World Scientific
www.worldscientific.com

AFFINE MODELS WITH STOCHASTIC MARKET PRICE OF RISK

RICCARDO REBONATO

EDHEC Business School, 10 Fleet Place, Ludgate
London EC4M 7RB, UK

EDHEC Risk Institute, 10 Fleet Place
London EC4M 7RB, UK
Riccardo.Rebonato@edhec.edu

In this paper we discuss the common shortcomings of a large class of essentially-affine models in the current monetary environment of repressed rates, and we present a class of reduced-form stochastic-market-risk affine models that can overcome these problems. In particular, we look at the extension of a popular doubly-mean-reverting Vasicek model, but the idea can be applied to all essentially-affine models. The model straddles the $\mathbb{P}$- and $\mathbb{Q}$-measures. By allowing for a market price of risk whose stochasticity is not fully spanned by the yield-curve state variables that enter the model specification, we break the deterministic link between the yield-curve-based return-predicting factors and the market price of risk, but we retain, on average, the observed statistical regularities reported in the literature. We discuss in detail how this approach relates to the recent work by Joslin *et al.* (2014) [S. Joslin, M. Priebsch & K. J. Singleton (2014) Risk premiums in dynamic term structure models with unspanned macro risk, *Journal of Finance* **LXIX** (3), 1197–1233]. We show that the parameters of the model can be estimated in a simple and robust manner using survey-like information; and that the model we propose affords a more plausible decomposition of observed market yields into expectations and risk premia during an important recent market event than the one produced by mainstream essentially-affine models.

Keywords: Affine modelling; market price of risk; predictions of interest rates.

1. The Motivation for This Paper

In this paper we pursue two goals. First, we intend to show that a large class of essentially-affine models with the market price of risk a deterministic (affine) function of the state variables produce a very similar and, in the current monetary conditions, similarly unreliable decomposition of the observed market yields in their expectation and risk-premium components. We explain why this is the case, and we analyze in detail in this light a market event (the "taper tantrum" of Spring 2013)

that constitutes a near-perfect real-life controlled experiment of the risk-premium/ expectation decomposition. To the extent that these models are used as ancillary tools to guide monetary decisions, their lack of reliability in recently experienced monetary conditions is important.

Second, we propose a reduced-form affine model that, by breaking the deterministic link between the market price of risk and the slope of the yield curve, can overcome the problems that we have identified in the first part of the paper. The model we introduce shares some of the insights found in the recent work by Joslin *et al.* (2014) (JPS in the following) about the variables needed to span bond excess returns, and can be regarded as a reduced-form version of their approach. One major difference is that the approach we propose that does not require knowledge on the $\mathbb{P}$-measure dynamics for the yield state variables that the JPS model requires and allows a more general and flexible specification of the market price of risk than the JPS model implies. Dispensing with the task of estimating the $\mathbb{P}$-measure reversion-speed matrix and reversion level for the yields greatly simplifies the parametrization of the model, because of the well-known quasi-unit-properties of rates. Thanks to its simplicity and parsimony, our approach can be implemented and calibrated very easily and robustly both to the market yield curve and to the yield covariance matrix (see Sec. 7). Its robustness decreases the risk of over-parametrization and facilitates the economic interpretation of the results.

Our main goal in formulating this model is to produce a robust and parsimonious tool for extracting risk premia and expectations. Since we do not estimate the $\mathbb{P}$-measure dynamics for the yield state variables, we obtain the required information from observed market prices and from projections (such as the "blue dots" published by the Federal Reserve Bank) of the real-world path of the short rate. We are aware that the interpretation of these projections as market expectations is not unproblematic. However, we show that the predictions for risk premia and expectations produced by the model here presented are likely be more reliable than those provided by many popular yield-curve-based essentially-affine models that eschew real-world projections and monetary guidance. We discuss in detail in Sec. 4 why we believe this to be the case. The model we propose could therefore help both in informing monetary policy and in the making of investment decisions.

Since our focus is on producing reliable estimates of expectations and risk premia in all market conditions (including conditions of very low rates), we extend our approach in Sec. 9 to incorporate the option-like nature of rates close to zero (see Black 1995). We do so by explicitly modeling the shadow rate. We can easily do so in the affine setting we use thanks to the accurate approximations recently introduced by Wu & Xia (2015). This allows us to determine a "confidence band" for our estimates of the risk premia.

Finally, the particular reduced-form implementation we present in this work can be seen as a particularly simple representative example of a whole class of reduced-form stochastic-market-price-of-risk models. A richer, yet still parsimonious, member of this class could, for instance, be a stochastic-market-price-of-risk model

in which principal components are used as state variables. We leave this development for future study.

2. The Spanning Conundrum

A number of empirical regression studies suggest that the excess returns reaped by investing in long-dated Treasury bonds and funding the purchase by borrowing at the short-maturity rate are predictable and state-dependent. In earlier studies, the return-predicting factor was identified as the slope of the yield curve (see, for example, Fama & Bliss 1987, Campbell & Shiller 1991).[a] More recent works suggest that a more complex pattern of forward rates (e.g. tent-shaped: see, for example, Cochrane & Piazzesi 2005, 2008, Hellerstein 2011, or bat-shaped: see, for example, Dai *et al.* 2004, Rebonato 2015, Villegas 2015) may have a greater explanatory power.[b] Whatever the exact shape of these yield-curve-based return-predicting factors,[c] for the future discussion it is important to remember (i) all these studies concur that the slope has a significant predictive power of excess returns and that (ii) more than three (linear combinations of) yields are needed in order to predict excess returns as efficiently as possible.[d]

All these approaches use as explanatory variables yield-curve-based quantities. However, Ludvigson & NG (2009) have recently pointed out that macrofinancial quantities (related to inflation and real economic activity) add significant explanatory power to the Cochrane–Piazzesi return-predicting factor (which the authors take as the most powerful yield-curve-based predictive benchmark). These findings have inspired the works by Radwanski (2010) and Cieslak & Povala (2010, 2013) amongst others, who combine yield-curve predictors with proxies for inflation

---

[a]More precisely, in Fama & Bliss (1987) the return-predicting factor for the n-maturity excess return was defined as the difference between the n-year forward rate and the 1-year excess return. Therefore, strictly speaking, a different return-predicting factor was applied to each individual excess return. (The different slopes, of course, are very similar, so it would not be difficult to construct a common factor out of the n highly correlated ones.) Similar considerations apply to the Campbell & Shiller's (1991) study, in that they used (different, but highly correlated) yield spreads as return-predicting factors.

[b]The tent-shaped and the bat-shaped return-predicting factors have virtually identical explanatory power for excess returns (as captured by the R^2 statistic). Rebonato (2015) argues that the optically striking differences between the two patterns are of little economic consequence, and can be traced to the quasi-collinearity of the regressors.

[c]We define as "yield-curve-based" a return-predicting factor made up exclusively of current or lagged yield-curve variables, such as (linear combinations of) yields or forward rates. When a macroeconomic variable [such as inflation in the case of Radwanski (2010)] is regressed on yields, and the right-hand-side variables with their coefficients are used as return-predicting factors, we are still dealing with a yield-curve-based approach to the prediction of excess returns. However, when the *residuals* of a regression on a macrofinancial variable [such as long-term inflation or the savings rate in the case of (Cieslak & Povala 2010)] form part of the return-predicting factor we no longer deal with a pure yield-curve-based prediction.

[d]Bauer & Hamilton (2015) express a dissenting view regarding point (ii); see, however, Cochrane (2015) for a rebuttal.

expectations and, by doing so, obtain very high R^2 statistics in their predictions of excess returns.

In the light of these findings it would seem obvious that an affine no-arbitrage term structure model designed to straddle the risk-neutral ($\mathbb{Q}$) and real-world ($\mathbb{P}$) measures should use as state variables both yield-curve quantities (such as principal components) and macrofinancial variables (such as expected inflation or some proxy for real economic activity). However, this creates a conundrum in the almost-universally-adopted essentially-affine framework — a conundrum which we discuss in some detail below, because it is central to our proposed approach.

As pointed out by, amongst others, Ludvigson & NG (2009), as long as the mapping from the state variables to the yields is invertible, then it is always possible to redefine the "co-ordinates" of a mixed yield-curve/macrofinancial model purely in terms of yields. More precisely, suppose that we start with n state variables, x_t^i, $i = 1, 2, \ldots, n$. Under the real-world measure we have

$$dx_t = \mathcal{K}^{\mathbb{P}}(\theta^{\mathbb{P}} - x_t)dt + Sdz_t, \tag{2.1}$$

where $\mathcal{K}^{\mathbb{P}}$ is the $n \times n$ reversion-speed matrix in the real-world (objective, $\mathbb{P}$) measure, $\theta^{\mathbb{P}}$ is the $n \times 1$ reversion-level vector, also in the real-world ($\mathbb{P}$) measure, x_t denotes the $n \times 1$ vector with the state variables, dz is an $n \times 1$ vector of independent Brownian increments and S is the $n \times n$ diffusion matrix. As usual, dz is defined on a filtered probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, that satisfies the usual conditions.

If the market price of risk vector, λ, is an affine function of the state variables,

$$\lambda_t = \lambda_0 + \Lambda x_t, \tag{2.2}$$

so is the short rate,

$$r_t = u_r + g^{\mathrm{T}} x_t. \tag{2.3}$$

It is well known (see e.g. Dai & Singleton 2000) that bond prices are given by

$$P_t^T = \exp^{A_t^T + (\mathbf{B}_t^T)^{\mathrm{T}} \mathbf{x}_0}, \tag{2.4}$$

for any $t \leq T$.[e]

This expression can be derived as the discounted expectation, under $\mathbb{Q}$, of the path of the short rate, in turn obtainable via Eq. (2.3) from

$$dx_t = \mathcal{K}^{\mathbb{P}}(\theta^{\mathbb{P}} - x_t)dt + Sdz + S(\lambda_0 + \Lambda x_t)dt \tag{2.5}$$

and the mapping afforded by Eq. (2.3). As well known, one can equivalently write

$$dx_t = \mathcal{K}^{\mathbb{Q}}(\theta^{\mathbb{Q}} - x_t)dt + Sdz, \tag{2.6}$$

where the terms $S\lambda_0$ and $S\Lambda x_t$ have been absorbed into a new reversion-speed matrix and reversion-level vector.

---

[e]The literature on affine models in finance in general, and term structure modeling in particular, is immense. For a recent review, see e.g. Piazzesi (2010). For a derivation of the result when the reversion speed matrix is invertible and diagonalizable, and for an expression for A_t^T and B_t^T, see Appendix A.

Since yields in this setting are affine functions of the state variables, as long as the linear mapping between the chosen state variables and the yields is invertible, we can always find n initial values for the n state variables, x_0^i, $i = 1, 2, \ldots, n$, such that n market yields are exactly recovered. Since the original state variables are linked to these n yields by an (assumed-invertible) affine transformation, we can just as well work with the n chosen yields as state variables. The market values of these chosen yields will be perfectly recovered, but that all the other yields only to within measurement error. To stress the privileged status of the state-variable yields, we will denote them by the symbol, $\widetilde{y}_t$. We stress again the importance of invertibility for the argument.

In this setting it follows that the (essentially-affine) market price of risk, which we had expressed as in Eq. (2.2), can be rewritten as

$$\lambda_t = \widetilde{\lambda}_0 + \widetilde{\Lambda}\widetilde{y}_t. \tag{2.7}$$

Suppose now, however, that the market price of risk *cannot* be fully expressed as an affine function of the special-yields state variables, but can instead be written as

$$\lambda_t = \widetilde{\lambda}_0 + \widetilde{\Lambda}\widetilde{y}_t + R w_t, \tag{2.8}$$

where w_t are m new nonyield-curve variables that are orthogonal to the special yields, $\widetilde{y}_t$.[f] If this is the case then the pricing equation for the state variables under $\mathbb{Q}$ becomes

$$
\begin{aligned}
d\widetilde{y}_t &= \mathcal{K}_y^{\mathbb{P}}(\theta_y^{\mathbb{P}} - \widetilde{y}_t)dt + S_y dz + S_y \lambda_t dt \\
&= \mathcal{K}_y^{\mathbb{P}}(\theta_y^{\mathbb{P}} - \widetilde{y}_t)dt + S_y dz + S_y(\widetilde{\lambda}_0 + \widetilde{\Lambda}\widetilde{y}_t + R w_t)dt.
\end{aligned}
\tag{2.9}
$$

Now, the terms $S_y\widetilde{\lambda}_0$ and $S_y\widetilde{\Lambda}\widetilde{y}_t$ can be absorbed into the new $\mathbb{Q}$-measure reversion-level vector and reversion-speed matrix exactly as before, but not so for the term $S_y R w_t$. One can now only write

$$d\widetilde{y}_t = \mathcal{K}_y^{\mathbb{Q}}(\theta_y^{\mathbb{Q}} - \widetilde{y}_t)dt + S_y dz + S_y R w_t dt. \tag{2.10}$$

Let us, however, augment the state vector from n to $n + m$ with the new unspanned variables w_t:

$$\{x_t\}_n \Rightarrow \{x_t'\}_{n+m} = \{x_t\}_n \cup \{w_t\}_m. \tag{2.11}$$

Using the augmented set as state variables, we can now absorb the term $S_y R w_t dt$ into a new reversion-speed matrix and a new reversion-speed level. And, of course, with the augmented set of state variables, the specification of the market price of risk becomes essentially affine again (in $\{x_t'\}_{n+m}$). This means that, using the

---

[f]With Eq. (2.8) one is implicitly saying that macroeconomic variables can affect the investors' utility function. This is *a priori* not implausible, as worsening economic conditions may increase risk aversion.

augmented state vector $\{x'_t\}_{n+m}$ we can write again

$$P_0^T = \exp^{A_0^T + (\mathbf{B}_0^T)^{\mathrm{T}} \mathbf{x}'_0}. \tag{2.12}$$

This being the case, by inversion we can obtain as before $n+m$ "special" yields, where "special" means again that they will be *exactly* priced and that they can be promoted to the status of state variables.

We have therefore transformed a problem with n yield state variables and an unspanned market of risk, to a problem with $n+m$ yield state variables and a fully spanned market price of risk (and therefore we are back to working with an essentially-affine purely-yield-curve-based model).

This is surprising, for several reasons: first, because this result seems to imply that, in a deterministic-volatility affine setting, there can be nothing more to the market price of risk than yield-curve variables. Second, because it seems to contradict the received wisdom that a small number of principal components give a very satisfactory description of the $\mathbb{Q}$-measure yield-curve dynamics.

Ludvigson & NG (2009) offer several possible explanations for this conundrum and we refer the reader to their Sec. 4 for a discussion. An additional explanation could be found along the following lines.

Consider the statistical estimation of excess returns. Once we have chosen the perfectly priced $n+m$ yields as explanatory variables, we are faced with the task of regressing the excess returns on this augmented (and now possibly rather large) set of state variables. Unfortunately, when we choose yields as state variables, we are choosing a highly collinear set of explanatory variables (of "co-ordinates") onto which to project excess returns. Since all yields are strongly correlated, to "pick up" information that was orthogonal to the original n yields and contained in the additional (but highly correlated) m yields [see Eq. (2.8)], the loadings of the return-predicting factor will necessarily amplify small intra-yield differences. No surprise, then, that both "tents" and "bats" return-predicting factors have such wild (and so optically different) shapes, with large positive and large negative coefficients, despite having almost identical predicting power [as documented in Rebonato (2015)]. Indeed, this amplification of small differences is one of the best "known statistical problems that arise when using linear combinations of multiple forward rates, such as the sensitivity to small measurement errors".[g] As Cieslak & Povala (2013) point out, "although the CP [Cochrane & Piazzesi] factor captures an important element of the risk premium, its coefficient estimates are noisy, creating a wedge between in- and out-of-sample forecasts".[h]

What if one had used principal components as the original state variables, i.e. linear combination of yields that are built to be orthogonal to each other? Suppose that n was chosen to be 3 and m to be 3 as well [Ludvigson & NG (2009) use up to eight linear combinations of macrofinancial quantities. We limit ourselves to

---

[g]Cieslak & Povala (2013, p. 3).
[h]See p. 32 of the work.

three for the sake of parsimony]. *In principle* the fourth, fifth and sixth principal components, which are orthogonal to the first three important components, could in principle pick up very efficiently the extra bond-return predictability. However, the problem now is that the estimation of high-order principal components is easily contaminated by noise and that they have much less stable loadings (a much less stable shape) than the first three.

From this argument it follows that one can look at the spanning problems from two different angles.

From one perspective one can claim that, *no matter how many yields one looks at*, there is information about excess returns that is not contained in these yields. In this view, the impossibility is, so to speak, irreducible. If this view is correct, there is something deeply inadequate and troubling about (invertible, deterministic-volatility) essentially-affine models.

Alternatively one can argue that, admittedly, three yields do not contain all the available information about excess returns and that more yields *in principle* could — and the argument above shows how they could be constructed. However, *in practice* these $n + m$ yields do not predict well, because they constitute a very poor set of coordinates onto which to project the excess-return information: in an $(n + m)$-dimensional space, the yield "axes" are almost collinear, and a small measurement error can translate into a large difference in the resulting projections onto the coordinate axes. And, as we argued, despite being orthogonal, principal components do not fare much better, because, for yield-curve-related problems, noise easily contaminates the components of high order.

It is important to note that in this view essentially-affine modeling with yield-curve-based variables is in (practical) troubles, but not affine modeling *tout court*. There is therefore nothing stopping the researcher from using a market price of risk of the form (2.8), and using the yield-curve-related and macroquantities as state variables, without translating them onto the equivalent special-yields coordinates.

3. The Approach by Joslin *et al.* (2014)

The discussion above puts into perspective the approach by Joslin *et al.* (2014), whose procedure we analyze more precisely in this section. For ease of reference to their original work, in this section we use the same symbols used in Joslin *et al.* (2014), even when the usage is slightly at odds with the (more common) notation we employ in the rest of the paper.

Joslin *et al.* (2014) use as state variables a set of quantities, Z_t,

$$dZ_t = \mathcal{K}^{\mathbb{P}}(Z_t - \theta^{\mathbb{P}})dt + S dz_t^{\mathbb{P}}, \tag{3.1}$$

which they split into a vector, $\mathcal{P}_t$, of yield-curve variables and a vector, $\mathcal{M}_t$, of macroeconomic variables. Importantly, they do *not* assume that the macroeconomic variables can be deterministically spanned by the yield-curve variables. This means

that in the projection

$$\mathcal{M}_t = \gamma_0 + \gamma_1 \mathcal{P}_t + OM_t \tag{3.2}$$

the residuals, OM_t, are not identically zero. This is in turn means that knowing the value of the (combination of) yields $\mathcal{P}_t$ does not convey everything there is to know about excess returns. So, they write for the $\mathbb{P}$-measure evolution of their state variables:

$$\begin{bmatrix} d\mathcal{P}_t \\ d\mathcal{M}_t \end{bmatrix} = \begin{bmatrix} \mathcal{K}_{\mathcal{PP}}^{\mathbb{P}} & \mathcal{K}_{\mathcal{PM}}^{\mathbb{P}} \\ \mathcal{K}_{\mathcal{MP}}^{\mathbb{P}} & \mathcal{K}_{\mathcal{MM}}^{\mathbb{P}} \end{bmatrix} \left(\begin{bmatrix} \theta_{\mathcal{P}}^{\mathbb{P}} \\ \theta_{\mathcal{M}}^{\mathbb{P}} \end{bmatrix} - \begin{bmatrix} \mathcal{P}_t \\ \mathcal{M}_t \end{bmatrix} \right) dt + \begin{bmatrix} S_{\mathcal{PP}} & S_{\mathcal{PM}} \\ S_{\mathcal{MP}} & S_{\mathcal{MM}} \end{bmatrix} \begin{bmatrix} dz_t^{\mathcal{P}} \\ dz_t^{\mathcal{M}} \end{bmatrix}. \tag{3.3}$$

Then Joslin *et al.* (2014) impose that the market price of risk, λ_t, should be an affine function of the *extended* variables, Z_t:

$$\lambda_t = \lambda_{0,Z} + \lambda_{1,Z} Z_t. \tag{3.4}$$

This makes their approach essentially affine, with the real-world expectations of (portfolios of) yields, $\mathcal{P}_t$, depending on the full set of variables $\{Z_t\} = \{\mathcal{P}_t\} + \{\mathcal{M}_t\}$. At this point, however, they impose that in the risk-neutral measure, $\mathbb{Q}$, the portfolios of yields should follow a mean-reverting process for the $\mathcal{P}$-variables only,[i] of the form:

$$d\mathcal{P}_t = \mathcal{K}_{\mathcal{P}}^{\mathbb{Q}}(\mathcal{P}_t - \theta_{\mathcal{P}}^{\mathbb{Q}})dt + S_{\mathcal{PP}}dz_t^{\mathbb{Q}}. \tag{3.5}$$

This implies

(1) the market yields, y_t^T, will be an affine function of the $\mathcal{P}_t$-variables only:

$$y_t^T = \alpha_t^T + (\beta_t^T)^{\mathrm{T}} \mathcal{P}_t; \tag{3.6}$$

(2) and also the short rate will be an affine function of the subset, $\mathcal{P}$, of variables that map exactly onto the yields only:

$$r_t = \xi_{0,y} + (\xi_{1,y})^{\mathrm{T}} \mathcal{P}_t. \tag{3.7}$$

As for the market price of risk, Joslin *et al.* (2014) assume that we know the real-world drifts for the $\mathcal{P}$-variables. Then the market price of risk for the principal components, $\Lambda_{\mathcal{P}}(Z_t)$, which depends on the whole set of variables, Z_t, will be given by the (scaled) difference between the real-world and risk-neutral drifts of the principal components:

$$\Lambda_{\mathcal{P}}(Z_t) = S_{\mathcal{PP}}^{-1}[\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t) - \mu_{\mathcal{P}}^{\mathbb{Q}}(\mathcal{P}_t)], \tag{3.8}$$

with $\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t)$ coming from Eq. (3.1) and $\mu_{\mathcal{P}}^{\mathbb{Q}}(\mathcal{P}_t)$ from Eq. (3.5). The important thing to note is that the $\mathbb{P}$-measure drift of the $\mathcal{P}$-variables, $\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t)$, (and hence their

---

[i]After rotation we can think of these $\mathcal{P}$-variables as principal components. Without loss of generality we will do so in what follows.

real-world expectations) depends on the full set of yield-curve and macrofinancial variables, and is given (to first order) by

$$\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t) = \mathcal{K}_{\mathcal{PP}}^{\mathbb{P}}(\theta_{\mathcal{P}}^{\mathbb{P}} - \mathcal{P}_t) + \mathcal{K}_{\mathcal{PM}}^{\mathbb{P}}(\theta_{\mathcal{M}}^{\mathbb{P}} - \mathcal{M}_t), \tag{3.9}$$

not just by

$$\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t) = \mathcal{K}_{\mathcal{PP}}^{\mathbb{P}}(\theta_{\mathcal{P}}^{\mathbb{P}} - \mathcal{P}_t). \tag{3.10}$$

This means that $\mathbb{P}$-measure expectations of yields, and hence excess returns, depend on the full set of yield-curve and macroeconomic variables.

We note that the approach by Joslin *et al.* (2014) rests on a "delicate cancellation". This can be seen as follows. In their model, investors form expectations in the real-world measure making use of yield-curve and nonyield-curve information [see the term $\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t)$]. However, their market price of risk must be such so as to exactly cancel the dependence on the macrofinancial variables present in the term $\mu_{\mathcal{P}}^{\mathbb{P}}(\mathcal{P}_t)$ when the $\mathbb{Q}$-measure drift is obtained:

$$\mu_{\mathcal{P}}^{\mathbb{Q}}(\mathcal{P}_t) = \mu_{\mathcal{P}}^{\mathbb{P}}(Z_t) - S_{\mathcal{PP}}^{-1}\Lambda_{\mathcal{P}}(Z_t). \tag{3.11}$$

This is, of course, possible, but it is difficult to justify on financial grounds. It is not clear, in fact, why certain features of investors' response to risk [namely, the nature and dependence on the state variables of their aversion to risk, as embodied in the market price of risk, $\Lambda_{\mathcal{P}}(Z_t)$] should be such exactly to cancel the $\mathcal{M}$-dependence of the $\mathbb{P}$-measure drift, $\mu_{\mathcal{P}}^{\mathbb{P}}(Z_t)$, of the chosen yields. See Eq. (3.8) again.

In our model we do not invoke such a "delicate cancellation", and, as we shall see, this will be one important point of difference between the approach we present below and the model by Joslin *et al.* (2014).

4. Practical Advantages of the Augmented-Variable Approach

Leaving the issue of the "delicate cancellation" to one side, approaches such as the one by Joslin *et al.* (2014) (and ours) can overcome an important shortcoming of yield-curve-based essentially-affine models. Since the overcoming of this limitation is central to our approach, we discuss it in some detail in this section.

4.1. *Similarity of yield-curve-based risk premia predictions*

To highlight the nature of the above-mentioned shortcoming, we look at the predictions of excess returns produced by three essentially-affine term-structure models, such as Adrian *et al.* (2013), D'Amico *et al.* (2010) and Rebonato *et al.* (2017) models, typical of current modeling practice. These three models were chosen because they use very different state variables and very different estimation methods of the market price of risk.[j] We also consider the predictions of excess returns produced

---

[j]The three approaches follow very different routes to estimating the market price of risk: one model (D'Amico *et al.* 2010) uses a Kalman filter technique to estimate the latent state variables in the

by a purely statistical estimation (multivariate regression against the slope of the yield curve).

To make the discussion more precise, let P_t^T be the time-t price of a T-maturity bond. If we assume that the price P_t^T is affected by n state variables,[k] and that investors form prices rationally,[l] the expected return from holding over a period Δt a (discount) bond of maturity T can in general be written as follows:

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{dP_t^T}{P_t^T}\right] = \left[r_t + \frac{1}{P_t^T}\sum_i \frac{\partial P_t^T}{\partial x_t^i}\sigma_i\lambda_i\right]dt, \tag{4.1}$$

where σ_i is the volatility of factor i and λ_i is the "market price of risk" associated with the same factor.

If one also assumes ("affine assumption") that

(i) the n state variables jointly evolve according to

$$d\mathbf{x}_t = \mathcal{K}(\boldsymbol{\theta} - \mathbf{x}_t)dt + Sd\mathbf{z}_t; \tag{4.2}$$

and (ii) there exist a scalar u_r and a vector $\mathbf{g}^T$ such that the "short" (riskless) rate is given by an affine function of the state variables:

$$r_t = u_r + \mathbf{g}^T\mathbf{x}_t, \tag{4.3}$$

then, as we have discussed (see again Dai & Singleton 2000), the bond price, $P_t^T = P(\tau)$, is given by[m]

$$P_t^T = e^{A_t^T + (\mathbf{B}_t^T)^{\mathrm{T}}\mathbf{x}_t}. \tag{4.4}$$

Under the affine assumption, Eq. (4.1) can be rewritten as

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{dP_t^{T_k}}{P_t^T}\right] = \left[r_t + \sum_i (B_t^{T_k})_i^{\mathrm{T}}\sigma_i\lambda_i\right]. \tag{4.5}$$

If one also assumes that the market price of risk is also an affine function of the same state variables (the "essentially-affine assumption"),

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_0 + \Lambda\mathbf{x}. \tag{4.6}$$

One can write for the expected excess return of a vector of bonds:

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{d\mathbf{P}_t^T}{P_t^T} - r_f\right] = (B_t^T)^{\mathrm{T}}\Sigma\boldsymbol{\lambda}. \tag{4.7}$$

$\mathbb{P}$-measure; another (the one of Adrian *et al.* 2013) uses OLS regressions to estimate, also in the $\mathbb{P}$-measure, the process for the now specified state variables; and the third, the PCA-based model by Rebonato *et al.* (2016), uses market prices and statistical estimates of excess returns.

[k]Without loss of generality, we assume in the following that the state variables are uncorrelated. This can always be achieved by a suitable rotation of axes.

[l]We therefore exclude, for instance, market segmentation and behavioral biases — or, rather, we assume that, if these deviations are present, pseudo-arbitrageurs always exist which can bring prices back to fundamentals.

[m]See Appendix A for the solution of this problem (i.e. for explicit expressions for A_t^T and B_t^T).

The empirical evidences regarding risk premia (market prices of risk) are that:

(1) the expectation hypothesis is rejected (forward prices are not unbiased estimators of future prices, and hence the λ_i are not all zero);

(2) the weak expectation hypothesis (existence of a *constant* risk premium) is rejected;

(3) what investors seek compensation for is *level* risk; and

(4) the magnitude and sign of the compensation is linked to the *slope* of the yield curve. See, in this respect, Campbell & Shiller (1991), Fama & Bliss (1987).[n]

If, without loss of generality, we choose as state variables principal components, which lend themselves to an interpretation as level, slope and curvature, then piece of evidence (4) implies that

$$\Lambda = \begin{bmatrix} 0 & a & 0 \\ 0 & b & 0 \\ 0 & c & 0 \end{bmatrix}. \tag{4.8}$$

Pieces of evidences (3) and (4) together imply that

$$\Lambda = \begin{bmatrix} 0 & a & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.9}$$

And, finally, piece of evidence (2) implies that

$$\boldsymbol{\lambda}_0 = 0. \tag{4.10}$$

Given this precise formalization of the problem, the shortcoming of essentially-affine models alluded to above can be clearly understood by looking at Fig. 1. This figure shows the risk premia for the 10-year yield obtained with three *prima facie* very different, but all yield-curve-based, affine term structure models (the curves labeled "ACM", "KW" and "Affine"), and from a purely statistical estimation of the risk premium obtained using a proxy for the yield-curve slope as state variable (curve labeled "slope"). Admittedly, the levels for the risk premia produced by the four estimation approaches are somewhat different.[o] However,

---

[n] As discussed, recent works (Cochrane & Piazzesi 2005, Cieslak & Povala 2010) suggest that more complex return-predicting factors may have better *ex ante* explanatory power. However, the consensus remains that the slope is certainly an important — albeit, possibly, not optimal — return-predicting factor. [Bauer & Hamilton (2015) actually argue that, out-of-sample, it remains the *only* robust factor.]

[o] As for the statistical estimation of the level of the excess return (i.e. the estimate that gives rise to the curve labeled "slope"), we find that the uncertainty associated with the intercept in the regression of excess returns against the second principal component is much larger than the uncertainty associated with the "beta" (slope) coefficient. Implicitly or explicitly, both the Adrian *et al.* (2013) and the D'Amico *et al.* (2010) models require the estimation of the real-world behavior for the quasi-unit-root process of the level of rates. The quasi-unit-root nature
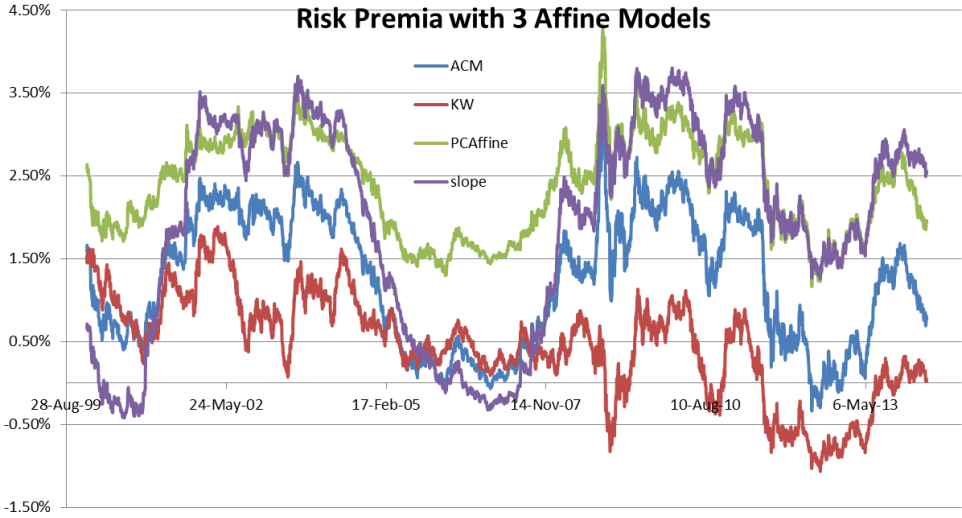
Fig. 1. The risk premia for the 10-year yield obtained by the Kim–Wright model (red line), the Rebonato–Saroka–Putyiatin affine model (green line) and from the regressions presented in Rebonato (2016) (blue line) for the period from 1999 to 2014. Note that the levels of the risk premia are rather different, but the changes are very strongly correlated. They are also clearly linked to the slope of the yield curve (purple line).

it is amazing how *changes* in the risk premia produced by the three models are very strongly correlated, and similar to what produced by the statistical regression.[P]

Why do these modeling approaches produce such similar estimates? The similarity of the model predictions with the statistical estimate of the risk premia suggests that the "training" of all the models implicitly discovers the slope dependence in the learning dataset. Then, the essentially-affine nature of these models and the choice of yield-curve-only quantities as state variables force a *deterministic* dependence of the market price of risk on the slope: simplifying somewhat, in all these essentially-affine, yield-curve-only models, when the slope is high, the market price of risk *must* be high. We argue in the next subsection that, while this regularity may well be valid on average, enforcing it (as yield-curve-based essentially-affine models must do) at all times can have perverse, and very unpleasant, consequences. We present our case by discussing in detail the rare — and, from the modeling point of view, precious — instance of a "real-life experiment" offered by the "taper tantrum" events of 2013.

---

of the process makes the estimate of the reversion level particularly difficult. This $\mathbb{P}$-measure reversion level, in turn, is the main determinant of the level of the risk premium. The level of the risk premium is therefore particularly difficult to estimate using relationships such as (3.8).
[P]See in this respect Joslin *et al.* (2014).

4.2. *The case of the "taper tantrum" and its significance*

To frame the discussion, it is important to give a synthetic description of the market developments that surrounded the "taper tantrum" events in the May–October 2013 period. Simplifying considerably, the pronouncements of Ben Bernanke (2013), then Chairman of the Federal Reserve Bank, were interpreted by the market as implying an earlier-than-previously-anticipated rise in rates. The important point is that, while uncertainty may well have been increased by the announcement, the announcement itself must have directly affected expectations. We can test whether the essentially-affine models mentioned above correctly (or, at least, plausibly) decomposed the market changes in yields in their expectation and risk premium components.[q]

The market interpretation of Mr. Bernanke's words caused yields in the 5-to-10-year maturity area to rise very sharply. Short-dated yields (say, of maturity of two years or shorter) barely moved, however, because, in early 2013 not even a very "hawkish" interpretation of Mr. Bernanke's words could have implied a substantial
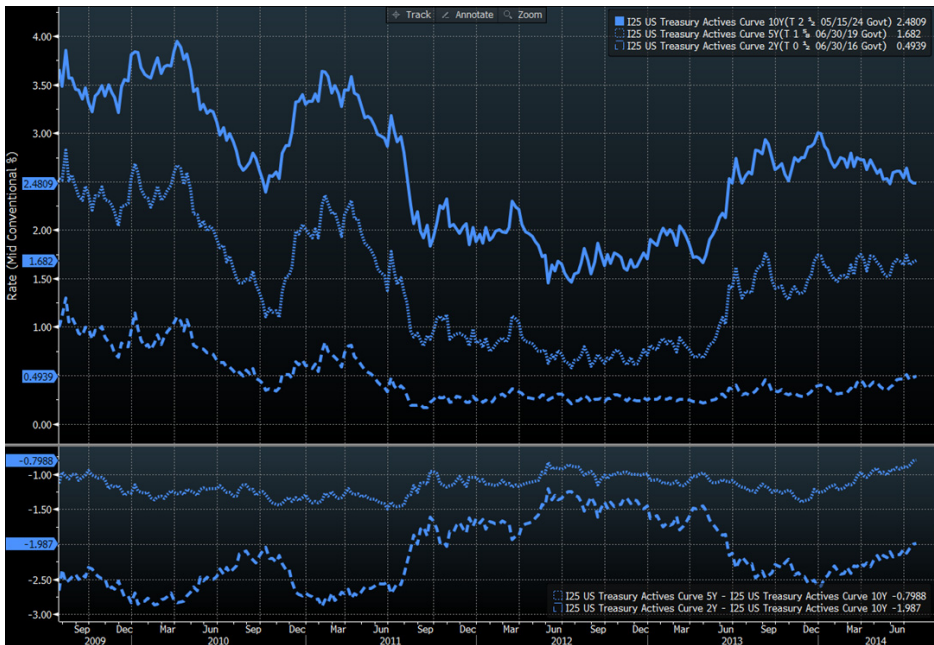
Source: Bloomberg.

Fig. 2. The behavior from September 2009 to June 2014 of the 2-year, 5-year and 10-year nominal yields of Treasuries, as proxied by the 2.5% Treasury maturing 15 May 2024, the 15/8% Treasury maturing 30 June 2019 and the 01/2% Treasury maturing 30 June 2016. The bottom half of the graph shows the difference between the 10-year proxy yield and the 5-year and the 2-year proxy yields, respectively.

[q]Since we look at yields below 10 years, changes in convexity play a small role in the decomposition.

rise in rates by early 2015. Therefore the effect on the yield curve of the "taper tantrum" was a marked steepening of the yield curve. See Fig. 2.

This mode of deformation of the yield curve was a direct result of the exceptional, but prolonged, monetary conditions prevailing at the time (the "ZIPR" — Zero Interest Rate Policy): it was because the very short end was pinned, that expectations of increased rates could only affect the medium-to-long end of the yield curve and produce a yield-curve steepening.[r] It is important to stress that in the pre-ZIRP times to which all the above-referred-to models have been trained, hawkish statements by central bankers would have typically caused a *flattening* of the yield curve (with the short end raising more than the long end).

Recall also that historically the steepening of the yield curve has been associated with the early phases of a recession, and early recessionary phases have been the periods with the highest Sharpe ratios for the excess-return strategy. See in this respect Naik *et al.* (2016).[s] Therefore the affine models (and the statistical analysis) trained on data before the ZIRP period "interpret" the steepening of the yield curve as a harbinger of high expected excess returns, and, *via the deterministic link between the market price of risk and the slope*, are "forced" to produce strongly increased risk premia.

The consequences of this are clearly reflected in Fig. 3, which shows the time series of the 10-year nominal expected excess returns from March 2013 to late summer 2014 produced by one of the models under discussion.[t] More precisely, the top curve (labeled "10YQ") shows the market 10-year yield; the curve labeled "RiskPrem" displays the contribution to the observed 10-year yield from the risk premium; the curve labeled "Slope [10-1]" shows the term premium predicted by regressing excess returns for the period 1971–2010 (see footnote U) against the yield-curve slope, proxied by the difference of the 10 minus the 1-year yields; and, finally, the bottom curve, labeled "10YP" depicts the $\mathbb{P}$-measure expectation of the 10-year yield — that is, the market 10-year yield minus the estimated risk premium.

Two observations jump from the page. First, the extreme closeness between the econometric predictions and the yield-curve-variable affine-model predictions of the risk premium. Second, the paradoxical result about expectations that the deterministic link between the market price of risk and the yield-curve slope produces: despite the fact that the "taper tantrum" surprise was about the earlier-than-expected *rising* of rates, expectations barely change after the announcement,[v] and almost all of

---

[r]Indeed, only a few weeks before the "taper tantrum" announcement, the median view expressed by the FOMC members on 20 March 2013 about the most likely level for the Fed funds was around 20 basis points for the end of 2014, and 1% by the end of 2015. The reader can refer to Target Fed funds rate, "blue dots", 20 March 2013.

[s]See p. 99, Exhibit 4.12 in particular.

[t]The graph refers to the Rebonato *et al.* (2017) model. Qualitatively very similar conclusions can be drawn by analyzing the responses of the other models under consideration.

[u]The data used for this part of the study are from the Gurkaynak *et al.*'s (2007) dataset.

[v]As it happens, a similar analysis conducted for the DAmico *et al.*'s (2010) model shows even a small decrease in rate expectations. Nothing too significant should be read into this, apart from
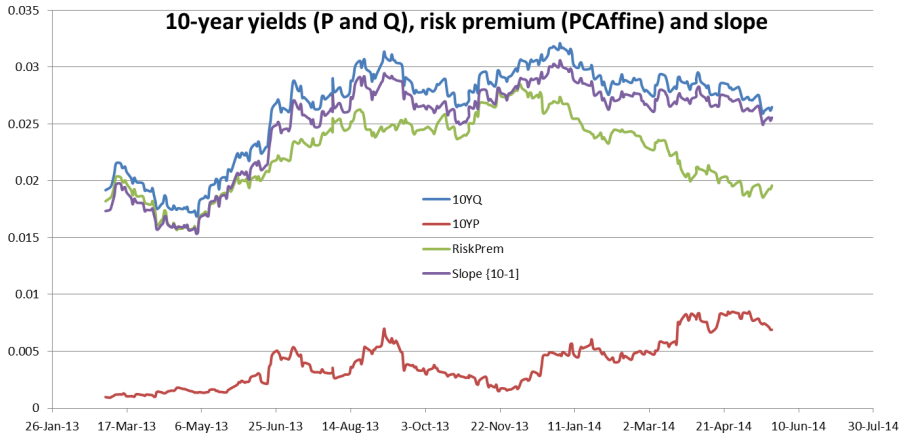
Fig. 3. The series of the 10-year nominal expected excess returns from March 2013 to late summer 2014 produced by the Rebonato *et al.*'s (2017) model: the top curve (labeled "10YQ") shows the market 10-year yield; the curve labeled "RiskPrem" displays the contribution to the observed 10-year yield from the risk premium; the curve labeled "Slope [10-1]" shows the term premium predicted by regressing excess returns for the period 1971–2010 against the yield-curve slope, proxied by the difference of the 10 minus the 1-year yields; and, finally, the bottom curve, labeled "10YP" depicts the $\mathbb{P}$-measure expectation of the 10-year yield — that is, the market 10-year yield minus the estimated risk premium.

the observed market changed is "interpreted" by the model(s) as an increase in risk premium.[w] Once again, this surprising behavior of the term premium (which implies near-zero expectations for the 10-year yield) stems from the high steepness of the yield curve in late 2013, which forces, via the inflexible link between the market price of risk and the slope of the yield curve, implausibly high expected excess returns.

Of course, the potential for "pathological" attributions is not confined to the "taper tantrum" period. As Ludvigson & NG (2009) have shown, yield-curve-only variables in general, and the slope in particular, do not span the full excess-return variability, and leave a substantial part of the predictability unexplained. There are therefore several periods when the predictive yield-curve variables (say, the slope) are in the same configuration, yet the most efficient predictions of excess returns (that is, a prediction that used both yield-curve and macroeconomic variables) would be significantly different.

As discussed in Sec. 3, Joslin *et al.* (2014) have provided an elegant modeling solution to this problem, by allowing a decoupling between the market price of risk and the shape of the yield curve. Their approach, however, presents substantial calibration challenges and, from the modeling perspective, relies on the "delicate

---

the fact the increase in risk premium forced by the steepening of the yield curve is so pronounced in this model that the risk premium becomes even higher than the market yield.

[w]Admittedly, the "taper tantrum" announcement increased uncertainty in the market, and it can therefore be argued that the risk premium should have increased as well. However, the fact still remains that a (possibly misinterpreted) statement about an earlier-than-expected lifting of rates should have changed (upwards) expectations.

cancellation" discussed at the end of Sec. 3. We therefore "borrow" their insight that the deterministic link between the yield-curve variables and the market price of risk must be, if not broken, certainly loosened. We do so via an extremely parsimonious reduced-form affine model, in which we assign a market price of risk which has a mean-reverting behavior, and which is not spanned by the chosen yield-curve variables. In our model, we want the slope/excess-return relationship to be still recovered *on average*, but not in a hard deterministic manner. So, at any given point in time, the curve could be steep, yet the market price of risk could be low (or even negative). Hopefully, the breakage of the rigid deterministic link between the shape of the yield curve and the market price of risk, but the simultaneous preservation of the statistical link between the two quantities, will give rise to more believable, and more useful, estimates of both risk premia and rate expectations.

5. The Model

In the simplest version of our model we use as state variables the short rate, its own stochastic reversion level and the market price of risk. To simplify the analysis, and in line with standard findings (see e.g. Cochrane & Piazzesi 2005, 2008, Adrian *et al.* 2013), we also assume in what follows that investors only seek compensation for the uncertainty about the level of rates, which we proxy in our approach as the long-term reversion level of the reversion level. We therefore assume that the risks associated with uncertainty in the short rate and, as in Cieslak & Povala (2010), in the market price of risk are not priced.

The model then is as follows:

$$dr_t^{\mathbb{Q}} = \kappa_r^{\mathbb{P}}[\theta_t - r_t]dt + \sigma_r dz_t^r, \tag{5.1}$$

$$d\theta_t^{\mathbb{Q}} = \kappa_\theta^P[\widehat{\theta}_t^{\mathbb{P}} - \theta_t]dt + \lambda_t \sigma_\theta dt + \sigma_\theta dz_t^\theta, \tag{5.2}$$

$$d\lambda_t = \kappa_\lambda[\widehat{\lambda}_t - \lambda_t]dt + \sigma_\lambda dz_t^\lambda, \tag{5.3}$$

where r_t, θ_t and λ_t are the time-t values of the short rate, its instantaneous reversion level (the "target rate") and the market price of risk, respectively; σ_r, σ_θ and σ_λ are the associated volatilities: $\widehat{\theta}_t^{\mathbb{P}}$ and $\widehat{\lambda}_t$ are the reversion levels of the "target rate" and of the market price of risk, respectively; and the increments dz_t^r, dz_t^θ and dz_t^λ are suitably correlated (see below). The model is fully specified once the initial state, r_0, θ_0 and λ_0, is given.

This model draws its inspiration from a simple version of the Taylor's rule, and is a variation on the theme of the well-known 'monetary' models and of similar "monetary" models, in which the short rate is attracted to a mean-reverting target rate.[x] For this interpretation to hold, it is usually assumed that, at least in the objective measure, $\kappa_r^P > \kappa_\theta^P$ and $\sigma_r < \sigma_\theta$.

---

[x]In more complex models of this family, the target rate can in turn be attracted to a mean-reverting process. We do not pursue this avenue here, in order to limit the number of parameters and to focus on the ability of the model to capture the essence of the problem at hand.

As stated, we have assumed that investors only seek compensation for the uncertainty about the long-term reversion level. It is for this reason that the reversion speed, $\kappa_r^{\mathbb{P}}$, of the short rate is the same in the real-world and in the risk-neutral measures. Also, the long-term reversion level, $\widehat{\theta}_t^{\mathbb{P}}$, of the "target rate" ($\theta_t$) that appears in the term in square brackets on the right-hand side of Eq. (5.2) is in the real-world measure. The increment, $d\theta_t^{\mathbb{Q}}$, is then in the risk-neutral measure because we are adding the risk compensation, $\lambda_t^\theta \sigma_\theta$. We could equivalently have written

$$d\theta_t^{\mathbb{Q}} = \kappa_\theta^P [\widehat{\theta}_t^{\mathbb{Q}} - \theta_t]dt + \sigma_\theta dz_t^r, \tag{5.4}$$

with

$$\widehat{\theta}_t^{\mathbb{Q}} = \frac{\sigma_\theta}{\kappa_\theta^P}\lambda_t. \tag{5.5}$$

The model can be easily solved analytically — the constitutive equations and their solutions are given in Appendix A.

6. Qualitative Behavior of the Model

In this section we present the qualitative behavior of the model, and we check whether it does behaves as it is designed to.

First of all, we show in Fig. 4 one typical evolution over a 30-year horizon of the three state variables with the following choice of parameters: $r_0 = -0.0042$, $\theta_0 = 0.0344$, $\lambda_0 = 0.0744$, $\kappa_r = 0.3437$, $\kappa_\theta = 0.085$, $\kappa_\lambda = 0.2816$, $\sigma_r = 0.0050$, $\sigma_\theta = 0.0157$, $\sigma_\lambda = 0.1200$, $\widehat{\theta}_t^P = 0.0350$, $\widehat{\lambda}_t = 0.1287$, $\rho_{r\theta} = 0.6$, $\rho_{r\lambda} = -0.05$ and $\rho_{\theta\lambda} = 0.64$.

Next, we show in Fig. 5 a sample of 32 paths showing the evolution over 10 years for the market price of risk $(\widehat{\lambda}_t = 0.1287, \lambda_0 = 0.0744, \kappa_\lambda = 0.2816$; see
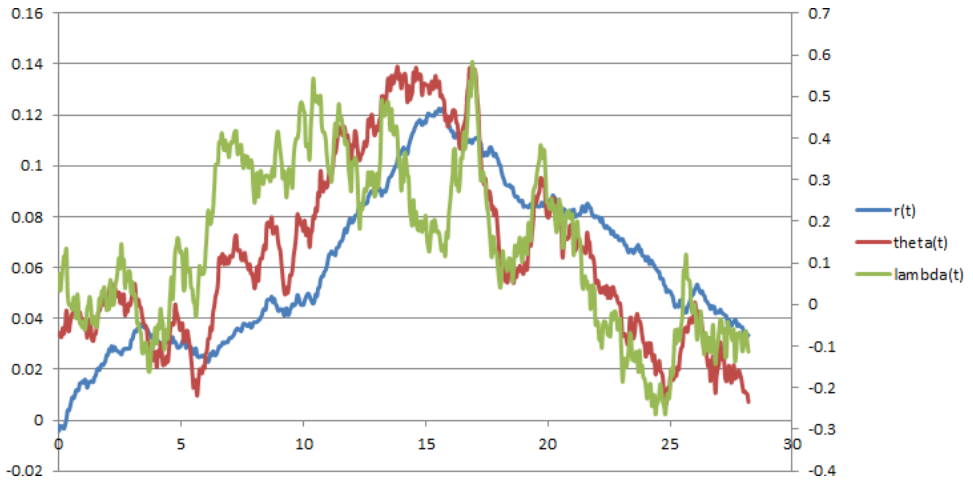

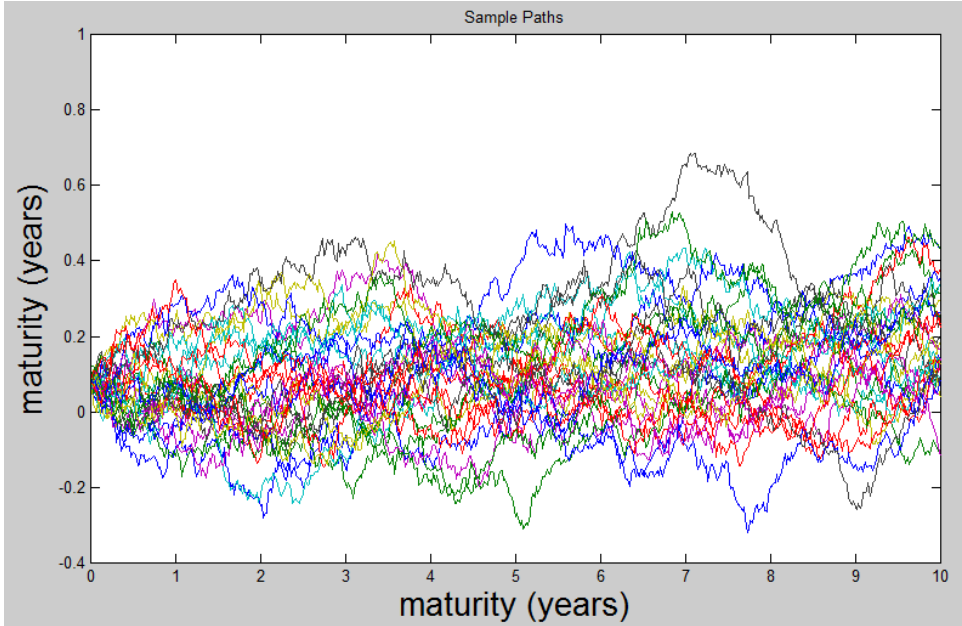Fig. 4. A typical evolution over a 30-year horizon of the three state variables.

Fig. 5. A sample of 32 paths showing the evolution over 10 years for the market price of risk. The model parameters used for the simulation were: $r_0 = -0.0042$, $\theta_0 = 0.0344$, $\lambda_0 = 0.0744$, $\kappa_r = 0.3437$, $\kappa_\theta = 0.085$, $\kappa_\lambda = 0.2816$, $\sigma_r = 0.0050$, $\sigma_\theta = 0.0157$, $\sigma_\lambda = 0.1200$, $\widehat{\theta}_t^P = 0.0350$, $\widehat{\lambda}_t = 0.1287$, $\rho_{r\theta} = 0.6$, $\rho_{r\lambda} = -0.05$ and $\rho_{\theta\lambda} = 0.64$.

the calibration section for the parameter choice). We note that, with the chosen parameters, (i) the market price of risk can assume both positive and negative values (as it is the case in reality); (ii) its distribution is centered around the range of values for the Sharpe ratio observed in statistical studies (see e.g. Naik *et al.* 2016); (iii) it is attracted to a positive reversion level, which has been set to the unconditional average Sharpe ratio of buy-long/fund/short "carry" strategies.

One of the distinguishing features of the approach presented here is its ability to generate a positive correlation between the slope of the yield curve and the market price of risk. For a suitable choice of the correlation matrix among the three state variables, the model can easily produce a significant correlation between the slope of the yield curve and the market price of risk. Figure 6 shows the correlated paths of the yield-curve slope and of the market price of risk obtained with a correlation coefficient, $\rho_{\theta\lambda}$, of 0.64 and a reversion speed for the market price of risk, κ_λ, of 0.2816.

Importantly, the same figure shows that at any point in time the market price of risk can be high or low even if the slope is low or high, respectively. In general, the higher the volatility of the market price of risk, and the lower its reversion speed, the greater its potential for assuming values other than what is implied by the
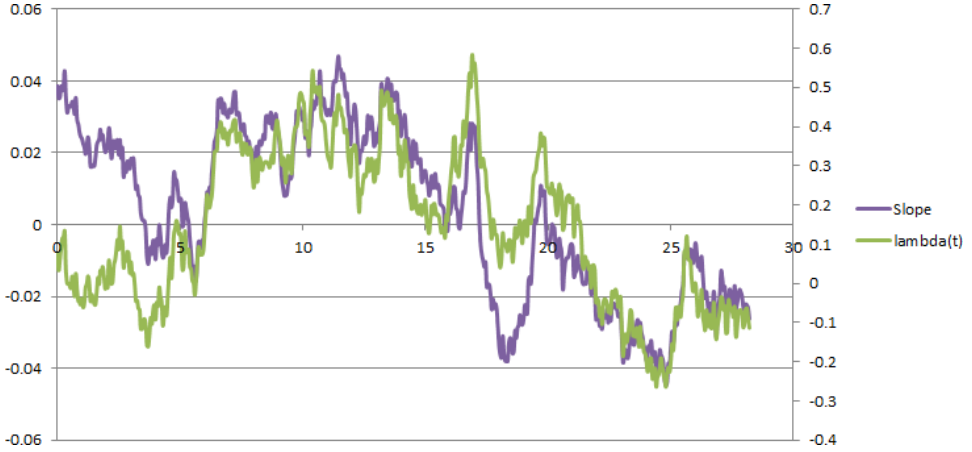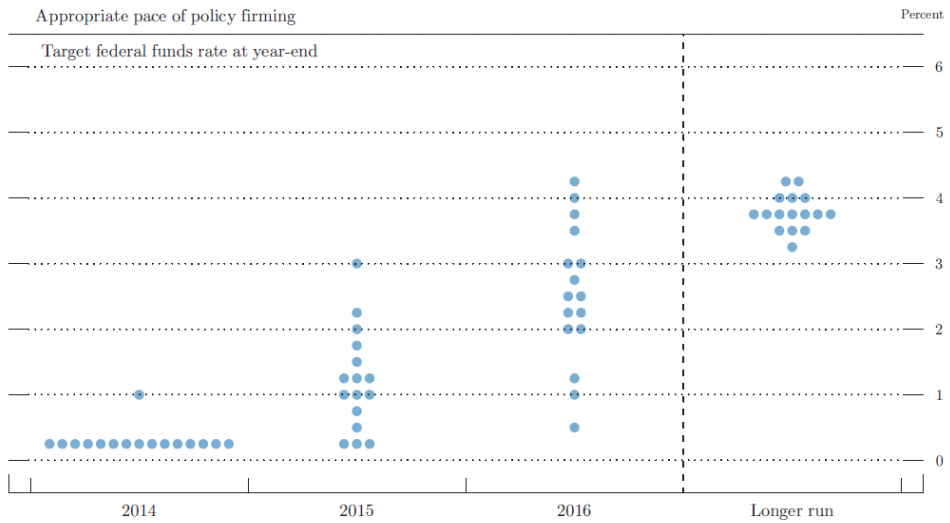
Fig. 6. A time series of the slope and of the market price of risk over a 30-year period. The model parameters used for the simulation were: $r_0 = -0.0042$, $\theta_0 = 0.0344$, $\lambda_0 = 0.0744$, $\kappa_r = 0.3437$, $\kappa_\theta = 0.085$, $\kappa_\lambda = 0.2816$, $\sigma_r = 0.0050$, $\sigma_\theta = 0.0157$, $\sigma_\lambda = 0.1200$, $\widehat{\theta}_t^P = 0.0350$, $\widehat{\lambda}_t = 0.1287$, $\rho_{r\theta} = 0.6$, $\rho_{r\lambda} = -0.05$ and $\rho_{\theta\lambda} = 0.64$.

deterministic dynamics, and to break the deterministic relationship between excess returns and the return-predicting factor.

7. Calibration of the Model

Any model that tries to account both for the $\mathbb{P}$- and the $\mathbb{Q}$-dynamics must add some source of information beside what is embedded in prices. As mentioned in the introduction, in our study we make use of real-world forecasts. These can be gleaned from a variety of sources. Christensen & Kwan (2014) discuss the expectations of federal funds rates from different surveys and from the FOMC participants' funds rate projections. Crump *et al.* (2016) discuss and combine several sources of information about expectations of rates (and other variables). For our purposes the right-hand panel of the "blue dots" (such as the cluster labeled "Longer Run" in Fig. 7) provided by the FOMC participants serves our purposes well, as it embodies the expectations of rate-setting body of the long-term level of the target rate.[y] We stress that any of the survey-based sources of information could be used for the study we describe.

[y] The "blue dots" have been published quarterly since early 2012. There is one dot for responding member. (The number of responding members is not identical in various meetings.) Each dot represents the estimate of one member of the FOMC of the most likely value for where the Fed funds will be at the end of the year when the question is posed, the following year and the year after that. Finally, there is a dot from each member that reflects that member's most likely estimate of the "long-run" level of the Fed funds rate. Each dot therefore represents a mode ("the most likely value"), not an average. This subtlety is not "innocuous", especially at the short end of the maturity spectrum, but should be of lesser relevance for the long-term estimates. The distinction between the sample mode and the average has not been taken into account in our study.

Fig. 7. Overview of FOMC participants' assessments of appropriate monetary policy. Each "dot" represents the estimate of one member of the committee of the "most likely" value for the target Fed funds rate on the projection dates on the x-axis.

Notes: In the upper panel, the height of each bar denotes the number of FOMC participants who judge that, under appropriate monetary policy, the first increase in the target Federal funds rate from its current rage 0–14% will occur in the specified calender year. In March 2014, the numbers of FOMC participants who indged that the first increase in the target federal funds rate would occur in 2014, 2015 and 2016 were, respectively, 1, 13 and 2. In the lower panel, each shaded circle indicates the value (rounded to the nearest 1/4% point) of an individual participant's judgment of the appropriate level of the target Federal funds rate at the end of the specified calender year or over the longer run.

In its most general formulation, the full specification of the model requires 11 parameters (three reversion speeds, two reversion levels, three volatilities and three correlations); on any given day, the three initial values of the state variables also have to be specified. This can raise concerns about the potential for overfitting, and about the robustness of the approach. The number of degrees of freedom, however, can be drastically reduced by proceeding as follows.

(1) The reversion level of the reversion level in the $\mathbb{P}$-measure, $\widehat{\theta}_t^{\mathbb{P}}$, can be required to be in the neighborhood of the "long-term blue dot" from the Fed quarterly report. See again Fig. 7.

(2) Since in a one-factor world the market price of risk equals the Sharpe ratio, the reversion level of the stochastic market price of risk can be set in the vicinity of the unconditional historical Sharpe ratio for the excess-return strategy.[z]

---

[z]As it is well known, the Sharpe ratio of the invest-long/fund-short US\$ Treasury strategy is mildly maturity-dependent. This statistic also shows a dependence on the precise historical period

(3) If the financial interpretation in terms of a target rate alluded to above is to be valid, the reversion speed of the short rate should be much higher than the reversion speed of the reversion level $\kappa_r^P > \kappa_r^\theta$. [See in this respect the discussion in Erekhinskiy (2013).]

(4) This also requires that the volatility of the short rate should be lower than the volatility of the reversion level: $\sigma_r < \sigma_\theta$. [See again Erekhinskiy (2013).]

(5) *A priori*, we cannot say much about the relative volatility of the market price of risk and of the target rate, and we will therefore let the calibration to the yield covariance matrix give us some guidance. However, we can have an order-of-magnitude estimate of the volatility of the market price of risk from the range of the historical variation of the Sharpe ratio over the 40 years in our dataset.

(6) In order to recover the "humped" shape of the term structure of yield volatilities [see Rebonato (2002) for a discussion of the origin of the "hump" in the context of the related US interest-rate swaption market], the correlation between the short rate and the reversion level, $\rho_{r,\theta}$, should be of the order of 0.4–0.8. Also in this respect see Erekhinskiy (2013).

(7) In order to recover the average (positive) dependence of the market price of risk on the slope, it is easy to calculate that we have to set the correlations $\rho_{r\lambda}$ and $\rho_{\theta\lambda}$ to have opposite signs, and $\rho_{\theta\lambda}$ to be positive (negative) for upward-(downward-)sloping yield curves.

(8) The initial value of the short rate must be in the vicinity of the current Fed funds rate.[aa]

Furthermore, to enhance robustness, we calibrate separately to the degrees of freedom (the reversion-speed matrix and volatility matrix) that control the covariance matrix (or the swaption prices), and to those that affect the shape of the yield curve (the reversion-level vector and the initial-state vector). As well known, the latter have no effect on the model covariance structure, but the former have a modest effect (through convexity) on the shape of the yield curve.[bb]

More precisely, the two-stage calibration is conducted as follows.

For the first phase we start from a statistically-determined yield-covariance matrix, Σ_{mkt}, and we equate it to the model covariance matrix, Σ_{mod}, that is immediately derived to be given by

$$\Sigma_{\text{mod}} = \mathbb{E}[dy\,dy^{\text{T}}] = B^{\text{T}}SS^{\text{T}}B,$$

used for its estimation. We have therefore allowed the nonlinear search over the reversion level of the market price of risk to find an optimal value in the range of the highest and lowest empirical Sharpe ratios.

[aa]This choice becomes more complex when we introduce the shadow rate. See the discussion in Sec. 9.

[bb]A joint calibration of the reversion speeds, reversion levels and volatilities to the yield curve and the yield-covariance matrix will in general produce a closer fit, but, for a finite maximum maturity, there is a risk that the reversion levels and the volatilities will "play against each other", with higher and higher reversion levels, which pull the yield curve upwards, compensated for by higher and higher volatilities — and hence higher convexity — which pull the yield curve down.

with the matrices S and B as per Eqs. (4.2)–(4.4). (The translation of the model Eqs. (5.1)–(5.3) to the canonical form (4.2) is provided in Appendix A.) The calibration is carried out by minimizing an appropriate distance between the model and the market covariance matrices. The distance is chosen to be the sum of the weighted squares of the differences between all the same-row, same-column elements of the two matrices. The weights are such that double importance is given to the diagonal elements of the matrix, because these are (squares of) the yield volatilities, which directly affect the convexity curvature of the yield curve. This stage of the calibration fixes the reversion-speed parameters (κ_r, κ_θ and κ_λ), the volatility parameters (σ_r, σ_θ and σ_λ) and the correlations.

This set of parameters are then kept fixed in the second stage of the calibration, where yield-curve information is brought to the fore, by fitting to the market yield curve the initial values of the state variables (r_0, θ_0 and λ_0) and the reversion levels $\widehat{\theta}_t^{\mathbb{P}}$ and $\widehat{\lambda}$. More precisely, we calculate the model yields as

$$y_{t,T}^{\mathrm{mod}} = -\frac{1}{T-t} \log P_t^T = \alpha_t^T + (\boldsymbol{\beta}_t^T)^{\mathrm{T}} \boldsymbol{x}, \tag{7.1}$$

with

$$\alpha_t^T = -\frac{1}{T-t} \log A_t^T, \tag{7.2}$$

$$\beta_t^T = -\frac{1}{T-t} \log B_t^T. \tag{7.3}$$

We then choose the initial values of the state variables and the parameters $\widehat{\theta}_t^{\mathbb{P}}$ and $\widehat{\lambda}$ in such a way as to minimize the unweighted sum of the squares of the differences between the market and the model yields. The market yields for discount bonds were obtained from the Federal Reserve data by Gurkaynak *et al.* (2006, 2007).

Of course, wherever appropriate, the soft or hard constraints expressed in the preceding bullet points are enforced in the optimization phase.

8. Results

Since our focus is the estimation of risk premia and expectations, and since this estimation always takes as a starting point the observed market yields, it is essential to ascertain that the yield contribution from convexity is well recovered, and that market yields are accurately reproduced by the our model. We explore these two aspects in this section.

8.1. *Quality of fit to the covariance matrix*

We calculated the market and model covariance matrices for changes in yields of maturities out to 10, 20 and 30 years [US\$, market covariance matrix sampled over 2000 business days, before and up to September 2014, data from Gurkaynak *et al.* (2007)]. Figures 8 and 9 show the market and the model yield covariance matrices
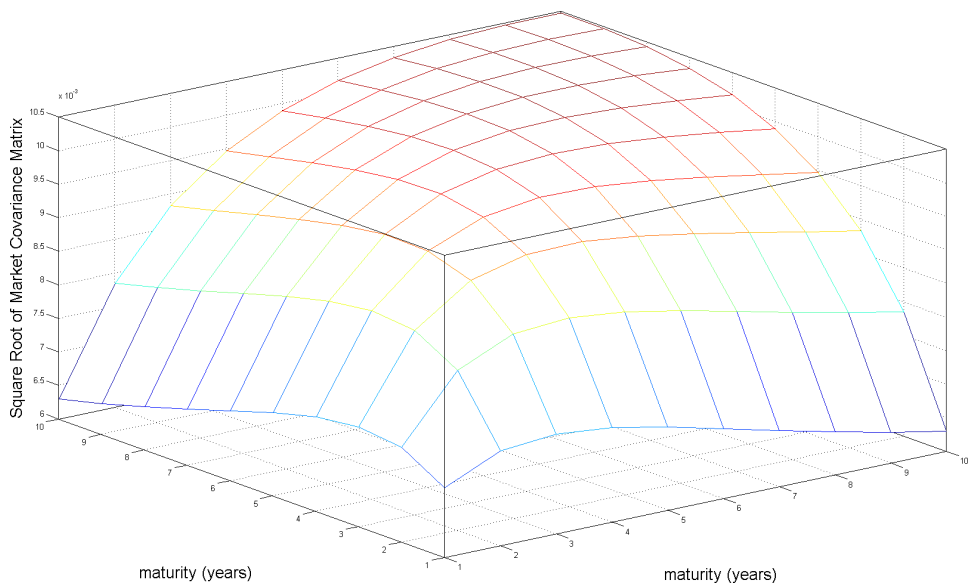
Fig. 8. The market yield covariance matrices for yields from 1 year to 20 years obtained from the data described in the text.
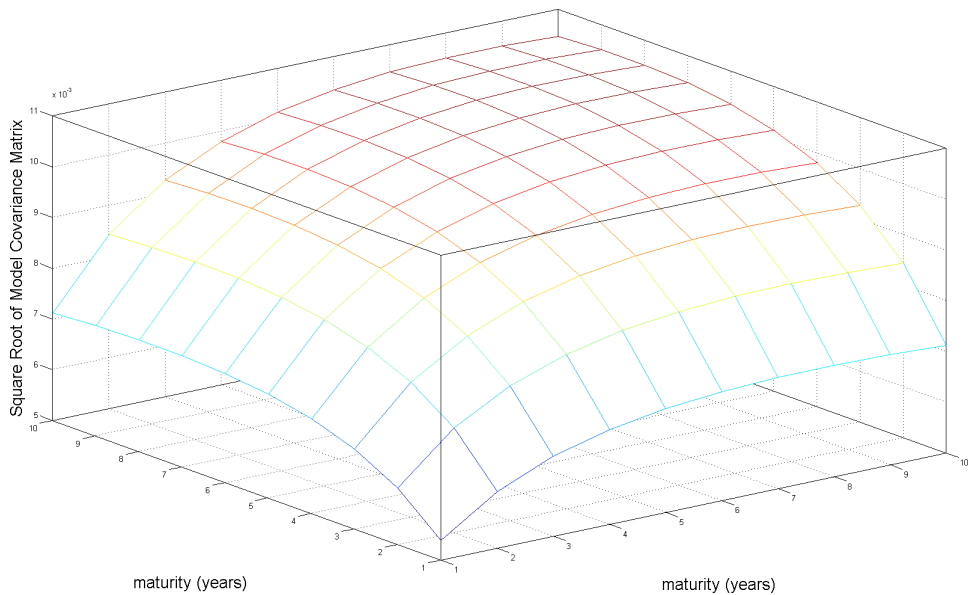

Fig. 9. The model yield covariance matrices for yields from 1 year to 20 years obtained with the parameters reported in the text.

for yields from 1 to 20 years. The model yield covariance matrix was obtained with a constrained choice of parameters, and the optimal sets of volatility and reversion speeds obtained using yields of maturities out to 10, 20 and 30 years are shown in Tab Cov:

$$
\begin{bmatrix}
 & 10y & 20y & 30y \\
\sigma_r & 0.0051 & 0.0050 & 0.0055 \\
\sigma_\theta & 0.0161 & 0.0157 & 0.0153 \\
\sigma_\lambda & 0.1200 & 0.1200 & 0.1200 \\
\kappa_r & 0.4105 & 0.3437 & 0.2737 \\
\kappa_\theta & 0.1001 & 0.0850 & 0.0667 \\
\kappa_\lambda & 0.4224 & 0.2816 & 0.2816
\end{bmatrix}.
\tag{8.1}
$$

Tab Cov: The parameters of the model as obtained from the calibration to the empirical covariance matrices among yield (changes) of maturity out to 10 (column labeled "10y"), 20 (column labeled "20y") or 30 years (column labeled "30y").

It is reassuring to note the stability of the optimal parameters obtained by fitting to the covariance matrices among yield (changes) of maturity out to 10, 20 or 30 years.

Figures 8 and 9 show that the recovery of the empirical covariance matrix, while far from perfect, was generally good. This is important for the decomposition of observed market yield into an expectation, a risk premium and a convexity component (which is one of the main goals of this study). As shown in Rebonato and Putiatyn (2017), any affine model that recovers correctly the yield curve and all the yield volatilities will always produce the correct theoretical convexity. Since the overall convexity contribution is relatively modest for all but the longest maturities, small mistakes in recovering the yield volatilities give rise to negligible mistakes in the overall yield decomposition. So, the imperfect recovery of the target covariance matrix will have but a modest effect on the estimation of expectations and risk premia.

8.2. *Quality of fit to the yield curve*

The attribution of the expectation and risk-premia components of the various yields would obviously fail if market yields were not accurately recovered by the approach we propose. In this subsection we therefore look carefully at the accuracy of the market yield-curve recovery.

The "market" yields were obtained from the data made available by Gurkaynak *et al.* (2007).

An example of the quality of the fit to the yield curve obtained by keeping the reversion-speed matrix and the volatility matrix fixed to the values obtained in the calibration to the covariance matrix, and by varying the initial state vector, $\mathbf{x}_0$, and

the reversion levels, $\widehat{\theta}_t^P$ and $\widehat{\lambda}$, is shown for three terminal maturities in Figs. 10–12. (The US$ curve used was the Treasury curve for May 2014.)

In all cases, the cross-sectional fit to the market yields was excellent. When the fit was carried out to different final maturities (10, 20 or 30 years) the resulting
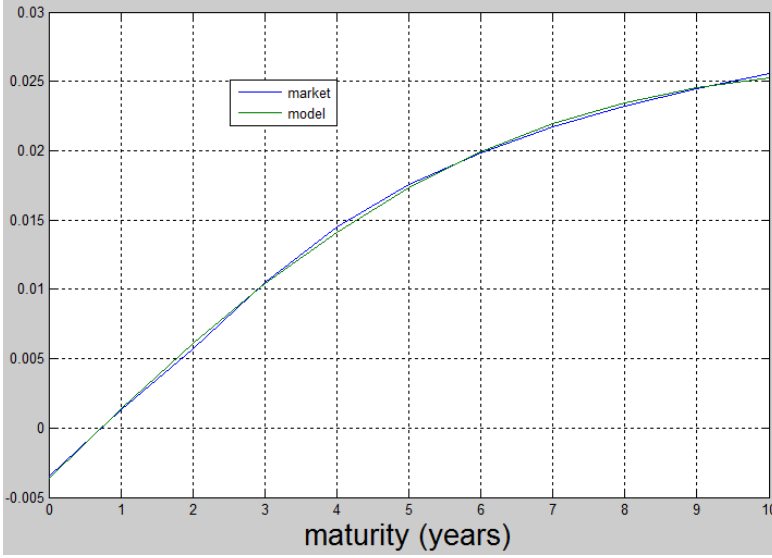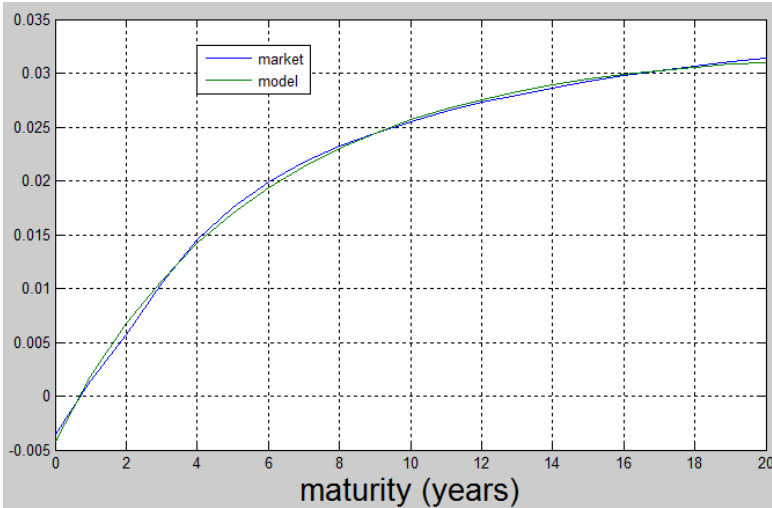

Fig. 10. The fit to the yield curve out to 10 years obtained by keeping the reversion-speed matrix and the volatility matrix fixed to the values obtained in the calibration to the covariance matrix, and by varying the initial state vector, $\mathbf{x}_0$, and the reversion levels, $\widehat{\theta}_t^P$ and $\widehat{\lambda}_t$. (For May 2014.)

Fig. 11. Same as Fig. 10, but with maturities extended out to 20 years.
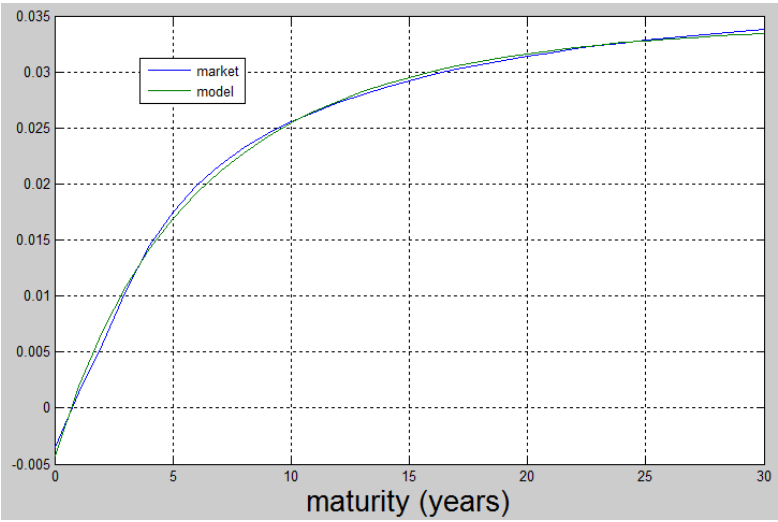
Fig. 12. Same as Fig. 10, but with maturities extended out to 30 years.

parameters (not reported for the sake of brevity) turned out to be very similar, irrespective of the final yield maturity to which the model was calibrated.

Figures 10–12 just provide a snapshot of the quality of the fit to the yield curve for one typical day. A more systematic picture of the quality of the yield-curve fits is provided by Figs. 13 to 16, which show the market and model yields and the pricing
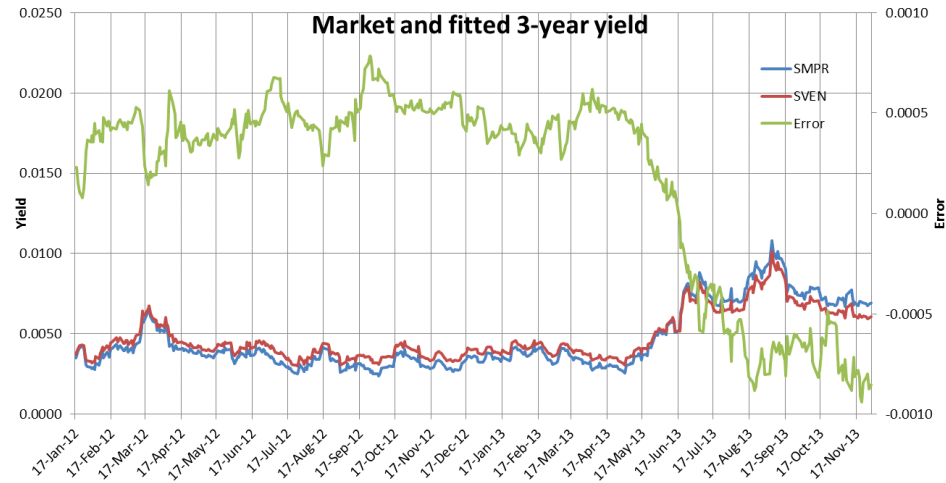

Fig. 13. Quality of the daily fits to the 3-year yield over a typical two-year period. The left-hand side y-axis shows the market and model yields and the right-hand side y-axis displays the error. We note in passing that the sharp change in pricing error occurs in correspondence of the "taper tantrum" discussed at length in Sec. 4.
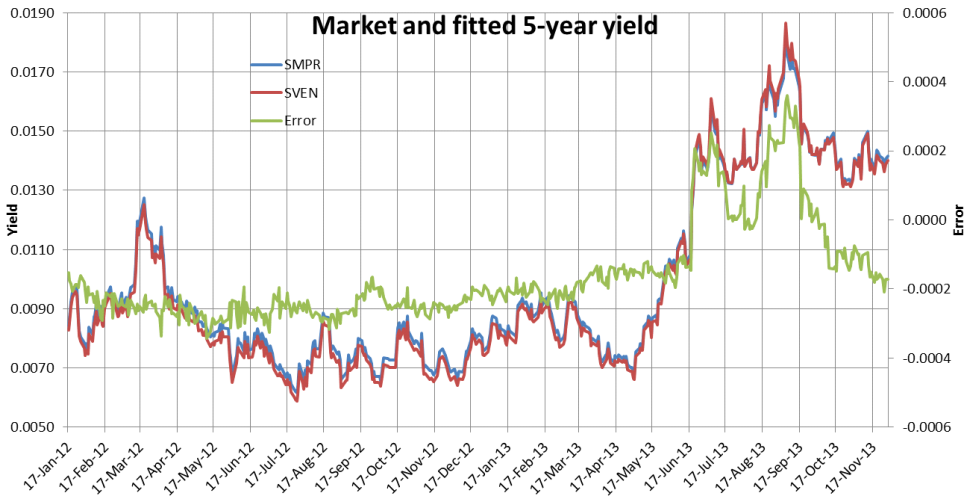
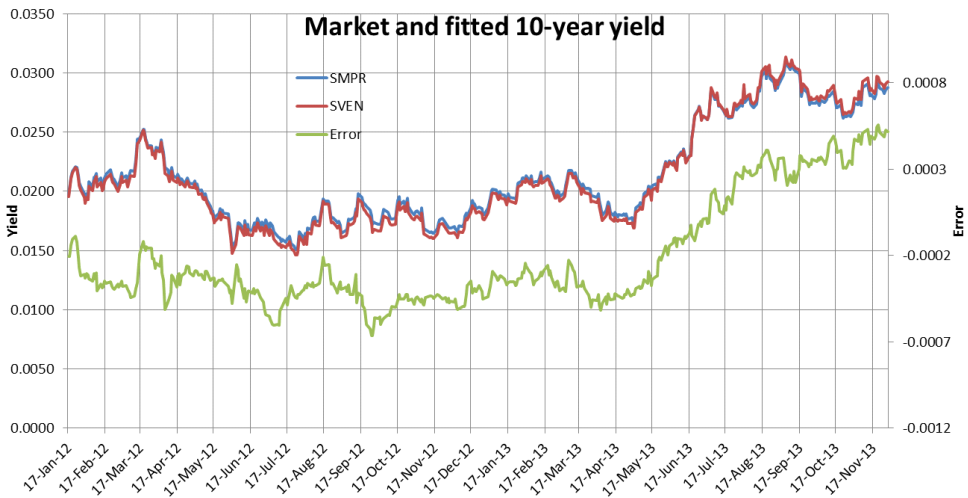Fig. 14. Same as Fig. 13 for the 5-year yield.

Fig. 15. Same as Fig. 13 for the 10-year yield.

errors for four reference yields (the 3-year, 5-year, 10-year and 30-year yields) over a two-year period. As apparent, the model always recovers very well (i.e. to within a few basis points) the market yields.

In sum, given the main focus of our work, the small magnitude of the pricing errors (differences between market and model yields) reported in this subsection gives us confidence that this source of error will not significantly affect our estimates of risk premia.
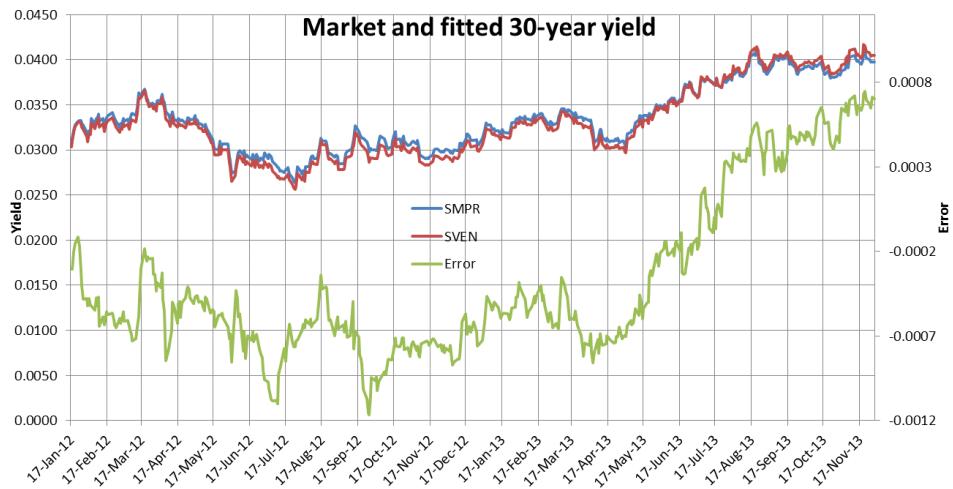
Fig. 16. Same as Fig. 13 for the 30-year yield.

8.3. *The risk premia predicted by the model*

Figure 17 shows the time series of the model predictions for the risk premia for several selected yields, and Fig. 18 shows the difference obtained in May 2014 between the market yields (top curve, labeled "$\mathbb{Q}$-measure") and the yields that would be "obtained" if investors had the same expectations as they do in the real world, but were risk-neutral (curve labeled "$\mathbb{P}$-measure").
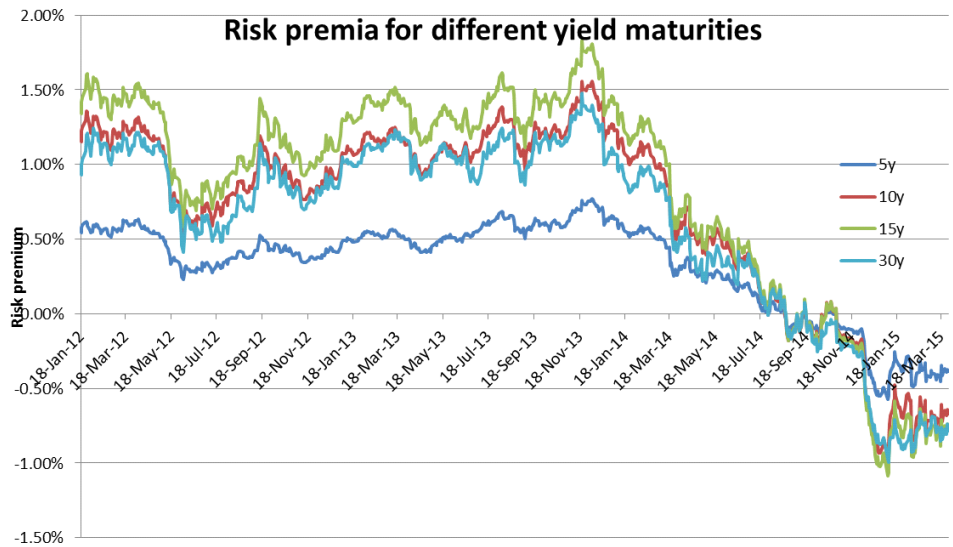

Fig. 17. Time series of the model predictions for the risk premia for the 5-, 10-, 15- and 30-year yields.
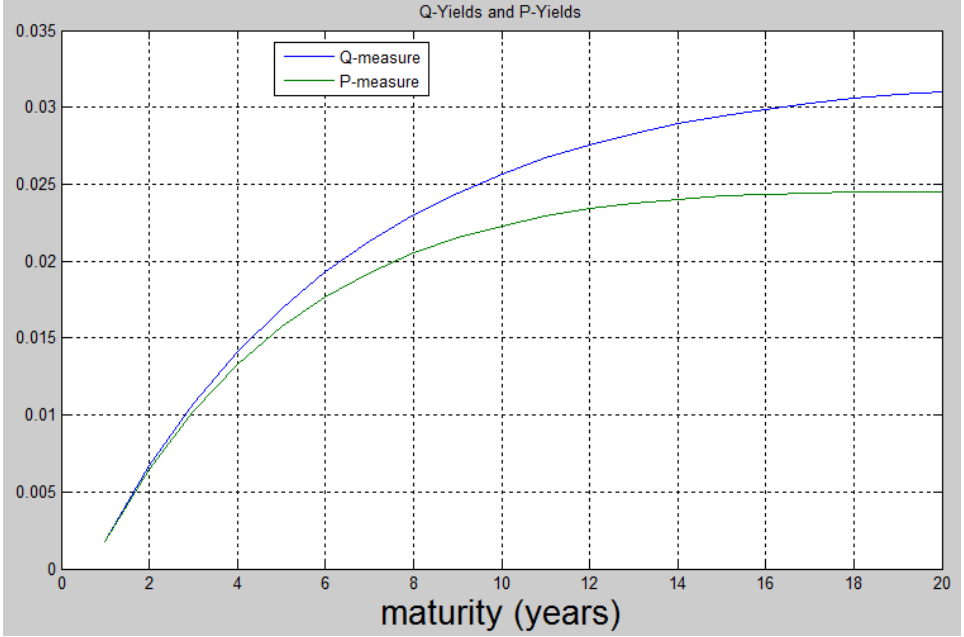
Fig. 18. The best-fit model yield curve (labeled "ℚ-measure") with the parameters discussed in the text and the best-fit yield curve after subtracting the term premia, for the US$ Treasury yield curve in May 2014.

We find the predictions about the magnitude of the expectation and risk-premium contribution to the observed market yields consistent with the consensus of declining risk premia in the last few years. See e.g. the work of D'Amico *et al.* (2010).

We discussed at length in Sec. 4.2 the significance of the "real-life experiment" provided by the "taper tantrum" events of 2013 in assessing the ability of essentially-affine models to recover expectations and risk premia convincingly. How does the model we have so far presented fare in this respect?

When we apply a magnifying glass to the "taper tantrum" period (see Fig. 19) we find the prediction for the model presented here more convincing than the estimates produced by the all-yield-curve variables models shown in Fig. 1. In particular, in the period following Chairman Bernanke's 2013 pronouncements, Figs. 13–16 clearly show that our model says that during this period risk premia for the 5-year yields did increase (as it is reasonable that they should have, given the heightened uncertainty), but that expectations increased more than the risk premia.

This should be contrasted with the predictions for risk premia and expectations of a model such as D'Amico *et al.* (2010): as Fig. 21 illustrates, their model predicts a minor *decline* in expectations after the "taper tantrum" announcement by the Chairman Bernanke, and a more-than compensating increase in risk
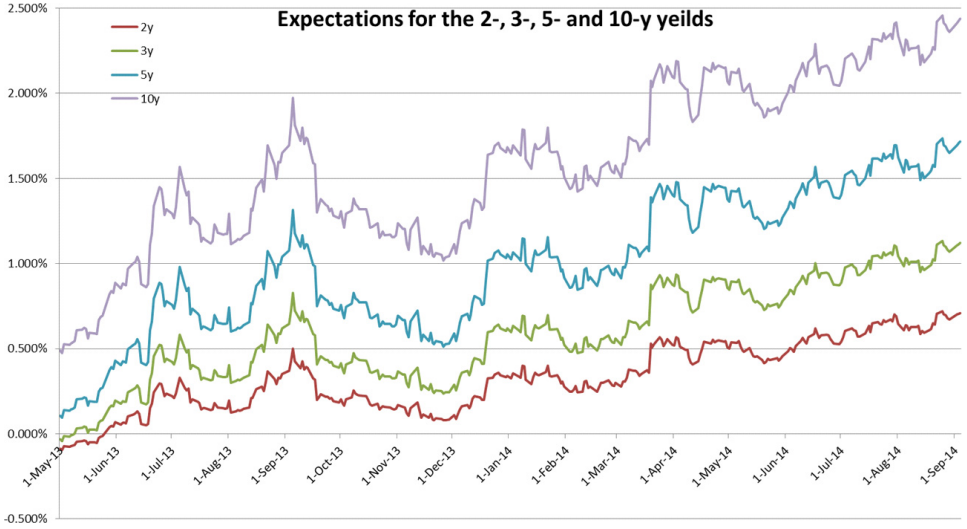
Fig. 19. The expectations for the 2-, 3-, 5- and 10-year yields produced by the model presented in this paper during the "taper tantrum" events discussed in Sec. 4 (mid-June 2013).

premium. Indeed, according to *all* the essentially-affine yield-curve-variable-only models discussed in Sec. 2, changes in risk premia drive the bulk of the variation in 5-year-plus market yields.[cc] See again Fig. 1. According to our model, changes in expectations play the more important role, and risk premia a minor part, but in the same direction. Given the exquisitely expectation-based nature of the "tapering" pronouncements, we find the decomposition produced by the model we propose more plausible.

As discussed, the origin of this enhanced plausibility is to be found in the breakage of the deterministic link between the slope of the yield curve and the market price of risk enforced by other models. Without our imposing that this should be the case (today's value of the market price of risk, λ_0, is one of the few quantities we do *not* preassign), we find the result that today's level of the market price of risk is significantly below its reversion level (and very significantly below the historical Sharpe ratio).

As for the reversion level of the market price of risk (which was also left unconstrained in the optimization), this turned out to be $\widehat{\lambda} = 0.128$. To gauge its reasonableness, we recall that, in a one-factor setting, the market price of risk should equal the Sharpe ratio. If we assume that the one-factor assumption is not too unreasonable, the value we find for the reversion level of the market price of risk

[cc]We choose the 5-year yield as a meaningful point of the yield curve because even the most dovish members of the FOMC believed that rates would have escaped the ZIRP regime at the time — as documented by the "blue dots" forward guidance at the time. The 5-year yield should therefore be definitely affected by changing expectations for the path of the Fed funds.
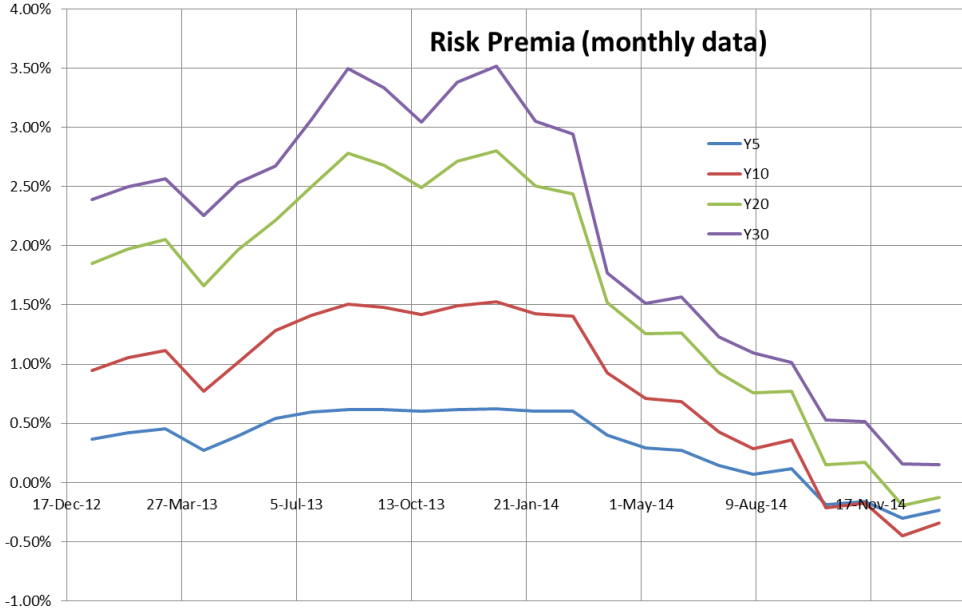
Fig. 20. The risk premia for 5-, 10-, 20- and 30-year yields produced by the D'Amico *et al.*'s (2010) model during the "taper tantrum" period. Note the sharp increase in the risk premia after the late-March "taper tantrum" announcement.

is towards the lower end of historical estimates of the Sharpe ratio for the "long-duration" bond-holding strategy. This would be consistent with the mid-2010s idea of a "New Normal", according to which real returns in all asset classes, and in fixed-income products in particular, will, on the secular horizon, turn out to be lower than they have been in the past. See again in this respect Fig. 18, which shows the best-fit model yield curve (labeled "$\mathbb{Q}$-measure"), and the best-fit yield curve after subtracting the term premia.

9. The Effect of the Shadow Rate

One of the main goals of the approach proposed here is the robust estimation of risk premia and expectations. Since extremely low rates have been a common feature in several developed-market currencies throughout the last decade, it is important to ascertain to what extent the predictions of risk premia are sensitive to different choices of the interest-rate "floor".

The intuition as to why a floor, K, should affect expectations and risk premia is straightforward: if we write a given forward rate as the sum of the expected future value of the short rate, $E_t[r_T]$, of the attending risk premium, rp_t^T, and of the option value, $\text{Opt}(K, T)$,

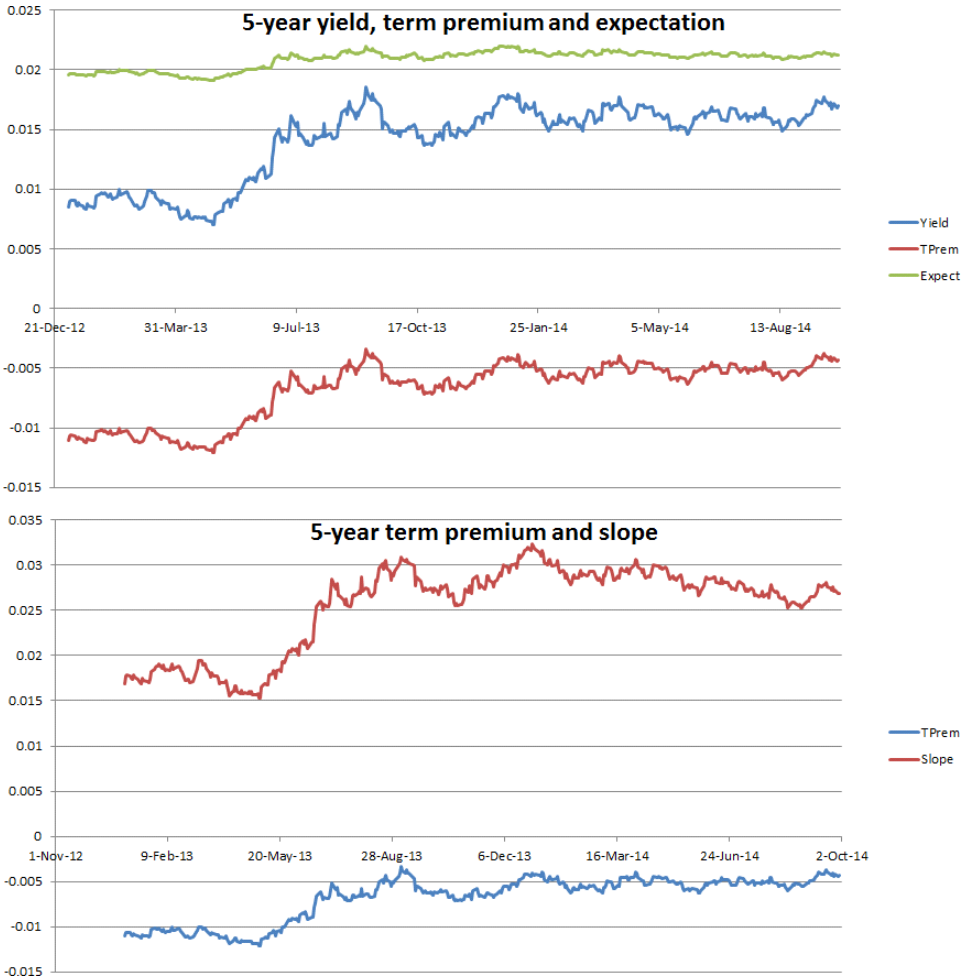$$f_t^T = E_t[r_T] + \text{rp}_t^T + \text{Opt}(K, T), \tag{9.1}$$

Fig. 21. The 5-year market yield produced by the D'Amico *et al.*'s (2010) model during the "taper tantrum" period (see the text for a discussion), decomposed into its expectation component (top curve in the upper panel) and its term premium (bottom curve in the top panel). The bottom panel compares the same term premium with the slope (top curve).

and we recall that the option value must always be nonnegative, we see that, for a given market-observed forward rate, the existence of a floor must lower the sum of the expectation and the risk premium.

These considerations become particularly interesting if one tries to link the estimate of these quantities to macrofinancial variables, and to the state of the economy. As Bomfin (2003) says, "investors and policy-makers alike would be considerably misled if they were to take the positive slope of [the yield curve built without the term $\text{Opt}(K, T)$] as an indication that market participants expect that economic activity is likely to increase in the future. In reality, the negative slope of
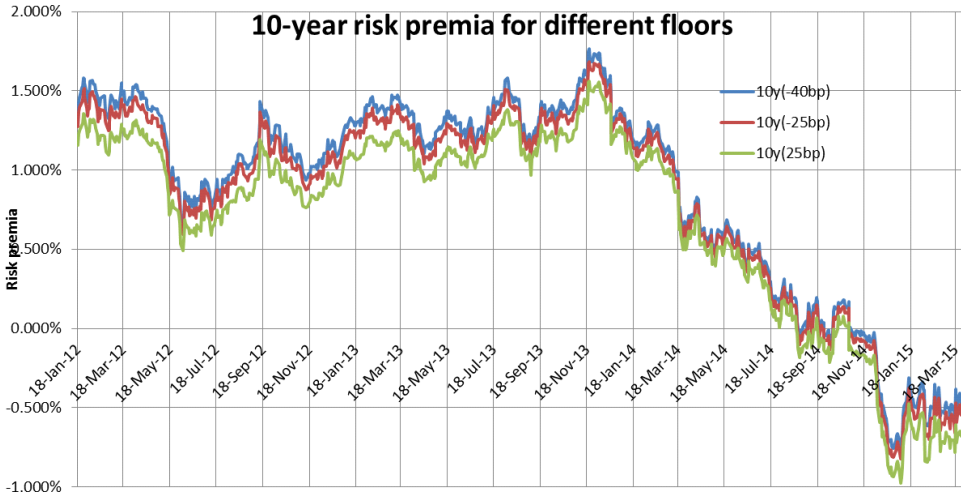
Fig. 22. Time series of the risk premia for the 10-year yield obtained using floors for the shadow rate located at −40 (blue line), −25 (red line) and +25 (green line) basis points.

the equilibrium [...] yield curve [built with the term $\mathrm{Opt}(K, T)$] suggests that the economy is expected to remain trapped in its low-activity equilibrium well into the future. *Thus we have a situation when an upward sloping (observed) yield curve is signalling expectations of a prolonged slump in the economy. This is exactly the opposite of the usual indicator property attributed to the yield curve!*"[dd]

We therefore show in Fig. 22 the time series of the 10-year risk premium in the presence of an interest-rate floor, where the associated option value was calculated using the approach in Xu & Wia (2015) for values of the floor of 25, −25 and −40 basis points.

A level of +25 basis points is the strike used in Wu & Xia (2015); as Government bond yields have recently become negative for some European sovereign issuers, we have also looked at the sensitivity of the results to different (lower) floor levels. There are some minor differences, but Fig. 22 shows that the same qualitative pattern shown in Fig. 17 is still clearly evident. More precisely, we found that the average, maximum and minimum differences in risk premia without a floor and for a floor of −25 basis points are 13, 31 and 2 basis points for the 5-year yield, respectively, and are 27, 46 and 6 basis points, respectively, for the 10-year yield. (The increase in "option value" for the longer yield, despite the fact that the 10-year yield is more "out of the money" than the five-year yield, stems from the interplay between nearness of the rate path to the floor, the level of the reversion level and the option time.)

[dd]Bomfin (2003, pp. 8–9), emphasis added.

Overall, we can therefore conclude that including the effect from the rate floor changes somewhat the attribution of the observed yields to risk premia and expectations, but does not change the qualitative conclusions drawn in the rest of the paper.

10. Conclusions

We have presented a stochastic-market-price-of-risk affine model that is in the spirit of the work by Joslin *et al.* (2014). One of the most appealing features of the approach we present is its ability to break the deterministic link between the market price of risk and the yield-curve state variables that essentially-affine, yield-curve-based-models are constrained to impose. Since our goal is to produce reliable and stable estimates of risk premia and expectations, we have shown that this degree of flexibility can be very important.

We find that the proposed model fits well the market covariance matrix, and that its pricing errors to market yields are very small. Apart from giving an intrinsic confidence in the reasonableness of the approach, these features are important for the reliable estimation of risk premia and expectations.

We analyzed in detail the predictions during the "taper tantrum" period produced by our model, and by models with a market price of risk described by a deterministic (affine) function of the state variables. We explained why it would be unwise to rely on the expectation/risk-premium decomposition afforded by these model in the abnormal conditions of rate repression observed after the Great Recession. We argued that the decomposition afforded by our model is in this respect more plausible.

The model is parsimonious, robust and easy to calibrate. It can be regarded as a particularly simple instance of a class of reduced-form stochastic volatility models. An obvious extension would be to use principal components as the yield-curve-based state variables. We leave this line of research for future work.

Appendix A. Solving the Model

Given a process for a vector, $\mathbf{x}_t$, of generic state variables

$$d\mathbf{x}_t = \mathcal{K}(\boldsymbol{\theta} - \mathbf{x}_t)dt + S d\mathbf{z}_t, \tag{A.1}$$

such that

$$r_t = u_r + \mathbf{g}^T \mathbf{x}_t, \tag{A.2}$$

the bond price, $P_t^T = P(\tau)$, is given by

$$P_t^T = e^{A_t^T + \mathbf{B}^T \mathbf{x}_t}, \tag{A.3}$$

with A_t^T and $\mathbf{B}^T$ obtained as the solutions of ordinary differential equations

$$\frac{dA(\tau)}{d\tau} = -u_r + \mathbf{B}_\tau^{\mathrm{T}} \mathcal{K}\boldsymbol{\theta} + \frac{1}{2}\mathbf{B}^{\mathrm{T}} S S^T \mathbf{B}_\tau^{\mathrm{T}}, \tag{A.4}$$

$$\frac{d\mathbf{B}(\tau)}{d\tau} = -\mathbf{g} - \mathcal{K}^{\mathrm{T}}\mathbf{B}_\tau, \tag{A.5}$$

with the initial conditions

$$A(0) = 0, \tag{A.6}$$

$$\mathbf{B}(0) = \mathbf{0}. \tag{A.7}$$

Assuming $\mathcal{K}$ to be invertible and diagonalizable as $\mathcal{K} = L\mathcal{L}L^{-1}$, one obtains

$$\mathbf{B}(\tau) = (e^{-\mathcal{K}^T\tau} - I_n)(\mathcal{K}^T)^{-1}\mathbf{g}, \tag{A.8}$$

$$\mathbf{B}_\tau^{\mathrm{T}} = \mathbf{g}^T\mathcal{K}^{-1}[e^{-\mathcal{K}\tau} - I_n], \tag{A.9}$$

$$A(\tau) = \mathrm{Int}_1 + \mathrm{Int}_2 + \mathrm{Int}_3, \tag{A.10}$$

with

$$\mathrm{Int}_1 = -u_r T, \tag{A.11}$$

$$\mathrm{Int}_2 = \mathbf{g}^T(L\mathcal{L}^{-1}D(T)\mathcal{L}L^{-1}\boldsymbol{\theta} - \boldsymbol{\theta}T), \tag{A.12}$$

$$\mathrm{Int}_3 = \mathrm{Int}_3^a + \mathrm{Int}_3^b + \mathrm{Int}_3^c + \mathrm{Int}_3^d, \tag{A.13}$$

$$\mathrm{Int}_3^a = \frac{1}{2}\mathbf{g}^T\mathcal{K}^{-1}LF(T)\mathcal{L}^{-1}L^T\mathbf{g}, \tag{A.14}$$

$$\mathrm{Int}_3^b = -\frac{1}{2}\mathbf{g}^T\mathcal{K}^{-1}C(L^{-1})^T D(T)\mathcal{L}^{-1}L^T\mathbf{g}, \tag{A.15}$$

$$\mathrm{Int}_3^c = -\frac{1}{2}\mathbf{g}^T\mathcal{K}^{-1}LD(T)L^{-1}C(\mathcal{K}^T)^{-1}\mathbf{g}, \tag{A.16}$$

$$\mathrm{Int}_3^d = \frac{1}{2}\mathbf{g}^T\mathcal{K}^{-1}C(\mathcal{K}^T)^{-1}\mathbf{g}T \tag{A.17}$$

with $M = L^{-1}C(L^{-1})^T$, $F = [f_{ij}]_n$ and

$$f_{ij} = m_{ij}\frac{1 - e^{(l_{ii}+l_{jj})T}}{l_{ii} + l_{jj}}, \tag{A.18}$$

$$D(T) = \mathrm{diag}\left[\frac{1 - e^{l_{ii}T}}{l_{ii}}\right]_n, \tag{A.19}$$

$$C = SS^T. \tag{A.20}$$

In the case of the model at hand, these equations specialize as follows. Let the model be

$$\begin{aligned}
dr_t &= \kappa_r(\theta_t - r_t)dt + \sigma_r dz_t^r, \\
d\theta_t &= \kappa_\theta(\Theta^\infty - \theta_t)dt + \lambda_t^\theta\sigma^\theta dt + \sigma^\theta dz_t^\theta
\end{aligned} \tag{A.21}$$

$$= [\kappa_\theta\Theta^\infty - \kappa_\theta\theta_t + \lambda_t^\theta\sigma^\theta]dt + \sigma^\theta dz_t^\theta, \tag{A.22}$$

$$d\lambda_t^\theta = \kappa_\lambda(\Lambda^\infty - \lambda_t^\theta)dt + \sigma^\lambda dz_t^\lambda, \tag{A.23}$$

with

$$E[dr d\theta] = \rho_{r\theta}, \tag{A.24}$$

$$E[dr d\lambda] = \rho_{r\lambda}, \tag{A.25}$$

$$E[d\theta d\lambda] = \rho_{\theta\lambda}. \tag{A.26}$$

Then, after numbering the state variables as $\lambda_t^\theta = x_t^1$, $\theta_t = x_t^2$ and $r_t = x_t^3$, one has

$$\mathcal{K} = \begin{bmatrix} \kappa_\lambda & 0 & 0 \\ -\sigma_\theta & \kappa_\theta & 0 \\ 0 & -\kappa_r & \kappa_r \end{bmatrix}, \tag{A.27}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \Lambda^\infty \\ \Theta^\infty + \dfrac{\sigma_\theta \Lambda^\infty}{\kappa_\theta} \\ \Theta^\infty + \dfrac{\sigma_\theta \Lambda^\infty}{\kappa_\theta} \end{bmatrix}. \tag{A.28}$$

The affine transformation from the state variables to the short rate is given by

$$u_r = 0, \tag{A.29}$$

$$\mathbf{g}^T = [0, 0, 1]. \tag{A.30}$$

For the matrix S one obtains

$$S = \begin{bmatrix} s_{11} & 0 & 0 \\ s_{21} & s_{22} & 0 \\ s_{31} & s_{32} & s_{33} \end{bmatrix}, \tag{A.31}$$

with

$$s_{11} = \sigma_\lambda, \tag{A.32}$$

$$s_{21} = \sigma_\theta \rho_{\lambda\theta}, \tag{A.33}$$

$$s_{22} = \sigma_\theta \sqrt{1 - \rho_{\lambda\theta}^2}, \tag{A.34}$$

$$s_{31} = \sigma_r \rho_{\lambda r}, \tag{A.35}$$

$$s_{32} = \sigma_r \frac{\rho_{\theta r} - \rho_{\lambda\theta}\rho_{\lambda r}}{\sqrt{1 - \rho_{\lambda\theta}^2}}, \tag{A.36}$$

$$s_{33} = \sigma_r \sqrt{1 - \rho_{\lambda r}^2 - \frac{(\rho_{\theta r} - \rho_{\lambda\theta}\rho_{\lambda r})^2}{1 - \rho_{\lambda\theta}^2}}. \tag{A.37}$$

References

T. Adrian, R. K. Crump & E. Moench (2013) Pricing the term structure with linear regressions. Staff Report No. 340, Federal Reserve Bank of New York, August 2008, revised 2013. Available at: http://www.newyorkfed.org/research/staff_reports/sr340.pdf [accessed 15 October 2014].

M. D. Bauer & J. D. Hamilton (2015) Robust risk premia, Working Paper, Federal Reserve Bank of San Francisco, September 2015. Available at: http://econweb.ucsd.edu/~jhamilto/bh_robust.pdf [accessed 27 November 2015].

B. S. Bernanke (2013) *The Economic Outlook*: *Testimony Before the Joint Economic Committee*, Washington, DC: US Congress. Available at: http://www.federalreserve.gov/newsevents/testimony/bernanke20130522a.htm [accessed 11 November 2014].

F. Black (1995) *Interest Rates as Options*, *The Journal of Finance* **50**, 1371–1376.

A. N. Bomfin (2003) "Interest Rates as Options": Assessing the markets' view of the liquidity trap. Working Paper, Board of Governors of the Federal Reserve System, Washington, DC. Available at: http://www.federalreserve.gov/pubs/feds/2003/200345/ [accessed 31 March 2015].

J. Y. Campbell & R. Shiller (1991) Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* **58**, 495–514.

J. H. E. Christensen & S. Kwan (2014) Assessing expectations of monetary policy. *FRBSF Economic Letter* 2014-27, 8 September 2014.

A. Cieslak & P. Povala (2010) Understanding bond risk premia. Working Paper, Kellogg School of Management, Wisconsin University and University of Lugano. Available at: https://www.gsb.stanford.edu/sites/default/files/documents/fin_01_11_CieslakAnna.pdf [accessed 5 May 2015].

A. Cieslak & P. Povala (2013) Expected returns in treasury bonds. Working Paper, Northwestern University and Birbeck College. [forthcoming in *Review of Financial Studies*].

J. H. Cochrane (2015) Comments on "Robust Risk Premia" by Michael Bauer and Jim Hamilton. Working Paper, University of Chicago, November 2015. Available at: http://faculty.chicagobooth.edu/john.cochrane/research/Papers/bauer_hamilton_comments.pdf [accessed 27 November 2015].

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

J. H. Cochrane & M. Piazzesi (2008) Decomposing the yield curve. University of Chicago and NBER Working Paper, March 2008.

R. K. Crump, S. Eusepi & E. Moench (2016) The term structure of expectations and bond yields, Federal Reserve Bank of New York Staff Report No. 775, May 2016.

S. D'Amico, D. H. Kim & M. Wei (2010) Tips from TIPS: The informational content of treasury inflation protected security prices. Working Paper, Finance and Economics Discussion Series, Federal Reserve Board, Washington, DC.

Q. Dai & K. J. Singleton (2000) Specification analysis of affine term structure models, *Journal of Finance* **55** (5), 1943–1978.

Q. Dai, K. J. Singleton & W. Yang (2004) Predictability of bond risk premia and affine term structure models. Working Paper, Stern School of Business.

A. Erekhinskiy (2013) *A Model for Interest Rates in Repressed Markets*. M.Sc. Thesis, University of Oxford, UK.

E. Fama & R. Bliss (1987) The information in long-maturity forward rates, *American Economic Review* **77**, 680–692.

R. S. Gurkaynak, B. Sack & J. H. Wright (2006) The US treasury yield curve: 1961 to the Present. Working Paper No. 2006-28, Division of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, DC.

R. S. Gurkaynak, B. Sack & J. H. Wright (2007) The US treasury yield curve: 1961 to the present, *Journal of Monetary Economics* **54**, 2291–2304. Available at: http://www.federalreserve.gov/econresdata/researchdata/feds200805.xls [accessed 27 January 2015].

R. Hellerstein (2011) Global bond risk premiums. Federal Reserve Bank of New York Staff Report No. 499, June. Available at: http://www.nyfedeconomists.org/research/staff_reports/ [accessed 4 June 2015].

S. Joslin, M. Priebsch & K. J. Singleton (2014) Risk premiums in dynamic term structure models with unspanned macro risks, *Journal of Finance* **LXIX** (3), 1197–1233.

S. C. Ludvigson & S. Ng (2009) Macro factors in bond risk premia, *Review of Financial Studies* **22**, 5027–5067. Available at: http://rfs.oxfordjournals.org/content/22/12/5027.abstract [accessed 6 May 2015].

V. Naik, M. Devarajan, A. Nowbilski, S. Page & N. Pedersen (2016) *Factor investing and asset allocation — A business cycle perspective.* CFA Research Foundation, London.

M. Piazzesi (2010) Affine term structure models. In: *Handbook of Financial Econometrics*, Vol. 1, 691–766. Amsterdam: Elsevier.

J. F. Radwanski (2010) Understanding bond risk premia uncovered by the term structure. Working Paper, Vienna Graduate School of Finance, Vienna University of Economics (VGSF). Available at: http://www.wu.ac.at/ruw/other/vgsf/students/Radwanski10Bonds.pdf [accessed 5 May 2015].

R. Rebonato (2002) Modern Pricing of Interest-Rate Derivatives: The LIBOR Market Model and Beyond, Princeton University Press, Princeton and Oxford.

R. Rebonato (2015) Return-predicitng factors for US treasuries: On the similarity of 'Tents' and 'Bats', *International Journal of Theoretical and Applied Finance* **18** (04), 1–14.

R. Rebonato (2016) Structural affine curve models for yield curve modeling. In: *Handbook of Fixed Income Securities* (P. Veronesi ed.), 239–264. Chichester: John Wiley & Sons.

R. Rebonato & V. Putiatyn (2017), The value of convexity: A theoretical and empirical investigation, forthcoming, accepted for publication in *Quantitative Finance.*

S. Rebonato, I. Saroka & V. Putiatyn (2017), Principal-component-based affine term structure models: Constraints and results, submitted to *Review of Financial Studies.*

D. Villegas (2015) *Bond Excess Returns Predictability and the Expectation Hypothesis* M.Sc. Thesis, University of Oxford, Hillary Term.

J. C. Wu & F. D. Xia (2015) Shadow federal funds rate. Available at: https://www.frbatlanta.org/cqer/researchcq/shadow_rate.aspx [accessed 2 April 2015].