (2000) with exponentially affine pricing kernels, model-free MVPs can be U-shaped when the model-based MVPs are monotonic.

REFERENCES

Acciaio, Beatrice, Mathias Beiglböck, Friedrich Penkner, and Walter Schachermayer, 2016, A model-free version of the fundamental theorem of asset pricing and the super-replication theorem, *Mathematical Finance* 26, 233–251.

Aït-Sahalia, Yacine, and Jefferson Duarte, 2003, Nonparametric option pricing under shape restrictions, *Journal of Econometrics* 16, 9–47.

Aït-Sahalia, Yacine, and Andrew W. Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* 53, 499–547.

Almeida, Caio, and René Garcia, 2017, Economic implications of nonlinear pricing kernels, *Management Science* 63, 3361–3380.

Alvarez, Fernando, and Urban Jermann, 2005, Using asset prices to measure the persistence of the marginal utility of wealth, *Econometrica* 73, 1977–2016.

Andersen, Torben G., Oleg Bondarenko, and Maria T. Gonzalez-Perez, 2015, Exploring return dynamics via corridor implied volatility, *Review of Financial Studies* 28, 2902–2945.

Backus, David K., Mikhail Chernov, and Stanley Zin, 2014, Affine term structure models and the forward premium anomaly, *Journal of Finance* 69, 51–99.

Bakshi, Gurdip, and Dilip Madan, 2008, Investor heterogeneity, aggregation, and the non-monotonicity of the aggregate marginal rate of substitution in the price of market-equity, Working paper, University of Maryland and Georgetown University.

Bansal, Ravi, and Bruce N. Lehmann, 1997, Growth-optimal portfolio restrictions in asset pricing models, *Macroeconomic Dynamics* 1, 333–354.

Bondarenko, Oleg, 2014, Why are put options so expensive?, *Quarterly Journal of Finance* 4, 1–50.

Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman, 2016, Misspecified recovery, *Journal of Finance* 71, 2493–2544.

Breeden, Douglas T., and Robert H. Litzenberger, 1978, Prices of state-contingent claims implicit in option prices, *Journal of Business* 51, 621–651.

Britten-Jones, Mark, and Anthony Neuberger, 2000, Option prices, implied price processes, and stochastic volatility, *Journal of Finance* 55, 839–866.

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 Keith Lewis, 2004, Corridor variance swaps, *Risk* 17, 67–72.

Carr, Peter, and Dilip Madan, 2001, Towards a theory of volatility trading, in Elyes Jouini, Jaksa Cvitanic, and Marek Musiela, eds.: *Handbooks in Mathematical Finance, Part Three: Risk Management and Hedging* (Cambridge University Press, New York).

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

Carr, Peter, and Jiming Yu, 2012, Risk, return, and Ross recovery, *Journal of Derivatives* 20, 38–59.

Chabi-Yo, Fousseni, 2012, Pricing kernels with stochastic skewness and volatility risk, *Management Science* 58, 624–640.

Chabi-Yo, Fousseni, Gurdip Bakshi, and Xiaohui Gao, 2015, A recovery that we can trust? Deducing and testing the restrictions of the recovery theorem, Working paper, University of Maryland and Fisher College of Business.

Chabi-Yo, Fousseni, René Garcia, and Eric Renault, 2008, State dependence can explain the risk aversion puzzle, *Review of Financial Studies* 21, 973–1011.

Chicago Board Options Exchange (CBOE), 2009, CBOE Volatility Index: VIX. Available at https://www.cboe.com/micro/vix/vixwhite.pdf.

Christensen, Timothy M., 2017, Nonparametric stochastic discount factor decomposition, *Econometrica* 85, 1501–1536.

Christoffersen, Peter, Steven Heston, and Kris Jacobs, 2013, Capturing option anomalies with a variance-dependent pricing kernel, *Review of Financial Studies* 26, 1963–2006.

Cochrane, John H., and Jesus Saa-Requejo, 2000, Beyond arbitrage: Good-deal asset price bounds in incomplete markets, *Journal of Political Economy* 108, 79–119.

Collin-Dufresne, Pierre, and Robert S. Goldstein, 2002, Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility, *Journal of Finance* 57, 1685–1730.

Curto, Raúl, and Lawrence Fialko, 1991, Recursiveness, positivity, and truncated moment problems, *Houston Journal of Mathematics* 17, 603–635.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Drmač, Zlatko, 2015, SVD of Hankel matrices in Vandermonde-Cauchy product form, *Electronic Transactions on Numerical Analysis* 44, 593–623.

Duffee, Gregory R., 2011, Information in (and not in) the term structure, *Review of Financial Studies* 24, 2895–2934.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Filipović, Damir, Eberhard Mayerhofer, and Paul Schneider, 2013, Transition density approximations for multivariate affine jump diffusion processes, *Journal of Econometrics* 176, 93–111.

Gruber, Peter, Claudio Tebaldi, and Fabio Trojani, 2015, The price of the smile and variance risk premia, Working paper, University of Lugano, Bocconi, University of Geneva, and the Swiss Finance Institute.

Hansen, Lars P., and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hansen, Lars P., and José Scheinkman, 2009, Long-term risk: An operator approach, *Econometrica* 77, 177–234.

Jackwerth, Jens, 2000, Recovering risk aversion from option prices and realized returns, *Review of Financial Studies* 13, 433–451.

Jensen, Christian S., David Lando, and Lasse H. Pedersen, 2017, Generalized recovery, *Journal of Financial Economics*, forthcoming.

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton, 2014, Risk premiums in dynamic term structure models with unspanned macro risks, *Journal of Finance* 69, 1197–1233.

Keller-Ressel, Martin, 2008, Affine processes, PhD thesis, Vienna University of Technology.

Keller-Ressel, Martin, and Eberhard Mayerhofer, 2015, Exponential moments of affine processes, *Annals of Applied Probability* 25, 714–752.

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

Lee, Roger, 2008, Corridor variance swap, Working paper, University of Chicago.

Lee, Roger, 2010, Gamma swap, in Rama Cont, ed.: *Encyclopedia of Quantitative Finance* (Wiley, Chichester).

Martin, Ian, 2013, Simple variance swaps, Working paper, Stanford University.

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

Neuberger, Anthony, 1994, The log contract, *Journal of Portfolio Management* 20, 74–80.

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

Orłowski, Piotr, Andras Sali, and Fabio Trojani, 2015, Arbitrage-free dispersion, Working paper, SFI, University of Geneva, and University of Lugano.

Qin, Likuan, and Vadim Linetsky, 2016, Positive eigenfunctions of Markovian pricing operators: Hansen-Scheinkman factorization, Ross recovery, and long-term pricing, *Operations Research* 64, 99–117.

Qin, Likuan, and Vadim Linetsky, 2017, Long-term risk: A martingale approach, *Econometrica* 85, 299–312.

Rosenberg, Joshua V., and Robert F. Engle, 2002, Empirical pricing kernels, *Journal of Financial Economics* 64, 341–372.

Ross, Steve, 2015, The recovery theorem, *Journal of Finance* 70, 615–648.
Schneider, Paul, 2015, Generalized risk premia, *Journal of Financial Economics* 116, 487–504.
Schneider, Paul, and Fabio Trojani, 2014, Fear trading, Working paper, USI Lugano and the Swiss Finance Institute.
Schneider, Paul, and Fabio Trojani, 2018, Divergence and the price of uncertainty, *Journal of Financial Econometrics*, forthcoming.
Snow, Karl N., 1991, Diagnosing asset pricing models using the distribution of asset returns, *Journal of Finance* 43, 955–983.
Song, Zhaogang, and Dacheng Xiu, 2016, A tale of two option markets: Pricing kernels and volatility risk, *Journal of Econometrics* 190, 176–196.
Turner, L. Richard, 1966, Inverse of the Vandermonde matrix with applications, Technical report, NASA.
Walden, Johan, 2017, Recovery with unbounded diffusion processes, *Review of Finance* 21, 1403–1444.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.