Internet Appendix for "Excess Volatility: Beyond Discount Rates"

Stefano Giglio Bryan Kelly

A. Affine Representation of Structural Models

The affine-$\mathbb{Q}$ representation is typically associated with reduced-form models, as in Duffie, Pan and Singleton (2000). However, many workhorse structural asset pricing models also feature affine $\mathbb{Q}$ dynamics. In this section we briefly review the $\mathbb{Q}$-dynamics of prevalent consumption-based models.

We begin with the long run risks models of Bansal and Yaron (2004), in which log consumption growth and its volatility follow linear dynamics. The log pricing kernel is approximately linear (the linearity is exact with unit EIS, and the linear approximation is extremely accurate, as shown in Dew-Becker and Giglio (2013)). In this model the log price and the log price-dividend ratio of all consumption or dividend strips are linear functions of the model's state variables (the persistent component of consumption growth x_t and the conditional variance of consumption growth σ_t^2). Prices of consumption and dividend strips therefore follow an exponentially affine specification (with heteroskedasticity).

A related paper, Drechsler and Yaron (2011), extends the model to match the variance risk premium. Dew-Becker et al. (2015) solve for the term structure of variance swaps in that model. In that model the $\mathbb{Q}-$dynamics of variance are linear, and the log pricing kernel is linear (under the standard approximation), and thus variance swaps also follow an affine structure. Note that in this paper the distribution of the shocks is not normal under $\mathbb{Q}$ (due to the presence of jumps), but this is irrelevant for the term structure of variance swaps because these are linear (not exponential) claims to future variance.

Next, we consider two time-varying rare disaster models, Gabaix (2012) and Wachter (2013). In Gabaix's model, the use of linearity-generating processes (LGP) implies that the (level) price-dividend ratio is linear for all dividend strips. While the LGP assumption buys tractability in modeling price-dividend ratios, the term structure of claims does not follow linear dynamics; the model therefore is not nested in the affine-$\mathbb{Q}$ class. In Wachter (2013), on the other hand, the prices and price-dividend ratios for consumption and dividend strips are loglinear in the disaster probability λ_t, which itself follows a (linear) square-root process. The model therefore follows in the category of exponential affine-$\mathbb{Q}$ models with heteroskedasticity, like long run risks.

Finally, the habit formation model of Campbell and Cochrane (1999) does not map directly into the affine specification, as discussed in Wachter (2005).

A number of papers have explored the relationship between learning and excess volatility, such as Timmermann (1993); Barsky and De Long (1993); Veronesi (1999); Pástor and Veronesi (2003, 2009*b*,*a*). In some (but not all) cases, such as in Barsky and De Long (1993), adding learning to the model preserves the affine-$\mathbb{Q}$ structure. In other cases, learning about model parameters induces non-linearities, with which we deal directly in Section IV.C..

B. Exponential-affine Models

The linear claim structure of Equation (4) is well suited for modeling variance swaps, inflation swaps, and related assets. It is less well suited to claims in other asset classes such as interest rates or credit default swaps, which are more naturally modeled as exponential-affine claims. In that case, prices are exponential rather than linear functions of x_t (itself linear in factors H_t as described by equation (5)):

The model restrictions and testing procedures we derived above also apply in the exponential-affine setting under two additional assumptions regarding the $\mathbb{Q}$ distribution of factor innovations, $\Gamma\epsilon_t^{\mathbb{Q}}$, in Equation (6). First, we require that $\epsilon_t^{\mathbb{Q}}$ follows a Gaussian distribution. Second, $\Gamma\epsilon_t^{\mathbb{Q}}$ must be homoskedastic or, alternatively, it needs to be heteroskedastic but its conditional volatility needs to be uncorrelated with the factors (as in unspanned volatility models).

In exponential-affine models, the price of a cumulative claim is:

$$p_{t,n} = E_t^{\mathbb{Q}}\left[\exp\left(x_{t+1} + ... + x_{t+n}\right)\right].$$

Interest rate claims are the leading example in this class, where r_t is the instantaneous interest rate and $x_t = -r_t$. Prices are then related to factors according to[43]

$$\log p_{t,n} = \mathbf{1}'\left[\rho^{\mathbb{Q}} + (\rho^{\mathbb{Q}})^2 + ... + (\rho^{\mathbb{Q}})^n\right] H_t + \text{constant}. \tag{20}$$

To see why, note that $x_{t+1} + ... + x_{t+n}$ has a conditional normal distribution under $\mathbb{Q}$, so that

$$\log p_{t,n} = \log E_t^{\mathbb{Q}}\left[\exp\left(x_{t+1} + ... + x_{t+n}\right)\right] = E_t^{\mathbb{Q}}\left[x_{t+1} + ... + x_{t+n}\right] + \frac{1}{2}V_t^{\mathbb{Q}}\left[x_{t+1} + ... + x_{t+n}\right]$$

Because of homoskedasticity, the variance term is constant, and therefore

$$\log p_{t,n} = E_t^{\mathbb{Q}}\left[x_{t+1} + ... + x_{t+n}\right] + \text{constant}$$

Other than a difference in the constant (which is irrelevant for variance tests), this is the same setup as in the linear case once prices are transformed from levels into logs.[44] We can therefore apply the same estimation and testing methodology in the exponential-affine case that we described for the benchmark affine case of the main text.

Throughout the paper we focus on the homoskedastic case for the following reasons. First, many of the asset classes we analyze (such as variance and inflation swaps) are typically modeled as claims to the level of x_t, in which case heteroskedasticity does not affect pricing. Conditional variance enters only in exponential models through the Jensen inequality term.

Second, conditional heteroskedasticity affects the loadings on the factors in exponential-affine models only to the extent that the factors themselves span the volatility of the errors. The term structure literature finds evidence of a large unspanned volatility component in interest rates (see, for example, Collin-Dufresne and Goldstein, 2002). So-called unspanned volatility models fix the loadings of bond prices on volatility factors to be zero. In this case, the factor loadings for log prices follow the same recursion as in standard homoskedastic models. For further discussion of the unspanned volatility case, see Collin-Dufresne and Goldstein (2002), Dai and Singleton (2003), Joslin (2006), Bikbov and Chernov (2009), and Creal and Wu (2015).

In the bond pricing literature, when the volatility of factor shocks is in fact spanned by prices, the magnitude of the effect on factor loadings is shown to be small. Nonetheless, spanned volatility models can potentially affect our variance ratio test and Appendix B.i. performs robustness tests that directly account for heteroskedasticity. Our main conclusion from this check is that heteroskedasticity

[43] A minor adaptation for the case of bonds is that powers of $\rho^{\mathbb{Q}}$ range from 0 to $n-1$ rather than from 1 to n, though this is inconsequential for our variance ratio test.

[44] For some claims it is preferable to model individual forwards with an affine-exponential form, which results in a similar relation under homoskedasticity and gaussianity: $\log f_{t,n} = \log E_t^{\mathbb{Q}}\left[\exp\left(x_{t+n}\right)\right] = \mathbf{1}'(\rho^{\mathbb{Q}})^n H_t +$ constant..

of factor innovations is not a central driver of our results.

B.i. Heteroskedasticity Adjustment in Exponential-affine Models

The exponential-affine model described in the previous section can be used to understand the effects of stochastic volatility on the model-predicted factor loadings. Below we analyze a detailed exponential-affine example applied to the variance swap market to demonstrate the robustness of our results to accounting for heteroskedastic factors. For some parameter configurations, heteroskedasticity artificially slows down the decay of factor loadings as maturity increases. Correcting for this effect can potentially reduce long maturity variance ratios. It is therefore important to quantify the magnitude of this adjustment. We turn to this estimation problem now.

Heteroskedastic Model. We consider a heteroskedastic affine term structure model to price forward claims on a cash flow $\exp\{x_t\}$, where x_t is linear in some latent factors: $x_t = \delta_0 + \mathbf{1}' H_t$. Assume that $\mathbb{P}$ dynamics of factors follow:

$$H_{t+1} = c + \rho^{\mathbb{P}} H_t + \Gamma_t \epsilon_{t+1}$$

Γ_t captures stochastic conditional volatility of the factors. Next, rather than directly assume $\mathbb{Q}$ distributions, we derive prices through a one-period stochastic discount factor of the form:

$$M_{t,t+1} = \exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1})$$

where the vector λ_t captures time-varying prices of risk of the different shocks.

For any forward asset on a cash flow x_t, with maturity $n+1$, we have the recursive equation:

$$f_{t,n+1} = E_t[\exp\{-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}\} f_{t+1,n}]$$

where the expectation E_t is taken under the physical measure. Next, conjecture that the forward price is an exponentially-affine function of the factors:

$$f_{t,n+1} = \exp\{a_{n+1} + b_{n+1} H_t\}$$

Taking logs:

$$\begin{aligned} a_{n+1} + b_{n+1} H_t &= \log E_t[\exp\{-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1} + a_n + b_n H_{t+1}\}] \\ &= -r_t - \frac{1}{2}\lambda_t'\lambda_t + a_n + b_n(c + \rho^{\mathbb{P}} H_t) + \frac{1}{2} V_t((-\lambda_t' + b_n \Gamma_t)\epsilon_{t+1}) \\ &= -r_t + a_n + b_n c + b_n \rho^{\mathbb{P}} H_t + \frac{1}{2} b_n \Gamma_t \Gamma_t b_n' - b_n' \Gamma_t \lambda_t. \end{aligned}$$

For the very first maturity forward price, $f_{t,1}$, we have:

$$\begin{aligned} a_1 + b_1 H_t &= \log E_t \exp\{-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1} + x_{t+1}\}\} \\ &= -r_t + \delta_0 + \mathbf{1}'c + \mathbf{1}'\rho^{\mathbb{P}} H_t + \frac{1}{2}\mathbf{1}'\Gamma_t\Gamma_t'\mathbf{1} - \mathbf{1}'\Gamma_t\lambda_t. \end{aligned}$$

In expressions for both $n = 1$ and for $n > 1$, we have the terms $\Gamma_t \Gamma_t'$ and $\Gamma \lambda_t$ that are functions of time-t information. To find an exponentially-affine solution, these terms need to be linear in the factors. Following the term structure literature, we assume that $\Gamma_t \Gamma_t'$ is linear in H_t (which makes the term $b_n \Gamma_t \Gamma_t b_n'$ also linear in H_t). In particular, we assume that:

$$V_t(H_{t+1}) = \Gamma_t \Gamma_t' = \Gamma\Gamma' \sigma_t^2 \qquad \text{and} \qquad \sigma_t^2 = a \cdot f_{1,t}$$

for some $a > 0$. This specification is easy to interpret in the empirical robustness test for the variance swap market that we perform below. The conditional variance of the variance swap factors is assumed to be proportional to the one-period swap price (essentially the VIX), capturing the intuition that as VIX increases, fluctuations in future variance will be more pronounced.

In the discount factor, λ_t is assumed to follow $\lambda_t = \Gamma_t^{-1}\Gamma(\lambda + \Lambda H_t)$. This makes the term $\Sigma_t \lambda_t$ also linear in H_t. In addition, if the risk-free rate is $r_t = a_0 + a_1 H_t$, the term a_1 would also enter the recursion for b_n. In what follows, we ignore risk-free rate variation as it plays a minor role in the pricing of variance swaps.

We can rewrite the expressions under $\mathbb{Q}$, using the same normalizations we have used in our main analysis: $\rho^{\mathbb{Q}} \equiv \rho^{\mathbb{P}} - \Gamma\Lambda$ (the VAR companion matrix under $\mathbb{Q}$) is diagonal, and $c^{\mathbb{Q}} \equiv c - \lambda\Gamma = 0$. We can then rewrite earlier expressions as:

$$a_{n+1} + b_{n+1} H_t = a_n + b_n \rho^{\mathbb{Q}} H_t + \frac{1}{2} b_n \Gamma\Gamma b_n' \sigma_t^2$$

$$a_1 + b_1 H_t = \delta_0 + \mathbf{1}'\rho^{\mathbb{Q}} H_t + \frac{1}{2}\mathbf{1}'\Gamma\Gamma'\mathbf{1}\sigma_t^2.$$

Because $\sigma_t^2 = a \cdot f_{1,t} = a\mathbf{1}'\rho^{\mathbb{Q}} H_t$, these expressions then become:

$$a_{n+1} + b_{n+1} H_t = a_n + b_n \rho^{\mathbb{Q}} H_t + \frac{1}{2} b_n \Gamma\Gamma b_n' (a\mathbf{1}'\rho^{\mathbb{Q}} H_t)$$

$$a_1 + b_1 H_t = \delta_0 + \mathbf{1}'\rho^{\mathbb{Q}} H_t + \frac{1}{2}\mathbf{1}'\Gamma\Gamma'\mathbf{1}(a\mathbf{1}'\rho^{\mathbb{Q}} H_t).$$

Finally, we can match coefficients on H_t, and obtain:

$$b_1 = \mathbf{1}\rho^{\mathbb{Q}} + \frac{1}{2}\mathbf{1}'\Gamma\Gamma'\mathbf{1}(a\mathbf{1}'\rho^{\mathbb{Q}})$$

$$b_{n+1} = b_n \rho^{\mathbb{Q}} + \frac{1}{2} b_n' \Gamma\Gamma b_n (a\mathbf{1}'\rho^{\mathbb{Q}}).$$

To learn about the magnitude of heteroskedasticity adjustments in the model coefficients, $\frac{1}{2}\mathbf{1}'\Gamma\Gamma'\mathbf{1}(a\mathbf{1}'\rho^{\mathbb{Q}})$ and $\frac{1}{2} b_n' \Gamma\Gamma b_n (a\mathbf{1}'\rho^{\mathbb{Q}})$, we proceed as follows. First, note that the conditional variance of the log cash flow in the model is (up to a constant):

$$V_t(x_{t+1}) = \mathbf{1}'\Gamma\Gamma'\mathbf{1}(a\mathbf{1}'\rho^{\mathbb{Q}} H_t) = \mathbf{1}'\Gamma\Gamma'\mathbf{1} a f_{t,1}$$

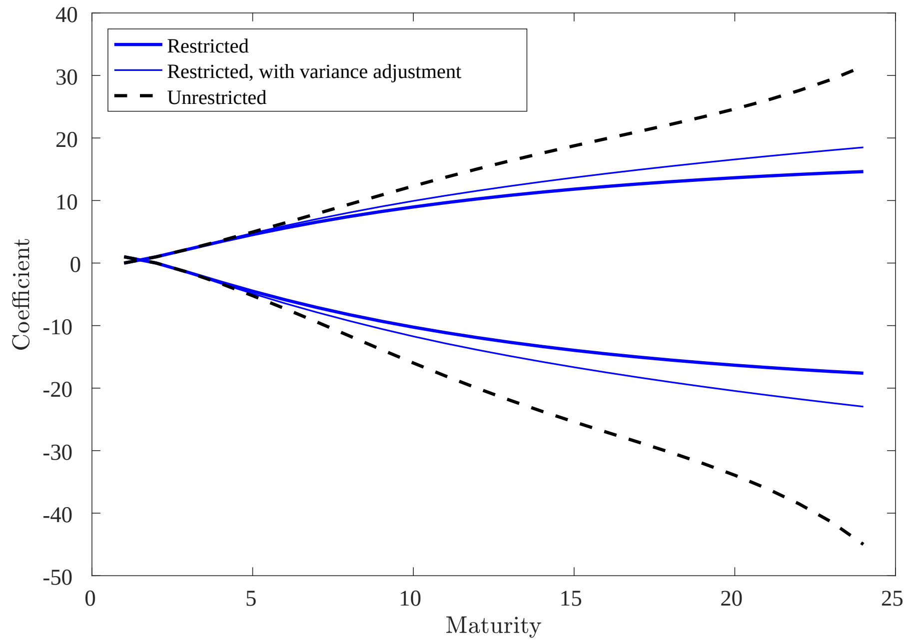
Therefore, regressing $V_t(x_{t+1})$ onto $f_{t,1}$ would yield an estimate of the term $\mathbf{1}'\Gamma\Gamma'\mathbf{1}a$. This would allow us to estimate the heteroskedasticity adjustment for b_1. Next, consider the conditional variance of the first log price (from the left-hand side of the equations above):

$$V_t(f_{t+1,1}) = b_1' \Gamma_t \Gamma_t b_1 = b_1' \Gamma\Gamma b_1 a f_{t,1}$$

The regression coefficient of $V_t(f_{t+1,1})$ onto $f_{t,1}$ yields an estimate of $b_1'\Gamma\Gamma b_1 a$, which we can use to adjust the coefficient b_2 for the effects of conditional volatility. Continuing the recursion, this allows us to compute the adjustment for all maturities.

Figure I reports the impact of heteroskedasticity adjustments on coefficient estimates in the vari-

Figure I Variance Swap Loadings (Homoskedastic vs Heteroskedastic model)



Note. The figure plots the loadings of prices of each maturity on the two factors (1-month and 2-month price). Dashed lines indicate loadings in the unrestricted model, solid lines indicate loadings in the restricted model. The thick line reports the coefficients under the homoskedasticity assumption, the thin line adjusts for heteroskedasticity.

ance swap market when applying an exponential-affine model. The figure highlights a number of facts. First, it looks very similar to Figure II, which is based on the affine model in levels, which demonstrates a basic robustness of our results to level-affine versus exponential-affine specifications. Second, the figure shows that adjusting for heteroskedasticity in the exponential-affine setting has a small effect on loading estimates, indicating that heteroskedasticity is not capable of rationalizing a the variance ratio of more than two at the 24-month maturity.

C. Risk-free Rate Variation

For many of the asset classes considered in this paper, time variation in the risk-free rate plays a minor role in determining the volatility of prices along the term structure, and is typically ignored in the literature (for example, Ait-Sahalia, Karaman and Mancini (2015) ignore risk-free rate variation when pricing variance swaps).

For other asset classes, interest rate variation plays a more important role. Here we show that in exponential-affine models where not only log cash flows x_t but also short-term rates r_t are linear functions of the factors, our test is valid even in the presence of (unmodeled) stochastic interest rates. Consider in particular a cumulative contract that pays all the cash flows at maturity, and has an upfront payment of the price. Then, we can write the price as:

$$p_{t,n} = E_t^{\mathbb{Q}}\left[\frac{e^{x_{t+1}+\ldots+x_{t+n}}}{e^{r_t+\ldots+r_{t+n-1}}}\right] = E_t^{\mathbb{Q}}\left[e^{y_{t+1}+\ldots+y_{t+n}}\right] \tag{21}$$

where $y_t = x_t - r_{t-1}$. If y_t is a linear function of the factors (for example because x_t and r_{t-1} are driven by the same factors), we can simply see this price as a claim to risk-free-adjusted cash flows y_t. Finally, remember that none of our analysis requires us to actually *observe* the cash flow (in this case y_t): it is enough to know that the price is determined according to an exponential-affine model in *some* cash flow y_t.

The argument also holds when all payments are exchanged at maturity, since in that case

$$p_{t,n} = E_t^{\mathbb{Q}} \left[\frac{E_t^{\mathbb{Q}}[e^{r_t+\ldots+r_{t+n}}]}{e^{r_t+\ldots+r_{t+n}}} e^{x_{t+1}+\ldots+x_{t+n}} \right]$$

which means that we can construct the price $\tilde{p}_{t,n} = p_{t,n}\delta_{t,n}$, where $\delta_{t,n}$ is the price of a risk-free bond with maturity n, and the adjusted price $\tilde{p}_{t,n}$ will have the same form as (21).

D. Example: Transformation of $\mathbb{P}$ Model to $\mathbb{Q}$ Model

This appendix provides a brief affine example illustrating how the risk-neutral, or $\mathbb{Q}$, measure representation of a model accounts for factors that drive time variation in risk premia. First we analyze a general two-factor affine specification. Then we specialize to the case where cash flows follow a one-factor model under $\mathbb{P}$, but due to time-varying risk premia, cash flows follow a two-factor model under $\mathbb{Q}$. For further details, we refer readers to Hamilton and Wu (2012).

Suppose that two factors given by vector H_t drive the physical dynamics of the economy (including both cash flows and risk premia). Assume the $\mathbb{P}$ dynamics of the factors is

$$H_{t+1} = c + \rho H_t + \Sigma u_{t+1}$$

where u_{t+1} is a two-dimensional vector of independent standard normals. The claim being priced is an n-period pure discount asset that has cash flow at maturity of X_{t+n} (where $X_t = \exp(\delta' H_t)$). Preferences are represented by the stochastic discount factor M_{t+1}, whose behavior depends on factor shocks u_{t+1} and risk price $\lambda_t = \Lambda H_t$ according to

$$M_{t+1} = \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t' u_{t+1} \right)$$

The claim price is a function of H_t and follows the price recursion

$$P_t(H_t) = E_t[P_{t+1}(H_{t+1})M_{t+1}] = \int P_{t+1}(H_{t+1})M_{t+1}\phi(H_{t+1};\mu_t,\Sigma\Sigma')dH_{t+1}. \tag{22}$$

In this example, the physical measure ($\mathbb{P}$) is described by the multivariate normal density function ϕ having mean $\mu_t = E_t^{\mathbb{P}}[H_{t+1}]$ and covariance matrix $\Sigma\Sigma'$.

To derive the equivalent "risk-neutral" pricing measure ($\mathbb{Q}$), we rewrite the price as

$$P_t(H_t) = E_t^{\mathbb{Q}}[P_{t+1}(H_{t+1})] \tag{23}$$

where $M_{t+1}\phi(H_{t+1};\mu_t,\Sigma\Sigma')$ is a transformation of the original probability measure into the new measure $\mathbb{Q}$. Like ϕ, $M_{t+1}\phi$ is a multivariate normal density. The mean of this density is $\mu_t^{\mathbb{Q}} = \mu_t - \Sigma\lambda_t$, and its variance is the same as ϕ's. Any claim whose price depends only on the factors H_t and can be represented by equation (22), can equivalently be represented by equation (23) where there is no explicit SDF/risk premium adjustment but where the mean of the factors has been additively shifted by an amount $-\Sigma\lambda_t$.

Under the original measure $\mathbb{P}$, the dynamics of H_t are linear, i.e.

$$\mu_t = E_t H_{t+1} = c + \rho H_t.$$

Importantly, the dynamics of H_t under the new measure $\mathbb{Q}$ remain linear. Given the form assumed for λ_t and the derived equation for $\mu_t^{\mathbb{Q}}$, we have

$$\mu_t^{\mathbb{Q}} = (c + \Sigma\lambda) + (\rho - \Sigma\Lambda)H_t.$$

Also note that the main persistence parameter relevant for term structure pricing is the $\mathbb{Q}$-persistence, which in this example is

$$\rho^{\mathbb{Q}} = \rho - \Sigma\Lambda. \tag{24}$$

D.i. *Cash Flows with One Factor Under $\mathbb{P}$, Two Factors Under $\mathbb{Q}$*

We now specialize from the preceding example to a case where cash flows follow a one-factor structure under $\mathbb{P}$, but a two-factor structure under $\mathbb{Q}$. First, assume that the two factors evolve autonomously, so that ρ is diagonal with elements ρ_1 and ρ_2, and also assume that the shocks to the two factors are entirely uncorrelated, so that $\Sigma_{12} = \Sigma_{21} = 0$. Next, suppose that physical cash flows are driven only by the first factor,

$$X_t = \exp(H_{1,t})$$

while risk prices are driven only by the second factor,

$$\lambda_t = H_{2,t}.$$

Now, the only role of $H_{2,t}$ is to describe time variation in the price of risk for the physical cash flow, X_t—it does not affect X_t's $\mathbb{P}$-dynamics directly. To summarize,

$$\delta = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \rho = \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}, \quad \text{and} \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}.$$

The dynamics of H_t under $\mathbb{Q}$ are immediate from (24), and the $\mathbb{Q}$-persistence matrix is

$$\rho^{\mathbb{Q}} = \begin{bmatrix} \rho_1 & -\sigma_1 \\ 0 & \rho_2 \end{bmatrix}.$$

In other words, under $\mathbb{P}$, cash flows are only driven by the first factor, and this evolves as an autonomous AR(1) process. But, under $\mathbb{Q}$, the evolution of the first factor is no longer autonomous (as seen from $\rho^{\mathbb{Q}}$) and instead follows a VAR(1). So cash flows are a one-factor model under $\mathbb{P}$, but a two-factor model under $\mathbb{Q}$. The difference stems from the fact that the $\mathbb{Q}$ measure incorporates variation in risk premia, which in this example is represented with an additional cash flow factor in the $\mathbb{Q}$ measure. In our paper, we estimated the dynamics of the $\mathbb{Q}$ measure directly. In this way, the null models throughout our analysis allow for any time variation in risk premia that is describable within the affine framework.

E. Model Testing

E.i. Bootstrap Inference for the Regression-based Test

We obtain bootstrap standard errors using the semiparametric bootstrap, and generate bootstrap samples from a null model where the affine restrictions hold. That is, under the null the variance ratio is one at *all* maturities.

Bootstrap standard errors are used to test the null hypothesis that the variance ratio at a given maturity $n > K$ is equal to one, or equivalently that the covariances of prices at maturity $n > K$ are consistent with the restricted (affine) model estimated from the vector of prices at maturities $1{:}K+1$.

The bootstrap proceeds as follows. We first estimate the null (affine) model using the first K maturities, i.e. we estimate the vector $\rho^{\mathbb{Q}}$ from the regression of $p_{t,K+1}$ on $P_{t,1:K}$. We then generate bootstrap samples from the affine model where $\rho^{\mathbb{Q}}$ determines the loadings of prices onto factors at *all* maturities: this model by definition features a variance ratio of one at all maturities in population. However, resampled measurement error is added to the prices in each bootstrap sample, so that in each sample the variance ratio will not be exactly one. We count how often the variance ratios estimated in each re-sample are as high as those we observe in the data. This provides a p-value for the one-sided test of the affine null that the variance ratio is equal to one.

In particular, we construct our bootstrap estimator using the semiparametric bootstrap procedure described in Davidson and MacKinnon (2004). First, we estimate the null model and construct fitted prices *under the null model* for each time t and all maturities n, $\hat{p}_{t,n}$. By construction, if one were to compute the variance ratio statistic using the panel of prices $\hat{p}_{t,n}$, one would find a variance ratio of exactly one at all maturities. To account for estimation error, we construct the panel of errors for each t and n, as:

$$\hat{\epsilon}_{t,n} = p_{t,n} - \hat{p}_{t,n}$$

Next, to account for the time-series correlation of the measurement errors $\hat{\epsilon}_{t,n}$, we estimate an AR(1) for the errors for each maturity:

$$\hat{\epsilon}_{t,n} = \gamma_n \hat{\epsilon}_{t-1,n} + \hat{u}_{t,n}$$

Each bootstrap sample is then generated by jointly resampling the error innovations $\hat{u}_{t,n}$ across maturities (to take into account also the cross-sectional correlation of measurement errors). Denote with tildes the quantities that are generated in each bootstrap sample; for example, the resampled error innovations u are denoted $\tilde{u}_{t,n}$. Using the estimated persistence $\hat{\gamma}_n$ for each maturity, together with the resampled error innovations $\tilde{u}_{t,n}$, we generate a panel of resampled errors $\tilde{\epsilon}_{t,n}$, again jointly across maturities. The panel of bootstrapped prices are then constructed as:

$$\tilde{p}_{t,n} = \hat{p}_{t,n} + \tilde{\epsilon}_{t,n}$$

Using the resampled term structure of prices $\tilde{p}_{t,n}$, we re-run our entire analysis in each bootstrap sample. In particular, in each sample we run both restricted and unrestricted regressions, and obtain a variance ratio statistic for that bootstrap sample. Importantly, we re-estimate the matrix $\rho^{\mathbb{Q}}$ in the bootstrap sample (obtaining a different estimated matrix $\tilde{\rho}^{\mathbb{Q}}$ in each bootstrap sample) to obtain the variance ratio test statistic. Because we re-estimate $\rho^{\mathbb{Q}}$ in each bootstrap sample, our procedure takes into account sampling uncertainty regarding the decay rate under $\mathbb{Q}$. We conduct all bootstrap inference using 1,000 bootstrap samples.

To sum up, the bootstrap samples add measurement error to a set of prices $\hat{p}_{t,n}$ that always satisfies the affine restrictions: in population, the variance ratio should be 1 at all maturities. The point of the bootstrap is to understand how likely it is, if the true data were generated by an affine model, we would observe variance ratios as high as we do in the data due only to the presence of estimation error in the regression coefficients of prices onto the short end prices $P_{t,1:K}$.

Note also that the semiparametric bootstrap does not need us to specify the $\mathbb{P}$ dynamics of the short-term prices $P_{t,1:K}$, just those dynamics of the errors (which are assumed to follow an AR(1) whose innovations are potentially cross-sectionally correlated).

We have also derived the analytical asymptotic distribution of the variance ratio statistic and compared this with the finite sample bootstrap-based inference, and they behave similarly in moderately sized samples. We find that bootstrap standard errors are more conservative in small samples and thus base our main analysis on these. Details for the derivation of the asymptotic distribution and its comparison with the bootstrap distribution are available upon request.

E.ii. Estimation and Inference for the Instrumented Regression-based Test

In this section we discuss estimation and testing for excess volatility using instrumental variables to handle potential measurement error at maturities 1 to K. We work under assumptions 1-3, but we replace Assumption 4REG with the following:
Assumption 4IV.

> Prices obey:
>
> $$P_{t,1:N} = \delta_0[1 \ ... \ N]' + [I_K \ \beta_{K+1} \ ... \ \beta_N]' H_t + \nu_t \tag{25}$$
>
> where H_t is a stationary process, and all variables satisfiy standard regularity conditions for OLS and Wald test consistency. In addition, there exist a set of K instruments Z_t for which $E[\nu_{t,j}|Z_t] = 0$ for $j = 1...K$.

Under Assumption 4IV, the instruments Z_t are exogenous to the short-maturity error ν_t, thus we can recover the coefficients $\beta_{K+1}...\beta_N$ using a standard IV regression. In particular, consider the population IV estimator (ignoring constants for simplicity):

$$\overline{\beta}_j = E[Z_t P'_{1:K,t}]^{-1} E[Z_t p_{j,t}],$$

therefore

$$\overline{\beta}_j = E[Z_t P'_{1:K,t}]^{-1} E[Z_t p_{j,t}] = E[Z_t(H'_t + \nu'_{1:K,t})]^{-1} E[Z_t(H'_t\beta_j + \nu'_{j,t})] = E[Z_t H'_t]^{-1} E[Z_t H'_t]\beta_j = \beta_j,$$

which is the basis of the usual consistency argument for IV. That is, the IV regression consistently estimates each β_j. From here, our variance ratio test can be constructed in the same way as the OLS regression-based test, and it is only a function of the $\beta_{K+1}, ..., \beta_N$ estimates and the sample variances of prices.

We construct bootstrap standard errors in the same way as the OLS-based regression test. In particular, we estimate the model using the IV-based regression approach discussed above, obtaining an estimate of $\rho^{\mathbb{Q}}$. We then generate bootstrap samples from the affine model where $\rho^{\mathbb{Q}}$ determines the loadings of prices onto factors at all maturities. As in the case of the OLS bootstrap, the short-end prices stand in for factors, but now they are projected onto the instruments to remove measurement error (corresponding to the first stage of the two-stage least square approach for IV). In this model, therefore, the short-end factors are free of measurement error, and the variance ratio is one at all maturities in population.

Like before, resampled measurement error is added to all prices in each bootstrap sample. Thus, in each bootstrap data set the variance ratio differ from one due to sampling variation. We resampe measurement error exactly as described for the OLS regression case. In each bootstrap sample, we

Table I
Simulated Variance Ratio Tests Under Correct Specification

	$\sigma_2^2/\sigma_1^2 = 0.25$			$\sigma_2^2/\sigma_1^2 = 0.10$			$\sigma_2^2/\sigma_1^2 = 0.05$			$\sigma_2^2/\sigma_1^2 = 0.01$		
ρ_2	5%	10%	$\frac{\text{Std}(VR)}{\text{BSE}(VR)}$	5%	10%	$\frac{\text{Std}(VR)}{\text{BSE}(VR)}$	5%	10%	$\frac{\text{Std}(VR)}{\text{BSE}(VR)}$	5%	10%	$\frac{\text{Std}(VR)}{\text{BSE}(VR)}$
0.9000	0.103	0.149	0.896	0.129	0.169	0.873	0.113	0.162	0.835	0.030	0.052	1.836
0.9500	0.043	0.097	0.876	0.078	0.123	0.871	0.094	0.130	0.862	0.136	0.167	0.786
0.9900	0.009	0.041	0.911	0.023	0.061	0.853	0.032	0.075	0.871	0.067	0.105	0.885
0.9990	0.019	0.037	1.182	0.024	0.054	1.052	0.033	0.069	0.951	0.071	0.116	0.942
0.9999	0.057	0.091	1.447	0.060	0.104	1.070	0.067	0.109	0.978	0.099	0.152	0.958

Note. Realized rejection rates across 5,000 simulations at 5% and 10% bootstrap critical values. $\frac{\text{Std}(VR)}{\text{BSE}(VR)}$ is the ratio of the standard deviation of 24-month variance ratio statistics to the median bootstrap standard error across simulations.

perform our instrumented IV estimation and finally we count how often the bootstrap variance ratios are as high as those we observe in the data. This delivers a p-value for the one-sided test of the affine null that the variance ratio is equal to one.

E.iii. Finite Sample Simulations

Our approach to inference for variance ratios relies on factor persistences estimated from prices on the short end of the term structure. A natural concern is that it may be hard to estimate the behavior of a small but very persistent factor from the short end alone. In other words, even when the model is correctly specified, one may be concerned that short end prices are not informative enough about the $\mathbb{Q}$-dynamics of low volatility/high persistence factors, and that this may lead to inappropriate inference. In this appendix, we show that this is not the case. Short maturity prices are sufficiently informative about low frequency $\mathbb{Q}$-dynamics so that our variance ratio tests always retain correct size. That is, when the null hypothesis is true, we reject the null approximately 5% of the time when we use a 5% critical value, we reject the null approximately 10% of the time when we use a 10% critical value, and so forth. In other words, our rejection of the affine model is not driven by our choice to estimate model parameters using short end prices.

Our estimation and inference procedure is well behaved because our bootstrap distribution for the variance ratio statistic takes into account sampling variation in the parameters estimated from the short end. If there are some parameters that are hard to accurately estimate (for example, the persistence parameter for a low variance factor), the variation in bootstrap samples fully accounts for this.

To understand the performance of our inference approach we conduct simulations. We generate term structures of prices with maturities up to 24 periods assuming a two-factor model. To keep the setting simple, we assume that the $\mathbb{P}$ and $\mathbb{Q}$ measures are the same, thus risk premia are zero. The first factor is the dominant factor and has variance $\sigma_1^2 = 1$ and persistence $\rho_1 = 0.75$. For the weaker second factor, we consider a gradually decreasing range of variance ($\sigma_2^2/\sigma_1^2 \to 0$) and a gradually increasing range of persistence ($\rho_2 \to 1$). We assume that measurement error is two percent of the unconditional price volatility for maturities greater than two. Based on 1,000 periods of simulated term structure prices, we estimate the model using the shortest maturities (1,2, and 3) and calculate the variance ratio statistic, its standard error, and its p-value for the 24-month claim. We generate 5,000 such samples at each set of parameters, and report summary statistics across simulations. We report the realized rejection rates based on 5% and 10% critical values of the test. We also report the ratio of the standard deviation of the variance ratio statistic to the median bootstrap standard error across simulation; which should be near one if the test is behaving appropriately. Results are shown in Table I.

Overall, finite sample inference behaves reasonably. The test seems to reject too *infrequently*, and the realized standard deviation of the variance ratio statistic tends to be slightly smaller than the asymptotic standard error. These facts indicate that the critical values that we use in our empirical analysis are slightly conservative.

F. Data Details and Asset-specific Modeling Considerations

In this section we show how each asset class considered maps into our linear or log-linear framework.

F.i. Variance Swaps and Related Variance Derivatives

As discussed in the text, the price of a variance swap follows:[45]

$$p_{t,n} = E_t^{\mathbb{Q}}\left[\sum_{j=1}^{n} RV_{t+j}\right]$$

We then model RV_t as a linear function of the factors, which immediately yields:

$$p_{t,n} = a_n + b_n' H_t \tag{26}$$

An attractive feature of the simple payoff structure of variance swaps is that dependence of prices on factors, $b_n' H_t$, is robust to many modifications of the factor model. For example, because the swap price is the expected value of the level of RV_t, having both prices and payoffs linear in the factors no longer requires Gaussianity. Any shock distribution with constant means implies the pricing structure in (26).

One important consideration to keep in mind is that because variances are non-negative, a homoskedastic linear Gaussian model is an imperfect description of RV_t. Stochastic variance is a standard feature in the bond and option pricing literatures, and a number of solutions exist that ensure positive variances. The most common solution is to use a CIR volatility process. In these models, the model innovations remain standard normal, but are multiplied by a volatility that scales with the factors (and hence with the level of volatility). The modified model takes the general form[46]

$$H_t = \rho H_{t-1} + \Sigma_{t-1} u_t$$

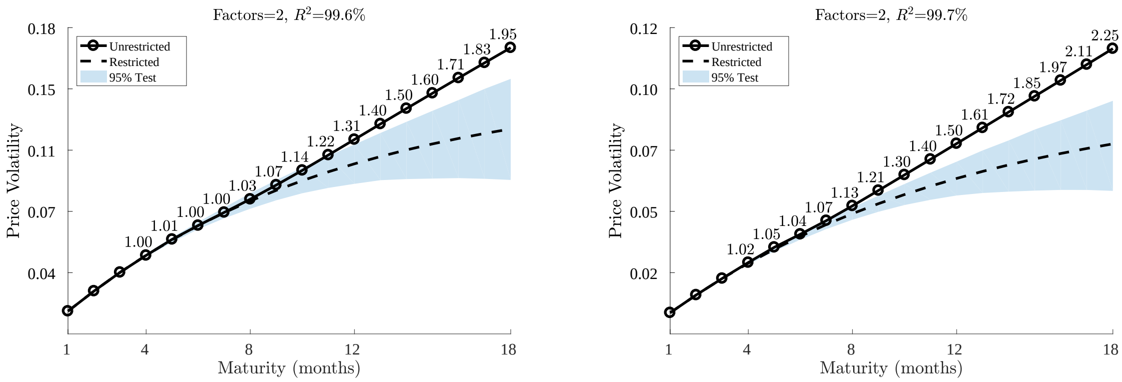
where Σ_{t-1} is a constant function of H_{t-1}. When the model is specified at a high enough frequency (going to continuous time in the limit), and assuming appropriate Feller conditions for the model parameters (see Dai and Singleton (2002)), the probability of variance going below zero tends to zero.

[45] We ignore risk-free rate variation, since its volatility and correlation with the variance swap payoff are small, following Ait-Sahalia, Karaman and Mancini (2015), Egloff, Leippold and Wu (2010), Dew-Becker et al. (2015).

[46] For infinitesimal time intervals, the variance may be constructed to maintain strictly positive variance while retaining the Gaussianity of factor innovations, u_t. In discrete time, this heteroskedastic Gaussian process does not perfectly rule out negative variances, but may be constructed to do so with probability arbitrarily close to one.

Figure II VIX term structure

(a) Apple IV (b) Citigroup IV



Note. See Figure I.

Note that this stochastic volatility case only affects the scale of the innovation u_t. Therefore, the expected *level* payoff in is unaffected, hence equation (26) is also unaffected. Different versions of this model are applied by Ait-Sahalia, Karaman and Mancini (2015), Egloff, Leippold and Wu (2010), Dew-Becker et al. (2015).

As discussed in the text, in some of our tests we take ATM implied variance as a proxy for the risk-neutral expected variance. This is motivated by the theoretical result of Carr and Lee (2009) who show that to a first-order approximation, ATM implied volatility corresponds to the price of a volatility swap (a claim to realized volatility). Perhaps more importantly, our use of ATM is also motivated by practical considerations. ATM volatility is more widely available, especially for long dated options, because it only requires one ATM option price to construct. The synthetic variance swap price, VIX^2, can be calculated for all of our option term structures but is less stable than ATM implied volatility due to its reliance on OTM option prices, of which fewer are available at long maturities.

Our analysis of the term structure of ATM implied variance uses the same model as for variance swaps, but sets $p_{t,n} = IV_t^n$, where IV_t^n is the n-maturity option-implied variance. To construct implied variances at constant monthly maturities from observed options (whose maturities are fixed in calendar time), we linearly interpolate the implied volatilities.

As a robustness check, we also construct the term structure of the VIX using option prices, following the SVI fitting procedure described in Dew-Becker et al. (2015). Note that we need to both interpolate and extrapolate the implied volatility curve (using the SVI model), and the relative scarcity of out-of-the-money options at long maturities can result in noisy VIX estimates. Also, for some of our options sample, there are not enough OTM options available to estimate the VIX at maturities above one year. We report the results using sample dates where the entire term structure up to 18 months is observed (for all contracts, we have between 1,000 and 2,000 days that can be used for estimation). Figure II shows that the variance ratios for the term structure of the VIX behave very similarly to the ones constructed for implied volatilities.

F.ii. Treasuries

Our development of the exponential-affine model for interest rates follows Hamilton and Wu (2012), who study the class of Gaussian affine term structure models developed by Vasicek (1977), Duffie, Kan et al. (1996), Dai and Singleton (2002), and Duffee (2002), and studied by many others.

In the Gaussian affine term structure model, bonds are claims on short-term interest rates. One-period log risk-free rate x_t is a linear function of the factors with factor dynamics under the pricing measure described by a VAR, just as in our main set-up. The price of a risk-free bond that pays \$1 after n periods is

$$P_{t,n} = E^{\mathbb{Q}}\left[\exp\left(-\sum_{j=1}^{n} x_{t+j}\right)\right]. \tag{27}$$

We assume that factor shocks are homoskedastic, so that $\Sigma_t = \Sigma$, following Hamilton and Wu (2012), which implies that the log bond price is

$$p_{t,n} \equiv \log P_{t,n} = a_n + b_n H_t.$$

The factor loading depends only on the persistence of the factors:

$$b_n = \mathbf{1}'(I + \rho^{\mathbb{Q}} + ... + (\rho^{\mathbb{Q}})^{n-1}). \tag{28}$$

The intercept is an inconsequential constant function of remaining model parameters, and drops out from all variance calculations.

F.iii. Credit Default Swaps

To model CDS spreads, we apply the reduced-form modeling of Duffie and Singleton (1999), in which the price of a defaultable bond is written in terms of a default intensity process λ_t and a process of loss given default L_t. The precise relationship between the price of the bond at time t, $P_{t,n}$, and the processes for λ_t and L_t does not directly map into our general framework of Section II..

However, Duffie and Singleton (1999) show that under the assumption of fractional recovery of market value in case of default, the price of a defaultable zero-coupon bond can be written as:

$$P_{t,n} = E_t^{\mathbb{Q}}\left[\exp(-\int_t^n R_s ds)\right]$$

with

$$R_s = r_s + \lambda_s L_s$$

where λ_t is the default intensity and L_t the loss given default. The defaultable bond can be modeled as a default-free bond with a default-adjusted interest rate. We assume that 1) r_s and $\lambda_s L_s$ are linear in the factors; 2) underlying factors are homoskedastic; and 3) coupons on the underlying defaultable bonds are small enough (relative to the default-adjusted interest rate) so that the yield of an n-maturity defaultable bond with coupon is close to an n-maturity zero-coupon defaultable bond. We can then write:

$$p_{t,n} = \log(P_{t,n}) = -ny_t^n = (a_r^n + a_{\lambda L}^n) + (b_r^n + b_{\lambda L}^n)H_t$$

while for the default-free bond with maturity n (with log yield y_F^n) we have:

$$-ny_{F,t}^n = a_r^n + b_r^n H_t.$$

To link the bond price to the observed CDS spread, we start from the approximate bond-CDS basis relation, that states

$$Z_t^n \simeq Y_t^n - Y_{F,t}^n$$

i.e. the CDS spread Z_t^n with maturity n is approximately equal to the yield of the bond Y_t^n of that maturity in excess of the corresponding risk-free rate Y_F^n with the same maturity.

Given that both Y_t^n and $Y_{F,t}^n$ are close to zero, we can write the yield spread to a first-order approximation as:

$$Y_t^n - Y_{F,t}^n \simeq \log(1 + Y_t^n) - \log(1 + Y_{F,t}^n) = y_t^n - y_{F,t}^n$$

so that:

$$nZ_t^n \simeq n(y_t^n - y_{F,t}^n) = -a_{\lambda L}^n - b_{\lambda L}^n H_t.$$

This representation allows us to focus on the cross-section of CDS spreads stripped of the risk-free rate dynamics, which will highlight the factor structure in default risk.

F.iv. Inflation Swaps

Inflation swaps are claims to future inflation where the the buyer commits to pay a predetermined amount $(1+p_{t,n})^n - 1$ and receives $[I(t+n)/I(t)] - 1$, where $I(t)$ is the price level index. Risk-neutral pricing implies:

$$(1 + p_{t,n})^n - 1 = E_t^{\mathbb{Q}}\left[\frac{I(t+n)}{I(t)} - 1\right].$$

Calling $\pi_t = \Delta \log I(t)$, and moving to continuous time, we can write:

$$P_{t,n} = e^{p_{t,n}n} = E_t^{\mathbb{Q}}\left[\exp(\int_t^{t+n} \pi_s ds)\right].$$

Just as in the case of bonds, we will have that log cumulative prices $n \cdot p_{t,n}$ will be linear in the factors:

$$n \cdot p_{t,n} = a_n + b_n H_t.$$

F.v. Commodity Futures

Call $F_{t,n}$ the price of a future with maturity n. As in Duffie (2010) and Casassus and Collin-Dufresne (2005), if S_t is the value of the underlying at time t, we have:

$$F_{t,n} = E_t^Q[S_{t+n}]$$

Now, if $X_t = \log(S_t)$, then we have:

$$F_{t,n} = E_t^Q[\exp\{X_{t+n}\}].$$

We can rewrite X_t as:

$$X_{t+n} = X_t + \sum_{s=1}^{n} x_{t+s}$$

with $x_t = \Delta X_t$. We may model these growth rates as functions of latent factors, so that $x_t = \delta_1' H_t$ and

$$F_{t,n} = E_t^Q[\exp\{X_t + \sum_{s=1}^{n} x_{t+s}\}].$$

We can therefore rewrite:

$$\frac{F_{t,n}}{S_t} = E_t^Q[\exp\{\sum_{s=1}^{n} x_{t+s}\}]$$

which has the standard exponential affine form. Note also that we can rewrite the expression for the futures without reference to the underlying, rescaling each future by the price of the first-maturity future:

$$F_{t,1} = S_t E_t^Q[\exp\{x_{t+1}\}]$$

so that:

$$\frac{F_{t,n}}{F_{t,1}} = \frac{E_t^Q[\exp\{\sum_{s=1}^n x_{t+s}\}]}{E_t^Q[\exp\{x_{t+1}\}]} \simeq E_t^Q[\exp\{\sum_{s=2}^n x_{t+s}\}]$$

This expression maps directly into our exponential-affine framework. Note finally that given the futures have fixed calendar time expiration dates, we linearly interpolate log future prices to obtain constant-maturity prices with monthly maturities.

F.vi. Quotes vs. Trades

In this paper we document excess volatility across a variety of asset classes, some of which are traded over the counter (like variance swaps and CDS), and some on exchanges (like options). Since for some of these markets we don't observe transaction prices, but only quotes, one may be worried that the excess volatility results may be driven by liquidity issues (stale prices, matrix prices, etc.) rather than true fundamental excess volatility.

There are several reasons why we believe our results are not due to liquidity issues. First, for our baseline term structure (variance swaps), the existing literature has already verified the accuracy of the quotes in our dataset against trades obtained from the DTCC. In particular, Dew-Becker et al. (2016) show that the median absolute pricing error is 1% of the transaction price. The variance swap quotes are therefore extremely accurate.

Second, one of the reasons to include measurement error in modeling prices (see Section IV.D.) is precisely to capture deviations from fundamentals due to liquidity (see Duffee, 2011). We allow for measurement error to be correlated over time and across maturities, therefore capturing different forms of liquidity effects on prices. In addition, we show that in several markets, measurement error volatility as large as the bid-ask spread (and therefore directly motivated by liquidity concerns) cannot explain the patterns of excess volatility we document.

Third, we have obtained transaction prices for 2013 from the DTCC. In addition to checking quotes against these prices (as in Dew-Becker et al. 2016), we can use these transaction prices to perform our variance ratio tests. Using transaction prices, we actually obtain a stronger variance ratio in the variance swap market: 2.68 at the 24 month maturity (significant at the 1% level). We have also performed a similar analysis using oil futures and options transaction prices from the CME, finding a statistically significant variance ratio of 4.6 at the longest maturity (24 months), higher than the value of 2.67 we found using quotes. Our transaction-based results therefore confirm our findings in the variance swap and commodities futures markets.

G. Additional Empirical Results

In Table II we report regression-based variance ratio tests in which we add one extra factor to each asset class and find broadly similar variance ratios to that in our baseline analysis. When adding extra factors, panel R^2s often reach above 99.9%. In order to add factors, we need to use more prices along the term structure to estimate the model. In some cases, the "short end" of the curve becomes almost the entire curve, as in the case of credit default swaps, and this reduces the tests power to detect affine violations. That is, if there are genuine affine violations at intermediate and long maturities, and some of these are mistakenly included in the definition of the short end, the violations will not be detected. We therefore expect the variance ratios to decline from adding factors. While we do see that long maturity variance ratios decrease for several term structures, the vast majority of them remain economically and statistically above one.

Table III describes summary statistics of estimated KF-MLE residuals from the alternative model that does not impose affine pricing restrictions. The first two columns of the table report the estimated parameters of the Σ matrix. The first parameter governs the ratio of residual standard deviation to raw price standard deviation. The second parameter governs the cross-sectional correlation among residuals. Columns 3 and 4 report the serial correlation of the measurement errors. For comparability across asset classes with different maturity structures, we average serial correlations among the short end of the curve (maturities 1 to $K + 1$) and the long end of the curve (maturities $K + 2$ and higher). The last two columns report the model R^2 at each maturity, again averaged between for the short end and the long end of the curve.

Table II

Variance Ratio Tests with additional factor

Asset	Baseline K	Baseline R^2				Extra factor K	Extra factor R^2			
				Panel A: Equity Variance						
			6m	12m	24m			6m	12m	24m
Variance Swaps	2	99.7	1.00	1.22**	2.15**	3	99.94	1.00	1.21**	2.16**
			12m	18m	24m			12m	18m	24m
Apple IV	2	99.3	1.21**	1.56**	2.01**	3	99.88	1.00	1.05**	1.14**
Citigroup IV	2	99.7	1.82**	3.17**	4.68**	3	99.93	1.00	1.01	0.99
STOXX 50 IV	2	99.4	1.22**	1.68**	2.27**	3	99.90	1.00	0.76	0.41
DAX IV	2	99.4	1.22**	1.68**	2.31**	3	99.75	1.00	1.11**	1.30**
				Panel B: Currency Variance						
			12m	18m	24m			12m	18m	24m
Euro IV	2	99.8	1.22**	1.65**	2.14**	3	99.93	1.00**	1.05**	1.11**
Yen IV	2	98.5	1.67	2.85*	4.57*	3	99.39	1.00	3.34*	6.51*
				Panel C: Interest Rates						
			20y	25y	30y			20y	25y	30y
Treasuries	3	99.9	1.20**	1.39**	1.64**	4	99.98	1.15**	1.31**	1.52**
				Panel D: Inflation						
			20y	25y	30y			20y	25y	30y
US Infl. Swaps	4	99.4	3.37**	5.54**	7.47**	5	99.63	1.30	1.60	1.75
EU Infl. Swaps	4	99.1	1.74**	2.45**	2.89**	5	99.45	1.21**	1.58**	1.83**
				Panel E: Commodities						
			6m	12m	24m			6m	12m	24m
Crude Oil Fut.	2	99.6	1.01**	1.19**	1.63**	3	99.95	1.00	1.11**	1.43**
Gold Fut.	2	99.5	1.04*	1.19**	1.53**	3	99.90	1.00	0.97	0.87
				Panel F: Credit						
			5y	7y	10y			5y	7y	10y
Brazil CDS	2	99.8	1.19**	1.64**	3.08**	3	99.95	1.00	1.00	1.06
Russia CDS	2	99.8	1.14**	1.46**	2.18**	3	99.99	1.00	1.02	1.05
GE CDS	2	99.5	1.12**	1.13**	1.45**	3	99.69	1.00	0.90	1.02
BoA CDS	2	99.7	1.06**	1.14**	1.38**	3	99.94	1.00	0.98	1.02

Note. The table reports regression-based variance ratio tests using an additional factor. The left panel reports the baseline results (as in table II), whereas the right panel adds one factor. Significance for the one-sided test that the variance ratio is greater than 1 at the 1% level is denoted by ** and at the 5% levels by *. Monthly maturities are denoted by "m" and annual maturities by "y."

Table III

SUMMARY OF KALMAN FILTER RESIDUALS

	Resid Vol./ Price Vol.	Cross-sec Corr.	Serial Correlation		R^2	
			Short End	Long End	Short End	Long End
Variance Swaps	0.07	0.02	0.77	0.77	99.5	99.7
Apple IV	0.10	0.00	0.75	0.63	98.8	99.5
Citigroup IV	0.07	0.00	0.37	-0.07	99.7	99.7
STOXX 50 IV	0.09	0.00	0.73	0.55	99.1	99.7
DAX IV	0.10	0.00	0.65	0.37	99.2	99.4
Euro IV	0.05	0.00	0.37	0.27	99.8	99.9
Yen IV	0.15	0.16	0.62	0.30	98.3	98.0
Treasuries	0.04	0.20	0.93	0.68	99.8	99.9
US Infl. Swaps	0.10	0.31	0.43	0.53	99.3	98.4
EU Infl. Swaps	0.09	0.29	0.25	0.22	99.2	98.1
Crude Oil Fut.	0.07	0.00	0.78	0.93	99.6	99.5
Gold Fut.	0.08	0.00	0.85	0.91	99.3	99.5
Brazil CDS	0.08	0.05	0.87	0.76	99.6	99.6
Russia CDS	0.05	0.00	0.92	0.83	99.8	99.9
GE CDS	0.09	0.00	0.37	0.39	99.4	99.3
BoA CDS	0.07	0.00	0.76	0.74	99.7	99.7

Note. Summary statistics of KF-MLE residuals.

Table IV
Simulations using KF-MLE measurement error

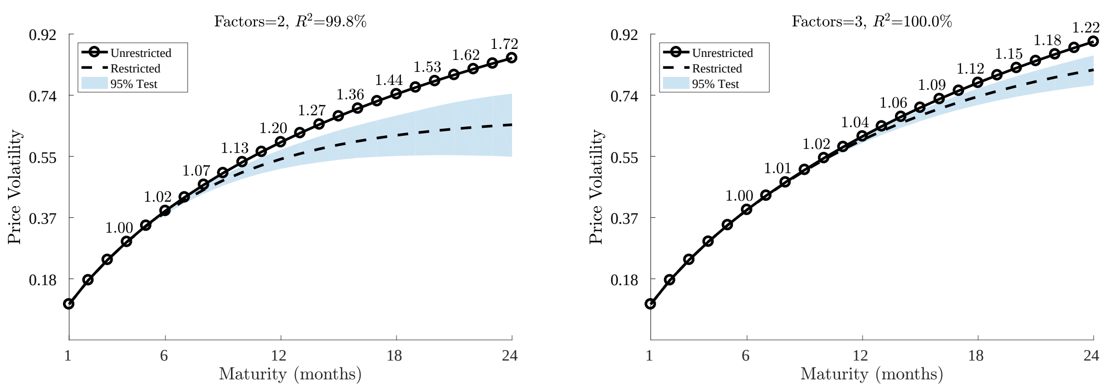
Asset	Data					
	Actual data			Affine + meas. err.		
	Panel A: Equity Variance					
	6m	12m	24m	6m	12m	24m
Variance Swaps	1.00	1.22**	2.15**	1.03	1.10**	1.18**
	12m	18m	24m	12m	18m	24m
Apple IV	1.21**	1.56**	2.01**	1.10**	1.19**	1.24**
Citigroup IV	1.82**	3.17**	4.68**	1.00	0.99	0.99
STOXX 50 IV	1.22**	1.68**	2.27**	1.00	1.00	1.00
DAX IV	1.22**	1.68**	2.31**	1.00	1.00	1.00
	Panel B: Currency Variance					
	12m	18m	24m	12m	18m	24m
Euro IV	1.22**	1.65**	2.14**	1.00	1.00	1.00
Yen IV	1.67	2.85*	4.57*	0.99	0.93	0.89
	Panel C: Interest Rates					
	20y	25y	30y	20y	25y	30y
Treasuries	1.20**	1.39**	1.64**	0.20	0.07	0.02
	Panel D: Inflation					
	20y	25y	30y	20y	25y	30y
US Infl. Swaps	3.37**	5.54**	7.47**	1.00	1.01	1.14**
EU Infl. Swaps	1.74**	2.45**	2.89**	1.00	0.94	0.98
	Panel E: Commodities					
	6m	12m	24m	6m	12m	24m
Crude Oil Fut.	1.01	1.19**	1.63**	1.00	1.00	1.00
Gold Fut.	1.04*	1.19**	1.53**	1.00	0.99	0.97
	Panel F: Credit					
	5y	7y	10y	5y	7y	10y
Brazil CDS	1.19**	1.64**	3.08**	0.76	0.51	0.27
Russia CDS	1.14**	1.46**	2.18**	1.01	1.02	1.03
GE CDS	1.12**	1.13**	1.45**	0.99	0.97	0.93
BoA CDS	1.06**	1.14**	1.38**	1.01	0.99	0.98

Note. The left panel reports the results for our regression-based variance ratio tests estimated on the actual data. The right panel reports the regression-based test estimated on simulated data. The simulated data is obtained as the fitted data from the affine model estimated from the short end of the term structure (which without measurement error would have variance ratio of 1 at all maturities), adding the measurement error estimated using KF-MLE. The number of factors is the same in both panels and equal to the benchmark case. Significance for the one-sided test that the variance ratio is greater than 1 at the 1% level is denoted by ** and at the 5% levels by *. Monthly maturities are denoted by "m" and annual maturities by "y."

Table IV reports results from a counterfactual analysis in which hypothetical prices are generated as follows. First, we estimate the $\mathbb{Q}$ dynamics from the short end of the curve exactly as in Section IV.D., generating a term structure of prices that satisfies the affine restrictions at all maturities by construction. Rather than adding simulated measurement error to the fitted prices, we add the estimated measurement errors from the unrestricted KF-MLE estimation. Next, we re-estimate regression-based variance ratios on the generated error-ridden prices. We find that estimated KF-MLE measurement error is unlikely to produce variance ratios as high as those we observe in the actual data. For most asset classes, variance ratios are just above or just below one. An interesting exception is the yield curve, where the variance ratio approaches zero at the long end. This occurs because $\mathbb{Q}$ dynamics estimated from the short end of the error-ridden hypothetical prices include an explosive root, hence the long end of the curve appears insufficiently volatile relative to the restricted model estimate.

H. Additional Simulations of Non-affine Models

Figure III Multifractal Variance Model



Note. Simulation of the multifractal volatility model as in Calvet and Fisher (2004), and variance ratio test with 2 factors (left panel) or 3 factors (right panel). See also Figure I.

This appendix presents results obtained by applying our variance ratio tests to additional non-affine term structures. Table V extends the analysis of non-linear logistic STAR model to allow for heteroskedastic shocks. The specifications are identical to those in Table IV and the shocks share the same unconditional shock variance. In Table V, however, the shocks follow a GARCH(1,1) process with parameters of $\alpha = 0.05$ and $\beta = 0.90$. The results and conclusions from Table IV are unaffected by the presence of heteroskedasticity.[47]

Next, we analyze the behavior of the variance ratio test for models with more complicated $\mathbb{Q}$-dynamics. In particular, Table VI reports results for various processes that additively combine a non-linear logistic STAR component (as in Section IV.C.) and an ARFIMA component (as in Section IV.B.). Because these specifications involve richer driving processes that the individual STAR and ARFIMA analyses in the main text, we allow the estimated affine model to have up to four factors. Again, this extended analysis does not change our conclusions from the main text.

Finally, we simulate the multifractal model of Calvet and Fisher (2004) for variance, and study the term structure of variance claims with up to 24 months maturity. We use the same parameterization of the variance process as in Calvet and Fisher (2004). Figure III shows that a 2-factor affine model

[47] All simulations in this section are generated by setting $\mathbb{P}$ and $\mathbb{Q}$ distributions to be the equal.

Table V
NON-LINEAR SPECIFICATION WITH HETEROSKEDASTICITY

		ρ=0.01			ρ=0.10			ρ=0.25		
γ	K	R^2	VR_{12}	VR_{24}	R^2	VR_{12}	VR_{24}	R^2	VR_{12}	VR_{24}
0.1	1.0	100.0	1.00	1.00	100.0	1.00	1.00	100.0	1.00	1.00
0.1	2.0	100.0	1.00	1.00	100.0	1.00	1.00	100.0	1.00	1.00
0.1	3.0	100.0	1.00	1.00	100.0	1.00	1.00	100.0	1.00	1.00
0.5	1.0	99.0	1.10	1.21	99.8	1.03	1.04	100.0	1.00	1.00
0.5	2.0	100.0	0.97	0.94	100.0	1.00	1.00	100.0	1.00	1.00
0.5	3.0	100.0	0.99	0.99	100.0	1.01	1.01	100.0	1.00	1.00
1.0	1.0	99.8	1.02	1.05	99.7	1.05	1.07	99.9	1.00	1.00
1.0	2.0	100.0	1.00	1.01	100.0	0.99	0.98	100.0	1.00	1.00
1.0	3.0	100.0	0.99	0.98	100.0	0.99	0.99	100.0	1.00	1.00
5.0	1.0	99.9	1.01	1.01	99.9	1.01	1.02	100.0	1.00	1.00
5.0	2.0	100.0	1.00	1.00	100.0	0.99	0.98	100.0	1.00	1.00
5.0	3.0	100.0	0.99	0.97	100.0	0.99	0.98	100.0	1.00	1.00

Note. Variance ratios and R^2 computed in simulations of a logistic STAR model with parameters γ and ρ. Shocks are GARCH(1,1) with parameters $\alpha = 0.05$ and $\beta = 0.90$, and with an unconditional standard deviation of one. K is the number of factors used in the variance ratio test. VR_{12} is the variance ratio at 12 months maturity, and VR_{24} is the test at 24 months.

generates a variance ratio of 1.7 at 24 months, and adding a third factor brings the variance ratio down to 1.2.

I. CHARACTERIZATION OF THE MODEL MISSPECIFICATION

In this section we propose a general characterization of the potential model misspecification. We start by proposing a recursive representation of the affine model; we then use it to derive in a general way the characteristics of the (non-affine) $\mathbb{Q}$ process that allow the model to perfectly fit the data.

I.i. A Convenient Recursive Representation

Under the assumptions of the affine model, it is easy to show that there exists a vector b such that:

$$f_{t,K+1} = b' F_{t,1:K}. \tag{29}$$

where $b = (b_1, ..., b_K)'$ is the coefficient in a projection of $f_{t,K+1}$ onto $F_{t,1:K}$. In this model, the projection is exact so there is no residual. The vector b depends on the $\mathbb{Q}$ dynamics $\rho^{\mathbb{Q}}$.

This equation only links maturities 1 through $K + 1$. We can derive a recursive relation that links the entire price curve to the short end in a convenient way. In particular, any two blocks of K consecutive forward prices with maturity shifted by one period (for example, $F_{t,1:K}$ and $F_{t,2:K+1}$) are

Table VI
Non-linear and Long Memory Mixture Models

		AR(1)=0.25			AR(1)=0.50			AR(1)=0.75		
d	K	R^2	VR_{12}	VR_{24}	R^2	VR_{12}	VR_{24}	R^2	VR_{12}	VR_{24}
Panel A: Non-linear component $\rho = 0.01$, $\gamma = 0.1$										
0.10	2	100.0	1.10	1.22	100.0	1.11	1.31	100.0	1.06	1.22
0.10	3	100.0	1.00	1.04	100.0	1.01	1.09	100.0	1.02	1.14
0.10	4	100.0	1.00	1.02	100.0	1.00	1.00	100.0	1.00	1.00
0.20	2	100.0	1.02	1.26	100.0	1.22	1.69	100.0	1.10	1.37
0.20	3	100.0	1.00	1.09	100.0	1.01	1.14	100.0	1.03	1.20
0.20	4	100.0	1.00	1.05	100.0	1.00	0.99	100.0	1.00	1.01
0.40	2	100.0	0.95	0.98	100.0	1.36	2.32	100.0	1.08	1.37
0.40	3	100.0	1.01	1.14	100.0	1.01	1.16	100.0	1.02	1.21
0.40	4	100.0	1.03	1.25	100.0	1.00	0.99	100.0	1.01	1.07
0.49	2	100.0	1.06	1.34	100.0	1.38	2.47	100.0	1.03	1.22
0.49	3	100.0	1.01	1.14	100.0	1.01	1.17	100.0	1.01	1.16
0.49	4	100.0	1.03	1.27	100.0	1.00	0.99	100.0	1.02	1.18
Panel B: Non-linear component $\rho = 0.01$, $\gamma = 0.5$										
0.10	2	99.9	1.04	1.16	100.0	1.04	1.16	100.0	1.06	1.22
0.10	3	100.0	1.01	1.09	100.0	1.01	1.07	100.0	1.01	1.10
0.10	4	100.0	1.02	1.13	100.0	1.02	1.13	100.0	1.02	1.11
0.20	2	99.9	1.03	1.14	99.9	1.06	1.21	100.0	1.08	1.28
0.20	3	100.0	1.01	1.10	100.0	1.00	1.07	100.0	1.02	1.13
0.20	4	100.0	1.02	1.13	100.0	1.02	1.13	100.0	1.02	1.13
0.40	2	99.9	1.03	1.17	99.9	1.16	1.53	100.0	1.11	1.47
0.40	3	100.0	1.01	1.13	100.0	1.01	1.13	100.0	1.03	1.24
0.40	4	100.0	1.03	1.19	100.0	1.02	1.16	100.0	1.02	1.18
0.49	2	99.8	1.05	1.25	99.9	1.21	1.74	100.0	1.12	1.55
0.49	3	100.0	1.01	1.15	100.0	1.02	1.19	100.0	1.04	1.27
0.49	4	100.0	1.03	1.24	100.0	1.02	1.13	100.0	1.02	1.19
Panel C: Non-linear component $\rho = 0.01$, $\gamma = 5.0$										
0.10	2	100.0	1.00	1.00	100.0	1.00	1.00	100.0	1.01	1.02
0.10	3	100.0	1.00	0.99	100.0	1.00	0.99	100.0	1.00	0.99
0.10	4	100.0	1.00	1.01	100.0	1.00	1.01	100.0	1.00	1.01
0.20	2	100.0	1.00	0.99	100.0	1.00	1.00	100.0	1.02	1.09
0.20	3	100.0	1.00	0.99	100.0	1.00	0.99	100.0	1.00	0.99
0.20	4	100.0	1.00	1.01	100.0	1.00	1.01	100.0	1.00	1.00
0.40	2	100.0	1.00	0.99	100.0	1.01	1.03	100.0	1.01	1.05
0.40	3	100.0	1.00	0.99	100.0	1.00	0.99	100.0	1.00	1.00
0.40	4	100.0	1.00	1.01	100.0	1.00	1.01	100.0	1.00	1.01
0.49	2	100.0	1.00	0.99	100.0	1.03	1.10	100.0	1.00	1.01
0.49	3	100.0	1.00	0.99	100.0	1.00	0.99	100.0	1.00	1.01
0.49	4	100.0	1.00	1.01	100.0	1.00	1.01	100.0	1.00	1.01

Note. Variance ratios and R^2 computed in simulations of a mixture model that is the sum of ARFIMA(1,d,0) and logistic STAR(ρ, γ) processes. K is the number of factors used in the variance ratio test. VR_{12} is the variance ratio at 12 months maturity, and VR_{24} is the test at 24 months.

linked by the equation:

$$(30) \qquad F_{t,j+1:K+j} = BF_{t,j:K+j-1}, \quad B = \begin{bmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ \vdots & & ... & & \\ 0 & 0 & 0 & ... & 1 \\ b_1 & b_2 & ... & b_{K-1} & b_K \end{bmatrix}.$$

By the definition of b in (29), the relationship in (30) holds for $j = 1$. It follows from the law of iterated expectations that (30) holds for $j = 2$ because

$$E_t^{\mathbb{Q}}[F_{t+1,2:K+1}] = BE_t^{\mathbb{Q}}[F_{t+1,1:K}] \Leftrightarrow F_{t,3:K+2} = BF_{t,2:K+1}.$$

A recursive argument therefore establishes (30). It pins down the price of any forward on the term structure with the prices at the K immediate neighboring maturities via the matrix B. Iteratively substituting (30) into itself implies

$$(31) \qquad F_{t,j+1:K+j} = BF_{t,j:K+j-1} = B^2 F_{t,j-1:K+j-2} = ... = B^j F_{t,1:K}.$$

The geometric recursion in (31) further shows that prices at any maturity are pinned down by *any* K prices, even those at distant maturities. In particular, the equation links any price to the "short-end" vector $F_{t,1:K}$, where the coefficients are entirely determined by the powers of B.

I.ii. Characterizing the Mispecification

We now use this algebra to propose a general characterization of our tests under model misspecification.

Our estimator assumes a K-factor affine-$\mathbb{Q}$ model of prices along the term structure. If this is *not* the true data generating process, then the population projection in Equation (29) becomes

$$(32) \qquad f_{t,K+1} = b' F_{t,1:K} + u_t$$

or, in analogy to the matrix recursion in (30),

$$(33) \qquad F_{t,2:K+1} = BF_{t,1:K} + U_t,$$

with B taking the same structure as earlier and $U_t = (0, ..., 0, u_t)'$. Equation (32) now contains a residual that is solely due to specification error.

Under misspecification, the coefficient B in (33) is no longer fixed and instead becomes specific to the maturities used in the projection. For other maturities, the projection coefficient generally takes a different value. This reflects the fact that cross-equation restrictions of the affine model in (31) are only satisfied when the model is correctly specified.

A key question is whether the violations of the cross-equation restrictions observed in the data can tell us anything about the nature of the model misspecification. We arrived at the no-arbitrage restrictions in (31) by iterating expectations in the price-on-price projection equation. Repeating this using the representation of Equation (32) and imposing the no-arbitrage condition that $E_t^{\mathbb{Q}}[f_{t+1,j}] = f_{t,j+1}$, we find for all $j > 1$ that

$$(34) \qquad F_{t,j+1:K+j} = B^j F_{t,1:K} + \sum_{l=0}^{j} B^l E_t^{\mathbb{Q}}[U_{t+l}].$$

Equation (34) is an exact representation of prices at all maturities regardless of misspecification (assuming there is no arbitrage). The first term on the right-hand side captures the variation in $F_{t,j+1:K+j}$ that is consistent with the affine model restrictions given projection (32). The second term captures the deviation from the model. We can decompose the behavior of this deviation by projecting it onto $F_{t,1:K}$. All elements of the vector U_{t+1} other than the first are zero, so we write this projection as

$$e_K \sum_{l=0}^{j} B^l E_t^{\mathbb{Q}}[u_{t+l}] = \gamma_{K+j} F_{t,1:K} + \zeta_{t,K+j},$$

where γ_{K+j} is a K-vector and $\zeta_{t,K+j}$ is scalar. This decomposition allows us to write (34) as

$$F_{t,j+1:K+j} = (B^j + \gamma_{K+j}) F_{t,1:K} + \zeta_{t,K+j} \tag{35}$$

where the projection residual $\zeta_{t,K+j}$ is orthogonal to the first K prices, $F_{t,1:K}$. When testing model restrictions, we estimate the unrestricted linear projection of $F_{t,j+1:K+j}$ on to $F_{t,1:K}$ in (35) and compare the estimated projection coefficient, $(B^j + \gamma_{K+j})$, to the affine-model-restricted coefficient, B^j.

The behavior of the unrestricted projection is informative about the nature of the misspecification. Two stark empirical facts emerge uniformly from data in all asset classes. First, the unrestricted linear factor model (35) provides an excellent fit of the data, with R^2 approaching 100%. Second, variance ratios are significantly greater than one.

Together, these facts provide insights about the behavior of the specification error term, $\sum_{l=0}^{j} B^l E_t^{\mathbb{Q}}[U_{t+l}]$. High variance ratios tell us that the total variation of the specification error, $Var(\sum_{l=0}^{j} B^l E_t^{\mathbb{Q}}[U_{t+l}])$, must be large. At the same time, an unrestricted R^2 approaching 100% means that the portion of the specification error that is uncorrelated with the short maturity prices, $Var(\zeta_{t,K+j})$, must be very small. In other words, the specification error must be nearly perfectly correlated with the factors from the short end. This is evidently the case, as high variance ratios are equivalent to the unrestricted projection coefficients being significantly larger in magnitude than the model restriction allows—the γ_{K+j} coefficients are far from zero (as found in Figure II).

J. Alternative Trading Strategy

In this section we propose an alternative trading strategy based on forward prices instead of cumulative prices. We first discuss the implementation of the forward-based strategy, then show the relation between the forward-based and cumulative-based trading strategies, and finally present the empirical results.

We focus on a one month holding period, both for explaining the trading strategy and in comparing with the cumulative claim strategy in the main text. Table VII describes two possible alternative investments strategies using forward contracts.

Strategy $\mathcal{A}$ buys a forward with maturity $N+1$ at price $f_{t,N+1}$. Strategy $\mathcal{B}$ buys a portfolio of forwards of maturities 2 to K, with weights γ_N, so that the cost of the portfolio is $\gamma_N' F_{t,2:K+1}$. After a one-month holding period, the value of the two portfolios becomes $f_{t+1,N}$ and $\gamma_N' F_{t+1,1:K}$, respectively. Neither trading strategy pays any cash flows during the holding period, because no forwards are actually held to maturity.

Under the null of the affine model, it is possible to choose γ_N such that

$$f_{t+1,N} = \gamma_N' F_{t+1,1:K}$$

Table VII

FORWARD-BASED TRADING STRATEGY

	Strategy $\mathcal{A}$		Strategy $\mathcal{B}$	
Date	Value	Cash Flows	Value	Cash Flows
t	$f_{t,N+1}$	0	$\gamma_N' F_{t,2:K+1}$	0
$t+1$	$f_{t+1,N}$	0	$\gamma_N' F_{t+1,1:K}$	0

Note. Portfolio $\mathcal{A}$ buys the $N+1$-maturity forward at a price of $f_{t,N+1}$. Portfolio $\mathcal{B}$ replicates $\mathcal{A}$ under the affine null model, buying all forward claims with maturities of $2, ..., K+1$ with the number of shares in each claim given by the vector γ_N.

Table VIII

CUMULATIVE-BASED TRADING STRATEGY

	Strategy $\mathcal{A}$		Strategy $\mathcal{B}$	
Date	Ongoing Value	Cash Flows	Ongoing Value	Cash Flows
t	$p_{t,N+1}$	0	$\beta_N' P_{t,2:K+1} + (1-\beta_N'\mathbf{1})p_{t,1}$	0
$t+1$	$p_{t+1,N}$	x_{t+1}	$\beta_N' P_{t+1,1:K}$	x_{t+1}

Note. Portfolio $\mathcal{A}$ buys the $N+1$-maturity claim at a price of $p_{t,N+1}$. Portfolio $\mathcal{B}$ replicates $\mathcal{A}$ under the affine null model, buying all cumulative claims with maturities of $2, ..., K+1$ with the number of shares in each claim given by the vector β_N, and buying $(1-\beta_N'\mathbf{1})$ shares of an 1-period claim.

or, in other words, that the two trading strategies have the same value at time $t+1$. A long-short position will therefore always unwind at zero profit or loss under the affine null. If no-arbitrage holds, the portfolio will not produce profits or loss at inception (at time t). Equivalently, if a trading strategy that trades $\mathcal{A}$ against $\mathcal{B}$ routinely buys the cheaper and sells the dearer to consistently produce positive profits, it represents a violation of no-arbitrage.

J.i. Forward Strategy Versus Cumulative Strategy

Here we show the relation between the forward-based strategy of Table VII and the cumulative-based strategy discussed in Section IV.E. (Table V). When the holding period is one month, the cumulative-based strategy simplifies to the one reported in table VIII.

First, β_N is chosen so that

$$p_{t+1,N} = \beta_N' P_{t+1,1:K},$$

which is guaranteed according to the affine model. Note that cumulative strategies produce a cash flow x_{t+1}, but that is the same across the two strategies $\mathcal{A}$ and $\mathcal{B}$, so that any long-short position in the two strategies will produce zero cash flow.

Next, we show that each of the two forwards-based strategies $\mathcal{A}$ and $\mathcal{B}$ is identical to the difference between adjacent-maturity cumulative-based trading strategies. In particular, a long position in strategy $\mathcal{A}$ for forward claims of maturity N is identical to a long position in Strategy $\mathcal{A}$ for cumulative claims of maturity N plus a short position in Strategy $\mathcal{A}$ for cumulative claims with maturity $N-1$. The same holds for Strategy $\mathcal{B}$.

Table IX shows the link formally for strategy $\mathcal{A}$ in the top panel and for strategy $\mathcal{B}$ in the bottom

Table IX
Cumulative-based trading strategy

	Strategy A, mat. $N+1$		Strategy A, mat. N		Difference	
Date	Value	Cash Flow	Value	Cash Flow	Value	Cash flow
t	$p_{t,N+1}$	0	$p_{t,N}$	0	$f_{t,N+1}$	0
$t+1$	$p_{t+1,N}$	x_{t+1}	$p_{t+1,N-1}$	x_{t+1}	$f_{t+1,N}$	0

	Strategy B, mat. $N+1$		Strategy B, mat. N		Difference	
Date	Value	Cash Flow	Value	Cash Flow	Value	Cash flow
t	$\beta_N' P_{t,2:K+1}$		$\beta_{N-1}' P_{t,2:K+1}$			
	$+(1-\beta_N'\mathbf{1})p_{t,1}$	0	$+(1-\beta_{N-1}'\mathbf{1})p_{t,1}$	0	$\gamma_N' F_{t,2:K+1}$	0
$t+1$	$\beta_N' P_{t+1,1:K}$	x_{t+1}	$\beta_{N-1}' P_{t+1,1:K}$	x_{t+1}	$\gamma_N' F_{t+1,1:K}$	0

Note. For each strategy ($\mathcal{A}$: top panel; $\mathcal{B}$: bottom panel), the table reports the value and cash flows of long position in the cumulative-based strategy with maturity $N+1$ (left), long position in the cumulative-based strategy with maturity N (middle), and the difference between the two (right), i.e. the value and cash flows of a long-short portfolio.

panel. For both strategies, the difference between the values and payoffs of the two cumulative-based strategies is equivalent to the corresponding forward-based strategy. The derivation proceeds as follows. For strategy $\mathcal{A}$, simply note that $p_{N,t} = f_{1,t} + \ldots + f_{N,t}$. For strategy $\mathcal{B}$, note that the affine model implies the following relation between cumulative prices and forwards at different maturities: $f_{t,K+j} = e_1' B^j F_{t,1:K}$, $p_{t,K+j} = e_1'(R + B + \ldots + B^j)R^{-1}P_{t,1:K}$, for R upper-triangular, e_1 a vector of zeros with 1 in the first position, and B a matrix that depends on the $\mathbb{Q}$ dynamics. This also implies that $\gamma_N = e_1' B^{N-K}$ and $\beta_N = e_1'(R + B + \ldots + B^j)R^{-1}$.

Finally, we have $P_{t,1:K} = R \cdot F_{t,1:K}$ and $P_{t,2:K} = R \cdot F_{t,2:K+1} + f_{t,1}$. Therefore the difference between the value of strategy $\mathcal{B}$ at maturity $N+1$ and at maturity N is:

$$\begin{aligned}
\text{Diff.} &= \beta_N' P_{t,2:K+1} + (1-\beta_N'\mathbf{1})p_{t,1} - \beta_{N-1}' P_{t,2:K+1} + (1-\beta_{N-1}'\mathbf{1})p_{t,1} \\
&= \beta_N'(R \cdot F_{t,2:K+1} + f_{t,1}) + (1-\beta_N'\mathbf{1})p_{t,1} - \beta_{N-1}'(R \cdot F_{t,2:K+1} + f_{t,1}) + (1-\beta_{N-1}'\mathbf{1})p_{t,1} \\
&= (\beta_N' - \beta_{N-1}')R \cdot F_{t,2:K+1} \\
&= \left[e_1'(R + B + \ldots + B^{N-K})R^{-1} - e_1'(R + B + \ldots + B^{N-K-1})R^{-1}\right] R \cdot F_{t,2:K+1} \\
&= e_1' B^{N-K} F_{t,2:K+1}.
\end{aligned}$$

This is precisely the same as $\gamma_N' F_{t,2:K+1}$, or the value of the forward-based portfolio of strategy $\mathcal{B}$ at time t. The same holds at maturity $t+1$, thus showing that the forward-based strategy is equivalent to a long-short position in the cumulative-based strategies with staggered maturities.

J.ii. Empirical results

Tables X and XI show the results of the forward-based trading strategies, both in-sample and out-of-sample, using different mispricing thresholds for trading. The tables also report the fraction of trades that yield a positive profit, separately by maturity.

The tables show several interesting results. First, the highest Sharpe ratios (and the highest fractions of positive profits) are concentrated at higher maturities. For forwards of 22 to 24 months

Table X
Forward-based trading strategy: in-sample results

	Threshold: 0%		Threshold: 50%		Threshold: 90%	
Maturity	SR	% pos. ret.	SR	% pos. ret.	SR	% pos. ret.
6	0.80	0.61	1.40	0.62	1.89	0.76
7	0.61	0.60	1.40	0.63	1.86	0.83
8	0.44	0.55	0.91	0.56	1.17	0.62
9	0.19	0.51	0.26	0.47	0.80	0.50
10	-0.05	0.45	-0.40	0.44	-0.02	0.46
11	-0.21	0.44	-0.27	0.48	0.08	0.52
12	-0.02	0.46	0.11	0.48	0.75	0.61
13	0.14	0.47	0.54	0.53	1.51	0.64
14	0.34	0.47	1.17	0.58	1.95	0.67
15	0.46	0.49	1.18	0.58	2.81	0.79
16	0.51	0.51	1.27	0.58	3.12	0.80
17	0.49	0.51	1.32	0.59	3.00	0.79
18	0.14	0.50	0.98	0.59	2.35	0.79
19	-0.22	0.46	0.51	0.54	1.99	0.78
20	0.10	0.48	0.14	0.52	0.30	0.50
21	0.90	0.58	0.93	0.61	1.88	0.75
22	1.54	0.65	1.87	0.74	2.95	0.93
23	1.72	0.70	2.58	0.84	3.40	0.92
24	1.79	0.73	2.92	0.91	3.14	0.92

Note. The table reports annualized Sharpe ratios for forward-based trading strategies that exploit mispricing relative to the affine-$\mathbb{Q}$ model in the variance swap market. The model is estimated using the entire sample. Each panel corresponds to a different level of mispricing (relative to the historical distribution) at which a trade is executed. In each panel, the first column reports the annualized Sharpe ratio of the strategy, the second panel reports the percentage of trades that produce a profit.

maturity, Sharpe ratios are above 1.5 even with a mispricing threshold for trading of 0% (that is, even trading when the mispricing appears small). For these maturities, a large fraction (around 70%) of the trades yield a positive profit. The results strengthen when the mispricing threshold is tightened so that the trade is only made when mispricings are sufficiently large. For the highest threshold we consider (right columns), more than 90% of the trades yield a positive profit for long-term forwards, generating a Sharpe ratio above 3.0. Profits also seem concentrated at higher maturities, but do not entirely disappear at shorter maturities.

Third, results are consistent when looking out-of-sample, highlighting the robustness of the trading strategy results. In fact, the results are slightly stronger in that case, possibly because the model is re-estimated each period using recent data, which may help capture some time variation in the coefficients of the $\mathbb{Q}$ model.

Finally, while some of these maturity-specific forwards have low Sharpe ratios individually, the cumulative-based trading strategies we presented in Section **??** look at the *joint* mispricing of portfolios of forwards, and therefore represent an alternative test of affine model violations.

Table XI
FORWARD-BASED TRADING STRATEGY: OUT-OF-SAMPLE RESULTS

	Threshold: 0%		Threshold: 50%		Threshold: 90%	
Maturity	SR	% pos. ret.	SR	% pos. ret.	SR	% pos. ret.
6	0.89	0.65	1.43	0.70	2.94	0.81
7	0.92	0.63	0.98	0.66	2.99	0.84
8	0.71	0.62	0.74	0.62	2.07	0.72
9	0.45	0.63	0.41	0.62	1.58	0.69
10	0.27	0.58	0.28	0.58	1.39	0.68
11	0.36	0.58	0.75	0.64	1.85	0.71
12	0.88	0.61	1.35	0.68	2.59	0.78
13	1.25	0.63	1.85	0.74	3.97	0.96
14	1.29	0.63	1.94	0.75	3.81	0.96
15	1.29	0.64	2.06	0.76	3.61	0.95
16	1.30	0.66	2.10	0.78	3.44	0.95
17	1.23	0.64	1.99	0.77	3.28	0.95
18	1.10	0.63	1.74	0.75	2.77	0.89
19	0.77	0.59	1.50	0.70	1.77	0.79
20	0.35	0.53	0.52	0.58	0.70	0.60
21	1.13	0.62	1.61	0.68	2.48	0.79
22	1.60	0.74	2.30	0.79	3.80	0.94
23	1.85	0.77	2.42	0.87	4.59	0.94
24	1.81	0.76	2.56	0.90	4.20	0.94

Note. The table reports annualized Sharpe ratios for forward-based trading strategies that exploit mispricing relative to the affine-$\mathbb{Q}$ model in the variance swap market. All strategies are implemented using information available to the investor at the time of the trade, and use a one month holding period ($n = 1$) for each trade. Each panel corresponds to a different level of mispricing (relative to the historical distribution) at which a trade is executed. In each panel, the first column reports the annualized Sharpe ratio of the strategy, the second panel reports the percentage of trades that produce a profit.

K. Affine Example With Measurement Error

Consider a swap with underlying cash flow process (under $\mathbb{P}$) defined as x_t that is the sum of two factors,

$$x_t = H_{1,t} + H_{2,t}$$

where

$$H_t = \begin{pmatrix} 0.3 & 0 \\ 0 & 0.98 \end{pmatrix} H_{t-1} + \begin{pmatrix} \sqrt{1-.3^2} & 0 \\ 0 & 0.2\sqrt{1-.98^2} \end{pmatrix} \epsilon_t$$

where ϵ_t is a bivariate standard normal. The form of the volatility matrix makes it easy to see that the unconditional standard deviations are 1.0 and 0.2 for the two factors, respectively. These dynamics include one factor that is highly volatile and has little persistence. The other has low volatility and high persistence. Under $\mathbb{P}$ the factors are independent.

Next, let risk premia be summarized by a matrix λ that is a wedge between the $\mathbb{P}$ and $\mathbb{Q}$ dynamic specifications. In particular,

$$\lambda = \begin{pmatrix} 0.46 & 0 \\ 0.1 & 0 \end{pmatrix}$$

which transforms the factor dynamics under $\mathbb{Q}$ to

$$H_t = \begin{pmatrix} 0.76 & 0 \\ 0.1 & 0.98 \end{pmatrix} H_{t-1} + \begin{pmatrix} \sqrt{1-.3^2} & 0 \\ 0 & 0.2\sqrt{1-.98^2} \end{pmatrix} \epsilon_t.$$

That is, risk premia depend on the factor with high volatility and low persistence. From here, true affine prices are given by the formulas in Section II. with $\rho^{\mathbb{Q}}$ equal to the $\mathbb{Q}$ persistence matrix in the equation above.

The final element of the model adds iid measurement error with standard deviation of 0.008 to each forward swap price.

K.i. Simulation Analysis

Next, we conduct a simulation analysis of this example model. We generate 1,000 datasets from the model each with T =5,000 observations and maturities up to 24 months. We then compute variance ratios on the simulated sample using the three estimators that we apply to the data in our paper:

1. OLS regression assuming no measurement error in short maturity prices but allowing for long maturity errors

2. KF-MLE, allowing for measurement error throughout the term structure

3. IV regression allowing for measurement error throughout the term structure. We simulate instruments by adding iid noise with standard deviation 0.1 to the true short-end prices, mimicking properties of the instruments we use for variance swaps in Section IV.D..

Below we show histograms of the 1,000 variance ratios estimated from each estimation technique at the 24-month maturity. The left panel reports regression-based variance ratios, middle shows KF-MLE, and right shows IV-based variance ratios. The regression-based test is biased due to measurement error, with an average variance ratio of 2.00 at 24 months with a standard error of 0.03 across simulations. However, the KF-MLE tests have an average variance ratio of 1.05 (standard error across sims of 0.14) and the IV-based average is 0.94 (standard error of 0.40). Thus, our tests that use alternatives to OLS appear unbiased in this example (though noisy in the case of the IV test).

Figure IV Histograms of 24-month Variance Ratios