Working Paper/Document de travail
2012-11

# Risk Premium, Variance Premium and the Maturity Structure of Uncertainty

by Bruno Feunou, Jean-Sébastien Fontaine,
Abderrahim Taamouti and Roméo Tédongap

www.bank-banque-canada.ca

Bank of Canada Working Paper 2012-11

March 2012

# Risk Premium, Variance Premium and the Maturity Structure of Uncertainty

by

**Bruno Feunou,[1] Jean-Sébastien Fontaine,[1]**
**Abderrahim Taamouti[2] and Roméo Tédongap[3]**

[1]Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
Correspondence: jsfontaine@bankofcanada.ca

[2]Universidad Carlos III de Madrid

[3]Stockholm School of Economics



ISSN 1701-9397

# Acknowledgements

We thank Nour Meddahi, Bo-Young Chang, Scott Hendry, Jorge Lopez, Alberto Romero and seminar participants at Duke University, the Bank of Canada, London School of Economics, CIRANO, the NFA 2011 conference and the CFE 2011 conference for their comments and suggestions.

## Abstract

Expected returns vary when investors face time-varying investment opportunities. Long-run risk models (Bansal and Yaron 2004) and no-arbitrage affine models (Duffie, Pan, and Singleton 2000) emphasize sources of risk that are not observable to the econometrician. We show that, for both classes of models, the term structure of risk implicit in option prices can reveal these risk factors ex-ante. Empirically, we construct the variance term structure implied in SP500 option prices. The variance term structure reveal two important drivers of the bond premium, the equity premium, and the variance premium, jointly. We also consider the term structure of higher-order risks as measured by skewness and kurtosis and still find that two factors are sufficient to summarize the information content from the term structure of risks. Overall, our results bode well for the ability of structural models to explain risk-returns trade-offs across different markets using only very few sources of risk.

*JEL classification: G12, G13*
*Bank classification: Financial markets; Asset pricing*



## Résumé

Les rendements espérés varient lorsque les occasions de placement qui s'offrent aux investisseurs ne sont pas constantes dans le temps. Les modèles avec risque de long terme (Bansal et Yaron, 2004) et les modèles affines basés sur l'absence d'arbitrage (Duffie, Pan et Singleton, 2000) mettent l'accent sur des sources de risque qui échappent à l'observation de l'économètre. Les auteurs montrent que, pour chacune de ces classes de modèles, la structure par terme des risques implicitement contenue dans les prix d'options peut révéler ces facteurs de risque *a priori*. De manière empirique, les auteurs construisent la structure par terme de la variance à partir des prix des options sur l'indice SP500. Cette structure par terme fait ressortir deux importants facteurs qui déterminent, conjointement, la prime obligataire, la prime sur actions et la prime liée à la variance. Les auteurs étudient par ailleurs la structure par terme des risques d'ordre supérieur quantifiés d'après les coefficients d'asymétrie et d'aplatissement. Là aussi, ils concluent que deux facteurs suffisent pour résumer le contenu informatif de la structure par terme des risques. Dans l'ensemble, les résultats obtenus augurent bien de la capacité des modèles structurels à rendre compte, à l'aide d'un nombre limité de facteurs, des arbitrages risque-rendement qui s'exercent sur différents marchés.

*Classification JEL : G12, G13*
*Classification de la Banque : Marchés financiers; Évaluation des actifs*

# 1 Introduction

Expected returns vary whenever investors face time-varying investment opportunities. For example, in Merton (1973), the premium between equilibrium expected returns on equity and the risk-free rate, $EP_t$, is proportional to the conditional variance of wealth, $\sigma_t^2$,

$$EP_t = \gamma \sigma_t^2, \tag{1}$$

where $\gamma$ is the coefficient of risk aversion. Unfortunately, the *ex-ante* conditional equity premium and variance are not directly observable to the econometrician.[1] In addition, recent equilibrium models imply that expected returns depend on other risk factors and, moreover, that the relationship between risks and returns varies across markets and across investment horizons. But, again, we do not observe expected returns and the relevant measures of risk. Hence, recent theoretical innovations only add to the challenges of empirical work and, therefore, the analysis of trade-offs between risks and returns remains a central question for researchers and practitioners in finance. Attesting to this, in his presidential address to the American Finance Association, John Cochrane refers to a "Multivariate Challenge" to returns predictability (Cochrane 2011). Specifically, he noted the abundance of empirical results linking one potential risk factor at a time to one type of return at a time. He laid particular emphasis on the fact that there is a strong common element underlying these relationships and asked "what is the factor structure of time-varying expected returns?"

This question is at the heart of our investigation. We build on the insight that option prices can provide model-free forward-looking measures of risks. Our main contribution is to show that the *term structure* of variance implicit in option prices can be used to reveal risk factors. Empirically, we find that (i) the variance term structure reveals two factors that drive significant variations in expected returns, (ii) the same two factors predict the bond, equity and variance premia, and (iii) the predictive content is particularly strong at short horizons, less than one year, where other popular predictors are relatively less effective, and (iv) the variance factors are related to its level and slope but they mix information from higher order principal components as well. Together, our results bode well for our ability to link risk-return trade-offs across different markets, and across horizons, within a unified theoretical framework. More immediately, the results show the rich information content of the term structure of option prices.

[1] This may explain why the empirical support for the theoretical prediction in Equation 1 is remarkably uneven.French et al. (1987), Campbell and Hentschel (1992), Ghysels et al. (2004), find a positive relation between volatility and expected returns. Turner et al. (1989), Glosten et al. (1993) and Nelson (1991) find a negative relation. Guo and Savickas (2006) finds that a positive relationship between index volatility and individual stock returns. Ludvigson and Ng (2005) find a strong positive contemporaneous relation between the conditional mean and conditional volatility and a strong negative lag-volatility-in-mean effect.

## 1.1 The variance term structure spans risk factors

We motivate our investigation in the context of Long-Run Risk economies (LRR) that match important stylized facts in finance. Specifically, we consider the class of affine LLR economies (Eraker 2008) that generalize the seminal paper of Bansal and Yaron (2004) to conditionally non-gaussian state variables. Risk-return trade-offs in these economies imply that the bond, equity and variance premia[2] at different investment horizons are linear functions of the *same* risk factors. In other words, all expected returns should exhibit a similar factor structure. The same result holds true in the broader family of affine no-arbitrage jump-diffusion models (Duffie, Pan, and Singleton 2000) that ignore structural cross-parameter restrictions. In any case, as in Merton's model above, the risk factors are unobservable to the econometrician.[3] Nonetheless, theory also predicts that the risk factors form a basis for the term structure of variance. In other words, a small number of linear combinations from the variance term structure, which can be measured from option prices, should span time-variations in expected returns.

## 1.2 The variance term structure predicts the bond and equity premium

Empirically, we follow the standard model-free approach from Bakshi and Madan (2000) to measure variance using options on SP500 futures across a range of maturities. We proceed in four steps. First, we show that the variance term structure exhibits a low-dimensional factor structure. Its first three principal components explain close to 95% of total variations, they have a systematic effect across maturities and each component can be interpreted as level, slope and curvature factors, respectively.

Second, we estimate how many factors from the variance term structure are sufficient to summarize its predictive content for bond and equity returns, *jointly*. We use the robust procedure of Cook and Setodji (2003). This dimension reduction procedure does not focus *a priori* on the leading principal components.[4] The test does not rely on any distributional assumption. It is also robust to departures from linearity. We find that two factors are sufficient to summarize the joint predictability of the bond and equity premia across maturities and across horizons.

In a third step, we estimate these factors via multivariate Reduced-Rank Regressions (RRR) of returns on the variance term structure - a generalization of multivariate OLS

[2]The variance premium is the difference between the expected variance under the historical measure and the risk-neutral measure, $\mathbb{Q}$, which is given by $VRP(t,\tau) = E_t^{\mathbb{Q}}[\sigma^2_{r,t+\tau}] - E_t[\sigma^2_{r,t+\tau}]$. This is analogous to the definition of the Equity Premium, $EP(t,\tau) = E_t[r_{t,t+\tau}] - E_t^{\mathbb{Q}}[r_{t,t+\tau}]$.

[3]The defining structure of affine LLR models combines recursive preference with small but persistent stochastic factors in the distribution of consumption growth. The latter are difficult to measure by construction. Similarly, no-arbitrage jump-diffusion models rely on unobservable factors that drive variations in the stochastic drift, volatility and jump intensity of the underlying process.

[4]Cochrane and Piazzesi (2005) provide an example in the context of bond returns where a small principal component of forward rates, which is typically ignored to explain the variations in forward rates themselves, plays an important role in predicting bond excess returns.

regressions to the case where the rank of the coefficient matrix corresponds to the number of linear combinations estimated above.[5] A rank-two coefficient matrix yields $R^2$s ranging from 5% to 7% for bond returns. This compares with $R^2$s ranging from 1% to 2% using the variance premium as reported in Mueller, Vedolin, and Zhou (2011). Next, the variance term structure predicts equity returns with $R^2$s that range from 3% to 6% at horizons between 1 and 12 month. The predictability is stronger at intermediate horizons and peaks for 3-month returns. This is consistent with Bollerslev, Tauchen, and Zhou (2009). They show that the variance premium predicts equity returns with comparable $R^2$s and that other popular predictors are typically less informative at these horizons. Importantly, their results also exhibit a peak at the 3-month horizon. Overall, the results confirm the information content of the variance term structure. Moreover, the reduced-rank restriction is supported in the data : allowing for more than two risk factors yields little statistical or economic gains.

Finally, we show formally that the variance term structure should also span the variance premium. We use this additional prediction as an out-of-sample check and ask whether the same two variance factors estimated to predict the bond and equity premia *only* can predict the variance premium *also.* We find that regressions of excess variance[6] on variance factors yield $R^2$s reaching up to 10% at the 6-month horizon with an inverted U-shape across horizons between one month and one year. Each of the factors plays an important role but at different horizons.

## 1.3 The factor structure extends to skewness and kurtosis

The variance term structure may fail to reveal all risk factors. This may arise if some factors do not affect the variance, or if the effects are small relative to the measurement errors in the variance or relative to the innovations in returns.[7] On the other hand, theory predicts that expected returns should nonetheless be linked with measures of higher-order risks such as asymmetry and tail thickness. In particular, all the cumulants of multi-horizon returns, including the variance, are affine.[8] Therefore, we can use the term structure of higher-order risks to discern further risk factors. Empirically, we construct model-free

[5]Estimation and inference in RRR models is available in closed-form. See Anderson (1951) and, more recently, Hansen (2008) as well as Reinsel and Velu (1998) for a textbook treatment.

[6] The excess variance, $xVR_{t,t+\tau}$, is defined relative to the Variance Premium in a way that is analogous to the definition of excess returns, $xR_{t,t+\tau}$, relative to the Equity Premium. We have that $xR_{t,t+\tau} = r_{t,t+\tau} - E_t^{\mathbb{Q}}[r_{t+1}]$ and $xVR_{t,t+\tau} = E_t^{\mathbb{Q}}[\sigma^2_{t,t+\tau}] - \sigma^2_{t,t+\tau}$, respectively.

[7]This is yet another similarity with the term structure of interest rates. In principle, yields can reveal all state variables related to the future behavior of the short rate. However, specific cases arise where some factors have small or no impact on interest rates and remain hidden. See Duffee (2011).

[8]Recall that the first cumulant corresponds to the mean, the second cumulant corresponds to the variance, the third cumulant corresponds to the third central moment and provides a measure of skewness, while the fourth cumulant corresponds to the fourth central moments minus 3 times the squared variance and provides a measure of the tails. The use of the cumulant-generating function to characterize the effect of higher-order cumulants on properties of asset prices is also suggested by Martin (2010). The cumulant term structure has been neglected in the literature.

measures of asymmetry and tail tichness based on cumulants (labeled as skewness and kurtosis hereafter). We find that each of the term structure of variance, skewness and kurtosis has a similar predictive content for the bond , equity and the variance premia. Importantly, each term structure's predictive content can be summarized by two factors. As above, we proceed in four steps. First, we document the factor structure. Second, we estimate the number of factors. Third, we estimate RRR models linking the skewness or kurtosis term structure, respectively, to the bond and equity returns. Finally, we use the same factors and confirm that the predictability extends to the variance premium. Two factors remain sufficient in every case. Strikingly, combining factors from the term structure of variance, skewness and kurtosis *does not* provide significant improvements in our ability to predict bond and equity returns. If anything, higher-order risk measures improve our ability to predict bond returns and may be correlated with flight-to-quality.

### 1.4 Literature

Christoffersen, Jacobs, and Chang (2011) review the vast literature that uses option-implied information in forecasting, including for returns predictability.[9] Our approach is most closely related to Bakshi, Panayotov, and Skoulakis (2011). They study the predictive content of the 1-month and 2-month forward variances for SP500 and Treasury bill returns.[10] We use a broader range of maturities, as well as higher-order moments, and consider the joint variations of expected returns across markets. Motivated by theory, we uncover the factor structure of option-implied variance and higher-order risk measures.

Leippold, Wu, and Egloff (2007), Amengual (2009), and Carr and Wu (2011) find that two factors are needed to describe the variance premium dynamics. We link these factors to the term structure of risk implicit in option prices. Bollerslev, Tauchen, and Zhou (2009), and Drechsler and Yaron (2011) ask whether the variance premium can predict the equity premium. Similarly, Zhou (2011) and Mueller, Vedolin, and Zhou (2011) ask whether the variance premium predict bond returns. But the variance premium is not observable to the econometrician and one must resort to proxies based on lagged observations to use in predictability regressions.[11] In contrast, we turn this view on its head and ask whether the ex-post excess variance is predictable using the same factors that drive the equity and the bond premium. The key insight from Bollerslev, Tauchen, and Zhou (2009) still holds: the variance premium is tightly linked to fundamental risk-return trade-offs. Backus, Chernov, and Martin (2010) find that the level of disaster risk required to explain the large unconditional equity premium is not consistent with the distribution implicit in option

[9] In particular, Ang et al. (2006) shows that option-implied market volatility is priced in the cross-section of equity returns. Chang, Christoffersen, and Jacobs (2011) shows that option-implied market skewness is priced in the cross-section of equity returns.

[10] Strictly speaking, they focus on the information content of payoffs contingent on the exponential of future integrated variance.

[11] The variance premium is unobservable because the conditional expectation of integrated variance under the historical probability measure is unobservable.

prices. We focus on conditional moments and provide further stylized facts from the option market. Constantinides and Ghosh (2011) use the risk-free rate and the price-dividend ratio to invert for the risk factors in the LRR specification of Bansal and Yaron (2004). Our empirical implementation is robust to misspecification of the underlying model.

The rest of the paper is organized as follow. Section 2 considers affine LRR economies and derives the multi-horizon cumulant-generating function of excess returns and excess variance. We then show how the term structure of uncertainty can be used to reveal fundamental risk factors. Section 3 introduces the data and measurement of risk from option prices. Section 4 evaluates the information content from the term structure of risk-neutral variance. Section 5 repeats the exercise by extending the information set to include the term structure of skewness and kurtosis. Section 6 concludes.

# 2 Variance Term Structure In Equilibrium

This Section motivate the empirical analysis that forms the core of the paper. We study the bond, equity and the variance premia within the broad family of affine general equilibrium models described in Eraker (2008). Equivalently, we could motivate our analysis within a reduced-form no-arbitrage representation of the economy in the family of asset pricing models with affine transforms (Duffie, Pan, and Singleton, 2000).[12] We focus on the distribution of multi-period returns under the risk-neutral and historical measure, $\mathbb{Q}$ and $\mathbb{P}$, respectively, via their cumulant-generating function. In particular, we derive expressions for the multi-horizon equity premium and bond premium. We also derive expressions for the conditional variance of returns across investment horizons. We then show how to recover the bond and equity premium from the term structure of variance.

## 2.1 Long-Run Risk Economies

Affine general equilibrium models build on the insights from the long-run risk literature and nest existing specifications where the mean and volatility of consumption growth are stochastic, possibly with jumps, and follow affine processes (e.g. Bansal and Yaron 2004, Bollerslev, Tauchen, and Zhou 2009, Drechsler and Yaron 2011). Consider an endowment economy where the representative agent's preference ordering over consumption paths can be represented by a recursive utility function of the Epstein-Zin-Weil form,

$$U_t = \left[ (1-\delta)\, C_t^{(1-\gamma)/\theta} + \delta \left( E_t \left[ U_{t+1}^{1-\gamma} \right] \right)^{1/\theta} \right]^{\theta/(1-\gamma)}, \tag{2}$$

with $\theta$ defined as,

$$\theta \equiv \frac{1-\gamma}{1-1/\psi},$$

[12] The essential component in the argument is that the joint Laplace transform of the state vector and of the change of measure is affine, or approximately so.Chamberlain (1988) provides an alternative argument based on a martingale representation argument. We thank Nour Meddahi for this suggestion.

where $\delta$ is the agent's subjective discount rate, $\psi$ measures the elasticity of intertemporal substitution and $\gamma$ determines risk aversion as well as the preference for intertemporal resolution of uncertainty. Next, assume that the joint dynamics of the (log) consumption growth process, $\Delta c_{t+1}$, and the $K$ state variables in the economy, $X_{t+1}$, have the following Laplace transform,

$$E_t\left[\exp\left(u\Delta c_{t+1} + v^\top X_{t+1}\right)\right] = \exp\left(F_0(u,v) + X_t^\top F_X(u,v)\right), \tag{3}$$

where the scalar function $F_0(u,v)$ and the vector function $F_X(u,v)$ describe the exogenous dynamics of the process $Y_{t+1}^\top \equiv \left(\Delta c_{t+1},\ X_{t+1}^\top\right)$ and must satisfy $F_0(0,0) = F_X(0,0) = 0$. As discussed above, this setting nests existing General Equilibrium models based on Epstein-Zinn-Weil preferences, with or without long-run risks.

Using the standard Campbell-Shiller approximation, $r_{t+1}^e = \kappa_0 + \kappa_1 w_{t+1} - w_t + \Delta c_{t+1}$, the wealth-consumption ratio is given by,

$$w_t = A_0 + A_X^\top X_t,$$

for values of $w_t$ near its steady-state. See appendix A.1.1 for details. We show that the change of measure, $M_{t,t+1}$, from the historical probability, $\mathbb{P}$, to the risk-neutral probability, $\mathbb{Q}$, is then given by:

$$N_{t,t+1} = \exp\left(H_0 + H_X^\top X_t - \gamma\Delta c_{t+1} - p_X^\top X_{t+1}\right), \tag{4}$$

where $H_0 = -F_0(-\gamma, -p_X)$, $H_X = -F_X(-\gamma, -p_X)$ and $p_X = (1-\theta)\kappa_1 A_X$. Lemma 1 characterizes the joint conditional distribution of returns and state variables under $\mathbb{P}$ and $\mathbb{Q}$, respectively.

**Lemma 1 Laplace transform of excess returns**

*If the representative agent has utility function given by Equation 2, and if the joint conditional Laplace transform of consumption growth $\Delta c_{t+1}$ and the remaining $K$ state variables $X_{t+1}$ are given by Equation 3, then the joint conditional Laplace transform of $X_{t+1}$ and of excess returns $r_{t+1}^e$ from a claim on aggregate consumption is given by*

$$E_t^{\mathbb{P}}\left[\exp\left(u\ xr_{t+1}^e + v^\top X_{t+1}\right)\right] = \exp\left(F_0^{\mathbb{P}}(u,v) + X_t^\top F_X^{\mathbb{P}}(u,v)\right),$$

*under the historical measure, $\mathbb{P}$, for a constant scalar $u$ and a $K$-dimensional vector $v$. Similarly, the corresponding conditional Laplace transform under the risk-neutral measure, $\mathbb{Q}$, is given by*

$$E_t^{\mathbb{Q}}\left[\exp\left(u\ xr_{t+1}^e + v^\top X_{t+1}\right)\right] = \exp\left(F_0^{\mathbb{Q}}(u,v) + X_t^\top F_X^{\mathbb{Q}}(u,v)\right),$$

*The coefficients are given in Appendix A.1.2.*

Lemma 1 shows that the conditional Laplace transform of excess returns is exponential-affine under $\mathbb{P}$ and $\mathbb{Q}$. Essentially, this follows from the choice of historical dynamics for the state vector, given in Equation 3, and from the fact that the change of measure given by Equation 4 is also exponential affine. Proposition 1 applies Lemma 1 repeatedly to characterizes multi-horizon excess returns. It establishes that the cumulant-generating function of multi-horizon excess returns is affine for any investment horizon $\tau$.

**Proposition 1 Cumulants of multi-horizon excess returns**

*The cumulant-generating function of excess returns from the claim on aggregate consumption over an investment horizon $\tau$,*

$$xr_{t,t+\tau} \equiv \sum_{j=1}^{\tau} xr_{t+j},$$

*is given by*

$$\log E_t^{\mathbb{P}} \left[ \exp \left( u \ xr_{t,t+\tau} \right) \right] = F_{r,0}^{\mathbb{P}} \left( u; \tau \right) + X_t^{\top} F_{r,X}^{\mathbb{P}} \left( u; \tau \right),$$

*under the $\mathbb{P}$ measure and by*

$$\log E_t^{\mathbb{Q}} \left[ \exp \left( u \ xr_{t,t+\tau} \right) \right] = F_{r,0}^{\mathbb{Q}} \left( u; \tau \right) + X_t^{\top} F_{r,X}^{\mathbb{Q}} \left( u; \tau \right),$$

*under the $\mathbb{Q}$ measure with coefficients given in Appendix A.1.2.*

## 2.2 Bond Premium, Equity Premium and Variance Premium

An immediate corollary of Proposition 1 is that the Bond Premium and the Equity Premium over any investment horizon $\tau$, $BP(t,\tau)$ and $EP(t,\tau)$, respectively, are affine. We have that,

$$\begin{aligned} BP(t,\tau) &\equiv E_t^{\mathbb{P}} \left[ xr_{t,t+\tau}^{b} \right] \\ &= \beta_{b,0}(\tau) + \beta_b(\tau)^{\top} X_t, \end{aligned} \tag{5}$$

and

$$\begin{aligned} EP(t,\tau) &\equiv E_t^{\mathbb{P}} \left[ xr_{t,t+\tau}^{e} \right] \\ &= \beta_{ep,0}(\tau) + \beta_{ep}(\tau)^{\top} X_t. \end{aligned} \tag{6}$$

The bond premium and the equity premium are linear in the state variables whenever $\theta \neq 1$ and $A_X \neq 0$. These conditions implies that $p_X \neq 0$ in Equation 4, and, therefore, that the pricing kernel varies with $X_t$. Intuitively, the first condition implies that the agent has preference over the intertemporal resolution of uncertainty (i.e. $\gamma \neq \psi$). The second condition implies that $X_{t+1}$ affects the conditional distribution of future consumption growth.[13] These two conditions are the fundamental ingredients of long-run risk models. The

[13]Strictly speaking, the prices of risk associated with innovations to $X_{t+1}$ may differ from zero, with $\gamma \neq \psi$,

price of risk parameters $p_X$ are generally left unrestricted in reduced-form representations.

Proposition 1 also implies that the Variance Premium over any investment horizon $\tau$, $VRP\left(t,\tau\right)$, is affine,

$$VRP\left(t,\tau\right) \equiv E_t^{\mathbb{Q}}\left[\sum_{j=1}^{\tau}\sigma_{t+j}^2\right] - E_t^{\mathbb{P}}\left[\sum_{j=1}^{\tau}\sigma_{t+j}^2\right] \tag{7}$$

$$= \beta_{vp,0}(\tau) + \beta_{vp}(\tau)^\top X_t, \tag{8}$$

where $\sigma_t^2 = Var_t(xr_{t,t+j})$. The coefficients $\beta_{vp,0}(\tau)$ and $\beta_{vp}(\tau)$ depend on the structure of the model. The Variance Premium is zero in a LRR economy when the second conditional moment of consumption is constant under both measures. Moreover, the Variance Premium differs from zero but remains constant whenever the volatility of consumption volatility is constant.[14]

### 2.3 Variance Term Structure

Equations 5 and 6 characterize the equilibrium risk-return trade-offs in a broad class of economies with long-run risks. Different LRR models emphasize different risk factors, $X_t$, and imply different patterns of risk loadings, $\beta_{ep,X}$ but the risk premium dynamics are linear in every case. The coefficients of that relationship could be estimated directly via OLS if the risk factors, $X_t$ were observable. This would provide a test to discriminate across different specifications, or serve as guidance to investors. However, the risk factors proposed in the literature, including in reduced-form specifications, are latent or difficult to measure. For example, the expected consumption growth (Bansal and Yaron 2004), the volatility of consumption volatility (Bollerslev, Tauchen, and Zhou 2009) or the time-varying jump intensity (Drechsler and Yaron (2011), Eraker 2008) all escape direct measurement.

In contrast, the term structure of risk-neutral variance can be measured from option prices. Moreover, Proposition 1 implies that the conditional variance of excess returns over an horizon $\tau$ is also affine and given by:

$$Var_t^{\mathbb{Q}}(\tau) = \beta_{vr,0}\left(\tau\right) + \beta_{vr}(\tau)^\top X_t, \tag{9}$$

with coefficients given in Appendix A.1.2. This implies that measures of variance at different maturities display a factor structure with dimension $K$. This is similar to interest rate models where yields at different maturities sum the contributions of the real rate, inflation and compensation for risk. In most models, these are determined by a small set of economic variables (e.g. wealth, technology, habits) that are often not observed directly, at least at

but with a constant wealth-consumption ratio (and risk premium) if $U_t/c_t$ varies with $X_{t+1}$. This arises in the knife-edge case where $\psi = 1$.

[14] Bollerslev et al. (2009) provide a results similar to Equation 7. Their solution contrasts with our since they rely on an additional log-linearization and they only cover the special case $\tau = 1$.

the desired frequency. But the unobservable economic variables can be revealed via their effects on yields. This important insight is applicable in our context.

## 2.4 Revealing Risk Factors

The risk-neutral variance can reveal the effect of risk factors. However, the measured risk-neutral variance differs from the true value, $Var_t^{\mathbb{Q}}(\tau) = \tilde{Var}_t^{\mathbb{Q}}(\tau) + \nu_t(\tau)$, where we assume that the measurement error, $\nu_t(\tau)$, is uncorrelated with $\tilde{Var}_t^{\mathbb{Q}}(\tau)$. In other words, in contrast with computation of bond yields from bond prices, measurement errors cannot be neglected when computing variance from option prices. Stacking measurements across horizons $\tau = \tau_1, \ldots, \tau_q$, and using Equation 9, we have that,

$$\tilde{Var}_t^{\mathbb{Q}} + \nu_t = B_{0,vr} + B_{vr} X_t$$

where the $q \times 1$ vector, $B_{0,vr}$, stacks the constant $\beta_{vr,0}(\tau)$, and the $q \times K$ matrix $B_{vr}$ stacks the corresponding coefficients $\beta_{vr}(\tau)^\top$. Note that we typically have more observations along the term structure than there are underlying factors (i.e., $q > K$). We can then write,

$$\tilde{X}_t = -\bar{B}_{vr} B_{0,vr} + \bar{B}_{vr} \tilde{Var}^{\mathbb{Q}}(t) + \bar{B}_{vr} \nu_t, \tag{10}$$

where the $K \times q$ matrix $\bar{B}_{vr} = (B_{vr}^\top B_{vr})^{-1} B_{vr}^\top$ is the left-inverse of $B_{vr}$.[15]

Using Equations 5 and 6, and stacking across horizons, we have that,

$$BP_t = \Pi_{bp,0} + \Pi_{bp} \tilde{Var}_t^{\mathbb{Q}} + \nu_t^{bp} \tag{11}$$

$$EP_t = \Pi_{ep,0} + \Pi_{ep} \tilde{Var}_t^{\mathbb{Q}} + \nu_t^{ep}, \tag{12}$$

so that we can use the variance term structure as a signal for the underlying risk factors. Each line of the vector $\Pi_{ep,0}$ and of the matrix $\Pi_{ep}$ is given by,

$$\begin{aligned} \Pi_{ep,0}(\tau) &= \beta_{ep,0}(\tau) - \beta_{ep}(\tau)^\top \bar{B}_{vr} B_{0,vr} \\ \Pi_{ep}(\tau) &= \beta_{ep}(\tau)^\top \bar{B}_{vr}, \end{aligned} \tag{13}$$

respectively. This will be crucial in the following. The definitions of $\Pi_{bp,0}$ and $\Pi_{bp}$ are analogous. In practice, we do not observe the bond premium or the equity premium, but

[15] The left-inverse exists since we consider cases with $q > K$ and $B_{vr}$ has full (column) rank. If the latter conditions is not satisfied, then the loadings of the conditional variance, $Var^{\mathbb{Q}}(t, \tau)$ on the risk factors $X_{t,k}$ are not linearly independent. This implies that less than $K$ linear combinations of the risk factors can be revealed from the variance term structure. In other words, some linear combinations of the risk factors are unspanned by the variance term structure. In this case, we redefine the risk vector in Equation 10 to be $X_t^{vr}$ that only contain those $K^{vr} < K$ linear combinations that are spanned. This issue also arises in the interest rate literature and as been discussed in Duffee (2011).

we can only measure ex-post excess returns,

$$xr^e_{t,t+\tau} = EP(t,\tau) + \epsilon^e_{t,t+\tau}$$
$$xr^b_{t,t+\tau} = BP(t,\tau) + \epsilon^b_{t,t+\tau},$$

which can be re-written as:

$$xr^b_{t+} = \Pi_{bp,0} + \Pi_{bp}\tilde{Var}^{\mathbb{Q}}_t + (\nu^{bp}_t + \epsilon^b_{t+}) \tag{14}$$
$$xr^e_{t+} = \Pi_{ep,0} + \Pi_{ep}\tilde{Var}^{\mathbb{Q}}_t + (\nu^{ep}_t + \epsilon^e_{t+}), \tag{15}$$

where the $xr_{t+}$ notation signals that we have stacked ex-post excess returns at different horizons.

Equations 14-15 form the basis of our empirical investigation below. But note that $\Pi_{bp}$ and $\Pi_{ep}$ have rank at most $K$ irrespective of the number of horizons $\tau$ used at estimation. This will be crucial in the following since it implies these equations cannot be estimated via standard OLS. Before we address this, the next Section introduces the data.

# 3 Data and Measurement

## 3.1 Excess Returns

We use the CRSP data set to compute end-of-the-month equity returns on the SP500 at horizons of 1, 2, 3, 6, 9 and 12 months. Longer-horizon returns are obtained from summing monthly returns. We use the Fama-Bliss zero coupon bond prices from CRSP to compute bond excess returns. Excess returns are computed using risk-free rates from CRSP.[16]

## 3.2 Excess Variance

As in the case of returns, longer-horizon realized variances are obtained from summing monthly realized variances.[17] We follow Britten-Jones and Neuberger (2000) to compute expected integrated variance under the risk-neutral measure (see Equation 7) from option prices. The excess variance is the difference between the realized variance under the historical measure and the ex-ante measure of conditional variance under the risk-neutral measure. This definition is completely analogous to the definition of excess returns. Explicitly, the excess variance is given by:

$$xv^e_{t,t+\tau} \equiv \tilde{E}^{\mathbb{Q}}_t \left( \sum_{j=1}^{\tau} \sigma^2_{r,t+j} \right) - \sum_{j=1}^{\tau} \sigma^2_{r,t+j}, \tag{16}$$

[16] The Fama-Bliss T-bill file covers maturities from 1 to 6 months. We use the 1-year rate from the Fama-Bliss zero-coupon files. The 9-month T-bill rate is interpolated when necessary.

[17] We thank Hao Zhou for making end-of-the-month SP500 realized variance data available on his web site.

where $\sigma^2_{r,t+j}$ is the realized variance in period $t+j$ and $\tilde{E}^{\mathbb{Q}}_t\left(\sum_{j=1}^{\tau}\sigma^2_{r,t+j}\right)$ is measured ex-ante from option prices.

### 3.3 Risk-Neutral Variance

We use the OptionMetrics database of European options written on the SP 500 index. We first construct a weekly sample of closing bid and ask prices observed each Wednesday. This mitigates the impact of intra-weekly patterns but includes 328,626 observations. Consistent with the extant literature, we restrict our sample to out-of-the-money call and put options. We also exclude observations with no bid prices (i.e. price is too low), options with less than 10 days to maturity, options with implied volatility above 70% and options with zero transaction volume. Finally, we exclude observations that violate lower and upper bounds on call and put prices. The OptionMetrics database supplies LIBOR and EuroDollar rates. To match an interest rate with each option maturity, we interpolate under the assumption of constant forward rates between available interest rate maturities. We also assume that the current dividend yield on the index is constant through the options' remaining maturities.[18] Finally, we restrict our attention to a monthly sample (see Appendix A.3). This yields 85,385 observations covering the period from January 1996 to October 2008. Table 1 contains the number of option contracts across maturity and moneyness groups. The sample provides a broad coverage of the moneyness spectrum at each maturity.

### 3.4 Summary Statistics

We then rely on the non-parametric approach of Bakshi and Madan (2000) to measure the conditional variance implicit in option prices at maturities of 1, 2, 3, 6, 9, 12, and 18 months. These correspond to the maturity categories available on the exchange (see Appendix A.4).[19] Table 2 provides summary statistics of variance across maturities. Risk-neutral variance is persistent with autocorrelation coefficients between 0.73 and 0.87 across maturities. The term structure is upward sloping on average but with an inverted U-shape. The volatility of risk-neutral variance peaks at 2 months and then gradually declines with maturity. In other words, the average variance of stock returns increases with maturity but, on the other hand, the conditional variance itself is less volatile for longer returns horizons. It is also more symmetric and has smoother tails for longer horizons.

### 3.5 Principal Components

Variance measures are highly correlated across maturities (not reported). For example, the correlation between 1-month and 2-month ahead risk-neutral variances (i.e. $Var^{\mathbb{Q}}(t,1)$ and

[18] See OptionMetrics documentation on the computation of the index dividend yield.

[19] We originally included the 24-month maturity category. However, its summary statistics contrast with the broad patterns drawn in other categories. For this maturity, risk-neutral variance is more skewed to the right, has fatter tails and is less persistent. Moreover, it is less correlated with other maturities. We consider these results a reflection of higher measurement errors and exclude this category in the following.

$Var^{\mathbb{Q}}(t,2)$) is 0.88 while the correlation between 1-month and 1-year ahead variances is 0.69. This suggests that a few systematic factors can explain most of the variation across maturities. Panel B of Table 2 reports the results from a Principal Component Analysis (PCA), which is a simple way to summarize this factor structure. The first three principal components explains 88%, 6% and 3% of the term structure of the risk-neutral variance, respectively, and together explain 97.4% of total variation.

These components reflect systematic variation across the variance term structure. The first component's loadings range from 0.31 to 0.44 with an inverted U shape across maturities. In other words, most of the variations in the risk-neutral variances can be summarized by a change in the level and curvature of its term structure. Next, the second component is similar to a slope factor. Its loadings increase, from -0.57 to 0.49, and pivot around zero near the 6-month maturity. The third component's loadings draw a curvature pattern. The correlation between the first component and a measure of the level, $L_t = \tilde{Var}^{\mathbb{Q}}(t,6)$, is 0.98, the correlation between the second component and a measure of the slope, $S_t = \tilde{Var}^{\mathbb{Q}}(t,18) - \tilde{Var}^{\mathbb{Q}}(t,1)$ is -0.90, and the correlation between the third component and a measure of the curvature, $C_t = 2\tilde{Var}^{\mathbb{Q}}(t,6) - \tilde{Var}^{\mathbb{Q}}(t,18) - \tilde{Var}^{\mathbb{Q}}(t,1)$, is 0.80.

# 4 Variance Risk-Returns Trade-Offs

Section 2 shows that a broad family of affine general equilibrium models, or affine reduced-form, models contains at its core the implication that a few linear combinations from the term structure of variance can be used to predict returns. Consistent with theory, Section 3 shows that the term structure of variance can be summarized by its leading principal components. This Section analyzes the relationship between the variance factors and the compensation for risk.

## 4.1 Estimating the number of factors in the variance term structure

We first ask how many linear combinations from the variance term structure summarize its information content for the bond and equity premia. In other words, we want to estimate the rank of the coefficient matrix, $\Pi$, in multivariate regressions with the following general form

$$Y_{t+} = \Pi \tilde{Var}_t^{\mathbb{Q}} + \Psi Z_t + \epsilon_{t+}. \tag{17}$$

where $Y_t$ is a vector of excess returns, $\tilde{Var}_t^{\mathbb{Q}}$ is a $q \times 1$ vector of risk-neutral variances and the vector $Z_t$ contains any other regressors, including the constant. This nests Equations 14 and 15. Recall that Equation 13 shows that $\Pi$ does not generally have full rank. The statistical literature on Sufficient Dimension Reduction provides a useful approach to estimating this rank.

Cook and Setodji (2003) introduces a model-free test of the null hypothesis that the rank is $r$ (i.e., $\mathrm{H}_0 : \mathrm{rank}\,\Pi = r$) against the alternative that the rank is strictly greater than $r$. The modified Cook and Setodji test-statistic, $\tilde{\Lambda}_r$, is available in closed-form and has a $\chi^2$ asymptotic distribution with known degrees of freedom. In particular, this test does not require Gaussian innovations in Equation 17. The test is also robust against departure from linearity.[20] Cook and Setodji (2003) propose the following iterated algorithm as an estimator for the rank of $\Pi$.

1. Initialize the null hypothesis with $\mathrm{H}_0^{(0)} : \mathrm{rank}\,\Pi = r^{(0)} = 0$.

2. For the hypothesis $\mathrm{H}_0^{(i)}$, compare the $\tilde{\Lambda}_{r^{(i)}}$ statistics with the chosen cut-off from the $\chi_g^2$ distribution, e.g., 5%.

3. If the probability of observing $\tilde{\Lambda}_{r^{(i)}}$ is lower than the cut-off, then reject the null, conclude that $\mathrm{rank}\,\Pi > r^{(i)}$, and repeat the test under a new null hypothesis where the rank is incremented, i.e., $r^{(i+1)} = r^{(i)} + 1$.

4. Otherwise, conclude that $\mathrm{rank}\,\Pi = r^{(i)}$. That is, there is insufficient evidence against $\mathrm{rank}\,\Pi = r^{(i)}$ but, yet, we have rejected $\mathrm{rank}\,\Pi < r^{(i)}$.

## 4.2 Estimating reduced-rank multivariate regressions

As stated above, Equations 14 and 15 form the basis of our empirical investigation and, for a given rank, $r$, they correspond to multivariate Reduced-Rank Regressions (RRR) for which estimators and the associated inference theory are available since at least Anderson (1951). In particular, for a given estimate of the rank, $r$, the $p \times q$ matrix, $\Pi$, can be rewritten as a product, $\Pi = A\Gamma$, where $A$ and $\Gamma$ have dimensions $(p \times r)$ and $(r \times q)$, respectively, and where $r < min(p, q)$.[21] Then, we can re-write Equation 17 as,

$$Y_t = A\Gamma \tilde{Var}_t^{\mathbb{Q}} + \Psi Z_t + \epsilon_t, \tag{18}$$

and the RRR estimators of $A$, $\Gamma$ and $\Psi$ are given from the solution to

$$\arg\min_{A,\Gamma,\Psi} \mathrm{trace}\left(\sum_{t=1}^{T} \epsilon_t \epsilon_t^\top\right), \tag{19}$$

[20] $\tilde{\Lambda}_r$, has a $\chi^2$ asymptotic distribution with $g$ degrees of freedom, where $\tilde{\Lambda}_r$ and $g$ depend on the data and are available in closed-form. If $E[Y_t|X_t]$ is not linear in $X_t$, in contrast with Equation 17, then inference about the rank of $\Pi$ from estimates of Equation 17 may still be used to form inference about the dimension of the Central Mean Subpace (CMS) of $Y_t|X_t$. A subspace $\mathcal{M}$ of $\mathbb{R}^q$ is a mean subspace of $Y_t|X_t$ if $E[Y_t|X_t]$ is a function of $M^\top X_t$ where the $q \times r$ matrix $M$ is a basis for $\mathcal{M}$. The CMS is the intersection of all mean subspaces. See Cook and Setodji (2003).

[21] See Reinsel and Velu (1998) for a textbook treatment of RRR and a discussion of existing applications in tests of asset pricing models (e.g. Bekker et al. (1996) and Zhou (1995)). Anderson (1999) provides a theory of inference under general (e.g. not Gaussian) conditions. Hansen (2008) provides a recent formulation of the estimator. The OLS regression emerges when $k = \min(p; q)$ or, trivially, when $r = 0$ and the regressors are irrelevant.

with closed-form expressions given in Appendix A.5. Note that that the estimated factors, $\hat{\Gamma}\tilde{Var}_t^{\mathbb{Q}}$, can be very different than the leading principal components of $\tilde{Var}_t^{\mathbb{Q}}$.[22] Finally, $A$ and $\Gamma$ are not separately identified, and we choose that rotation which yields orthogonal factors. This is analogous to the standard identification choice in Principal Component Analysis.

## 4.3 The advantages of reduced-rank regressions

Our methodological approach imposes the factor structure predicted by theory but remains agnostic regarding other structural assumptions. This approach is in line with Cochrane (2011) who emphasizes the need to uncover the factor structure behind time-varying expected returns. It is also closely related to Cochrane and Piazzesi (2008) who show that a single factor from forward rates is sufficient to summarize the predictability of bonds with different maturities.

In the same spirit, we test the joint hypothesis of linearity and reduced-rank structure without any other joint hypothesis about the number and the dynamics of state variables, the conditional distribution of shocks, or the preference of the representative agent. Otherwise, the test will over-reject the null hypothesis of a given low number of factors, even if it holds in the data, when these maintained hypotheses are not supported by the data. Similarly, estimation based on the Kalman filter will be severely biased if the maintained structural or distributional assumptions are not supported in the data. In contrast, our approach does need additional hypotheses but, instead, exploits the fundamentally multivariate nature of the problem.[23]

## 4.4 Predictability Results

### *4.4.1 Excess Returns Predictability*

Formally, we consider different versions of a joint model for the bond, equity premium, and variance premia,

$$xr_{t+} = \Pi_0 + A\Gamma\ \tilde{Var}_t^{\mathbb{Q}} + \epsilon_{t+}, \tag{20}$$

where we stack Equations 14 and 15. Line-by-line estimation is not feasible when $A\Gamma$ does not have full rank. Panel A of Table 3 displays the $p$-values associated with the Cook-Setodji statistics, $\tilde{\Lambda}_r$, for different ranks ranging from 1 to $q$. The tests reject that $\operatorname{rank}\Pi = 0$ or

[22] See, for example, the discussion by Dennis Cook in his Fisher Lecture (Cook 2007) and in particular, this quote from Cox (1968) "... there is no logical reason why the dependent variable should not be closely tied to the least important principal component [of the predictors]." (Cochrane and Piazzesi 2005) is a case in point in Finance in the context of bond returns predictability. Their returns-forecasting factor is a linear combination of forward rates that is only weakly spanned by the leading principal components of forward rates.

[23] In particular, our testing and estimation procedure could not be applied to each line of Equation 17 separately.

rank $\Pi = 1$. But we do not reject that rank $\Pi = 2$. The results suggest that 2 risk factors are sufficient to summarize the predictive content of the variance term structure.

Panel B reports the $R^2$s of predictability regressions of bond excess returns across different rank hypotheses. In particular, the $R^2$s in the case where the rank is $r = 2$ are 7.3%, 6.6%, 5.9% and 5.5% for annual returns on bonds with 2, 3, 4, and 5 years to maturity, respectively. Compare this with the case where $r = 7$ and where the model corresponds to standard OLS predictive regressions. The $R^2$s in this case are 11.5%, 10.1%, 8.8% and 7.9%, respectively. Similarly, Panel C reports $R^2$s for equity returns predictability. For $r = 2$, the $R^2$s are 3.1% and 6.3% for 1-month and 2-month excess returns. It then declines smoothly to 3.6% at the 12-month horizon. In all cases, there is little gain from increasing the rank from $r = 2$ to $r = 7$ given the large increase in the number of parameters.[24]

Estimation of the 14 unrestricted univariate regressions on 7 variance measures uses 98 parameters. In contrast, allowing for a factor structure in expected returns is parsimonious and yields disciplined results. Estimation of the multivariate system with only two linear combinations of variance reduces the number of parameters to 42. It is also more informative relative to the OLS. Standard OLS inference, based on $F$-statistics, rejects the null hypothesis that the variance term structure is irrelevant (unreported). The Cook-Setodji statistic above also leads to a rejection that the rank is $r = 0$. But OLS misses the factor structure in expected returns. In contrast, following the Cook-Setodji procedure, we conclude that two factors are sufficient and that the increased predictive power of unrestricted regressions ($r = 7$) can be attributed to sampling variability.

#### *4.4.2 Excess Variance Predictability*

Equation 7 relates the variance premium to $X_t$ and provides a revealing way to check whether the estimated risk factors truly reflect compensation for risks. We can write the variance risk premium in terms of the variance term structure,

$$xv^e_{t+} = \Pi_{vrp,0} + \Pi_{vrp}\ \tilde{Var}^{\mathbb{Q}}_t + (\nu^{vrp}_t + \epsilon^v_{t+}), \tag{21}$$

where the definitions of $\Pi_{vrp,0}$ and $\Pi_{vrp}$ are analogous to those given in Equation 13 for excess returns and $xv^e_{t,t+\tau}$ is the ex-post excess variance over an horizon $\tau$. Theory predicts that the same risk factors can be used to predict excess returns and excess variance. We can then combine Equation 21 with the linear combinations of variance estimated above, $\hat{\Gamma}\tilde{Var}^{\mathbb{Q}}_t$, and check that they also predict excess variance. This is akin to an out-of-sample robustness check since the excess variance was not used to estimate these factors.

Specifically, Table 4 reports estimates and $R^2$s from the following OLS regressions,

$$xv^e_{t+}(\tau) = \Pi_{vrp,0}(\tau) + a_{1,vrp}(\tau)\hat{\Gamma}_1\ \tilde{Var}^{\mathbb{Q}}_t + a_{2,vrp}(\tau)\hat{\Gamma}_2\ \tilde{Var}^{\mathbb{Q}}_t + \epsilon_{t+}(\tau), \tag{22}$$

[24] We do not report estimates of $A$ and $\Gamma$ since the orthonormal rotation used for estimation has no special economic meaning.

where we use estimates of $\hat{\Gamma}$ obtained above in the case with $r = 2$. The results are striking. Together, the two linear combinations that were estimated to predict the bond and equity premia also predict the variance premium with $R^2$s ranging from 6.2%, 9.5%, 9,0% and 10.1% at horizons of 1, 2, 3 and 6 months, respectively, and then to 8.7% and 2.7% at horizons of 9 and 12 months, respectively. Looking at individual coefficients shows that each of the estimated linear combinations plays an important role. The first plays a significant role in the variations of the variance premium at relatively short horizons, up to three months ahead, while the second linear combination plays a significant role at longer horizons, beyond three months.

It may appear tempting to use Equation 21 along with bond and equity returns in a RRR regression. However, the excess variance equation presents an econometric difficulty. The measurement errors in excess variance that arise because we measure $\tilde{E}_t^{\mathbb{Q}}\left(\sum_{j=1}^{\tau}\sigma_{t+j}^2\right)$ from option prices are correlated with the measurement errors in $\tilde{Var}_t^{\mathbb{Q}}$, which is also obtained from option prices. Therefore, this equation cannot be used directly at estimation.[25]

# 5 Term Structure of Higher-Order Cumulants

We show that measures of higher order risks can also be used to reveal risk factors. Empirically, we find that, the skewness and kurtosis term structures predict the bond premium, the equity premium and the variance premium. Their predictive content is similar to that of the variance term structure and can be summarized by 2 risk factors. Consistent with theory, combining measures of variance, skewness and of kurtosis improves predictability only marginally and, strikingly, the predictive content of this broad information set can still be summarized by two factors.

## 5.1 Higher-Order Cumulants in Equilibrium

The variance term structure may fail to reveal all risk factors. This may arise if some factors do not affect the variance, or if their effects are small relative to the measurement errors in the variance or to the innovations in returns. It may be possible to increase the efficiency of our estimates and parse the variance term structure to find additional factors. But this neglects low-hanging fruits. An alternative way is to broaden the information set include other measurements where the effect of other risk factors may be more easily measured. Looking back, Proposition 1 implies that every cumulant[26] of returns is affine in the state

[25] Stambaugh (1988) provides an example where measurement errors due to bid-ask spreads in bond prices leads to over-rejection of small factor structure and wrongly favors larger factor structure (his Section 4.4, p.58). Cochrane and Piazzesi (2005) provide a similar example. They use a single factor from forward rates to study bond returns. They show that the single-factor restriction is rejected statistically but that deviations from a single-factor structure are economically insignificant.

[26] Recall that the first cumulant corresponds to the mean, the second cumulant corresponds to the variance, the third cumulant corresponds to the third central moment and provides a measure of skewness, while the fourth cumulant corresponds to the fourth central moments minus 3 times the squared variance and provides a measure of the tails.

vector,

$$M_{t,n}^{\mathbb{Q}}(\tau) = \beta_{n,0}(\tau) + X_t^{\top}\beta_{n,X}(\tau),$$

for any returns horizon $\tau$, and where coefficients depend on the underlying model.[27] Then, an argument similar to Section 2.4 shows that higher-order cumulants can also be used reveal $X_t$,

$$\tilde{X}_t = -\bar{B}_n B_{0,n} + \bar{B}_n \tilde{M_{t,n}}^{\mathbb{Q}} + \bar{B}_n \nu_{n,t}. \tag{23}$$

In the following, we follow a path parallel to the previous section and construct model-free measures of returns cumulants of order 3 and 4(see Appendix A.4). We also exchange a slight abuse of terminology for ease in the exposition and label these cumulants skewness and kurtosis, respectively.[28]

## 5.2 Summary Statistics and Factor Structure

Panel A and Panel B of Table 5 presents summary statistics of the conditional skewness and kurtosis of returns, respectively. The average distribution of returns implicit in index option is left-skewed and has fat tails. The average skewness lies below zero and slopes downward with the horizon. On the other hand, the average tail is fatter at longer horizons. Skewness and kurtosis are persistent, especially at intermediate horizons.

The correlation matrices (Panel C and Panel D) suggest a low-dimensional factor structure as in the case of risk-neutral variance. Panel E and Panel F present PCA results for the term structure of skewness and kurtosis, respectively. The first three principal components of skewness explain 67%, 15% and 12% of total variations, respectively, and together explain 93%. Similarly, the first three principal components of kurtosis explains 65%, 19% and 12% of total variation, respectively. As for the variance, the loadings of reveal that the leading components of skewness and kurtosis have a systematic effect on their respective term structure.

## 5.3 Predictability results

We estimate different variations of the following multivariate regression,

$$xr_{t+} = \Pi_0 + A\Gamma\ F_t + \epsilon_{t+} \tag{24}$$

where, as above, $xr_{t+}$ stacks 4 excess bond returns and 6 excess equity returns. We consider different combinations of the variance, skewness and of kurtosis term structure to construct the regressors, $F_t$.

[27]The scalar coefficient, $\beta_{n,0}(\tau)$, and the vector coefficient, $\beta_{n,X}(\tau)$, are defined as

$$\beta_{n,0}(\tau) = \mathcal{D}^n F_{r,0}^{\mathbb{Q}}(0;\tau) \quad \text{and} \quad \beta_{n,X}(\tau) = \mathcal{D}^n F_{r,X}^{\mathbb{Q}}(0;\tau),$$

where the matrix jacobian operator $\mathcal{D}^n$ is defined in Appendix A.1. These can typically be computed in closed-form, up to the usual recursions on $\tau$.

[28]The conventional measure of skewness and kurtosis are not affine in the risk factors.

### 5.3.1 *Excess returns with skewness or kurtoris*

We first consider each term structure separately. Panel A of Table 6 presents results. First, model $V(2)$ uses the term structure of variance as predictors (i.e., $F_t = \tilde{Var}_t^{\mathbb{Q}}$). This reproduces a subset of the results presented above (Table 3) and provides a point of comparison for models using skewness or kurtosis as predictors. Second, Model $S(2)$ only includes the term structure of skewness (i.e., $F_t = \tilde{Skew}_t^{\mathbb{Q}}$). Third, model $K(2)$ only includes the term structure of kurtosis (i.e., $F_t = \tilde{Kurt}_t^{\mathbb{Q}}$). In model $S(2)$, the $p$-value is 6.1% for the null that $r = 1$ and 38.2% for the null that $r = 2$. Similarly, for the $K(2)$ model, the $p$-value is 7.9% for the null that $r = 1$ and 32.2% for the null that $r = 2$. Hence, the test based on each of these higher moments come close to reject the rank-one restrictions in favor of a higher rank while the rank-two restrictions is clearly not rejected. Nonetheless, we report estimation results based on $r = 2$ for comparison because more general models combining information from different term structures consistently reject the case $r = 1$ (see below). The results show that the ability to predict bond and equity excess returns, as measured by the $R^2$s, is strikingly similar whether we use any one of the variance, skewness and kurtosis term structures. This is consistent with theory. If anything, skewness and kurtosis appear to be slightly more informative about bond returns while variance appears to be slightly more informative about equity returns. We stress that this does not imply that only the variance matters. The term structure of risk-neutral variance combines information about historical variance, skewness and kurtosis (Bakshi and Madan 2006), and changes in the prices of risk.

### 5.3.2 *Combining variance, skewness and kurtosis term structure*

The VSK(2,2) model combines the two risk factors estimated separately from each of the variance, skewness and kurtosis term structures. Hence, this uses 6 predictors and asks whether these risk factors add up to more than two factors when combined in the same model. The evidence is unambiguous. The $p$-value is 1.1% for the null that $r = 1$ and 32.6% for the null that $r = 2$. Again, this is consistent with theory. The predictive content available from the term structure of different risk measures is broadly overlapping. As expected, estimation in the case $r = 2$ yields $R^2$s that are very close to the highest value obtain above. Of course, we could (at least) reach these values by setting $r = 6$. What is unexpected is that we can summarize these 6 risk factors into two with little loss of predictive ability.

The VSK(2,2) is a second-stage estimation that uses factors obtained in a first-stage procedure. Next, we introduce model VSK(7,2) that combines the entire variance, skewness and kurtosis term structures in a single RRR step. This is an alternative way to ask whether the risk factors measured from different term structures add up to more than two factors. Model VSK(7,2) model is estimated in one step but, on the other hand, it is more exposed to over-fitting given the large number of regressors. Nonetheless, these model yield consistent evidence. The $p$-value is 1.4% for the null that $r = 1$ and 9.7% for the null that $r = 2$. The $p$-value has decreased substantially but Cook and Setodji (2003) report that this test tends to over-reject when the number of predictors and regressors is particularly high like in model VSK(7,2). This bias our result toward concluding in favor of a greater number of factor. Nonetheless, there is a substantial increase in predictability in the case $r = 2$ when we combine all the risk measures. $R^2$s now range from 17% to 22% in the case of bond returns (compare to the 9%-10% of more parsimonious models) and from 6% to 18% in the case of equity returns (compare to the 3%-8%). The next section uses the variance premium as an out-of-sample check.

#### *5.3.3 Excess variance*

We check that the in-sample predictability obtained from bond and stock returns extends to the variance premium. Panel A of Table 6 presents results of excess variance predictability regressions. The results are broadly consistent across all models, the $R^2$s have an inverted U-shape across horizons, reaching a maximum close to 10% at intermediate horizons between 3 and 6 months. This holds whether the risk factors were extracted from the variance, skewness or kurtosis term structure. Once again, the theoretical prediction is supported in the data. In particular, there is no improvement in excess variance predictability for the VSK(7,7) model. Hence, this out-of-sample exercise suggests that some of the increased excess returns predictability obtained above for the VSK(7,7) model is due to in-sample over-fitting.

## 6 Conclusion

We find that the term structure of risks can be used to reveal risk factors that are important drivers of bond premium, equity premium and variance premium variations. Consistent with theory, we find that a small number of factors, two, summarize the relationship between the equity premium, the bond premium and the variance implicit in option prices. The Long-Run Risk literature emphasizes slowly moving factors that affect the future conditional distribution of consumption growth. But, almost by construction, these factors are difficult to measure from the macro data. Similarly, reduced-form parametrizations of the stock returns process introduce latent variations in stochastic volatility or jump intensity. In each case, the risk-return trade-offs are difficult to measure and present a challenge to the econometrician. On the other hand, model-free measures of risk-neutral variance, and higher-order moments, are available from option prices.

Our results open several avenues for future research. First, does the predictive content from the term structure of option prices extend to other markets? In particular, is the valuation of individual firms' corporate bonds and equities related to the same risk factors? Similarly, are the risk premia implicit in other derivative markets (e.g., interest rate or

FX derivative markets) related to the risk factor from index options? Second, how can we reconcile the factor structure common to the variance, skewness and kurtosis term structure with its predictive content for returns within a reduced-form form asset pricing specification? Finally, given an appropriate reduced-form specification that matches the stylized facts uncovered here, what equilibrium model can relate these facts to preferences and economic fundamentals?

## References


Amengual, D. (2009). The term structure of variance risk premia. Princeton University.

Anderson, T. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *The Annals of Mathematical Statistics 22*, 327–351.

Anderson, T. (1999). Asymptotic distribution of the reduced rank regression estimator under general conditions. *The Annals of Statistics 27*, 1141–1154.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The cross-section of volatility and expected returns. *The Journal of Finance 61*(1).

Backus, D., M. Chernov, and I. Martin (2010). Disasters implied by equity index options. *The Journal of Finance*. forthcoming.

Bakshi, G., N. Kapadia, and D. Madan (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies 16*, 101–143.

Bakshi, G. and D. Madan (2000). Spanning and derivative-security valuation. *Journal of Financial Economics 58*, 205–238.

Bakshi, G. and D. Madan (2006). A theory of volatility spreads. *Management Science 52*, 1945–1956.

Bakshi, G., G. Panayotov, and G. Skoulakis (2011). Improving the predictability of real economic activity and asset returns with forward variances inferred from option portfolios. *Journal of Financial Economics 100*(3), 475 – 495.

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

Bekker, P., P. Dobbelstein, and T. Wansbeek (1996). The APT model as a reduced-rank regression. *Journal of Business and Economics Statistics 14*, 199–202.

Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *RFS 22*(11), 4463–4492.

Britten-Jones, M. and A. Neuberger (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance 55*(2), 839–866.

Campbell, J. and L. Hentschel (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics 31*, 281–331.

Carr, P. and L. Wu (2011). Variance risk premiums. *Review of Financial Studies 22*(3), 1311–1341.

Chamberlain, G. (1988). Asset pricing in multiperiod securities markets. *Econometrica 56*(6), pp. 1283–1300.

Chang, B.-Y., P. Christoffersen, and K. Jacobs (2011). Market skewness risk and the cross-section of stock returns. *JFE*. forthcoming.

Christoffersen, P., K. Jacobs, and B.-Y. Chang (2011). Forecasting with option-implied information. Volume 2 of *Handbook of Economic Forecasting*. Elsevier.

Cochrane, J. and M. Piazzesi (2008). Decomposing the yield curve. Graduate School of Business, University of Chicago.

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

Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*(1), 138–160.

Constantinides, G. M. and A. Ghosh (2011). Asset pricing tests with long run risks in consumption growth. *The Review of Asset Pricing Studies 1.*

Cook, R. (2007). Fisher lecture : Dimension reduction in regression. *Statistical Science 22.*

Cook, R. D. and C. M. Setodji (2003). A model-free test for reduced rank in multivariate regression. *Journal of the American Statistical Association 98*(462), 340–351.

Darolles, S., C. Gourieroux, and J. Jasiak (2006). Structural laplace transform and compound autoregressive models. *Journal of Time Serie Analysis 27*, 477–503.

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

Duffee, G. (2011). Information in (and not in) the term structure. *Review of Financial Studies. forthcoming.*

Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica 68*(6), 1343–1376.

Eraker, B. (2008). Affine general equilibrium models. *Management Science 54*, 2068–2080.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987). Expected stock returns and volatility. *Journal of Financial Economics 19*, 3–30.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). There is a risk-return trade-off after all. *Journal of Financial Economics 76*, 509–548.

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

Guo, H. and R. Savickas (2006). Idiosyncratic volatility, stock market volatility, and expected stock returns. *Journal of Business and Economic Statistics 24*(1), 43–56.

Hansen, P. (2008). Reduced-rank regression: A useful determinant identity. *Journal of Statistical Planning and Inference 138.* Special Issue in Honor of Theodore Wilbur Anderson, Jr.

Leippold, M., L. Wu, and D. Egloff (2007). Variance risk dynamics, variance risk premia, and optimal variance swap investments.

Ludvigson, S. C. and S. Ng (2005). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics 83*, 171–222.

Martin, I. (2010). Consumption-based asset pricing with higher cumulants. NBER 16153.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica 41*, 867–887.

Mueller, P., A. Vedolin, and H. Zhou (2011). Short-run bond risk premia. Technical report, London School of Economics.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*, 347–370.

Reinsel, G. and R. Velu (1998). *Multivariate Reduced-Rank Regression*, Volume 136 of *Lectures Notes in Statistics*. Springer.

Stambaugh, R. F. (1988). The information in forward rates: Implications for models of the term structure. *Journal of Financial Economics 21*(1), 41 – 70.

Turner, C. M., R. Startz, and C. R. Nelson (1989). A Markov model of heteroskedasticity, risk and learning in the stock market. *Journal of Financial Economics 25*, 3–22.

Zhou, G. (1995). Small sample rank tests with applications to asset pricing. *Journal of Empirical Finance 2*, 71–93.

Zhou, H. (2011). Variance risk premia, asset predictability puzzles and macroeconomic uncertainty. Working Paper.

# A Appendix

## A.1 The Distribution of Multi-horizon returns in Equilibrium

### *A.1.1 Affine general equilibrium models*

We consider an Affine General Equilibrium Model (AGEM) similar to Eraker (2008). Suppose that the state of the economy can be summarized by a Markov process $Y_{t+1} \equiv \left(\Delta c_{t+1}, X_{t+1}^{\top}\right)^{\top}$ where $\Delta c_{t+1}$ is the consumption growth process and $X_{t+1}$ is a vector of $K$ (observed and unobserved) state variables independent of consumption growth. The moment-generating function of this state vector under the physical measure is given by

$$E_t\left[\exp\left(x\Delta c_{t+1} + y^{\top} X_{t+1}\right)\right] = \exp\left(F_0\left(x, y\right) + X_t^{\top} F_X\left(x, y\right)\right),$$

where the scalar function $F_0\left(x, y\right)$ and the vector function $F_X\left(x, y\right)$ describe the exogenous dynamics of the vector process $Z_{t+1}$. Assume, further, that the representative agent has recursive preferences of Epstein-Zin-Weil type. Consequently, the logarithm of the intertemporal marginal rate of substitution is given by

$$s_{t,t+1} = \theta \ln \delta - \frac{\theta}{\psi}\Delta c_{t+1} - (1-\theta)\, r_{t+1}, \tag{25}$$

where $r_{t+1}$ is the return to the aggregate consumption claim. Using the standard Campbell-Shiller approximation, $r_{t+1} = \kappa_0 + \kappa_1 w_{t+1} - w_t + \Delta c_{t+1}$, the log price-consumption ratio $w_t$ can be well-approximated by an affine function of the vector state variable $X_t$ as

$$w_t = A_0 + A_X^{\top} X_t, \tag{26}$$

where the scalar coefficient $A_0$, and the vector coefficient $A_X$ depend on model and preference parameters. Solving for these coefficients is standard in the literature. The (log) stochastic discount factor can then be re-written as

$$\begin{aligned} s_{t,t+1} = \theta \ln \delta &- (1-\theta)\left(\kappa_0 + (\kappa_1 - 1)\, A_0 - A_X^{\top} X_t\right) \\ &- \gamma \Delta c_{t+1} - (1-\theta)\, \kappa_1 A_X^{\top} X_{t+1}, \end{aligned} \tag{27}$$

and the model-implied log risk-free rate is given by,

$$r_{f,t+1} = B_0 + B_X^{\top} X_t, \tag{28}$$

where the scalar coefficient $B_0$ and the vector coefficient $B_X$ depend on the exogenous dynamics and preference parameters,

$$B_0 = -\theta \ln \delta + (1-\theta)\left(\kappa_0 + (\kappa_1 - 1)\, A_0\right) - F_0\left(-\gamma, -(1-\theta)\, \kappa_1 A_X\right) \tag{29}$$

$$B_X = -(1-\theta)\, A_X - F_X\left(-\gamma, -(1-\theta)\, \kappa_1 A_X\right). \tag{30}$$

It follows that, in this economy, the change-of-measure from the historical probability to the risk-neutral probability is given by

$$M_{t,t+1} = \exp\left(s_{t,t+1} + r_{f,t+1}\right) = \exp\left(H_0 + H_X^{\top} X_t - \gamma \Delta c_{t+1} - p_X^{\top} X_{t+1}\right), \tag{31}$$

where

$$H_0 = -F_0\left(-\gamma, -p_X\right), \quad H_X = -F_X\left(-\gamma, -p_X\right) \quad \text{and} \quad p_X = (1-\theta)\, \kappa_1 A_X. \tag{32}$$

### *A.1.2 Cumulants term structure*

To compute risk-neutral cumulants of the excess return, $r_{t+1}^{e}$, from the claim on aggregate consumption, it is sufficient to know the moment-generating function of the vector process $\left(r_{t+1}^{e}, X_{t+1}^{\top}\right)^{\top}$ under the risk-neutral measure. This moment-generating function is given by

$$E_t^{\mathbb{Q}}\left[\exp\left(x\ r_{t+1}^{e} + y^{\top} X_{t+1}\right)\right] = \exp\left(F_{r,0}^{\mathbb{Q}}\left(x, y\right) + X_t^{\top} F_{r,X}^{\mathbb{Q}}\left(x, y\right)\right) \tag{33}$$

where the scalar function $F^{\mathbb{Q}}_{r,0}(x,y)$ and the vector function $F^{\mathbb{Q}}_{r,X}(x,y)$ are defined by

$$
\begin{aligned}
F^{\mathbb{Q}}_{r,0}(x,y) &= H_0 - xG_0 + F_0\left(-\gamma + x, -p_x + y + x\kappa_1 A_X\right) \\
F^{\mathbb{Q}}_{r,X}(x,y) &= H_X - xG_X + F_X\left(-\gamma + x, -p_x + y + x\kappa_1 A_X\right),
\end{aligned}
\tag{34}
$$

and $r^e_{t+1}$ is given by

$$
r^e_{t+1} = r_{t+1} - \mu^{\mathbb{Q}}_t = -G_0 - G_X^\top X_t + \Delta c_{t+1} + \kappa_1 A_X^\top X_{t+1}, \tag{35}
$$

where $\mu^{\mathbb{Q}}_t = E^{\mathbb{Q}}_t[r_{t+1}]$ is given by

$$
\mu^{\mathbb{Q}}_t = \kappa_0 + (\kappa_1 - 1)A_0 + G_0 + (G_X - A_X)^\top X_t, \tag{36}
$$

with coefficients,

$$
G_0 = \mathcal{D}F_0(-\gamma, -p_X)\begin{pmatrix} 1 \\ \kappa_1 A_X \end{pmatrix} \quad \text{and} \quad G_X = \mathcal{D}F_X(-\gamma, -p_X)\begin{pmatrix} 1 \\ \kappa_1 A_X \end{pmatrix}. \tag{37}
$$

The operator $\mathcal{D}$ defines the Jacobian matrix of a real matrix function of a matrix of real variables.[29] Formally, for a given function $\Upsilon$ defined over $\mathbb{R}^m \times \mathbb{R}^n$ and with values in $\mathbb{R}^p \times \mathbb{R}^q$, which associates to the $m \times n$ matrix $\xi$ the $p \times q$ matrix $\Upsilon(\xi)$, we have that $\mathcal{D}\Upsilon(\xi)$ is the $pq \times mn$ matrix defined by

$$
\mathcal{D}\Upsilon(\xi) = \frac{\partial vec(\Upsilon(\xi))}{\partial vec(\xi)^\top} \quad \text{and} \quad \mathcal{D}\Upsilon(\xi^*) = \left.\frac{\partial vec(\Upsilon(\xi))}{\partial vec(\xi)^\top}\right|_{\xi=\xi^*}, \tag{38}
$$

and we also define the operator $\mathcal{D}_i$ for which the derivative is taken with respect to the $i$th argument of the function $\Upsilon$.

To derive the term-structure of all risk-neutral moments, it is sufficient to compute the conditional moment-generating function of aggregate returns, given by,

$$
E^{\mathbb{Q}}_t\left[\exp\left(x\sum_{j=1}^{\tau} r^e_{t+j}\right)\right] = \exp\left(F^{\mathbb{Q}}_{r,0}(x;\tau) + X_t^\top F^{\mathbb{Q}}_{r,X}(x;\tau)\right) \tag{39}
$$

where the sequence of functions $F^{\mathbb{Q}}_{r,0}(x;\tau)$ and $F^{\mathbb{Q}}_{r,X}(x;\tau)$ satisfy the following recursions,

$$
\begin{aligned}
F^{\mathbb{Q}}_{r,0}(x;\tau) &= F^{\mathbb{Q}}_{r,0}(x;\tau-1) + F^{\mathbb{Q}}_{r,0}\left(x, F^{\mathbb{Q}}_{r,X}(x;\tau-1)\right) \\
F^{\mathbb{Q}}_{r,X}(x;\tau) &= F^{\mathbb{Q}}_{r,X}\left(x, F^{\mathbb{Q}}_{r,X}(x;\tau-1)\right)
\end{aligned}
\tag{40}
$$

with initial conditions $F^{\mathbb{Q}}_{r,0}(x;1) = F^{\mathbb{Q}}_{r,0}(x,0)$ and $F^{\mathbb{Q}}_{r,X}(x;1) = F^{\mathbb{Q}}_{r,X}(x,0)$. Then, the $n$th order cumulants of excess returns denoted, $M^{\mathbb{Q}}_n(t,\tau)$, is the derivative of the log moment-generating function of aggregate returns with respect to $x$, and evaluated at $x = 0$,

$$
M^{\mathbb{Q}}_n(t,\tau) = \beta_{n,0}(\tau) + X_t^\top \beta_{n,X}(\tau), \tag{41}
$$

where

$$
\beta_{n,0}(\tau) = \mathcal{D}^n F^{\mathbb{Q}}_{r,0}(0;\tau) \quad \text{and} \quad \beta_{n,X}(\tau) = \mathcal{D}^n F^{\mathbb{Q}}_{r,X}(0;\tau). \tag{42}
$$

In particular,

$$
\begin{aligned}
\beta_{vr,0}(\tau) = {} & \beta_{vr,0}(\tau-1) + \mathcal{D}_2 F^{\mathbb{Q}}_0(0,0)\,\beta_{vr,X}(\tau-1) \\
& + \left(\begin{pmatrix} 1 \\ \beta_{1X}(\tau-1) \end{pmatrix}\right)^\top \mathcal{D}^2 F^{\mathbb{Q}}_0(0,0)\begin{pmatrix} 1 \\ \beta_{ep,X}(\tau-1) \end{pmatrix}
\end{aligned}
\tag{43}
$$

[29]See e.g. See Magnus and Neudecker (1988), Ch. 9, Sec. 4, p. 173.

with $\beta_{vr,0}(1) = \mathcal{D}_1^2 F_0^{\mathbb{Q}}(0,0)$, for the drift coefficient, and

$$\begin{aligned} \beta_{vr,X}(\tau) &= \mathcal{D}_2 F_X^{\mathbb{Q}}(0,0)\, \beta_{vr,X}(\tau-1) \\ &\quad + \left( \begin{pmatrix} 1 \\ \beta_{ep,X}(\tau-1) \end{pmatrix} \otimes I_K \right)^{\top} \mathcal{D}^2 F_X^{\mathbb{Q}}(0,0) \begin{pmatrix} 1 \\ \beta_{ep,X}(\tau-1), \end{pmatrix} \end{aligned} \tag{44}$$

with $\beta_{vr,X}(1) = \mathcal{D}_1^2 F_X^{\mathbb{Q}}(0,0)$, for the slope coefficient, and where

$$\begin{aligned} \beta_{ep,0}(\tau) &= \beta_{ep,0}(\tau-1) + \mathcal{D} F_0^{\mathbb{Q}}(0,0) \begin{pmatrix} 1 \\ \beta_{ep,X}(\tau-1) \end{pmatrix} \\ \beta_{ep,X}(\tau) &= \mathcal{D} F_X^{\mathbb{Q}}(0,0) \begin{pmatrix} 1 \\ \beta_{ep,X}(\tau-1) \end{pmatrix} \end{aligned} \tag{45}$$

with $\beta_{ep,0}(1) = \mathcal{D}_1 F_0^{\mathbb{Q}}(0,0)$ and $\beta_{ep,X}(1) = \mathcal{D}_1 F_X^{\mathbb{Q}}(0,0)$.

## A.2 Affine Reduced-Form Models

Discrete time affine specifications of the return process have the following general Laplace transform of excess returns (Darolles, Gourieroux, and Jasiak 2006),

$$E_t\left[\exp\left(x\, r_{t+1}^e + y^{\top} X_{t+1}\right)\right] = \exp\left(A(x,y)^{\top} Z_t + B(x,y)\right), \tag{46}$$

and the risk-free interest rate is defined as,

$$i_t = \rho_0 + \rho_1' X_t. \tag{47}$$

Similarly, under the risk-neutral measure, $\mathbb{Q}$, affine models have the following general representation,

$$E_t^{\mathbb{Q}}\left[\exp\left(x\, r_{t+1}^e + y^{\top} X_{t+1}\right)\right] = \exp\left(A^{\mathbb{Q}}(x,y)^{\top} X_t + B^{\mathbb{Q}}(x,y)\right) \tag{48}$$

where

$$\begin{aligned} A^{\mathbb{Q}}(x,y) &= A(u+\gamma, v+\Gamma) - A(\gamma,\Gamma) \\ B^{\mathbb{Q}}(x,y) &= B(u+\gamma, v+\Gamma) - B(\gamma,\Gamma). \end{aligned}$$

The parameters $\gamma$ and $\Gamma$ characterize the conditional state-price density, $Z_{t+1}$,

$$M_{t+1} = \exp\left(\gamma\, r_{t+1}^e + \Gamma^{\top} F_t + \theta_t\right), \tag{49}$$

and $Z_{t+1}$ must satisfy,

$$\begin{aligned} E_t\left[M_{t+1}\right] &= \exp(-i_t) \\ E_t\left[M_{t+1}\exp(r_{t+1}^e)\right] &= \exp(-i_t) \end{aligned} \tag{50}$$

which together imply that

$$\begin{aligned} \theta_t &= -A(\gamma,\Gamma)^{\top} X_t - B(\gamma,\Gamma) - i_t \\ 0 &= A(1+\gamma,\Gamma)^{\top} X_t + B(1+\gamma,\Gamma) - A(\gamma,\Gamma)^{\top} X_t - B(\gamma,\Gamma). \end{aligned} \tag{51}$$

If follows easily that multi-horizon returns have the following cumulant-generating function,

$$E_t\left[\exp\left(x\, r_{t,t+\tau}^e\right)\right] = \exp\left(C(x,\tau)^{\top} X_t + D(x,\tau)\right), \tag{52}$$

where

$$
\begin{aligned}
C(x, \tau+1) &= A(x, C(x, \tau)) \\
D(x, \tau+1) &= B(x, C(x, \tau)) + D(x, \tau),
\end{aligned}
\tag{53}
$$

and

$$
C(u, 0) = 0, \;\; D(u, 0) = 0. \tag{54}
$$

The analog results follow under $\mathbb{Q}$. Combining the cumulant-generating function under $\mathbb{P}$ and $\mathbb{Q}$ allows for the computation of the bond, equity and variance premium by computing the appropriate $n^{\text{th}}$-order Jacobian matrix as in section A.1.2 above.

## A.3 Constructing A Monthly Sample

Option settlement dates follow a regular pattern though time: contracts are available for 3 successive months, then for the next 3 months in the March, June, September, December cycle and, finally for the next two months in the June and December semi-annual cycle. This leads to maturity groups with 1, 2 or 3 months remaining to settlement and then between 3 and 6, between 6 and 9, between 9 and 12 months, between 12 and 18 and between 18 and 24 months remaining to settlement. We group option prices at the monthly frequency using their maturity date, so that enough observations are available within each group to construct non-parametric measures. To see why this is a natural strategy, note first that each contract settles on the third Friday of a month. Consider, then, all observations intervening between two successive (monthly) settlement dates. Each of these observations can be unambiguously attributed to one maturity date. Moreover, within that period, each contract will be attributed to the same maturity group.[30] While a higher number of observations reduce sampling errors in our estimates of risk-neutral moments, it may also increase noise if there is large within-month time-variation in the distribution of stock returns at given maturities. To mitigate this effect, we always use the most recent observation when the same contract (i.e. same maturity and strike price) is observed more than once.

## A.4 Cumulants

We rely on the non-parametric approach of Bakshi and Madan (2000) to measure the conditional variance implicit in option prices. Any twice-differentiable payoff, $H(S(t+\tau))$, contingent on the future stock price, $S(t+\tau)$, can be replicated by a portfolio of stock options. The portfolio allocations across option strikes are specific to each payoff $H$ and given by derivatives of the payoff function evaluated at the corresponding strike price. Following Bakshi and Madan, we take

$$
H(S(t+\tau)) \equiv (r^e_{t,t+\tau})^n = \ln\left(\left(\frac{S(t+\tau)}{(S(t)}\right)^n\right),
$$

so that the fair value, at time $t$, of a contract paying the second moments of returns over the next $\tau$ periods ahead, $V_2^{\mathbb{Q}}(t,\tau) \equiv E_t^{\mathbb{Q}}[e^{-r\tau}\left(r^e_{t,t+\tau}\right)^2]$, is given by

$$
\frac{V_2^{\mathbb{Q}}(t,\tau)}{2} = \int_0^{S(t)} \frac{1-\ln(K/S(t))}{K^2} P(t,\tau,K)dK + \int_{S(t)}^{\infty} \frac{1-\log(K/S(t))}{K^2} C(t,\tau,K)dK,
$$

and can be directly computed from the relevant European call and put option prices, $C(t,\tau,K)$ and $P(t,\tau,K)$, with maturity $\tau$ and strike price $K$. Finally, the risk-neutral variance at maturity $\tau$ is given by

$$
Var^{\mathbb{Q}}(t,\tau) = e^{r\tau} V_2^{\mathbb{Q}}(t,\tau) - \mu^{\mathbb{Q}}(t,\tau)^2,
$$

[30] Take any contract, on any observation date. This contract is assigned to the 1-month maturity group if its settlement date occurs on the following third-Friday, to the 2-month group if it occurs on the next to following third-Friday, etc. This grouping does not change until we reach the next settlement date.

where we follow Bakshi et al. (2003) to compute $\mu^{\mathbb{Q}}(t,\tau)$. Similarly, option-implied risk-neutral returns cumulants are given by

$$
\begin{aligned}
M_1^{\mathbb{Q}}(t,\tau) &\equiv \mu^{\mathbb{Q}}(t,\tau) \approx e^{r\tau} - 1 - \frac{e^{r\tau}}{2} V_2^{\mathbb{Q}}(t,\tau) - \frac{e^{r\tau}}{6} V_3^{\mathbb{Q}}(t,\tau) - \frac{e^{r\tau}}{24} V_4^{\mathbb{Q}}(t,\tau) \\
M_2^{\mathbb{Q}}(t,\tau) &\equiv Var^{\mathbb{Q}}(t,\tau) = e^{r\tau} V_2^{\mathbb{Q}}(t,\tau) - \mu^{\mathbb{Q}}(t,\tau)^2 \\
M_3^{\mathbb{Q}}(t,\tau) &= e^{r\tau} V_3^{\mathbb{Q}}(t,\tau) - 3\mu^{\mathbb{Q}}(t,\tau) e^{r\tau} V_2^{\mathbb{Q}}(t,\tau) + 2\mu^{\mathbb{Q}}(t,\tau)^3 \\
M_4^{\mathbb{Q}}(t,\tau) &= e^{r\tau} V_4^{\mathbb{Q}}(t,\tau) - 4\mu^{\mathbb{Q}}(t,\tau) e^{r\tau} V_3^{\mathbb{Q}}(t,\tau) + 6\mu^{\mathbb{Q}}(t,\tau)^2 e^{r\tau} V_2^{\mathbb{Q}}(t,\tau) - 3\mu^{\mathbb{Q}}(t,\tau)^4,
\end{aligned}
$$

where we closely followed Bakshi et al. (2003) in the computation of $\mu^{\mathbb{Q}}$. Recall that the first cumulant is the mean, the second cumulant is the variance, the third cumulant is the third centered moment, and the fourth cumulant is the fourth centered moment minus 3 times the squared variance.

## A.5 Reduced-Rank Regressions

A multivariate reduced-rank regression model can be written as

$$
Y_t = A\Gamma^\top F_t + \Psi Z_t + \epsilon_t \quad t = 1, \dots, T, \tag{55}
$$

where $A$ and $\Gamma$ have size $(p \times K)$ and $(q \times K)$, respectively. The RRR estimators are given from the solution to

$$
\min_{A,\Gamma,\Psi} \left| \sum_{t=1}^{T} \epsilon_t \epsilon_t' \right|, \tag{56}
$$

and closed-form expressions are given in Theorem 5 of Hansen (2008). In his notation, define the moment matrix,

$$
M_{yf} = T^{-1} \sum_{t=1}^{T} Y_t F_t^\top, \tag{57}
$$

and define the matrices $M_{yy}$, $M_{yz}$, $M_{ff}$ similarly. Also, define

$$
\begin{aligned}
S_{yy} &= M_{yy} - M_{yz} M_{zz}^{-1} M_{zy} \qquad (58) \\
S_{yf} &= M_{yf} - M_{yz} M_{zz}^{-1} M_{zf},
\end{aligned}
$$

and define $S_{ff}$ and $S_{yf} = S_{fy}^\top$ similarly. Then, the estimator of $A$, $\Gamma$ and of $\Psi$ are given by,

$$
\begin{aligned}
\hat{\Gamma}\top &= [\hat{v}_1, \dots, \hat{v}_K]\phi \qquad (59) \\
\hat{A} &= S_{y,f} \hat{B} (\hat{B}^\top S_{ff} \hat{B})^{-1} \\
\hat{\Psi} &= M_{yz} M_{zz}^{-1} - \hat{A}\hat{B} M_{fz} M_{zz}^{-1} \qquad (60)
\end{aligned}
$$

where $[\hat{v}_1, \dots, \hat{v}_K]$ are the eigenvectors corresponding to the largest $K$ eigenvalues of,

$$
|\lambda S_{ff} - S_{fy} S_{yy}^{-1} S_{yf}| = 0, \tag{61}
$$

and $\phi$ is an arbitrary $(K \times K)$ matrix with full rank. It is a normalization device and corresponds to the choice of a particular basis for the subspace spanned by the rows of $\hat{\Gamma}$.

Table 1: **Option Sample Summary Statistics**

Number of observations (out-of-the-money puts and calls) in each maturity (months) and moneyness (K/S) group. SP 500 futures option data from January 1996 to October 2008.

| | Moneyness | | | | | | |
|---|---|---|---|---|---|---|---|
| Maturity | $< 0.90$ | $0.90 - 0.95$ | $0.95 - 0.975$ | $0.975 - 1$ | $1 - 1.025$ | $1.025 - 1.05$ | $> 1.05$ |
| 1 | 3173 | 3498 | 2229 | 2435 | 2429 | 2178 | 2638 |
| 2 | 4849 | 3350 | 2115 | 2423 | 2435 | 2098 | 3938 |
| 3 | 3077 | 1789 | 1151 | 1423 | 1371 | 1029 | 2649 |
| 6 | 4248 | 1694 | 987 | 1056 | 917 | 789 | 2957 |
| 9 | 2679 | 1020 | 635 | 645 | 484 | 405 | 2049 |
| 12 | 1621 | 598 | 368 | 417 | 375 | 264 | 1507 |
| 18 | 1504 | 500 | 279 | 313 | 267 | 169 | 1107 |
| 24 | 890 | 259 | 176 | 235 | 149 | 103 | 703 |

Table 2: **Risk-Neutral Variance Summary Statistics**

Summary statistics of conditional risk-neutral variance across maturities from 1 to 18 months (Panel A) and loadings from principal component analysis of risk-neutral variance (Panel B). Risk-neutral variance measures at each maturity constructed using the model-free method of Bakshi and Madan (2000). Option data from January 1996 to October 2008.

Panel A Summary Statistics

| | **1** | **2** | **3** | **6** | **9** | **12** | **18** |
|---|---|---|---|---|---|---|---|
| Mean | 0.037 | 0.045 | 0.046 | 0.049 | 0.047 | 0.044 | 0.044 |
| Std. Dev. | 0.024 | 0.027 | 0.027 | 0.026 | 0.022 | 0.021 | 0.022 |
| Skewness | 1.484 | 1.193 | 1.047 | 0.888 | 0.549 | 0.847 | 0.478 |
| Kurtosis | 5.332 | 4.066 | 3.725 | 3.579 | 2.497 | 3.559 | 2.932 |
| $\rho(1)$ | 0.738 | 0.730 | 0.788 | 0.820 | 0.871 | 0.812 | 0.809 |

Panel B Summary Statistics

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|
| | 0.36 | 0.49 | -0.75 | -0.23 | 0.10 | -0.05 | -0.03 |
| | 0.44 | 0.38 | 0.33 | 0.16 | -0.41 | -0.06 | 0.60 |
| | 0.43 | 0.20 | 0.28 | 0.12 | -0.07 | 0.52 | -0.63 |
| Loadings | 0.42 | -0.06 | 0.26 | 0.02 | 0.32 | -0.76 | -0.27 |
| | 0.35 | -0.28 | -0.01 | 0.15 | 0.70 | 0.37 | 0.40 |
| | 0.31 | -0.42 | 0.08 | -0.81 | -0.23 | 0.09 | 0.06 |
| | 0.31 | -0.57 | -0.41 | 0.48 | -0.42 | -0.07 | -0.06 |
| $R^2$ | 0.88 | 0.06 | 0.03 | 0.02 | 0.01 | 0.00 | 0.00 |
| Cum. $R^2$ | 0.88 | 0.94 | 0.97 | 0.99 | 0.99 | 1.00 | 1.00 |

Table 3: **Excess Return and the Variance Term Structure**

Rank test $p$-values and $R^2$s in multivariate regressions, $Y_t = \Pi_0 + \Pi F_t + \epsilon_t$ where each component of $Y_t$ is an excess bond or equity returns, $xr_{t,t+\tau}$, and where $F_t = \{\hat{Var}^{\mathbb{Q}}(t,\tau)\}_{\tau=1,\ldots,q}$ is a $q \times 1$ vector of risk-neutral variance measures. We consider annual excess returns for bonds with maturities of 2, 3, 4 and 5 years, and SP 500 excess returns at horizons 1, 3, 6, 9 and 12 months. Panel A displays $p$-values associated with the Cook and Setodji modified statistics, $\tilde{\Lambda}_r$, in a test of the null hypothesis that the rank of the matrix $\Pi$ is $r$. Panel B displays the $R^2$ associated with each of the individual bond returns predictability regression obtained via multivariate reduced-rank regression (RRR) estimation but for different hypothesis on the rank of the matrix $\Pi$. Panel C displays the $R^2$ associated with each of the individual equity returns predictability regression. Risk-neutral variance measures at each maturity constructed using the model-free method of Bakshi and Madan (2000). Monthly Returns and Option data from January 1996 to October 2008.

Panel A - Rank test $p$-values

| | $H_0:\ r=0$ | $H_0:\ r=1$ | $H_0:\ r=2$ | $H_0:\ r=3$ | $H_0:\ r=4$ | $H_0:\ r=5$ | $H_0:\ r=6$ |
|---|---|---|---|---|---|---|---|
| $p$-val | 0.0 | 4.3 | 22.9 | 64.8 | 82.5 | 81.4 | 73.0 |

Panel B - Bond returns $R^2$s

| | $r=1$ | $r=2$ | $r=3$ | $r=4$ | $r=5$ | $r=6$ | $r=7$ |
|---|---|---|---|---|---|---|---|
| 2 | 7.3 | 7.3 | 9.2 | 11.1 | 11.4 | 11.4 | 11.5 |
| 3 | 6.6 | 6.6 | 7.8 | 9.6 | 9.9 | 10.0 | 10.1 |
| 4 | 5.7 | 5.9 | 6.6 | 8.2 | 8.7 | 8.7 | 8.8 |
| 5 | 5.0 | 5.5 | 5.8 | 7.3 | 7.8 | 7.9 | 8.0 |

Panel C - Equity returns $R^2$s

| | $r=1$ | $r=2$ | $r=3$ | $r=4$ | $r=5$ | $r=6$ | $r=7$ |
|---|---|---|---|---|---|---|---|
| 1 | 1.9 | 3.1 | 3.1 | 3.3 | 3.4 | 3.5 | 3.7 |
| 2 | 4.0 | 6.3 | 7.2 | 8.8 | 9.2 | 9.2 | 9.2 |
| 3 | 5.4 | 6.3 | 7.5 | 10.7 | 11.1 | 11.1 | 11.3 |
| 6 | 3.3 | 5.3 | 7.6 | 9.0 | 9.0 | 9.1 | 9.6 |
| 9 | 3.5 | 4.2 | 7.9 | 10.1 | 10.1 | 10.1 | 10.3 |
| 12 | 3.5 | 3.6 | 10.5 | 11.0 | 11.0 | 11.0 | 11.1 |

Table 4: **Excess Variance Predictability**

Results from multi-horizon predictability regressions of the excess variance over an horizon of of $\tau$, $xv_{t,t+\tau}$, with $\tau = 1, 2, 3, 6, 9$ and 12 months, respectively. The predictors include a constant and $\hat{\Gamma} F_t$, the risk factors obtained from the multivariate reduced-rank regression of bond and equity excess returns on the variance term structure (See Table 3). Newey-West t-statistics with lags corresponding to the investment horizon plus 3 months in parenthesis and $R^2$ reported in percentage. Risk-neutral variance measures at each maturity constructed using the model-free method of Bakshi and Madan (2000). Monthly Variance and Option data from January 1996 to October 2008.

| | 1 | 2 | 3 | 6 | 9 | 12 |
|---|---|---|---|---|---|---|
| $\hat{\Gamma}_1 \tilde{Var}_t^{\mathbb{Q}}$ | -0.011 | -0.012 | -0.010 | -0.009 | -0.003 | -0.003 |
| | (-2.15) | (-1.94) | (-1.55) | (-1.35) | (-0.49) | (-0.42) |
| $\hat{\Gamma}_2 \tilde{Var}_t^{\mathbb{Q}}$ | -0.005 | -0.007 | -0.008 | -0.008 | -0.009 | 0.005 |
| | (-1.23) | (-1.78) | (-1.74) | (-2.21) | (-2.51) | (1.56) |
| $R^2$ | 6.2 | 9.5 | 9.0 | 10.1 | 8.7 | 2.7 |

Table 5: **Summary Statistics - Term Structure of Higher Order Moments**

Panel A and Panel B report summary statistics of risk-neutral cumulants 3 and 4, respectively, across maturities from 1 to 24 months. Panel C and Panel D report the corresponding correlation matrix. Panel E and Panel F report the loadings and explanatory power of each component from a principal component analysis (PCA). Cumulant measures at each maturity constructed using the model-free method of Bakshi and Madan (2000). Option data from January 1996 to October 2008.

Panel A - Skewness summary statistics

| | **1** | **2** | **3** | **6** | **9** | **12** | **18** |
|---|---|---|---|---|---|---|---|
| Mean | -0.003 | -0.005 | -0.007 | -0.011 | -0.012 | -0.011 | -0.015 |
| Std. Dev. | 0.002 | 0.004 | 0.005 | 0.007 | 0.007 | 0.008 | 0.010 |
| Skewness | -2.036 | -1.555 | -1.184 | -1.083 | -1.164 | -1.866 | -1.677 |
| Kurtosis | 8.009 | 5.307 | 3.745 | 3.575 | 4.279 | 8.786 | 7.234 |
| $\rho(1)$ | 0.408 | 0.690 | 0.659 | 0.804 | 0.742 | 0.615 | 0.577 |

Panel B - Kurtosis summary statistics

| | **1** | **2** | **3** | **6** | **9** | **12** | **18** |
|---|---|---|---|---|---|---|---|
| Mean | 0.001 | 0.002 | 0.003 | 0.006 | 0.008 | 0.008 | 0.012 |
| Std. Dev. | 0.001 | 0.002 | 0.003 | 0.005 | 0.006 | 0.008 | 0.010 |
| Skewness | 1.626 | 1.662 | 1.292 | 1.151 | 1.245 | 2.847 | 2.483 |
| Kurtosis | 5.811 | 5.550 | 3.942 | 3.637 | 4.705 | 14.979 | 11.668 |
| $\rho(1)$ | 0.509 | 0.670 | 0.622 | 0.785 | 0.711 | 0.608 | 0.472 |

Panel C - Skewness correlations

| **1** | **2** | **3** | **6** | **9** | **12** | **18** |
|---|---|---|---|---|---|---|
| 1.00 | 0.76 | 0.66 | 0.62 | 0.51 | 0.20 | 0.18 |
| 0.76 | 1.00 | 0.87 | 0.87 | 0.66 | 0.42 | 0.37 |
| 0.66 | 0.87 | 1.00 | 0.83 | 0.69 | 0.40 | 0.45 |
| 0.62 | 0.87 | 0.83 | 1.00 | 0.79 | 0.62 | 0.57 |
| 0.51 | 0.66 | 0.69 | 0.79 | 1.00 | 0.45 | 0.66 |
| 0.20 | 0.42 | 0.40 | 0.62 | 0.45 | 1.00 | 0.49 |
| 0.18 | 0.37 | 0.45 | 0.57 | 0.66 | 0.49 | 1.00 |

Panel D - Kurtosis correlations

| **1** | **2** | **3** | **6** | **9** | **12** | **18** |
|---|---|---|---|---|---|---|
| 1.00 | 0.75 | 0.65 | 0.62 | 0.54 | 0.25 | 0.22 |
| 0.75 | 1.00 | 0.81 | 0.83 | 0.61 | 0.38 | 0.34 |
| 0.65 | 0.81 | 1.00 | 0.78 | 0.65 | 0.35 | 0.41 |
| 0.62 | 0.83 | 0.78 | 1.00 | 0.73 | 0.61 | 0.53 |
| 0.54 | 0.61 | 0.65 | 0.73 | 1.00 | 0.30 | 0.57 |
| 0.25 | 0.38 | 0.35 | 0.61 | 0.30 | 1.00 | 0.40 |
| 0.22 | 0.34 | 0.41 | 0.53 | 0.57 | 0.40 | 1.00 |

Panel E - Skewness PCA

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|
| | 0.08 | -0.18 | 0.11 | -0.09 | 0.11 | -0.79 | 0.55 |
| | 0.21 | -0.33 | 0.14 | -0.29 | 0.01 | -0.39 | -0.77 |
| | 0.26 | -0.31 | 0.21 | -0.34 | 0.70 | 0.39 | 0.18 |
| Loadings | 0.46 | -0.41 | 0.06 | -0.20 | -0.67 | 0.26 | 0.25 |
| | 0.43 | -0.12 | 0.33 | 0.82 | 0.12 | -0.01 | -0.07 |
| | 0.38 | -0.10 | -0.89 | 0.11 | 0.18 | -0.07 | -0.02 |
| | 0.58 | 0.76 | 0.14 | -0.25 | -0.00 | -0.07 | -0.01 |
| $R^2$ | 0.67 | 0.15 | 0.12 | 0.04 | 0.01 | 0.01 | 0.00 |
| Cum. $R^2$ | 0.67 | 0.82 | 0.93 | 0.97 | 0.99 | 1.00 | 1.00 |

Panel F - Kurtosis PCA

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|
| | 0.03 | -0.02 | 0.08 | 0.07 | 0.10 | 0.39 | -0.91 |
| | 0.08 | -0.07 | 0.20 | 0.29 | 0.13 | 0.83 | 0.41 |
| | 0.12 | -0.06 | 0.26 | 0.35 | 0.82 | -0.33 | 0.01 |
| Loadings | 0.32 | -0.26 | 0.43 | 0.55 | -0.53 | -0.23 | -0.06 |
| | 0.33 | 0.04 | 0.67 | -0.66 | 0.03 | 0.02 | 0.03 |
| | 0.40 | -0.81 | -0.36 | -0.21 | 0.10 | 0.03 | 0.01 |
| | 0.78 | 0.52 | -0.34 | 0.08 | 0.01 | 0.02 | -0.00 |
| $R^2$ | 0.65 | 0.19 | 0.12 | 0.03 | 0.01 | 0.00 | 0.00 |
| Cum. $R^2$ | 0.65 | 0.84 | 0.96 | 0.99 | 1.00 | 1.00 | 1.00 |

Table 6: **Predictive Content of Higher Order Moments**

Panel A displays $p$-values of rank tests and $R^2$s in multivariate regressions, $Y_t = \Pi_0 + \Pi F_t + \epsilon_t$ where each component of $Y_t$ is an excess bond or equity returns, $xr_{t,t+\tau}$, and where $F_t$ with different combinations of risk-neutral variance, skewness and of kurtosis at horizons 1, 3, 6, 9, 12 and 18 months. Model $V(2)$ assumes $r = 2$ and includes the term structure of variance in $F_t$. Models $S(2)$ and $K(2)$ also assume $r = 2$. Model $S(2)$ includes the term structure of skewness and model $K(2)$ includes the term structure of kurtosis. Model VSK(2,2) combines the two risk factors estimated from each of the variance, skewness and kurtosis term structure and assumes $r = 2$. Model VSK(7,2) assumes $r = 2$ and combines all measures from the variance, skewness and kurtosis term structure. We report $p$-values associated with the Cook and Setodji modified statistics, $\tilde{\Lambda}_r$, for tests of the null hypothesis that the rank of the matrix $\Pi$ is $r = 2$ and the of the null that the rank is $r = 3$. Panel B displays $R^2$s from multi-horizon predictability regressions of the excess variance, $xv_{t,t+\tau}$ on a constant and $\hat{\Gamma} F_t$, the risk factors obtained from the multivariate reduced-rank regression. Annual excess returns for bonds with maturities of 2, 3, 4 and 5 years, SP 500 excess returns at horizons 1, 3, 6, 9 and 12 months, and excess variance at horizons 1, 3, 6, 9 and 12 months. Risk-neutral variance, skewness and kurtosis measures at each maturity constructed using the model-free method of Bakshi and Madan (2000). Monthly returns, realized variance and option data from January 1996 to October 2008.

Panel A - Bond and equity premium

| Model | Rank Test H0 | | Bond Maturity | | | | Equity return horizons | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $H0: r = 1$ | $H0: r = 2$ | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 6 | 9 | 12 |
| $V(2)$ | 4.3 | 22.9 | 7.3 | 6.6 | 5.9 | 5.5 | 3.1 | 6.3 | 6.3 | 5.3 | 4.2 | 3.6 |
| $S(2)$ | 6.1 | 38.2 | 9.1 | 9.8 | 9.7 | 9.5 | 2.8 | 6.3 | 4.9 | 2.9 | 2.3 | 1.1 |
| $K(2)$ | 7.9 | 32.2 | 10.2 | 11.0 | 10.8 | 10.6 | 2.3 | 4.6 | 5.4 | 6.0 | 4.1 | 3.6 |
| $VSK(2,2)$ | 1.1 | 32.6 | 10.1 | 10.0 | 9.4 | 9.0 | 3.7 | 7.9 | 6.5 | 6.4 | 4.6 | 3.4 |
| $VSK(7,2)$ | 1.4 | 9.7 | 21.6 | 20.3 | 18.2 | 16.5 | 5.8 | 16.4 | 16.6 | 17.4 | 18.9 | 17.9 |

Panel B - Variance premium

| Model | Returns horizons | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 6 | 9 | 12 |
| $V(2)$ | 6.2 | 9.5 | 9.0 | 10.1 | 9.7 | 2.7 |
| $S(2)$ | 1.1 | 6.0 | 6.5 | 9.8 | 7.5 | 0.2 |
| $K(2)$ | 2.0 | 5.6 | 5.5 | 9.0 | 5.7 | 0.1 |
| $VSK(2,2)$ | 5.8 | 9.8 | 9.5 | 11.2 | 9.5 | 1.2 |
| $VSK(7,2)$ | 1.7 | 4.3 | 4.7 | 9.7 | 8.8 | 2.6 |