WORKING PAPER

DANMARKS NATIONALBANK

7 OCTOBER 2019 — NO. 142

Modeling Persistent Interest Rates with Volatility-Induced Stationarity

Anne Lundgaard Hansen
alh@econ.ku..dk

DANMARKS NATIONALBANK AND
UNIVERSITY OF COPENHAGEN

WORKING PAPER — DANMARKS NATIONALBANK
7 OCTOBER 2019 — NO. 142

Modeling Persistent Interest Rates with Volatility-Induced Stationarity

Abstract
It is well-known that interest rates are extremely persistent, yet they are best modeled and understood as stationary processes. These properties are contradictory in the workhorse Gaussian affine term structure model in which persistent data often result in unit roots that imply non-stationarity. We resolve this puzzle by proposing a term structure model with volatility-induced stationarity. Our model employs a level-dependent conditional volatility that maintains stationarity despite the presence of unit roots in the characteristic polynomial corresponding to the conditional mean. The model is consistent with key characteristics of U.S. Treasury data and obtains term premia that are economically plausible and consistent with survey data. Compared to the Gaussian affine term structure model, we improve out-of-sample forecasting of the yield curve.

Resume
Det er velkendt at renter er ekstremt persistente, men bedst modelleres og forsås som stationære processer. Disse egenskaber er modsigende i de populære Gaussiske affine rentemodeller, hvor persistente data ofte resulterer i enhedsrødder, der medfører ikke-stationaritet. Vi løser denne uoverensstemmelse ved at foreslå en rentemodel med volatilitetsinduceret stationaritet. Vores model introducerer en niveauafhængig betinget variance der sørger for at processen er stationær på trods af en enhedsrod i det karakteristiske polynomie svarende til den betingede middelværdi. Modellen er konsistent med vigtige karakteriska i rentedata fra USA og estimerer økonomisk plausible risikopræmier. Vi forbedrer også prædiktion af rentekurven sammenholdt med den Gaussiske affine rentemodel.

Key words
Yield curve; unit root; persistence problem; volatility-induced stationarity; macro-finance term structure model; level-dependent conditional volatility.

JEL classification
E43; E44; G12.

Acknowledgements
I thank Torben G. Andersen, Luca Benzoni, Heino Bohn Nielsen, Anders Rahbek, Guillaume Roussellet, Viktor Todorov, and seminar and conference participants at the 11th International Conference on Computational and Financial Econometrics; the 1st International Conference on Quantitative Finance and Financial Econometrics; the 2017 Society of Financial Econometrics Workshop at Kellogg School of Management, Northwestern University; Danmarks Nationalbank; University of Copenhagen; and the 12th Annual Society of Financial Econometrics Conference.

All remaining errors are my own.

Modeling Persistent Interest Rates with Volatility-Induced Stationarity

Anne Lundgaard Hansen †

October 7, 2019

Latest version available here

Abstract

It is well-known that interest rates are extremely persistent, yet they are best modeled and understood as stationary processes. These properties are contradictory in the workhorse Gaussian affine term structure model in which persistent data often result in unit roots that imply non-stationarity. We resolve this puzzle by proposing a term structure model with volatility-induced stationarity. Our model employs a level-dependent conditional volatility that maintains stationarity despite the presence of unit roots in the characteristic polynomial corresponding to the conditional mean. The model is consistent with key characteristics of interest rate data. In an empirical macro-finance application, we obtain term premia that are economically plausible and consistent with survey data. Compared to the Gaussian affine term structure model, we improve out-of-sample forecasting of the yield curve. Our empirical evidence suggests that volatility-induced stationarity is unspanned by the yield curve.

Keywords: Yield curve, unit root, persistence problem, volatility-induced stationarity, macro-finance term structure model, level-dependent conditional volatility.
JEL classification: E43, E44, G12.

†University of Copenhagen, Department of Economics. *E-mail:* alh@econ.ku.dk. *Website:* http://sites.google.com/view/annelundgaardhansen.
I thank Torben G. Andersen, Luca Benzoni, Heino Bohn Nielsen, Anders Rahbek, Guillaume Roussellet, Viktor Todorov, and seminar and conference participants at the 34th Annual Congress of the EEA, 12th Annual SoFiE conference, 11th International Conference on Computational and Financial Econometrics, 1st International Conference on Quantitative Finance and Financial Econometrics, 2017 SoFiE Workshop at Kellogg School of Management, Danmarks Nationalbank, and University of Copenhagen. All remaining errors are my own.

1 Introduction

Many macro-finance term structure models are specified by vector autoregressive (VAR) models with Gaussian and homoskedastic shocks. While these models are celebrated for their tractability, they are inconsistent with key characteristics of U.S. Treasury yield data. This paper introduces a novel class of discrete-time term structure models that can generate yield curve dynamics supported by the data. Specifically, we bridge macro-finance term structure modeling with the double-autoregressive (DAR) model studied in Ling (2004) and Nielsen and Rahbek (2014).

We are motivated by three stylized facts of nominal bond yields that standard VAR models fail to accommodate. First, U.S. Treasury bond yields are extremely persistent and formal tests often fail to reject the presence of unit roots. When VAR models are presented with highly persistent data, they imply a sharp distinction between I(0) and I(1) models. While I(0) models are stationary, they fail to match the degree of persistence in the yield data (Goliński and Zaffaroni, 2016). On the other hand, I(1) models are sufficiently persistent but non-stationary, which is counterfactual from both theoretical and empirical viewpoints (Beechey et al., 2009). Second, the data exhibit periods of rapid changes perhaps marking the beginning and end of monetary policy cycles. Finally, interest rates exhibit time-varying conditional volatility.[1]

To accommodate these stylized facts, we develop a term structure model with multivariate DAR dynamics. The DAR model is a vector autoregression with conditional volatility that depends on lagged levels of the process. Thus, in particular, the model is consistent with conditionally heteroskedastic interest rates. To fix ideas, consider the univariate DAR model from Ling (2004):

$$x_t = \phi x_{t-1} + \left(\omega + \psi x_{t-1}^2\right)^{1/2} z_t \tag{1}$$

with $z_t \sim$ i.i.d. $\mathcal{N}(0,1)$. A crucial feature of the model is that stationarity is not ruled out by the presence of a unit root, $\phi = 1$. Instead, the stationarity condition depends on both the conditional mean through ϕ and the conditional variance through ψ. Therefore, the

[1]Heteroskedastic interest rates have been acknowledged by non-Gaussian affine term structure models (Dai and Singleton, 2000), which have been studied in discrete time by Le et al. (2010). However, these models fail to capture the first two stylized facts that we emphasize. We provide a formal comparison between our approach and the affine framework in the paper.

model is said to exhibit volatility-induced stationarity. By allowing for unit roots without implying non-stationarity, the DAR model is consistent with the stylized fact that interest rate data are persistent but best described by stationary processes. The DAR model can also generate the jump behavior of interest rates by a sequence of shocks with the same sign. These shocks accumulate in the conditional variance so that when a shock of the opposite sign arrives, it is weighted by a large conditional variance that pushes the process rapidly downwards. Thus, the DAR model is consistent with the stylized facts of interest rate data.

We present an empirical application with a macro-finance state vector consisting of the one-month U.S. Treasury bill rate, the ten-year Treasury bond yield, and measures of inflation and real activity. The data guide a model with reduced rank in the autoregressive coefficient and long-run equilibria given by the yield spread and a Taylor rule. We therefore implement the DAR model with these features and benchmark our results against the cointegrated VAR model.[2] Our empirical analysis shows that the misspecification of the VAR model has both econometric and economic consequences that can be alleviated by the DAR model. Econometrically, we show that the DAR model passes misspecification tests of the standardized residuals that the corresponding VAR model fails. Economically, we emphasize the well-known problem that VAR models distort model-implied term premia because they fail to match the persistence of the data with a stationary model. In the following, we illustrate this so-called persistence problem and how the DAR model can remedy this limitation of linear models.

Term premia are here defined by the residuals of the yield curve that are not explained by the expectations hypothesis, which asserts that yields of long maturities are determined by expected future short rates only. Thus, term premia can be estimated based on model-implied forecasts of the short rate. In the stationary VAR, these forecasts quickly revert to the unconditional mean defined by the model. Therefore, the expectations hypothesis of the stationary VAR predicts nearly constant yields, and virtually all variation in the yield curve is assigned to term premia. This issue, named the persistence problem, has been recognized by Jardet et al. (2013), Kozicki and Tinsley (2001), and Shiller (1979). In the cointegrated VAR model with the yield spread as cointegrating relation, forecasts of future short rates converge to a level that is proportional to the ten-year yield. In result,

[2]Term structure modeling based on the cointegrated VAR model has been considered in Chernov and Creal (2019).

the expectations hypothesis explains most of the variation in the yield curve, resulting in nearly constant term premia.[3]

In sum, the VAR framework can either generate term premia that are approximately proportional to the yield curve or constant. Interestingly, the DAR model can generate a richer set of term structure decompositions than the VAR models by reconciling unit roots and stationarity. In particular, the DAR model predicts term premia that are time-varying, but much less correlated with the yield curve than explained by the stationary VAR model. Indeed, we find that the DAR model matches expected future short rates as measured by the Survey of Professional Forecasters better than the stationary and cointegrated VAR models.

Given these promising results related to the modeling of macro-finance dynamics, we embed the DAR model into a macro-finance term structure model with no-arbitrage restrictions. Assuming a standard exponential-linear stochastic discount factor preserves the DAR model under the pricing measure. We propose a quadratic approximation to facilitate analytical computation of no-arbitrage bond yields. Our model obtains an in-sample fit of the yield curve comparable to the Gaussian affine term structure model (GATSM) that is based on the VAR model. In fact, the quadratic component of our bond yield formula that is generated from volatility-induced stationarity explains practically no variation in the yield curve. This result can be interpreted as volatility-induced stationarity being unspanned by the yield curve, which is consistent with the literature on unspanned stochastic volatility (USV) (Collin-Dufresne and Goldstein, 2002, 2009, Creal and Wu, 2015, Joslin, 2017). In contrast, the DAR term structure model does outperform the GATSM in terms of out-of-sample performance across almost all maturities from one to ten years and forecasting horizons of 3, 6, and 12 months. Importantly, the DAR model also outperforms the random walk, which is a competitive benchmark for standard term structure models (Duffee, 2002).

Volatility-induced stationarity in interest rate data was first studied by Conley et al. (1997) who consider Markov diffusion models with constant volatility elasticity as in the CKLS model in Chan et al. (1992). Conley et al. (1997) apply these models to overnight effective federal funds rates and conclude that *"when interest rates are high, local mean reversion is small and the mechanism for inducing stationarity is the increased volatility"*. Nicolau (2005) also shows that the federal funds rate can be modelled by a process that

[3]This observation also explains why the cointegrated VAR has been used to test the expectations hypothesis, e.g., in Campbell and Shiller (1987), Hall et al. (1992), and Shea (1992).

exhibits volatility-induced stationarity. Nielsen and Rahbek (2014) extend these analyses by modeling two interest rates, namely the one- and three-month Treasury bill rates, allowing for reduced rank. Their implementation, however, does not impose no-arbitrage restrictions. This paper contributes to this literature by (i) proposing a no-arbitrage model for the entire term structure and (ii) allowing for more than two factors, e.g., the usual level, slope, and curvature factors of the yield curve as suggested by Litterman and Scheinkman (1991) and macroeconomic factors as in the macro-finance term structure literature (Ang and Piazzesi, 2003, Ang et al., 2006, Diebold et al., 2005, Duffee, 2006, Hördahl et al., 2006, Joslin et al., 2014, Rudebusch and Wu, 2008). We are the first to suggest that the persistence problem can be resolved by volatility-induced stationarity.

Other methodologies have been suggested to overcome the persistence problem. One strand of literature focuses on the well-known statistical problem that the autoregressive parameter of stationary VAR models is downwardly biased in small samples when data are persistent. To tackle this problem, Kim and Orphanides (2007) and Kim and Orphanides (2012) augment the data with survey forecasts and Bauer et al. (2014) suggest a bias-correction that results in stable term premia. This approach is conceptually different from that taken in this paper in which linear dynamics is abandoned to introduce nonlinearity in the form of volatility-induced stationarity. In result, we show empirically that term premia implied by our model have different properties from those obtained by Bauer et al. (2014). Abbritti et al. (2016) and Goliński and Zaffaroni (2016) suggest that long memory represents a realistic, intermediate case between I(0) and I(1) GATSMs. Along these lines, Jardet et al. (2013) consider near-cointegration implemented by averaging the parameter estimates of the stationary and cointegrated VAR models. The resulting term premia coincide with those of the DAR model during the zero-lower bound regime, but differ elsewhere.

The paper is structured as follows. Section 2 introduces the DAR model and discusses how unit roots can be reconciled with stationary dynamics through volatility-induced stationarity. The empirical analysis of the DAR model using a set of macro-finance risk factors is conducted in Section 3. In Section 4, we embed the DAR process into a no-arbitrage term structure model and use this model to assess the implications of volatility-induced stationarity on the yield curve. Finally, Section 5 compares the volatility specification of the DAR model to other specifications from the literature. Conclusions follow in Section 6.

2 Double Autoregressive Models

DAR models specify both the conditional mean and the conditional variance in terms of lagged levels of the process. The conditional mean is equivalent to that of the VAR, while the conditional variance can be specified based on various multivariate GARCH models. In general, the p-dimensional DAR model with one lag in both the conditional mean and the conditional variance is given by

$$\begin{aligned} X_{t+1} &= \mu + \Phi X_t + \Omega_{t+1}^{1/2}\varepsilon_{t+1}, \\ \Omega_{t+1} &= f(X_t), \end{aligned}$$

where $\varepsilon_{t+1} \sim$ i.i.d. $\mathcal{N}(0, \mathrm{I}_p)$ and $f : \mathbb{R}^p \to \mathbb{R}^{p\times p}$ is a function that maps the levels of the process X_t into a symmetric and positive definite matrix. Next, we specify the conditional variance, Ω_{t+1}.[4]

2.1 Conditional Variance Specification

We specify the conditional variance such that (i) symmetry and positive definiteness is imposed by construction; (ii) the number of parameters is feasible for estimation; and (iii) we can establish time series properties of the model. The BEKK ARCH model in Engle and Kroner (1995) specified in levels rather than residuals satisfies these requirements. The resulting DAR model is given by

$$\begin{aligned} X_{t+1} &= \mu + \Phi X_t + \Omega_{t+1}^{1/2}\varepsilon_{t+1}, \\ \Omega_{t+1} &= \Sigma_0\Sigma_0' + \Sigma_1 X_t X_t' \Sigma_1', \\ \varepsilon_{t+1} &\sim \text{i.i.d. } \mathcal{N}(0, \mathrm{I}_p). \end{aligned} \tag{2}$$

where Σ_0 is lower triangular with strictly positive elements on the diagonal. This unique Cholesky factor ensures that the conditional variance matrix is positive definite without imposing further parameter restrictions. The $p \times p$ matrix Σ_1 determines the sensitivity of the conditional volatility to the level of the process realized in the previous period. In the special case where all elements of this matrix are zero, the model reduces to a VAR.

[4]Note that we adopt the notation from the GARCH literature and denote the conditional variance given X_t by Ω_{t+1}.

Economically, the model allows uncertainty as measured by the conditional variance to increase with the levels of the yield curve factors. By including macroeconomic variables in the model, we can accommodate the hypothesis that a higher inflation rate increases uncertainty about monetary policy (Ball, 1992, Fischer and Modigliani, 1978, Friedman, 1977, Logue and Willett, 1976).[5] Also, the model is consistent with Hayford (2000) who finds that inflation Granger causes unemployment uncertainty. Due to these economic channels of heteroskedasticity, we present a macro-finance empirical application in Section 3.

2.2 Stationarity Condition and Time Series Properties

Nielsen and Rahbek (2014) show that X_t given by (2) is globally stationary and geometrically ergodic if the Lyapunov exponents are strictly negative, i.e.,

$$\gamma\left(\Phi, \Sigma_1\right) = \lim_{\xi\to\infty}\left[\mathbb{E}\left(\log \|\prod_{t=1}^{\xi}(\Phi + e_t)\|\right)\right] < 0, \tag{3}$$

where e_t is a $p \times p$ matrix that is i.i.d. normally distributed with mean zero and covariance matrix given by $\Sigma_1 \otimes \Sigma_1$. We note that this condition is determined by both the conditional mean through Φ and the conditional variance through Σ_1. Thus, stationarity can be induced by both the mean and the variance.

The stationarity condition in (3) motivates a classification of the DAR model into four cases: (i) non-stationary I(1) models, (ii) models that are stationary due to the conditional volatility only, (iii) models with both mean- and volatility-induced stationarity, and (iv) stationary models without volatility-induced stationarity, that is I(0) models.[6] To characterize the properties of the DAR model, we look at these cases separately.

(i) I(1) models

Assume that the characteristic polynomial corresponding to the conditional mean exhibits one or more unit roots. In addition, suppose that the parameter Σ_1 in the conditional variance is not sufficiently large to ensure stationarity through a strictly negative top-Lyapunov exponent. Formally, let $\lambda_1, \ldots, \lambda_p$ denote the eigenvalues of Φ. Then, $|\lambda_i| = 1$ for $i = 1, \ldots, q \leq p$, $|\lambda_j| < 1$ for $j = q + 1, \ldots, p$, and $\gamma(\Phi, \Sigma_1) \geq 0$. If the number of unit roots equals the dimension of the model, i.e., if $q = p$, then the model can be made

[5]This hypothesis has been tested comprehensively in the literature, see for instance Chang (2012), Fountas (2010), Golob (1994), Hartmann and Herwartz (2012), and Kim and Lin (2012).

[6]We abstract from cases where the characteristic polynomial corresponding to the conditional mean has solutions outside the unit circle although (3) can be satisfied under such cases.

stationary by first differencing. Otherwise, the rank of $\Phi - \mathrm{I}_p$ is reduced to $r = p - q$, and Φ can be parametrized by $\Phi = I_p + \alpha\beta'$, where α and β are $p \times r$ matrices. Furthermore, if $\beta' X_t$ is stationary given an initial distribution, we say that X_t is cointegrated with cointegrating vector $\beta \neq 0$ as defined by Johansen (1995) for the I(1) VAR model, i.e., for $\Sigma_1 = 0_{p\times p}$. Due to the presence of unit roots, the constant term μ is aggregated into a linear trend if not restricted appropriately as a constant in the cointegrating relations.

(ii) Purely volatility-induced stationary models

The DAR model is purely volatility-induced stationary if the conditional mean exhibits one or more unit roots, but the top-Lyapunov exponent is strictly negative due to a sufficiently large level effect in the conditional variance: $|\lambda_i| = 1$ for some $i = 1, \dots, q \leq p$, $|\lambda_j| < 1$ for $j = q+1, \dots, p$, and $\gamma(\Phi, \Sigma_1) < 0$. Thus, the model is stationary despite the presence of unit roots because of the dynamics of the conditional volatility. Crucially, we note that if the conditional volatility was not time-varying, the model would belong to the non-stationary class of I(1) models described in case (i).

In general, the model does not have any finite unconditional moments. Thus, stationarity is not equivalent to mean-reversion in the traditional sense, where the process reverts back to a level given by the unconditional mean. Instead, the process will tend to spend most time at the level at which the conditional variance is low, i.e., at zero. What happens as the process moves away from zero, say, due to a series of positive shocks? Increasing values of X_t accumulate in the conditional variance so that the stochastic component becomes larger as the process moves farther away from zero. Since the error term is normally distributed and thus symmetric, a negative innovation will eventually arrive, which pushes the process downwards. In this way, the process can quickly return to its stable level. It will take another series of innovations of the same sign for the process to repeat this pattern. Theoretically, nothing prevents that the innovation will continue to be positive such that the process never returns towards its stable level. However, this event happens with zero probability because the innovation term is Gaussian.[7] Thus, it is not a relevant concern for the empirical application of the model.

Finally, Nielsen and Rahbek (2014) show that due to volatility-induced stationarity, the constant term μ does not accumulate into a linear trend as is the case in I(1) models.

[7] For a sequence of i.i.d. variables $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T$, where ε_i is symmetrically distributed with mean zero, $\Pr(\varepsilon_1 > 0, \varepsilon_2 > 0, \dots, \varepsilon_T > 0) = 0.5^T$.

(iii) Mean- and volatility-induced stationary models

Suppose that all eigenvalues of Φ are inside the unit circle, $|\lambda_i| < 1$ for all $i = 1, \ldots, p$, and the conditional variance is level-dependent. For empirically relevant values of Σ_1, it will be the case that $\gamma(\Phi, \Sigma_1) < 0$ and the model is stationary.[8] Stationarity is ensured jointly by Φ and Σ_1, hence, the model exhibits both mean- and volatility-induced stationarity. The unconditional mean, variance, and autocovariance are given by

$$\begin{aligned}
\mathbb{E}\left(X_t\right) &= \left(\mathrm{I}_p - \Phi\right)^{-1} \mu, \\
\operatorname{vec}\left[\operatorname{Var}\left(X_t\right)\right] &= \left(\mathrm{I}_{p^2} - \Phi \otimes \Phi - \Sigma_1 \otimes \Sigma_1\right)^{-1} \operatorname{vec}\left(\Sigma_0 \Sigma_0' - \Sigma_1 \left(\mathrm{I}_p - \Phi\right)^{-1} \mu\mu' \left(\mathrm{I}_p - \Phi'\right)^{-1} \Sigma_1'\right), \\
\operatorname{Cov}\left(X_t, X_{t-m}\right) &= \Phi^m \operatorname{Var}(X_t).
\end{aligned}$$

To understand the model dynamics intuitively, consider a case where $\mathbb{E}(X_t) > 0$ such that the level of mean-reversion is different and exceeds the level at which the model has low conditional variance. When the process is at zero, the stochastic component is small due to a small conditional variance, and the process is mainly controlled by the conditional mean that drives the process towards its unconditional mean. Thus, the transition from levels close to zero to the unconditional mean resembles that of the stationary VAR model. Above the unconditional mean, however, the stochastic component is attributed more weight as the conditional variance is increased by the process being away from zero. In the case of a negative shock, the process will quickly revert to the unconditional mean as both the conditional mean and stochastic component drive the process downwards. A positive shock, on the other hand, implies that the conditional mean and conditional variance work in opposite directions.

(iv) I(0) models

Let the eigenvalues of Φ be inside the unit circle and the conditional variance be constant: $\Sigma_1 = 0_{p\times p}$ and $|\lambda_i| < 1$ for all $i = 1, \ldots, p$. Then, the DAR model reduces to the stationary VAR whose properties are well-known. Since the conditional variance is constant, the model is stationary purely due to the absence of unit roots in the characteristic polynomial corresponding to the conditional mean.

[8] Ling (2004) shows in the univariate case that extremely large values of Σ_1 can result in non-stationarity. We will not pay attention to this empirically irrelevant case.

2.3 Numerical Illustrations

We illustrate these properties numerically by simulating paths of the model in each of the four cases. Consider a sample of length $T = 300$ generated from univariate DAR models with the following parameter values:

(i) I(1) model: $(\mu,\ \Phi,\ \Sigma_0,\ \Sigma_1) = (0.01,\ 1,\ 0.1,\ 0)$.

(ii) purely volatility-induced stationary model: $(\mu,\ \Phi,\ \Sigma_0,\ \Sigma_1) = (0.01,\ 1,\ 0.1,\ 0.3)$.

(iii) Mean- and volatility-induced stationary model: $(\mu,\ \Phi,\ \Sigma_0,\ \Sigma_1) = (0.01,\ \phi,\ 0.1,\ 0.3)$.

(iv) I(0) model: $(\mu,\ \Phi,\ \Sigma_0,\ \Sigma_1) = (0.01,\ \phi,\ 0.1,\ 0)$.

We repeat the simulation exercise for two different values of the autoregressive coefficient in the cases (iii) and (iv): $\phi = 0.99$, which is close to the unit-root case and an empirically relevant value, and $\phi = 0.95$ to illustrate the model when the mean-reversion effect is stronger. Results are shown in Figure 1. The I(1) model is a random walk and the I(0) model is a stationary AR process that fluctuates around the unconditional mean illustrated by the dotted blue line. When pure volatility-induced stationarity is present, the process tends to spend most time around zero, where volatility is low. As the process moves away from zero, volatility increases which can generate spikes as those observed around $t = 225$ and $t = 275$ in the simulation sample. However, as soon as a negative innovation is realized, the process returns quickly to more stable levels. The dynamics of the mean- and volatility-induced stationary model, case (iii), depends crucially on the autoregressive parameter. When $\phi = 0.99$, i.e., close to unity, the process behaves almost like the purely volatility-induced stationary model, see panel (a). With a stronger degree of mean-reversion, see panel (b), the DAR model resembles the I(0) model.

2.4 Likelihood

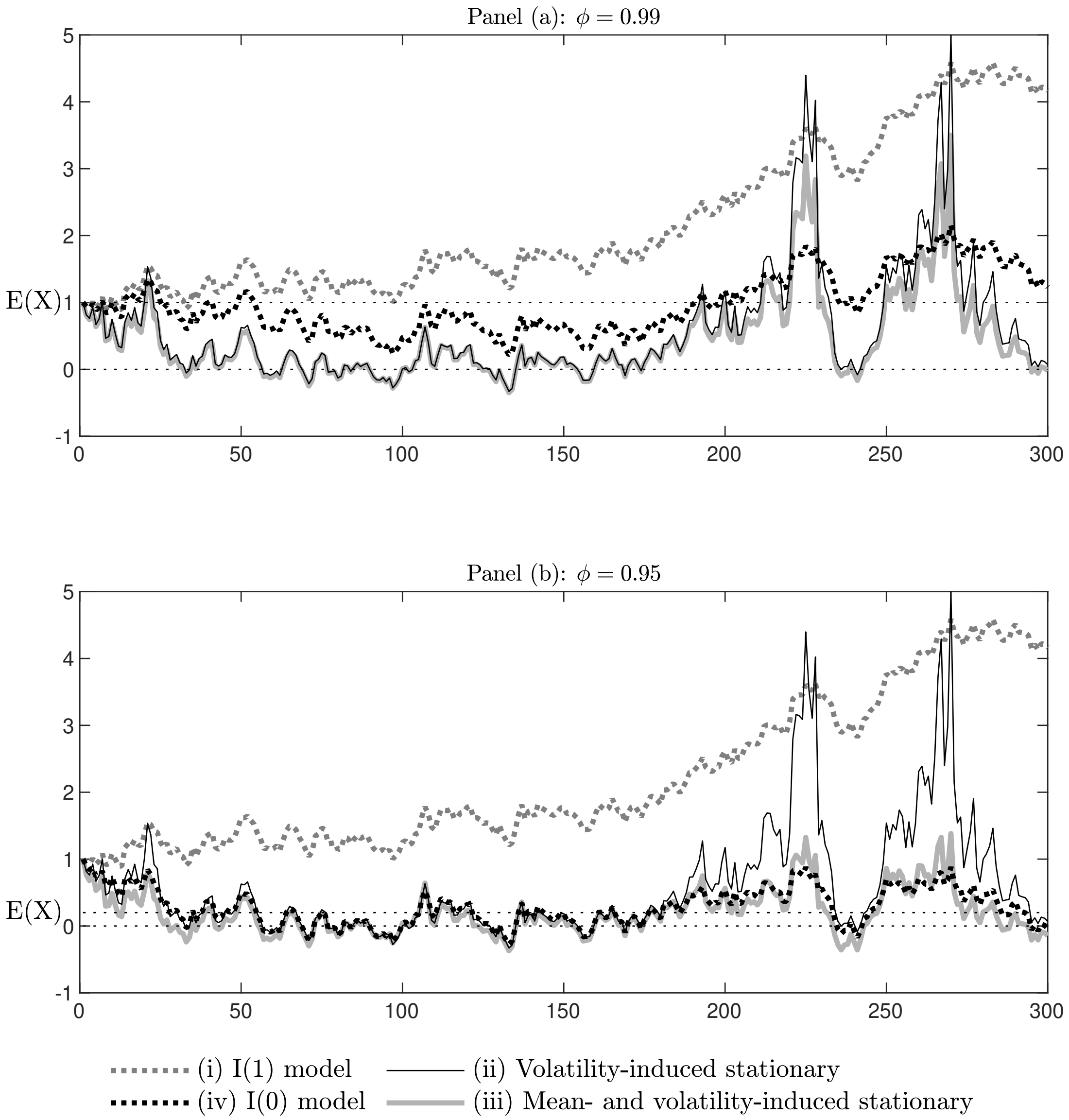
The process X_t in (2) is conditionally Gaussian given X_{t-1} with conditional mean and variance equal to

$$\mathbb{E}_{t-1}(X_t) = \mu + \Phi X_{t-1},$$
$$\mathrm{Var}_{t-1}(X_t) = \Omega_t = \Sigma_0\Sigma_0' + \Sigma_1 X_{t-1}X_{t-1}'\Sigma_1'.$$

Thus, the log-likelihood function is given up to a constant by

$$\mathcal{L}(\Theta^{\mathbb{P}}) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log|\Omega_t| + (X_t - \mu - \Phi X_{t-1})'\Omega_t^{-1}(X_t - \mu - \Phi X_{t-1})\right],$$

Figure 1: Simulated Path of Univariate DAR Models



Note: Simulated paths of sample length $T = 300$ generated by univariate DAR models with parameters $(\mu, \Phi, \Sigma_0, \Sigma_1)$ given by $(0.01, 1, 0.1, 0)$ in case (i), $(0.01, 1, 0.1, 0.3)$ in case (ii), $(0.01, \phi, 0.1, 0.3)$ in case (iii), and $(0.01, \phi, 0.1, 0)$ in case (iv) with $\phi = 0.99$ in Panel (a) and $\phi = 0.95$ in Panel (b).

where we note that Ω_t is a function of Σ_0 and Σ_1 given in (2) and the parameters are given by $\Theta^{\mathbb{P}} = \{\mu, \Phi, \Sigma_0, \Sigma_1\}$. Consistency and asymptotic normality of the maximum likelihood estimator has been established in the univariate case by Ling (2004); in a bivariate model under certain parameter restrictions by Nielsen and Rahbek (2014); and in the multivariate setting but with a diagonal conditional covariance matrix in Zhu et al. (2017). Since there are no results available for the general multivariate specification, we confirm by simulations that the maximum likelihood estimators exhibit reasonable properties, i.e., are approximately Gaussian and centered around their true values.[9]

3 Empirical Analysis

For the empirical analysis, we focus on the purely volatility-induced stationary model versus the special case when the conditional volatility is constant, i.e., the I(1) VAR model. We consider a macro-finance setting for two reasons. Fist, macroeconomic variables are important predictors of term premia (Joslin et al., 2014, Wright, 2011). Second, the volatility specification in (2) has an economic motivation involving inflation rates and unemployment as discussed in Section 2.1. In particular, we model the short and long ends of the yield curve (r_t, R_t) and two macroeconomic measures interpreted as respectively inflation (π_t) and real activity (g_t). Let X_t be the vector containing these variables, $X_t = [r_t, R_t, \pi_t, g_t]'$.

3.1 Data

We use monthly data between January 1985 and December 2016 measured end-of-month. The short rate is the one-month U.S. Treasury Bill rate from the Fama Treasury Bills Term Structure Files available at CRSP. We define the long rate by the ten-year U.S. Treasury bond yield from Gürkaynak et al. (2007).

The macroeconomic variables are constructed following the approach in Ang and Piazzesi (2003) and Goliński and Zaffaroni (2016). The inflation measure, π_t, is given by the first principal component of standardized series of CPI and PPI data from the U.S. Bureau of Labor Statistics. The measure of real activity, g_t, is the first principal component of standardized data on the unemployment and employment growth rates from the U.S. Bureau of Labor Statistics; the industrial production index from Federal

[9]The simulation results are available upon request.

Table 1: Interpretation of the Macroeconomic Factors

	Explained (pct)	Empirical correlation coefficients					
		CPI	PPI	UNEMP	EMP	PROD	HELP
π_t	85.49	0.92	0.92	-	-	-	-
g_t	65.75	-	-	-0.71	0.93	0.71	0.87

Note: The percentage of variation in respectively inflation data (CPI, PPI) and data related to real activity (unemployment rates, UNEMP, employment growth rates, EMP, the production index, PROD, and the HELP index, HELP) explained by respectively the inflation measure (π_t) and the real-activity measure (g_t). Empirical correlation coefficients are shown as well.

Reserve Economic Data; and the help-wanted-advertising-in-newspapers (HELP) index from Barnichon (2010).

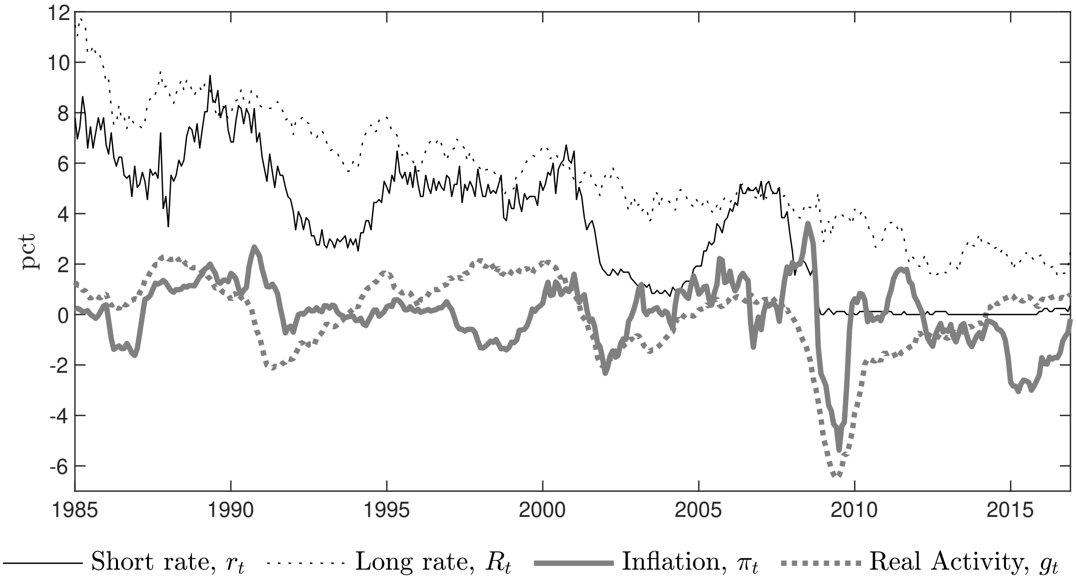
Table 1 details how π_t and g_t correlate with the underlying observed data as well as the fraction of total variation they capture. The inflation variable, π_t, is highly correlated with both inflation measurements and explains 85 pct of the total variation in these data. The variable measuring real activity, g_t, correlates strongest with employment growth rate and the HELP index. Correlation with the unemployment rate is negative as expected. Our measure of real activity captures 66 pct of the variation in the underlying observables.

The data for $X_t = [r_t, R_t, \pi_t, g_t]'$ are exhibited in Figure 2. The series appear extremely persistent and use of conventional unit-root and stationarity tests indeed identify unit roots, see Table 2. Therefore, modeling these data using a VAR model involves the implicit assumption that interest rates are generated by non-stationary processes, which is puzzling from both theoretical and empirical standpoints. To alleviate this problem, we propose the DAR model for these persistent and stationary data.

3.2 Model Specification and Estimation

To achieve a well-specified model, we allow for an extensive lag structure in the conditional mean. We find that this generalization is sufficient to match the data and thus we leave the conditional variance as specified in (2). The resulting generalized DAR model is given

Figure 2: Monthly Interest Rates and Macroeconomic Factors



Note: The interest rates, r_t and R_t, are the 1-month Treasury bill rate and the 10-year Treasury bond yield. The inflation measure, π_t, is the first principal component of CPI and PPI rates. Real activity, g_t, is measured the first principal component of the unemployment rate, the growth rate of employment, and the industrial production and HELP indices.

Table 2: Testing for Unit Roots

	Null hypothesis	r_t	R_t	π_t	g_t
ADF test	unit root	2.16	3.19	-16.98	-14.83
		[0.44]	[0.17]	[0.00]	[0.01]
KPSS test	stationarity	1.49	1.89	0.39	0.43
		[0.00]	[0.00]	[0.08]	[0.06]

Note: Augmented Dickey-Fuller and KPSS tests for respectively unit roots and stationarity. P-values in brackets.

by

$$\begin{aligned}
X_{t+1} &= \mu + \Phi X_t + \sum_{k=1}^{K} \Gamma_k \Delta X_{t-k} + \Omega_{t+1}^{1/2} \varepsilon_{t+1}, \\
\Omega_{t+1} &= \Sigma_0 \Sigma_0' + \Sigma_1 X_t X_t' \Sigma_1', \\
\varepsilon_{t+1} &\sim \text{i.i.d. } \mathcal{N}_p(0, \mathrm{I}_p)
\end{aligned} \tag{4}$$

for $K \geq 1$.[10] As we focus on the purely volatility-induced stationary case, we decompose $\Phi = \mathrm{I}_p + \alpha\beta'$, where α and β are of dimension $p \times r$ with the rank of Φ satisfying $0 \leq r \leq p$. The parameters of the model are $\Theta^{\mathbb{P}} = \{\mu, \alpha, \beta, \Gamma_1, \Gamma_2, \ldots, \Gamma_K, \Sigma_0, \Sigma_1\}$.

The data suggest a reduced rank of $r = 2$ and a lag length of $K = 3$, when specification testing is conducted with use of conventional methods for VAR models. With these choices, the DAR model appears to be well-specified. In fact, compared to the corresponding cointegrated VAR (CVAR) model that appears as the special case when $\Sigma_1 = 0_{p\times p}$, the DAR model removes autocorrelation and improves normality tests of the standardized residuals, see Table 3. The DAR model obtains the lowest AIC value and the likelihood values of the models are significantly different when compared by a LR test. Moreover, we note that the estimated top-Lyapunov exponent in the DAR is strictly negative.[11] Therefore, the process is indeed volatility-induced stationary. Further estimation details and parameter estimates are provided in Appendix A.

Long-Run Equilibria

The long-run equilibria in the DAR model are estimated, up to a constant, by

$$\begin{aligned} \hat{\beta}_1' X_t &\propto r_t - 3.672\pi_t - 1.681 g_t, \\ \hat{\beta}_2' X_t &\propto R_t - r_t, \end{aligned}$$

see Table 9 in Appendix. One relation is the spread between long and short rates as in Hall et al. (1992). The other is given by the short rate, inflation, and real activity. Since the short rate follows the federal funds rate closely, this relation mimics the dual mandate of the Federal Reserve (Fed). In addition, the signs of the estimates are intuitive: a low

[10] The associated top-Lyapunov exponent is given by

$$\gamma = \lim_{\xi\to\infty} \left[\mathbb{E}\left(\log \| \prod_{t=1}^{\xi} (\tilde{\Phi} + \tilde{e}_t) \| \right) \right],$$

where $\tilde{e}_t$ has dimension $p(K+1) \times p(K+1)$ and is i.i.d. normal with mean zero and covariance matrix equal to $\tilde{\Sigma}_1 \otimes \tilde{\Sigma}_1$. $\tilde{\Phi}$ and $\tilde{\Sigma}_1$ are defined by

$$\tilde{\Phi} = \begin{pmatrix} \Phi + \Gamma_1 & \Gamma_1 - \Gamma_2 & \ldots & \Gamma_K - \Gamma_{K-1} & -\Gamma_K \\ \mathrm{I}_p & 0_{p\times p} & \ldots & 0_{p\times p} & 0_{p\times p} \\ 0_{p\times p} & \mathrm{I}_p & & 0_{p\times p} & 0_{p\times p} \\ \vdots & & \ddots & & \vdots \\ 0_{p\times p} & 0_{p\times p} & \ldots & \mathrm{I}_p & 0_{p\times p} \end{pmatrix}, \quad \tilde{\Sigma}_1 = \begin{pmatrix} \Sigma_1 & 0_{p\times pK} \\ 0_{pK\times p} & 0_{pK\times pK} \end{pmatrix}.$$

[11] The Lyapunov exponents are obtained by the efficient and numerically stable algorithm described in Nielsen and Rahbek (2014).

Table 3: Misspecification Testing

	DAR				CVAR			
Log-likelihood	6986				6823			
AIC	-13806				-13507			
Top-Lyapunov	-0.004							
	r_t	R_t	π_t	g_t	r_t	R_t	π_t	g_t
Ljung-Box	7.71	1.67	0.61	4.96	3.72	1.48	1.05	5.75
	[0.10]	[0.80]	[0.96]	[0.29]	[0.44]	[0.83]	[0.90]	[0.22]
Engle's ARCH	23.81	4.37	2.53	1.58	36.15	9.92	20.82	0.60
	[0.00]	[0.04]	[0.11]	[0.21]	[0.00]	[0.00]	[0.00]	[0.44]
Kolmogorov-Smirnov	0.07	0.06	0.06	0.06	0.09	0.06	0.10	0.05
	[0.05]	[0.09]	[0.12]	[0.16]	[0.00]	[0.12]	[0.00]	[0.12]

Note: Log-likelihood values, Akaike information criteria (AIC), and top-Lyapunov exponent of the DAR model. Residual specification tests: Ljung-Box test of no autocorrelation. Engle's test of no ARCH effects. Kolmogorov-Smirnov test of standard normal distribution. P-values in brackets.

interest rate is associated with high levels of inflation and real activity. Also, we note that the estimated adjustment matrix places more weight on the yield spread relation compared to the dual mandate, see appendix. These results are practically identical for the CVAR model.

3.3 Conditional Volatilities

Model-implied conditional variances are exhibited in Figure 3. In the DAR model, conditional variances are highly time-varying and fluctuate around the constant levels estimated by the CVAR model. With exception of the short rate, which is subject to money market noise (Piazzesi, 2005) and institutional effects (Hilton, 2005), the factors exhibit countercyclical volatility. In particular, there are pronounced spikes at the outbreak of the financial crisis in 2007/08. Volatilities are small and nearly constant during the zero-lower bound regime in the aftermath of the crisis.

The DAR model allows all variables to exhibit volatility-induced stationarity and furthermore, the conditional heteroskedasticity can be driven by all variables. This general setting allows us to make statements about (i) which variables exhibit volatility-induced stationarity? and (ii) what variables drive this feature? From Figure 3, the short rate stands out with a highly volatile conditional variance that ranges from high and rapidly changing to low and stable. From the estimation results provided in Appendix A, the conditional variance of the short rate is given by

$$\mathrm{Var}_t\left(r_{t+1}\right) = \left(95.7r_t + 18.9R_t + 15.2\pi_t\right)^2,$$

where insignificant coefficients are suppressed. Thus, volatility-induced stationarity in the short rate is mainly driven by the short rate itself, but also by the long rate and inflation.

Figure 3 also shows the conditional correlations in the DAR model. We note that these are time-varying, which suggests that the flexibility offered by DAR models in terms of time-varying conditional correlations in contrast to $A_p(p)$ models is indeed necessary to fit the data. The short rate correlates positively with the long rate through the majority of the sample implying that the monetary transmission mechanism from short to long rates works in normal times. However, the correlation becomes negative following recessionary periods.

3.4 Term Premia

Next, we show that volatility-induced stationarity impacts model-implied term premia. Term premia are defined by an accounting identity that decomposes the bond yield, $Y_{t,n}$, into the yield that would prevail if investors were risk neutral, $\tilde{Y}_{t,n}$, and a residual, the term premium, $\mathrm{TP}_{t,n}$: $Y_{t,n} = \tilde{Y}_{t,n} + \mathrm{TP}_{t,n}$. By definition,

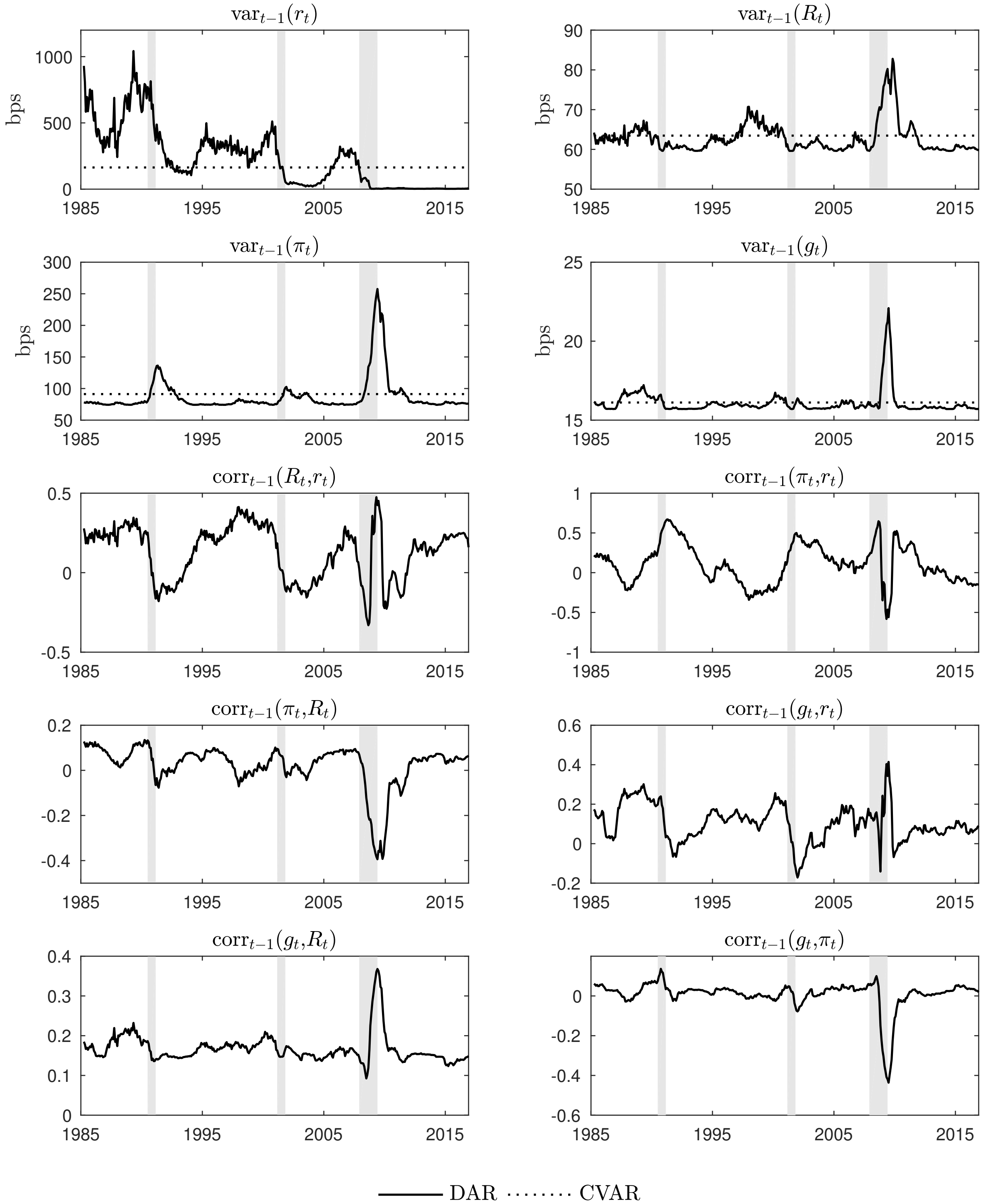
$$\tilde{Y}_{t,n} = -\frac{1}{n} \log \mathbb{E}_t\left(\exp\left[-\sum_{i=0}^{n-1} r_{t+i}\right]\right), \tag{5}$$

where $\mathbb{E}_t(\cdot)$ is the conditional expectation given the filtration at time t under physical probabilities and r_t is the short rate. Using observed yields for $Y_{t,n}$, the term premium follows by computing $\tilde{Y}_{t,n}$.[12]

Model-implied term premia with maturity of ten years, $n = 120$, are shown in Figure 4. Besides comparing the DAR and CVAR models, we also report term premia implied by the stationary VAR model. The models agree that term premia are countercyclical, which is consistent with the intuition on how risk premia behave. However, the DAR model implies stronger cyclicality than the VAR models.

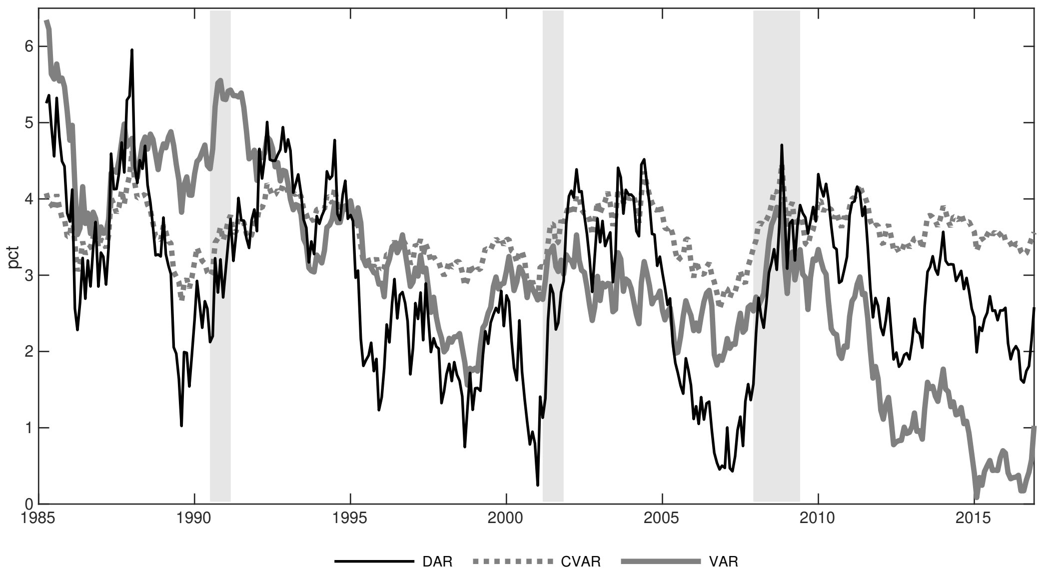
[12]The expectation in (5) can either be simulated or approximated by the method that will explained in Section 4.2, where the $\mathbb{Q}$-parameters are replaced by the corresponding parameters under the $\mathbb{P}$-measure. Here, we report term premia obtained by approximation.

Figure 3: Estimated Conditional Variances and Correlations



Note: Estimated conditional variances (in bps) in the DAR and CVAR models. Conditional correlations are reported for the DAR model. Shaded areas mark recessionary periods defined by NBER.

Figure 4: Ten-Year Term Premia



Note: Ten-year term premia implied by the DAR, CVAR, and stationary VAR models. The ten-year yield is plotted for reference. Shaded areas mark recessionary periods defined by NBER.

We make two observations regarding the VAR models. First, the stationary VAR model implies that term premia are downward-sloping corresponding to the down-trending long rate. In fact, the correlation between the ten-year term premium of the stationary VAR model and the ten-year yield is 0.91. This high correlation reflects the strong mean reversion of stationary VAR models that implies that $\tilde{Y}_{t,n}$ is nearly constant. In result, almost all of the variation in yields is attributed to term premia. This persistence problem has previously been recognized in the literature (Abbritti et al., 2016, Goliński and Zaffaroni, 2016, Jardet et al., 2013, Kozicki and Tinsley, 2001, Shiller, 1979). Second, we note that the CVAR model predicts an almost constant ten-year term premium in the range of 3-4 pct throughout the entire sample even as the ten-year yield falls toward 2 pct towards the end of the sample. It is, however, counterintuitive that the term premium is well above the yield itself for a long period of time, as it means that investors expect future short rates to become highly negative. The stable term premium in the CVAR model is a direct implication of the model result that the short rate adjusts to the yield spread as a long-run stable relation. In turn, $\tilde{Y}_{t,n}$ converges to the long rate such that no residual variation can be assigned to the term premium.

The flexibility of the DAR model allows for term premia that are more time-varying than those of the CVAR model, but not close to perfectly correlated with yields as in the

Table 4: Matching Survey Expectations

	DAR	CVAR	VAR
3M	**34.55**	37.34	39.15
6M	**44.92**	53.63	58.54
12M	**74.45**	91.98	106.09

Note: Root mean squared errors in bps of forecasts of the short rate in the DAR, CVAR, and stationary VAR models compared to the expected 3-month Treasury bill rate from the Survey of Professional Forecasters. Forecasting horizons are 3, 6, and 12 months. Lowest errors across models are boldfaced.

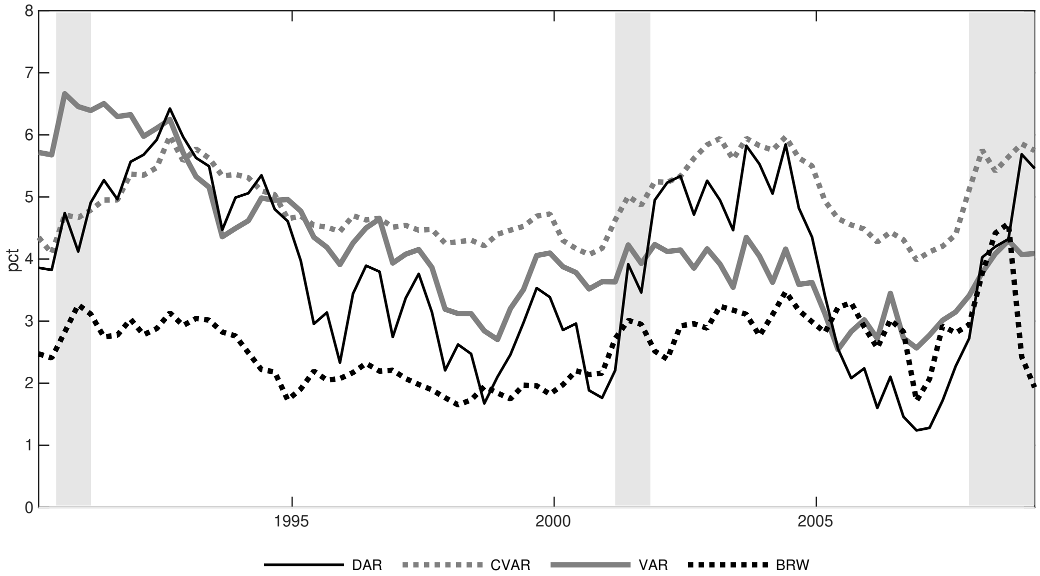
stationary VAR model. To numerically evaluate the DAR model's ability to decompose interest rates, we compare model-implied expectations of short rates with market expectations measured by survey forecasts. The survey data are from the Survey of Professional Forecasters (SPF) conducted by the Federal Reserve Bank of Philadelphia on a quarterly basis. We use median forecasts of the three-month Treasury Bill rate as a proxy for the short rate. We compare these data to expectations computed by respectively the DAR model and the cointegrated and stationary VAR models. Table 4 compares root mean squared errors between model-implied and survey expectations for forecasting horizons of 3, 6, and 12 months. The results unambiguously show that volatility-induced stationarity help matching market expectations. We interpret this result as an indication that the DAR model provides a more accurate term structure decomposition than the cointegrated and stationary VAR models.

Alternative Solutions to the Persistence Problem

Bauer, Rudebusch, and Wu (2012) (hereafter BRW) suggest that the persistence problem of the stationary VAR model can be resolved by correcting for the well-known downward bias in the autoregressive coefficient matrix. Figure 5 compares model-implied five-by-five year forward term premia of the DAR, CVAR, and stationary VAR models compared to those in BRW, which are available at the quarterly frequency from 1990:Q1 to 2009:Q1 from the AEA website associated with the paper.[13] Descriptive statistics of the forward term premia are given in Table 5.

[13]BRW consider a macro-finance term structure model where the factors are given by the first three principal components of the yield curve along with two unspanned macro risks constructed by smoothed inflation and GDP growth data.

Figure 5: Five-by-Five Year Forward Term Premia



Note: Five-by-five year forward term premia implied by the DAR model and the bias-corrected I(0) VAR model in Bauer et al. (2012) (BRW). Shaded areas mark recessionary periods defined by NBER. Data are in quarterly frequency.

For the considered sample and at the quarterly frequency, the BRW forward term premia are as stable as those of the CVAR both with an empirical standard deviation of 0.6. Therefore, we expect that the BRW model encounters the same problem as the CVAR model if extrapolated into the zero-lower bound regime. Moreover, the BRW forward term premia are negatively correlated with the forward rate. To the extent that higher levels of yields are associated with more volatility, we would expect the correlation to be positive as predicted by the DAR and VAR models.

The persistence problem is also considered in Jardet et al. (2013), who suggest an averaging estimator that combines the parameter estimates of the stationary VAR and CVAR models. We adopt their weighting scheme to combine our estimated VAR models, which give term premia as depicted in Figure 6.[14] The term premia estimated by the averaging model are in-between those of the stationary VAR and CVAR models and thus highly stable. In result, this model cannot produce term premia that are either below or above the estimates of the VAR models. The averaging model will therefore differ from the DAR model in most of the sample per construction. An exception is during the

[14] Jardet et al. (2013) chooses a weighting scheme such that the forecasting error of the future path of short rates is minimized. In result, the stationary VAR estimates are weighted by 0.2617 which implies a weight on the CVAR model equal to 0.7383.

Table 5: Empirical Standard Deviations and Correlations of Five-by-Five Year Forward Term Premia

		DAR	CVAR	VAR	BRW
Standard deviation		1.39	0.59	1.06	0.60
Correlation with forward rate		0.50	0.09	0.89	-0.09
Correlation matrix:					
	DAR	1			
	CVAR	0.85	1		
	VAR	0.70	0.33	1	
	BRW	0.38	0.60	0.20	1

Note: Empirical standard deviations of five-by-five-year forward term premia implied by the DAR, CVAR, and VAR models and the bias-corrected I(0) VAR model in Bauer et al. (2012) (BRW). Correlations with the forward rate as well as correlations between the models are reported as well. All data are in quarterly frequency from 1990:Q1 to 2009:Q1.

zero-lower bound regime, where the term premia of the DAR model and averaging model coincide.

4 Volatility-Induced Stationary Term Structure Modeling

This section casts the DAR model analyzed thus far into a macro-finance term structure model. We consider a four-factor term structure model with the observable state vector $X_t = [r_t, R_t, \pi_t, g_t]'$ whose dynamics is given by (4).

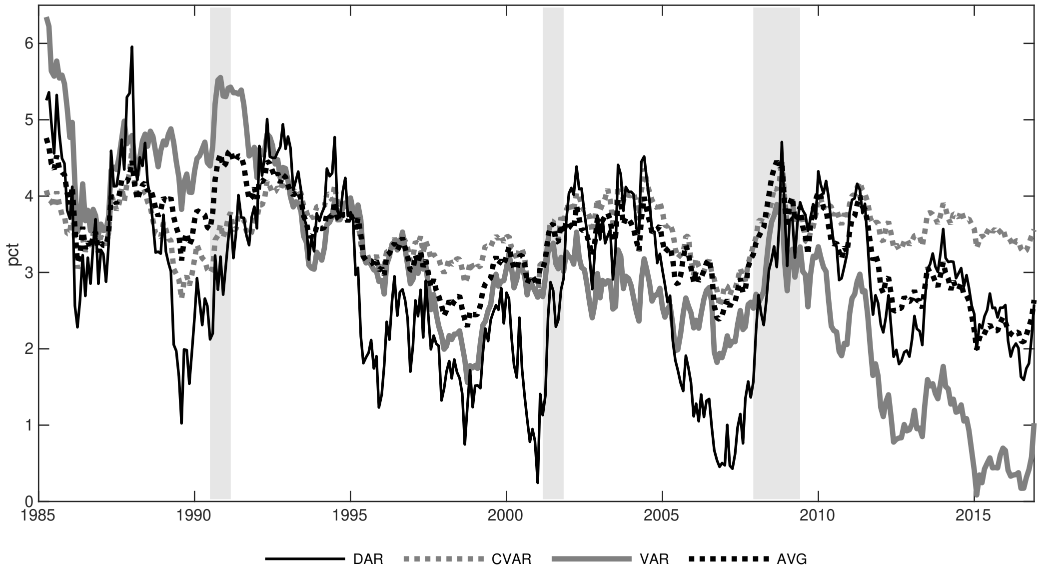
4.1 Stochastic Discount Factor and $\mathbb{Q}$-Dynamics

We adopt the standard linear-exponential stochastic discount factor given by

$$\mathcal{M}_{t+1} = \exp\left(-r_t - \frac{1}{2}\Lambda_t'\Omega_{t+1}\Lambda_t - \Lambda_t'\Omega_{t+1}^{1/2}\varepsilon_{t+1}\right), \tag{6}$$

where Λ_t is the market price of risk with risk measured by the conditional variance, Ω_{t+1}. We specify the market price of risk such that the factor dynamics under the risk-neutral $\mathbb{Q}$-measure follows a DAR model. Moreover, to reduce the number of parameters we treat lagged variables as unspanned factors as in Joslin et al. (2013). Thus, the lag structure determines the dynamics under the real-world measure but is not priced in the term

Figure 6: Ten-Year Term Premia



Note: Ten-year term premia implied by the DAR, CVAR, and stationary VAR models along with the model averaging estimator (AVG) that combines the parameters of the VAR and CVAR models with weights equal to respectively 0.2617 and 0.7383. The ten-year yield is plotted for reference. Shaded areas mark recessionary periods defined by NBER.

structure cross-section. The market price of risk is defined by:

$$\Lambda_t = \Omega_{t+1}^{-1} \left[(\mu - \mu^{\mathbb{Q}}) + \left(\Phi - \Phi^{\mathbb{Q}} \right) X_t + \sum_{k=1}^{K} \Gamma_k \Delta X_{t-k} \right], \tag{7}$$

with risk-neutral $\mathbb{Q}$-dynamics given by the following DAR model:

$$\begin{aligned} X_{t+1} &= \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t + \Omega_{t+1}^{1/2} \varepsilon_{t+1}^{\mathbb{Q}}, \\ \Omega_{t+1} &= \Sigma_0 \Sigma_0' + \Sigma_1 X_t X_t' \Sigma_1', \\ \varepsilon_{t+1}^{\mathbb{Q}} &\sim \text{i.i.d. } \mathcal{N}_p(0, \mathrm{I}_p). \end{aligned} \tag{8}$$

We note that the market price of risk is time-t measurable as Ω_{t+1} depends on X_t. Finally, per construction of the state vector, the short rate and the state vector are related by $r_t = \iota_1' X_t$ where ι_1' is a unit vector with one in the first entry. With these assumptions, the special case where $\Sigma_1 = 0_{p \times p}$ corresponds to the GATSM based on the CVAR model rather than the stationary VAR model that is standard in GATSMs. In the following, we use the acronym GATSM to describe the model based on the CVAR specification that is nested in our model.

4.2 Bond Pricing

The no-arbitrage price of a zero-coupon bond with $n+1$ periods to maturity is given by

$$P_{t,n+1} = \mathbb{E}_t\left(\mathcal{M}_{t+1}P_{t+1,n}\right),$$

where $\mathbb{E}_t(\cdot)$ denotes the conditional expectation given $\mathcal{F}_t = \{X_t, X_{t-1}, \ldots, X_1\}$ under real-world probabilities. Our model does not admit a closed-form bond price expression that satisfies this equation. Instead, we propose an exponential-quadratic approximation that allows the conditional covariance matrix to affect bond yields. This is similar to the GATSM in which the closed-form solution depends on the constant conditional variance, see Ang and Piazzesi (2003). Also, we make sure that for $\Sigma_1 = 0_{p\times p}$, bond yield computation must coincide with the solution of the GATSM. Appendix B shows that such an approximation can be obtained by controlling the dynamics of the conditional variance under the $\mathbb{Q}$-measure. The resulting approximation is given by

$$P_{t,n} = \exp\left(A_n + B_n' X_t + C_n' \mathrm{vec}\left(X_t X_t'\right)\right), \tag{9}$$

where

$$\begin{aligned}
A_n &= A_{n-1} + B_{n-1}'\mu^{\mathbb{Q}} + C_{n-1}'\left(\mathrm{vec}\left(\mu^{\mathbb{Q}}\mu^{\mathbb{Q}'}\right) + \mathrm{vec}\left(\Sigma_0\Sigma_0'\right)\right) + \frac{1}{2}B_{n-1}'\Sigma_0\Sigma_0' B_{n-1} \\
B_n' &= -\iota_1 + B_{n-1}'\Phi^{\mathbb{Q}} + C_{n-1}'\left(\Phi^{\mathbb{Q}}\otimes\mu^{\mathbb{Q}} + \mu^{\mathbb{Q}}\otimes\Phi^{\mathbb{Q}}\right) \\
C_n' &= C_{n-1}'\left(\Phi^{\mathbb{Q}}\otimes\Phi^{\mathbb{Q}} + \Sigma_1\otimes\Sigma_1\right) + \frac{1}{2}\left(\left[B_{n-1}'\Sigma_1\right]\otimes\left[B_{n-1}'\Sigma_1\right]\right)
\end{aligned}$$

initiated at $n=0$ with $A_0 = 0$, $B_0 = 0_{p\times 1}$, $C_0 = 0_{p^2\times 1}$.[15] The associated approximated bond yield is given by

$$Y_{t,n} = -\frac{1}{n}\log\left(P_{t,n}\right) = -\frac{1}{n}A_n - \frac{1}{n}B_n' X_t - \frac{1}{n}C_n'\mathrm{vec}\left(X_t X_t'\right). \tag{10}$$

The approximated bond yield expression is similar to the solution of the class of quadratic term structure models (QTSMs) studied in Leippold and Wu (2002), Ahn et al. (2002), and Realdon (2006). Thus, the DAR term structure model and the QTSM can produce similar shapes of the yield curve. However, the source of the quadratic term and thus the loading recursions are highly different across the two model frameworks:

[15] To evaluate the approximation error, we proxy the exact solution by averaging 10,000,000 paths of $\exp\left(-\sum_{i=0}^{n-1}\iota_1'\hat{X}_t\right)$, where $\hat{X}_t$ is simulated under the $\mathbb{Q}$-measure according to (8). Then, this simulated bond price is converted to yields. We repeat this procedure for all months of January in the sample using parameter values reported in Table 6. The approximation error is largest for the ten-year yield, for which the average absolute error is 38 bps corresponding to 6.94 pct of the average ten-year yield level.

Table 6: Estimated $\mathbb{Q}$-Dynamics

	DAR term structure model				GATSM			
$(\mu^{\mathbb{Q}})' \times 100$	-0.007 (0.001)	0.001 (0.000)	-0.046 (0.003)	-0.018 (0.003)	-0.010 (0.006)	0.001 (0.000)	-0.029 (0.034)	-0.002 (0.026)
$\Phi^{\mathbb{Q}}$	0.758 (0.018)	0.242 (0.000)	0.093 (0.090)	0.225 (0.078)	0.905 (0.001)	0.111 (0.000)	0.013 (0.003)	0.076 (0.002)
	-0.008 (0.016)	1.007 (0.000)	-0.001 (0.082)	-0.003 (0.070)	-0.010 (0.000)	1.009 (0.000)	0.001 (0.003)	-0.002 (0.002)
	0.886 (0.010)	-0.756 (0.000)	0.477 (0.049)	-0.887 (0.043)	0.156 (0.000)	-0.104 (0.000)	0.879 (0.002)	-0.145 (0.001)
	0.807 (0.019)	-0.738 (0.000)	-0.438 (0.093)	0.138 (0.080)	0.168 (0.001)	-0.168 (0.000)	-0.089 (0.003)	0.783 (0.002)

Note: Estimated parameters related to the $\mathbb{Q}$-dynamics, $\mu^{\mathbb{Q}}$ and $\Phi^{\mathbb{Q}}$, in the DAR term structure model and the GATSM. The parameters are estimated given $\hat{\Theta}^{\mathbb{P}} = \{\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\Sigma}_0, \hat{\Sigma}_1, \hat{\Gamma}_1, \hat{\Gamma}_2, \hat{\Gamma}_3\}$ from Section 3. Standard errors in paranthesis.

whereas the quadratic bond yield in our model stems from the variance specification in the DAR model, the QTSM imposes this non-linearity through a quadratic specification of the short rate. This difference is particularly highlighted in the macro-finance model considered in this paper in which the short rate is a factor itself. In this setting, the short-rate specification is linear per construction and the QTSM reduces in this case to the GATSM.

4.3 Estimation

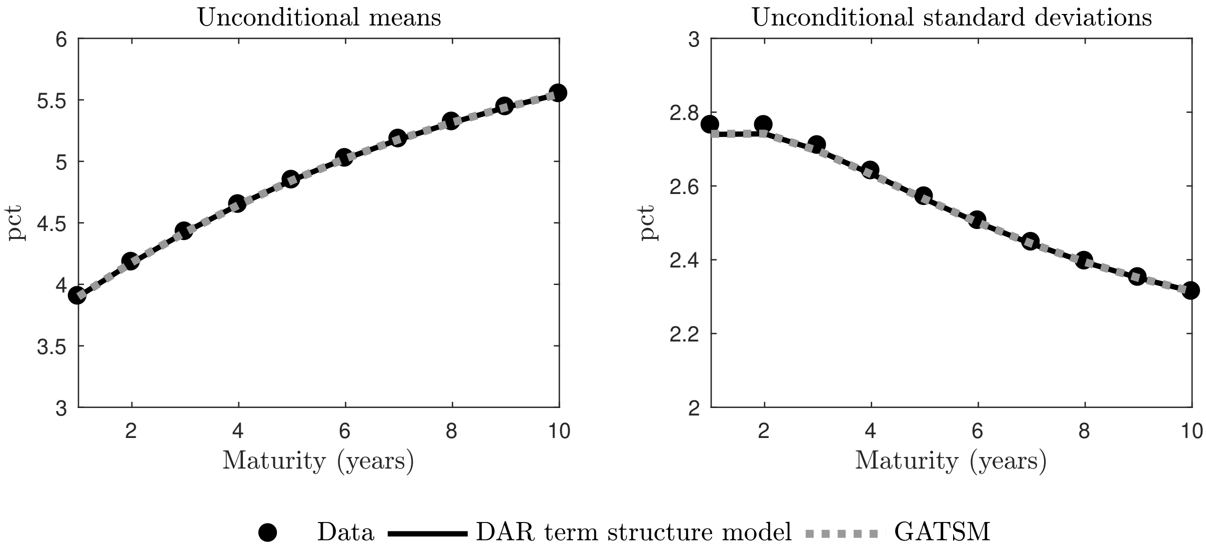
We estimate $\mu^{\mathbb{Q}}$ and $\Phi^{\mathbb{Q}}$ by non-linear least squares given the parameters obtained for the factor dynamics in Section 3, $\hat{\Theta}^{\mathbb{P}} = \{\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\Sigma}_0, \hat{\Sigma}_1, \hat{\Gamma}_1, \hat{\Gamma}_2, \hat{\Gamma}_3\}$.[16] The data are U.S. Treasury bonds yields from Gürkaynak et al. (2007) with maturities $n = 1, 2, \ldots, 10$ years. The results are shown in Table 6. The DAR model remains stationary under the $\mathbb{Q}$-measure with a top-Lyapunov exponent of -0.017.

4.4 In-Sample Fit

Figure 7 shows how the DAR term structure model matches the unconditional first and second empirical moments of the yield curve cross-section compared to the GATSM. We observe that the models fit the yield curve and moreover, that the models are practi-

[16] This two-step estimation method is a common approach in the macro-finance term structure literature, see for instance Ang and Piazzesi (2003), Ang et al. (2006), and Wright (2011).

Figure 7: In-Sample Yield Curve Fit



Note: Empirical unconditional mean and standard deviation of observed and model-implied yields in the DAR term structure model and GATSM.

cally identical in this respect. To explain this similarity, we compare estimated factor loadings A_n and B_n in Figure 8. Indeed, the DAR term structure model does not imply different loadings from those obtained by the GATSM. In Figure 9, we plot the component that prices volatility-induced stationarity, $-n^{-1}C_n' \text{vec}\,(X_t X_t')$ from (10). This component accounts for a very small part of yields across all maturities, which confirms that the quadratic component generated by volatility-induced stationarity is not priced by the yield curve. Since the time-varying conditional variance of the DAR model only affects bond yields through a convexity effect, this finding is consistent with Joslin and Konchitchki (2018) who show that convexity effects under the $\mathbb{Q}$-measure are small.

4.5 Out-of-Sample Performance

We assess the out-of-sample performance through a rolling-window forecasting exercise. In particular, we estimate the models with one lag in the factor dynamics as in (2), for the sample from January 1985 to December 2005 ($T = 252$). Using these estimated models, the yield curve is forecasted 3, 6, and 12 months ahead. We repeat this procedure by re-estimating the models based on a rolling-window sample of length $T = 252$ from January 2006 to December 2015. This period contains events that are difficult to forecast including the financial crisis of 2007/08 and the zero-lower bound regime. Root mean squared

Figure 8: Factor Loadings

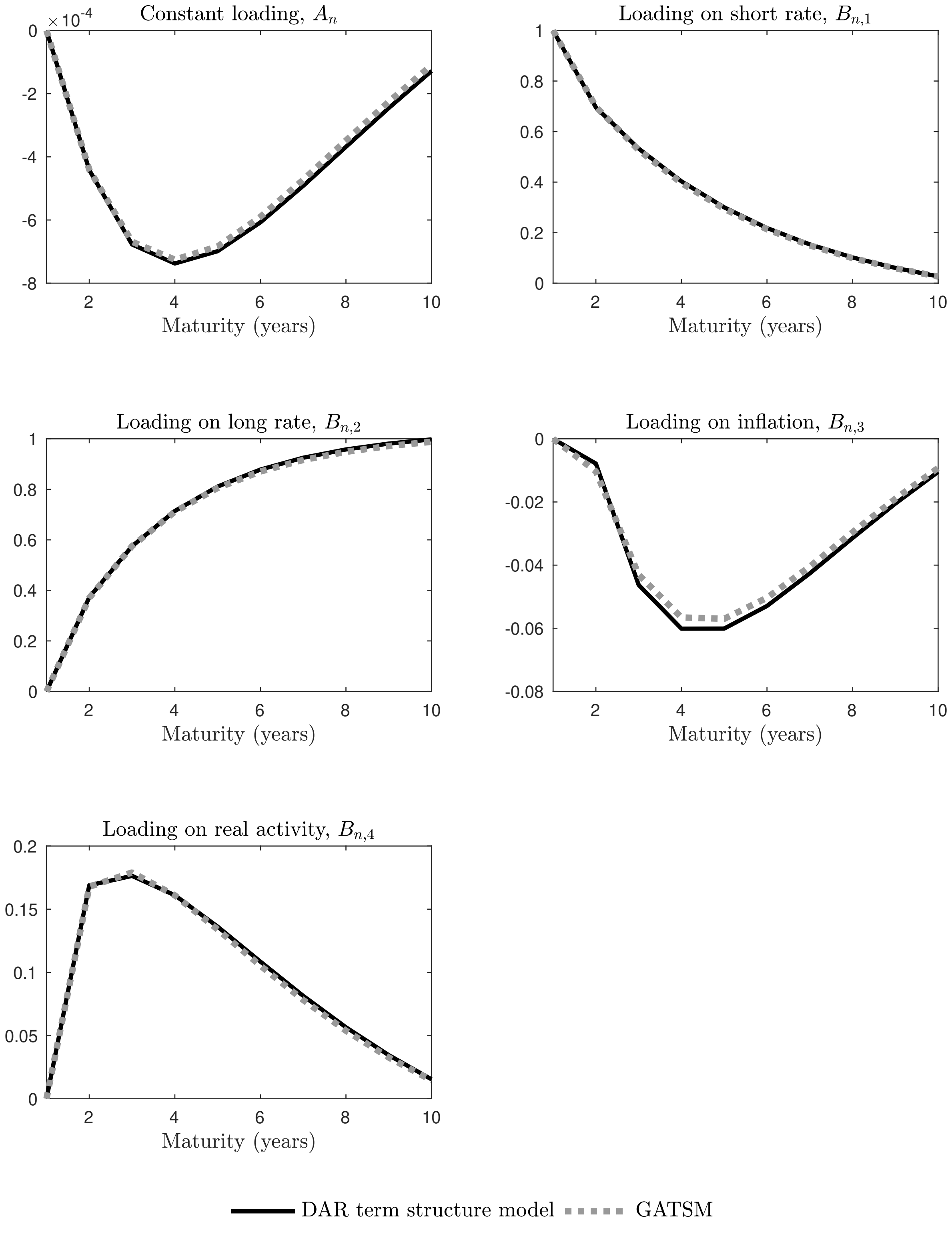
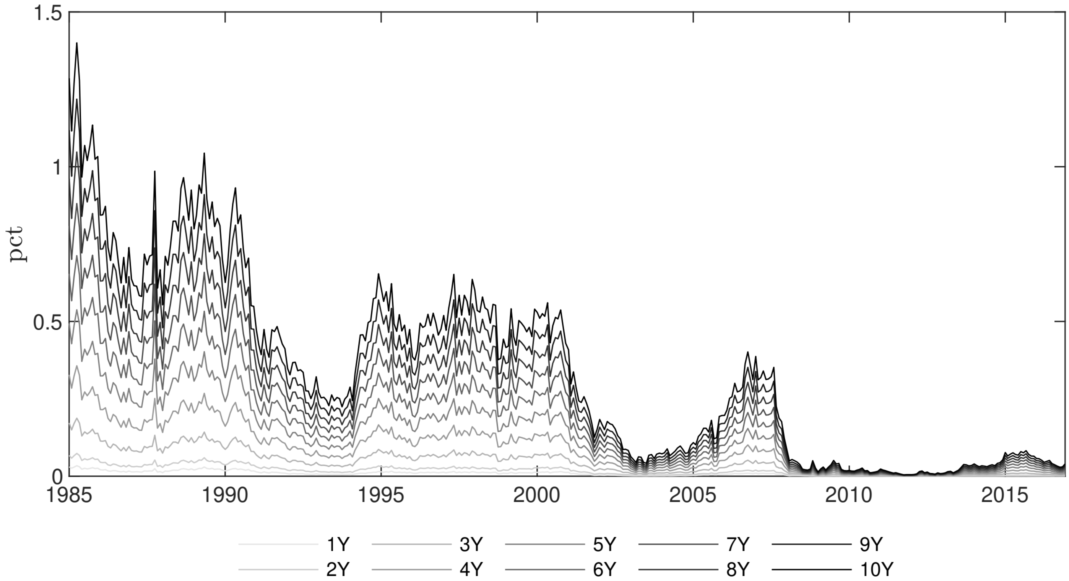


Figure 9: Quadratic Component of Model-Implied Bond Yield



Note: The component $-n^{-1}C_n' \text{vec}\,(X_t X_t')$ from (10). Reported in pct.

errors from this exercise are presented in Table 7 along with random walk forecasts. The DAR term structure model outperforms both the GATSM and the random walk almost uniformly across all maturities. The exceptions are forecasts of the 10-year yield and the 12-month ahead forecast of the 9-year yield. The differences between the models' forecasting performance are larger for shorter maturities reflecting that volatility-induced stationarity is generated by the short end of the yield curve. Thus, volatility-induced stationarity clearly improves out-of-sample forecasting of the yield curve.

5 Volatility Specifications in the Term Structure Literature

We conclude the paper by comparing the volatility specification of the DAR model in (2) with models that are common in the discrete-time term structure literature. The standard specification is the Poisson-Gamma mixture from Le et al. (2010), which underpins the discrete-time equivalents of the $A_m(p)$ models in Dai and Singleton (2000). We also consider a related class of models given by Wishart autoregressive models.

Table 7: Out-of-Sample Performance

	DAR model			GATSM			Random Walk		
	3M	6M	12M	3M	6M	12M	3M	6M	12M
Average	**58.06**	**71.77**	**87.42**	69.07	88.21	102.74	63.37	79.48	125.61
1Y	65.90	**85.18**	**113.43**	99.57	131.34	168.83	**65.53**	102.71	175.08
2Y	**63.23**	**78.86**	**99.29**	86.35	111.69	136.38	64.60	91.40	156.03
3Y	**61.77**	**75.03**	**90.18**	77.96	99.04	115.84	64.20	83.29	140.76
4Y	**60.24**	**72.83**	**85.93**	72.16	90.75	103.14	63.82	78.08	129.29
5Y	**58.33**	**70.88**	**83.53**	67.52	84.74	94.88	63.45	75.12	120.86
6Y	**56.33**	**69.02**	**81.94**	63.50	79.94	89.03	63.12	73.64	114.60
7Y	**54.55**	**67.39**	**80.76**	59.96	75.87	84.56	62.80	73.00	109.85
8Y	**53.24**	**66.16**	**79.95**	56.90	72.36	80.96	62.47	72.73	106.11
9Y	**52.64**	**65.56**	79.46	54.35	69.37	**78.05**	62.07	72.54	103.05
10Y	54.36	**66.78**	79.72	**52.42**	66.95	**75.77**	61.60	72.28	100.44

Note: Root mean squared errors from forecasting the yield curve using the DAR term structure model and the GATSM. The models are estimated on a rolling window starting with the sample from January 1985 to December 2005. Forecasts by the random walk are reported for reference. The minimum value obtained for each forecast horizon and maturity is boldfaced. Reported in bps.

Poisson-Gamma mixtures: discrete-time $A_m(p)$ models

Starting with the seminal work of Duffie and Kan (1996), a substantial literature[17] has considered term structure models with conditional variance of the form

$$\Omega_{t+1} = V \begin{pmatrix} \alpha_1 + \beta_1' Z_t & & 0 \\ & \ddots & \\ 0 & & \alpha_p + \beta_p' Z_t \end{pmatrix} V', \tag{11}$$

where Z_t is a vector of length m with $0 \leq m \leq p$. The conditional density of $Z_{i,t}$, $i = 1, \ldots, m$, given Z_{t-1} is defined as a Poisson mixture of standard gamma distributions, which ensures that $Z_{i,t} > 0$. The models allow m variables to drive the time-varying conditional variance of all factors, which gives rise to the notation $A_m(p)$ for p-factor models.

[17] In the discrete-time setting see Darolles et al. (2006), Gourieroux and Jasiak (2006), Gourieroux et al. (2002), Le et al. (2010). Continuous-time models are detailed in Dai and Singleton (2000) which contains further references.

We note that the DAR model can also restrict the number of factors that determine the volatility of all factors by adding structure to the matrix Σ_1.

The Poisson-Gamma mixture model requires the volatility factors to be strictly positive processes. Le et al. (2010) argue that this assumption is consistent with models of habit formation as in Campbell and Cochrane (1999) and Wachter (2006) and the long-run risk model in Bansal and Yaron (2004). In the former framework, surplus consumption is argued to be central for the dynamics of risk premia, while Z_t is given by a standard deviation in the latter. However, more recent literature shows that variations in inflation and real activity predict risk premia (Joslin et al., 2014). As these macro risks can be both positive and negative, the $A_m(p)$ models are not well-suited for modeling this feature of the data. In contrast, our volatility specification accommodates volatility factors that are conditionally Gaussian, hence can take values on the entire real axis.

The structure in (11) implies that conditional volatilities are diagonal up to a time-invariant linear transformation. Gourieroux and Sufana (2011) show that this assumption is not necessary and limits the generality of the $A_m(p)$ models. Dai and Singleton (2000) recognize a trade-off in the $A_m(p)$ models between the structure of correlations and the number of volatility factors m. At the extreme, $A_p(p)$ models do not allow for conditionally correlated factors and restrict their unconditional correlation to be non-negative. The additional flexibility of our DAR model relieves this tension as we allow factors to be conditionally correlated regardless of the number of volatility factors. Furthermore, given that unconditional moments exist, i.e., when the DAR model belongs to case (iii), the unconditional correlation between factors can be both positive and negative.

Wishart autoregressive processes

Gourieroux et al. (2009) and Gourieroux and Sufana (2011) propose a volatility specification based on the Wishart autoregressive (WAR) process:

$$\begin{aligned}\Omega_{t+1} &= \sum_{k=1}^{K} z_{k,t} z_{k,t}', \\ z_{k,t} &= \Psi z_{k,t-1} + \Sigma_z \epsilon_{k,t}, \quad \epsilon_{k,t} \sim \mathcal{N}(0, \mathrm{I}_p).\end{aligned}$$

Since the conditional volatility is a sum of quadratic forms, the specification ensures symmetry and positive definiteness. Our volatility specification resembles a WAR process with $K = 1$. There are, however, two important distinctions: (i) we allow the conditional volatility to exhibit a time-invariant level equal to $\Sigma_0 \Sigma_0'$, and (ii) the outer product in our specification is given by conditional Gaussian variables weighted by Σ_1. These features facilitate volatility-induced stationarity.[18].

6 Conclusion

This paper presented a novel class of macro-finance term structure models based on the double-autoregressive model. The dynamic model is consistent with key stylized facts of interest rate data that the VAR framework fails to accommodate. A defining feature of our model is that it exhibits volatility-induced stationarity implying that the conditional variance of the model can maintain stationarity even in the presence of unit roots in the conditional mean. We showed that this property is important for decomposing the term structure into expected future short rates and term premia. We embedded the DAR model into a no-arbitrage term structure model and provided an approximation for computing model-implied bond yields analytically. Volatility-induced stationarity helps forecasting bond yields. However, compared to the GATSM based on a VAR model, there are no in-sample improvement of the DAR model. This can be interpreted as evidence that volatility-induced stationarity is unspanned by the yield curve. Thus, our findings are consistent with the notion of unspanned stochastic volatility. Future work may focus on volatility-induced term structure models in which the conditional volatility is unspanned by construction.

[18]Stationarity of the WAR process requires that the eigenvalues of Ψ are inside the unit circle (Gourieroux et al., 2009)

References

Abbritti, M., Gil-Alana, L.A., Lovcha, Y., Moreno, A., 2016. Term Structure Persistence. Journal of Financial Econometrics 14, 331–352.

Ahn, D.H., Dittmar, R.F., Gallant, A.R., 2002. Quadratic term structure models: Theory and evidence. The Review of Financial Studies 15, 243–288.

Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50, 745–787.

Ang, A., Piazzesi, M., Wei, M., 2006. What Does the Yield Curve Tell us about GDP Growth? Journal of Econometrics 131, 359–403.

Ball, L., 1992. Why does high inflation raise inflation uncertainty? Journal of Monetary Economics 29, 371–388.

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

Barnichon, R., 2010. Building a composite Help-Wanted Index. Economics Letters 109, 175–178.

Bauer, M.D., Rudebusch, G.D., Wu, J.C., 2012. Correcting Estimation Bias in Dynamic Term Structure Models. Journal of Business and Economic Statistics 30, 454–467.

Bauer, M.D., Rudebusch, G.D., Wu, J.C., 2014. Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset: Comment. American Economic Reivew 101, 1514–1534.

Beechey, M., Hjalmarsson, E., Österholm, P., 2009. Testing the expectations hypothesis when interest rates are near integrated. Journal of Banking and Finance 33, 934–943.

Campbell, J.Y., Cochrane, J.H., 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior.

Campbell, J.Y., Shiller, R.J., 1987. Cointegration and Tests of Present Value Models. Journal of Political Economy 95, 1062–1088.

Cavaliere, G., Rahbek, A., Taylor, A.M.R., 2014. Bootstrap Determination of the Co-Integration Rank in Heteroskedastic VAR Models. Econometric Reviews 33, 606–650.

Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B., 1992. An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. Journal of Finance XLVII.

Chang, K.L., 2012. The impacts of regime-switching structures and fat-tailed characteristics on the relationship between inflation and inflation uncertainty. Journal of Macroeconomics 34, 523–536.

Chernov, M., Creal, D., 2019. The PPP view of multihorizon currency risk premiums. Working Paper .

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

Collin-Dufresne, P., Goldstein, R.S., Jones, C.S., 2009. Can interest rate volatility be extracted from the cross-section of bond yields? Journal of Financial Economics 94, 47–66.

Conley, T.G., Hansen, L.P., Luttmer, E.G.J., Scheinkman, J.A., 1997. Short-Term Interest Rates as Subordinated Diffusions. The Review of Financial Studies 10, 525–577.

Creal, D.D., Wu, J.C., 2015. Estimation of Affine Term Structure Models with Spanned or Unspanned Stochastic Volatility. Journal of Econometrics 185, 60–81.

Dai, Q., Singleton, K.J., 2000. Specification Analysis of Affine Term Structure Models. Journal of Finance 55, 1943–1978.

Darolles, S., Gourieroux, C., Jasiak, J., 2006. Structural Laplace transform and compound autoregressive models. Journal of Time Series Analysis 27, 477–503.

Diebold, F.X., Piazzesi, M., Rudebusch, G.D., 2005. Modeling Bond Yields in Finance and Macroeconomics. American Economic Review: Papers and Proceedings 95, 415–420.

Duffee, G.R., 2002. Term Premia and Interest Rate Forecasts in Affine Models. Journal of Finance 57.

Duffee, G.R., 2006. Term structure estimation without using latent factors. Journal of Financial Economics 79, 507–536.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical Finance 6, 379–406.

Engle, R.F., Kroner, K.F., 1995. Multivariate Simultaneous Generalized ARCH. Econometric Theory 11, 122–150.

Fischer, S., Modigliani, F., 1978. Towards an understanding of the real effects and costs of inflation. Weltwirtschaftliches Archiv 114, 810–833.

Fountas, S., 2010. Inflation, inflation uncertainty and growth: Are they related? Economic Modelling 27, 896–899.

Friedman, M., 1977. Nobel Lecture: Inflation and Unemployment. Journal of Political Economy 85, 451–472.

Goliński, A., Zaffaroni, P., 2016. Long memory affine term structure models. Journal of Econometrics 191, 33–56.

Golob, J.E., 1994. Does Inflation Uncertainty Increase with Inflation? Q-III.

Gourieroux, C., Jasiak, J., 2006. Autoregressive Gamma Processes. Journal of Forecasting 25, 129–152.

Gourieroux, C., Jasiak, J., Sufana, R., 2009. The Wishart autoregressive process of multivariate stochastic volatility. Journal of Econometrics 150, 167–181.

Gourieroux, C., Monfort, A., Polimenis, V., 2002. Affine term structure models. Centre de Recherche en Economie et Statistique, Working Papers .

Gourieroux, C., Sufana, R., 2011. Discrete time Wishart term structure models. Journal of Economic Dynamics and Control 35, 815–824.

Gürkaynak, R.S., Sack, B., Wright, J.H., 2007. The U.S. Treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304.

Hall, A., Anderson, H., Granger, C., 1992. A Cointegration Analyisis of Treasury Bill Yields. The Review of Economics and Statistics 74, 116–126.

Hartmann, M., Herwartz, H., 2012. Causal relations between inflation and inflation uncertainty—Cross sectional evidence in favour of the Friedman–Ball hypothesis. Economic Letters 115, 144–147.

Hayford, M.D., 2000. Inflation Uncertainty, Unemployment Uncertainty and Economic Activity. Journal of Macroeconomics 22, 315–329.

Hilton, S., 2005. Trends in Federal Funds Rate Volatility. Fed NY, Current Issues in Economics and Finance 11.

Hördahl, P., Tristani, O., Vestin, D., 2006. A joint econometric model of macroeconomic and term-structure dynamics. Journal of Econometrics 131, 405–444.

Jardet, C., Monfort, A., Pegoraro, F., 2013. No-arbitrage Near-Cointegrated VAR(p) term structure models, term premia and GDP growth. Journal of Banking & Finance 37, 389–402.

Johansen, S., 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. Econometrica 59, 1551–1580.

Johansen, S., 1995. Likelihood-based Inference in Cointegated Vector Autoregressive Models. Oxford University Press, New York.

Johansen, S., 1996. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press.

Joslin, S., 2017. Can Unspanned Stochastic Volatility Models Explain the Cross Section of Bond Volatilities? Management Science .

Joslin, S., Konchitchki, Y., 2018. Interest rate volatility, the yield curve, and the macroeconomy. Journal of Financial Economics 128, 344–362.

Joslin, S., Le, A., Singleton, K.J., 2013. Gaussian Macro-Finance Term Structure Models with Lags. Journal of Financial Econometrics 11, 581–609.

Joslin, S., Priebsch, M., Singleton, K.J., 2014. Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks. Journal of Finance 69, 1197–1233.

Kim, D.H., Lin, S.C., 2012. Inflation and Inflation Volatility Revisited. International Finance 15, 327–345.

Kim, D.H., Orphanides, A., 2007. The Bond Market Term Premium: What is It, and How Can we Measure It? BIS Quarterly Review , 27–40.

Kim, D.H., Orphanides, A., 2012. Term Structure Estimation with Survey Data on Interest Rate Forecasts. Journal of Financial and Quantitative Analysis 47, 241–272.

Kozicki, S., Tinsley, P.A., 2001. Shifting endpoints in the term structure of interest rates. Journal of Monetary Economics 47, 613–652.

Le, A., Singleton, K.J., Dai, Q., 2010. Discrete-time affine term structure models with generalized market prices of risk. Review of Financial Studies 23, 2184–2227.

Leippold, M., Wu, L., 2002. Class Asset Pricing under the Quadratic. The Journal of Financial and Quantitative Analysis 37, 271–295.

Ling, S., 2004. Estimation and testing of stationarity for double-autoregressive models. Journal of Statistics 66, 63–78.

Litterman, R., Scheinkman, J., 1991. Common Factors Affecting Bond Returns. Journal of Fixed Income 1, 54–61.

Logue, D.E., Willett, T.D., 1976. A Note on the Relation between the Rate and Variability of Inflation. Economica 43, 151–158.

Nicolau, J., 2005. Processes with volatility-induced stationarity: an application for interest rates. Statistica Neerlandica 59, 376–396.

Nielsen, H.B., Rahbek, A., 2014. Unit root vector autoregression with volatility induced stationarity. Journal of Empirical Finance 29, 144–167.

Piazzesi, M., 2005. Bond yields and the Federal Reserve. Journal of Political Economy 113, 311–344.

Realdon, M., 2006. Quadratic Term Structure Models in Discrete Time. Finance Research Letters 3, 277–289.

Rudebusch, G., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. The Economic Journal 118, 906–926.

Shea, G.S., 1992. Benchmarking the Expectations Hypothesis of the Interest-Rate Term Structure: An Analysis of Cointegration Vectors. Journal of Business & Economic Statistics 10, 347–366.

Shiller, R., 1979. The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure. The Journal of Political Economy 87, 1190–1219.

Wachter, J.A., 2006. A consumption-based model of the term structure of interest rates. Journal of Financial Economics 79, 365–399.

Wright, J.H., 2011. Term Premiums and Inflation Uncertainty: Empirical Evidence from an International Dataset. American Economic Reivew 101, 1514–1534.

Zhu, H., Zhang, X., Liang, X., Li, Y., 2017. On a vector double autoregressive model. Statistics & Probability Letters 129, 85–96.

A Specification and Estimation of the DAR

We specify the model in (4) by use of conventional methods. In particular, the rank is determined by the likelihood-ratio test in Johansen (1991), which is based on the VAR, with critical values obtained using the wild bootstrap procedure in Cavaliere et al. (2014). The test model is specified with a restricted constant, i.e., a constant term appears in the cointegrating relations only.[19] The lag structure of the test is specified by general-to-specific LR tests, information criteria, and misspecification tests. For the choice of 3-months lags, the residuals are not autocorrelated according to univariate Ljung-Box tests, see the right panel of Table 3 in the body of the paper. The likelihood-ratio test, for which results are reported in Table 8, suggests a reduced rank of $r = 2$. We interpret these findings as indications that the DAR model in (4) may be well-specified by $r = 2$ and lag length $K = 3$ as well. The left panel of Table 3 confirms this presumption.

Table 8: Rank Testing

$r \leq 0$	$r \leq 1$	$r \leq 2$	$r \leq 3$
90.89	32.62	15.99	5.06
[0.00]	[0.03]	[0.07]	[0.11]

Note: Likelihood-ratio test of the null $r \leq \{0, 1, 2, 3\}$ against $r = p$. P-values obtained with the wild bootstrap in brackets.

We estimate (4) with $r = 2$ and $K = 3$ by maximum likelihood under just-identifying restrictions. Then, we impose further restrictions to obtain models with economically sensible interpretations. The restrictions are imposed sequentially starting with setting the most insignificant estimates to zero first in the relations $\beta' X_t$ and then in the adjustment matrix. At each step, the restrictions are tested by LR tests and the short-run coefficient estimates are compared. Parameter estimates are given in Tables 9 and 10.

[19] If a constant is unrestricted, the cointegrated VAR model implies that the data contains a linear trend (Johansen, 1996).

Table 9: Parameter Estimates Related to the Conditional Mean

	DAR				CVAR			
μ' ($\times 10^{-4}$)	-0.041	0.288	-0.273	-0.244	-0.637	-0.215	-0.122	-0.241
	(0.060)	(0.217)	(0.263)	(0.106)	(0.448)	(0.119)	(0.425)	(0.095)
α'	0	0	0.017	0.008	-0.016	0	0.017	0.009
			(0.008)	(0.003)	(0.008)		(0.009)	(0.003)
	0	-0.031	0	0	0.050	0	-0.033	0
		(0.021)			(0.027)		(0.017)	
β'	1	0	-3.672	-1.681	1	0	-3.097	2.174
			(1.150)	(0.611)			(0.911)	(0.715)
	-1	1	0	0	-1	1	0	0
Γ_1	-0.175	0.059	0.027	0.036	-0.427	0.183	0.014	-0.064
	(0.064)	(0.034)	(0.022)	(0.039)	(0.076)	(0.088)	(0.060)	(0.161)
	-0.026	0.027	0.150	0.140	-0.028	0.034	0.148	0.069
	(0.034)	(0.055)	(0.050)	(0.102)	(0.032)	(0.059)	(0.054)	(0.100)
	0.039	0.032	0.444	-0.087	0.020	0.024	0.449	0.029
	(0.032)	(0.056)	(0.059)	(0.113)	(0.030)	(0.056)	(0.077)	(0.140)
	0.027	0.010	0.068	0.244	0.020	0.021	0.065	0.236
	(0.016)	(0.027)	(0.022)	(0.048)	(0.015)	(0.027)	(0.023)	(0.049)
Γ_2	-0.028	-0.032	0.011	0.014	-0.125	0.164	0.035	0.527
	(0.072)	(0.023)	(0.019)	(0.045)	(0.082)	(0.080)	(0.065)	(0.165)
	-0.035	-0.119	0.064	0.048	-0.025	-0.117	0.057	0.005
	(0.040)	(0.055)	(0.044)	(0.106)	(0.039)	(0.053)	(0.048)	(0.114)
	0.036	-0.063	-0.162	0.292	0.015	-0.035	-0.124	0.480
	(0.032)	(0.049)	(0.066)	(0.118)	(0.033)	(0.052)	(0.079)	(0.137)
	0.008	-0.015	0.025	0.187	0.010	-0.009	0.017	0.193
	(0.017)	(0.025)	(0.029)	(0.053)	(0.016)	(0.025)	(0.033)	(0.049)
Γ_3	-0.061	-0.022	-0.0359	0.061	0.030	-0.064	-0.190	-0.070
	(0.056)	(0.022)	(0.0220)	(0.038)	(0.074)	(0.083)	(0.062)	(0.141)
	-0.045	0.065	-0.090	0.001	-0.026	0.056	-0.078	-0.063
	(0.034)	(0.058)	(0.043)	(0.096)	(0.034)	(0.061)	(0.044)	(0.099)
	0.007	0.064	0.122	-0.084	0.009	0.066	0.063	0.050
	(0.029)	(0.051)	(0.056)	(0.113)	(0.030)	(0.056)	(0.070)	(0.111)
	-0.039	-0.015	0.034	0.288	-0.035	-0.017	0.032	0.278
	(0.015)	(0.025)	(0.025)	(0.052)	(0.014)	(0.024)	(0.028)	(0.053)

Note: Estimates of parameters related to the conditional mean of the DAR and CVAR models. Standard errors in parentheses.

Table 10: Parameter Estimates of the Volatility Dynamics

	DAR				CVAR			
Σ_0 ($\times 10^{-3}$)	0.045 (0.010)				0.369 (0.022)			
	0.049 (0.037)	0.217 (0.012)			0.029 (0.015)	0.228 (0.010)		
	-0.016 (0.040)	0.022 (0.016)	0.247 (0.017)		0.020 (0.013)	-0.007 (0.016)	0.275 (0.016)	
	0.012 (0.015)	0.014 (0.007)	0.002 (0.009)	0.113 (9,995)	0.013 (0.005)	0.017 (0.005)	-0.001 (0.008)	0.114 (0.005)
Σ_1	-0.096 (0.013)	-0.019 (0.006)	-0.015 (0.006)	-0.008 (0.005)				
	-0.015 (0.014)	0.007 (0.010)	0.012 (0.011)	-0.028 (0.012)				
	-0.006 (0.018)	-0.007 (0.011)	-0.012 (0.018)	0.076 (0.016)				
	-0.002 (0.006)	-0.000 (0.004)	-0.008 (0.007)	-0.0074 (0.007)				

Note: Estimates of parameters related to the conditional volatility in the DAR and CVAR models. Standard errors in parentheses.

B Bond Yield Approximation

Define $\varepsilon_t^{\mathbb{Q}} = \Omega_t^{1/2}\varepsilon_t^{\mathbb{Q}}$. From the factor dynamics under the $\mathbb{Q}$-measure in (8), write

$$\begin{aligned} X_t X_t' =& \mu^{\mathbb{Q}}\mu^{\mathbb{Q}'} + \mu^{\mathbb{Q}} X_{t-1}' \Phi^{\mathbb{Q}'} + \Phi^{\mathbb{Q}} X_{t-1}\mu^{\mathbb{Q}'} + \Phi^{\mathbb{Q}} X_{t-1} X_{t-1}' \Phi^{\mathbb{Q}'} + \left(\mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1}\right)\varepsilon_t^{\mathbb{Q}'} \\ &+ \varepsilon_t^{\mathbb{Q}}\left(\mu^{\mathbb{Q}'} + X_{t-1}'\Phi^{\mathbb{Q}'}\right) + \varepsilon_t^{\mathbb{Q}}\varepsilon_t^{\mathbb{Q}'}, \end{aligned}$$

or by using the vectorization operator, $\mathrm{vec}\,(A)$, that stacks the columns of the matrix A into a vector and its relation with the Kronecker product denoted $\otimes$,

$$\begin{aligned} \mathrm{vec}\,(X_t X_t') =& \mathrm{vec}\left(\mu^{\mathbb{Q}}\mu^{\mathbb{Q}'}\right) + \left(\Phi^{\mathbb{Q}} \otimes \mu^{\mathbb{Q}} + \mu^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right) X_{t-1} + \left(\Phi^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right)\mathrm{vec}\left(X_{t-1}X_{t-1}'\right) \\ &+ \left(I_4 \otimes \left(\mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1}\right) + \left(\mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1}\right) \otimes I_4\right)\varepsilon_t^{\mathbb{Q}} + \mathrm{vec}\left(\varepsilon_t^{\mathbb{Q}}\varepsilon_t^{\mathbb{Q}'}\right). \end{aligned}$$

Next, we compute the conditional expectation given $\mathcal{F}_{t-1} = \{X_{t-1}, \ldots, X_1\}$ under $\mathbb{Q}$-measure probabilities, $\mathbb{E}_{t-1}^{\mathbb{Q}}(\cdot)$, of this expression. It follows that

$$\begin{aligned} \mathbb{E}_{t-1}^{\mathbb{Q}}\left(\mathrm{vec}\,(X_t X_t')\right) =& \mathrm{vec}\left(\mu^{\mathbb{Q}}\mu^{\mathbb{Q}'}\right) + \left(\Phi^{\mathbb{Q}} \otimes \mu^{\mathbb{Q}} + \mu^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right) X_{t-1} + \left(\Phi^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right)\mathrm{vec}\left(X_{t-1}X_{t-1}'\right) \\ &+ \mathrm{vec}\,(\Omega_t). \end{aligned}$$

To derive a bond yield expression in closed-form, we introduce the following approximation:

$$\mathrm{vec}\,(X_t X_t') \approx \mathrm{vec}\left(\mu^{\mathbb{Q}}\mu^{\mathbb{Q}'}\right) + \left(\Phi^{\mathbb{Q}} \otimes \mu^{\mathbb{Q}} + \mu^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right) X_{t-1} + \left(\Phi^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right)\mathrm{vec}\left(X_{t-1}X_{t-1}'\right) + \mathrm{vec}\,(\Omega_t),$$

where $\approx$ denotes an equality that is valid only approximately. Given this equation, the zero-coupon bond price takes the form

$$P_{t,n+1} = \exp\left(A_{n+1} + B_{n+1}' X_t + C_{n+1}'\mathrm{vec}\,(X_t X_t')\right).$$

It is straightforward to prove this claim and derive recursive formulas for the loadings:

$$\begin{aligned} \log P_{t,n+1} =& -r_t + A_n + B_n'(\mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t) + C_n'\mathrm{vec}\left(\mu^{\mathbb{Q}}\mu^{\mathbb{Q}'}\right) + C_n'\left(\Phi^{\mathbb{Q}} \otimes \mu^{\mathbb{Q}} + \mu^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right) X_t \\ &+ C_n'\left(\Phi^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right)\mathrm{vec}\,(X_t X_t') + C_n'\mathrm{vec}\,(\Omega_{t+1}) + \log \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(B_n'\varepsilon_{t+1}^{\mathbb{Q}}\right)\right] \\ =& -r_t + A_n + B_n'(\mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t) + C_n'\mathrm{vec}\left(\mu^{\mathbb{Q}}\mu^{\mathbb{Q}'}\right) + C_n'\left(\Phi^{\mathbb{Q}} \otimes \mu^{\mathbb{Q}} + \mu^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right) X_t \\ &+ C_n'\left(\Phi^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}}\right)\mathrm{vec}\,(X_t X_t') + C_n'\mathrm{vec}\,(\Omega_{t+1}) + \frac{1}{2}B_n'\Omega_{t+1}B_n. \end{aligned}$$

Gathering terms result in factor loading recursions given by

$$\begin{aligned}
A_{n+1} &= A_n + B_n' \mu^{\mathbb{Q}} + C_n' \left(\operatorname{vec}\left(\mu^{\mathbb{Q}} \mu^{\mathbb{Q}\prime} \right) + \operatorname{vec}\left(\Sigma_0 \Sigma_0' \right) \right) + \frac{1}{2} B_n' \Sigma_0 \Sigma_0' B_n \\
B_{n+1}' &= -\iota_1 + B_n' \Phi^{\mathbb{Q}} + C_n' \left(\Phi^{\mathbb{Q}} \otimes \mu^{\mathbb{Q}} + \mu^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}} \right) \\
C_{n+1}' &= C_n' \left(\Phi^{\mathbb{Q}} \otimes \Phi^{\mathbb{Q}} + \Sigma_1 \otimes \Sigma_1 \right) + \frac{1}{2} \left([B_n' \Sigma_1] \otimes [B_n' \Sigma_1] \right).
\end{aligned}$$

To be consistent with $r_t = \iota_1' X_t$, the recursions are initiated at $n = 0$ with $A_0 = 0$, $B_0 = \mathbf{0}_{p \times 1}$, $C_0 = \mathbf{0}_{p^2 \times 1}$.

DANMARKS NATIONALBANK
HAVNEGADE 5
DK-1093 COPENHAGEN K
WWW.NATIONALBANKEN.DK

DANMARKS
NATIONALBANK

As a general rule, Working Papers are not translated, but are available in the original language used by the contributor.

Danmarks Nationalbank's Working Papers are published in PDF format at www.nationalbanken.dk. A free electronic subscription is also available at this Website. The subscriber receives an e-mail notification whenever a new Working Paper is published.

Please direct any enquiries to Danmarks Nationalbank, Communications, *kommunikation@nationalbanken.dk*