UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Ph.D. Thesis

Anne Lundgaard Hansen

Yield Curve Dynamics

Persistence, Volatility, and the Real Economy

Advisors: Heino Bohn Nielsen and Anders Rahbek
Date of Submission: April 30, 2020

Contents

Preface

This thesis was written as part of my graduate studies at the Department of Economics at University of Copenhagen and Danmarks Nationalbank. I thank these institutions for giving me the opportunity and financial support to pursue my research and a graduate degree. I especially thank my advisors Heino Bohn Nielsen and Anders Rahbek for their guidance and support. I am also grateful to Rasmus Søndergaard Pedersen and Søren Hove Ravn for their comments on my work. To Emiliano Santoro, thank you for your invaluable advice and support during my journey on the international job market for economists.

In the spring 2018, I had the great pleasure of visiting the Department of Finance at Kellogg School of Management at Northwestern University. I am indebted to Torben G. Andersen and Viktor Todorov for inviting me, for their willingness to spend time to discuss my work, and more generally for their support. I have learned a lot from our many discussions both during my visit and on several occasions afterwards around the world. My visit was financially supported through grants from Knud Høgaards Fond, Augustinus Fonden, Julie von Müllens Fond, Incentive Fonden, and Etly og Jørgen Stjerngrens Fond for which I am grateful.

During my graduate studies, I also visited the Bank of England as a Ph.D. Intern. I thank Mike Joyce and Iryna Kaminska for granting me this opportunity and for our great collaboration. I also thank my colleagues at the bank for making me feel welcome as part of the team.

My interest in yield curve modeling was sparked while I was studying towards my bachelor's degree at Aarhus University. I am indebted to Martin Møller Andreasen for introducing me to this topic and supervising me in my first attempt to work with these models in my bachelor's thesis. I also thank Martin for invaluable comments on the second and third chapters in this thesis.

Writing this thesis would not have been nearly as enjoyable had it not been for fellow graduate students, some of whom has turned into lifetime friends. To Philipp Kless, Marcus Mølbak Ingholt, Simon Lund Hetland, Nick Fabrin Nielsen, and Patrick Thöni, thank you for many good times spent in- and outside the office. I also thank Martin Thyrsgaard for salutary lunch breaks during my visit in Chicago. I hope to see you all in Charlotte.

Finally, I am immensely grateful to my family and friends. To my parents, thank you for your endless support and encouragement. I thank my brother for inspiring me to pursue a Ph.D. and in so many other ways. To Phaél, thank you for always believing in me and keeping me on the *focus program*.

Anne Lundgaard Hansen
Copenhagen, April 2020

Summary

Interest rates vary with time horizons. This relationship, known as the term structure of interest rates or the yield curve, contains information about market expectations on future interest rates, inflation, and economic activity; risk attitudes; and recession probabilities. Understanding yield curve dynamics is thus crucial for monetary policy makers and investors to respond appropriately to fluctuations in financial markets and the economy.

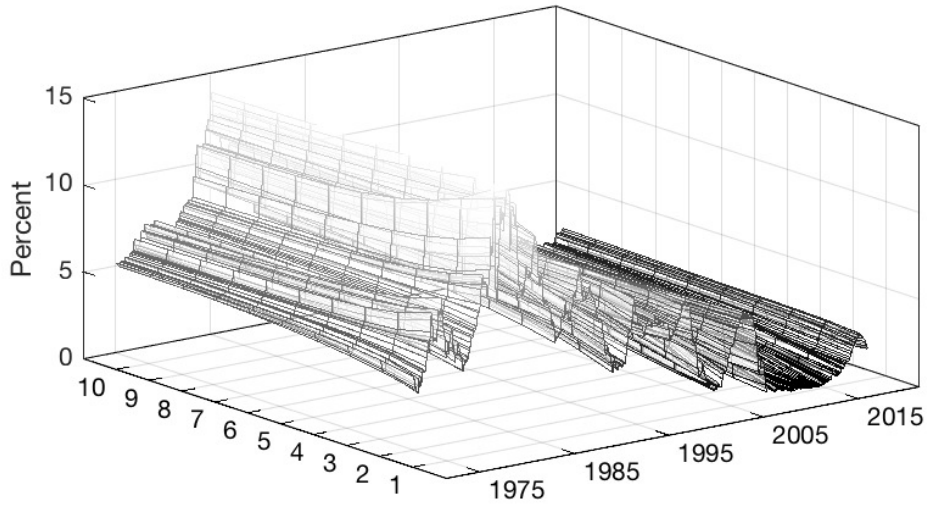
This thesis addresses key challenges for modeling and interpreting yield curve dynamics. Through three self-contained chapters, I present new methodologies and empirical insights related to the time-series properties of bond yields, risk factors in bond markets, and implications for monetary policy. To illustrate the central challenges in dynamic term structure modeling, consider the term structure of U.S. Treasury bond yields from 1971 to 2019 in Figure 1. I highlight two features of these data:

(i) The time series are highly persistent such that yields only change by small amounts between observations.

(ii) There is high correlation in the cross-section such that yields for different maturities strongly co-move.

The persistent behavior of interest rates poses a challenge for the workhorse models in the literature. These so-called Gaussian affine term structure models are based on stationary, linear dynamics that cannot capture the degree of persistence observed in the data. The econometrician can enforce the linear model to be sufficiently persistent to match the data by imposing a unit root, but with the result of non-stationarity. This conundrum is the focal point of **Chapter 1** ("Modeling Persistent Interest Rates with Volatility-Induced Stationarity"). I contribute to a literature that argues that linear dynamic models estimate unreliable decompositions of long-term bond yields into market expectations on future short-term interest rates and term premia. This so-called persistence problem arises because the decomposition is based on the ability of the model to forecast future short-term yields. In a linear and stationary model such forecasts converge at a faster rate than suggested by the data. Therefore, model-implied expectations to future short-term interest rates are excessively stable and in turn, model-implied term premia are excessively correlated with

Figure 1: U.S. Treasury Bond Yield Curve



Notes: The figure shows the U.S. Treasury bond yield curve measured end-of-month from September 1971 to August 2019 for maturities of one to ten years. Data are from Gürkaynak, Sack, and Wright (2007).

the yield curve. The opposite problem arises in linear models with unit roots, i.e., the variation in expectations mimics the variation in bond yields and term premia are excessively stable.

To accommodate the persistence problem, I introduce the double autoregressive (DAR) process from Ling (2004) and Nielsen and Rahbek (2014) in a term structure model. The DAR process has a level-dependent conditional volatility, which ensures that the model can remain stationary even in the presence of a unit root in the characteristic polynomial corresponding to the conditional mean. This feature, which has been coined volatility-induced stationarity, is consistent with important characteristics of yield curve data. Empirically, I show that the DAR model estimates term premia that are economically plausible and consistent with survey measures of market expectations. I also show that the model exhibits improved out-of-sample forecasting performance compared with the Gaussian affine term structure model.

The second challenge - the fact that yields of different maturities are highly correlated - has motivated the use of low-dimensional factor models, where the yield curve is assumed to be driven by a few latent factors (Duffie and Kan, 1996, Litterman and Scheinkman, 1991). These models can match the yield curve data to a high precision, but they offer no insights on the nature of these factors and thus the risk factors that drive bond markets.

Chapter 2 ("Yield Curve Volatility and Macroeconomic Risk") contributes to a literature that explores the fundamental drivers behind the latent yield curve factors. Starting with the seminal work by Ang and Piazzesi (2003), many contributions have provided insights using Gaussian affine term structure models. Using structural vector autoregressive models and associated variance decompositions, these models can answer, what share of the variation in bond yields can be attributed to macroeconomic shocks? A central assumption in the established literature is that the amount of variation to be explained, i.e., the volatility, is constant. I argue that this assumption is critical as it implies that the share of variation attributed to macroeconomic shocks is constant over time per construction. To overcome this deficiency, I extend the canonical model with time-varying volatility given by the multivariate BEKK-GARCH model in Engle and Kroner (1995) and a variance risk premium introduced by Monfort and Pegoraro (2012). I show that my model can be estimated by a generalization of the important method developed by Joslin, Singleton, and Zhu (2011).

I apply my model to characterize the variation in the U.S. Treasury bond yield curve over the period from 1971 to 2019 in terms of shocks to inflation and the unemployment gap. I show that the relationship between the yield curve and these macroeconomic risks has varied over time, and that this time-variation exhibits interesting patterns: First, the macroeconomic contribution to short-term yield volatility was high in the end 1970s, low during the Great Moderation starting in the mid-1980s, and high again after the financial crisis. Second, I find that market expectations on future short-term yields have increasingly been related to the macroeconomy after the Great Moderation. Third, I document that deflation fears increased term premia during the financial crisis. Finally, I zoom in on the spring 2019, where the yield curve inverted with long-term yields below short-term yields. I show that macroeconomic shocks do not explain the inversion and argue that the inversion is unlikely to be a warning of an imminent recession.

Time-varying volatility plays a vital role in both Chapters 1 and 2, and thus in the solution to some of the challenges in term structure modeling. However, modeling bond market volatility has proven difficult in traditional continuous-time affine term structure models with stochastic volatility (Jacobs and Karoui, 2009). This difficulty arises because these models rely on a linear relationship between bond yield levels and bond market volatility, which appear at odds with the data (Andersen and Benzoni, 2010, Collin-Dufresne, Goldstein, and Jones, 2009).

In **Chapter 3** ("A Joint Model for the Term Structures of Interest Rates and Realized Volatility"), I propose a new method for modeling the term structure of

bond market volatility. Specifically, I present a framework for the joint modeling of bond yield levels and realized volatility constructed from high-frequency data. My approach relies on results showing that realized volatility can be interpreted by the sum of conditional volatility and a mean-zero error. From this identity, I construct a measurement equation that links the realized yield covariance matrix to a conditional covariance matrix implied by a multivariate GARCH-type model. These ideas are integrated into a no-arbitrage term structure model with a low-dimensional latent state vector. I derive closed-form solutions for no-arbitrage bond yields, term premia, conditional yield curve volatility, and multi-step ahead forecasts of both yields and realized volatility. Furthermore, I develop an algorithm for filtering latent factors derived from the basic principles of the standard, linear Kalman (1960) filter. My filter is exact in contrast with existing non-linear Kalman filters. Empirically, I show that my model describes both the U.S. Treasury yield curve and the associated realized covariance matrix with high precision. I also present encouraging out-of-sample results for forecasting realized yield variances and covariances with multi-step ahead horizons. Finally, I use my model to show that the risk-neutral dynamics extracted from first and second moments of the yield curve do not describe the pricing of interest-rate derivatives.

Altogether, this thesis contributes with new methods to solve some of the challenges that prevail in the literature on modeling yield curve dynamics. These methods provide new insights into the interpretation of movements in the yield curve.

Summary in Danish

Renter varierer med tidshorisonter. Dette forhold, kaldt rentestrukturen eller rentekurven, indeholder information om markedsforventninger til fremtidige renter, inflation og økonomisk aktivitet; risikopræferencer og sandsynlighed for recession. Det er således vigtigt for centralbanksøkonomer og investorer at forstå rentekurvedynamikker for at kunne respondere hensigtsmæssigt på bevægelser i finansielle markeder og økonomier.

Denne afhandling adresserer væsentlige udfordringer ved modellering og fortolkning af rentekurvedynamikker. Igennem tre selvstændige kapitler præsenterer jeg nye metoder og empirisk indblik relateret til tidsrækkeegenskaber for obligationsrenter, risikofaktorer i obligationsmarkeder og implikationer for pengepolitik. For at illustrere de centrale problematikker i dynamisk rentestruktursmodellering, betragtes rentekurven for amerikanske statsobligationer fra 1971 til 2019 i Figur 1 (se side viii). Jeg fremhæver to karakteristika i disse data:

(i) Tidsrækkerne er ekstremt persistente, det vil sige at renterne kun ændrer sig meget lidt mellem observationstidspunkter.

(ii) Der er høj korrelation på tværs af rentekurven, det vil sige at renter for nærtliggende udløbstidspunkter følger hinanden tæt.

Rentekurvens persistente adfærd er en udfordring for de gængse modeller i litteraturen. Disse såkaldte Gaussisk affine rentestruktursmodeller er baseret på stationære, lineær dynamikker, som ikke kan opfange den grad af persistens, som data viser. Økonometrikeren kan påtvinge den lineære model til at være tilstrækkelig persistent til at beskrive data ved at indføre en enhedsrod, men dette resulterer i ikke-stationaritet. Dette problem er omdrejningspunktet i **Kapitel 1** ("Modeling Persistent Interest Rates with Volatility-Induced Stationarity"). Jeg bidrager til en litteratur, der argumenterer for, at lineære dynamiske modeller estimerer upålidelige dekompositioner af langsigtede renter i henholdsvis forventninger til fremtidige korte renter og risikopræmier. Dette såkaldte persistensproblem opstår, fordi dekompositionen forlader sig på modellens evne til at fremskrive korte renter. En lineær og stationær model medfører, at disse fremskrivninger konvergerer med en hurtigere hastighed end antydet i data. Derfor er modelimplicitte forventede fremtidige renter for stabile og omvendt, er modelimplicitte risikopræmier for korrelerede med

rentestrukturen. Det modsatte problem opstår i lineære modeller med enhedsrødder, nemlig at variationen i forventninger efterligner variationen i renterne og risikopræmier er for stabile.

For at imødekomme persistensproblemet introducerer jeg den dobbelte autoregressive (DAR) proces fra Ling (2004) samt Nielsen og Rahbek (2014) i en rentestruktursmodel. DAR processen har en niveaubestemt betinget volatilitet, som sikrer, at modellen kan forblive stationær selv under tilstedeværelsen af en enhedsrod i det karakteristiske polynomium svarende til den betingede middelværdi. Denne egenskab, som kaldes volatilitetsinduceret stationaritet, er konsistent med vigtige karakteristika i rentestruktursdata. Jeg viser empirisk, at modellen estimerer risikopræmier, der er økonomisk plausible og konsistente med spørgeundersøgelsesdata for markedsforventninger. Jeg viser også, at modellen viser forbedret evne til at fremskrive hele rentestrukturen i sammenligning den Gaussisk affine rentestruktursmodel.

Den anden udfordring - det faktum at renter til forskellige udløbstidspunkter er stærkt korrelerede - har motiveret brugen af lavdimensionale faktormodeller, hvor rentekurven antages at være drevet af få latente faktorer (Duffie og Kan, 1996, Litterman og Scheinkman, 1991). Disse modeller kan beskrive rentekurven med en høj præcision, men giver ingen information om, hvad de latente faktorer indeholder og dermed, hvilke risikofaktorer der driver obligationsmarkedet.

Kapitel 2 ("Yield Curve Volatility and Macroeconomic Risk") bidrager til en litteratur, der undersøger det fundamentale drivværk bag latente faktorer i rentekurven. Med afsæt i det grundlæggende arbejde af Ang og Piazzesi (2003), har mange forsøgt at besvare dette spørgsmål ved hjælp af Gaussisk affine rentemodeller. Ved brug af strukturelle vektorautoregressive modeller og dertilhørende variansdekompositioner kan disse modeller besvare, hvor stor en andel af variationen i renter kan forklares med observerbare makroøkonomiske variable? Det er en central antagelse i den etablerede litteratur, at mængden af variation der ønskes forklares, det vil sige volatiliteten, er konstant. Jeg argumenterer for, at denne antagelse er kritisk idet den medfører, at andelen af variationen der tillægges makroøkonomiske stød, er konstant over tid per konstruktion. For at overvinde denne utilstrækkelighed udvider jeg den traditionelle model med tidsvarierende volatilitet givet ved den multivariate BEKK-GARCH model i Engle og Kroner (1995) og en variansrisikopræmie introduceret af Monfort og Pegoraro (2012). Jeg viser, at min model kan estimeres ved en udvidelse af den vigtige metode udviklet af Joslin, Singleton og Zhu (2011).

Jeg anvender min model til at karakterisere, hvordan variationen i rentestrukturen i amerikanske statsobligationer kan tillægges stød til inflation og arbejdsløshedsgabet i perioden fra 1971 til 2019. Jeg viser, at sammenhængen mellem rentestrukturen og disse makroøkonomiske risikofaktorer har varieret over tid, og at denne tidsvariation udviser interessante mønstre. For eksempel var det makroøkonomiske bidrag til variationen i kortsigtede renter høj i slutningen af 1970'erne, lav i løbet af Den Store Moderation, startende fra midt 1980'erne, og høj igen efter den finansielle krise. Jeg viser også, at markedsforventningerne til fremtidige kortsigtede renter i en stadigt højere grad tillægges udviklingen i makroøkonomiske variable efter Den Store Moderation. Dernæst dokumenterer jeg, at deflationsfrygt øgede risikopræmierne i løbet af den finansielle krise. Endelig zoomer jeg ind på foråret 2019, hvor rentekurven inverterede med langsigtede renter lavere end kortsigtede renter. Jeg viser, at makroøkonomiske stød ikke kan forklare inverteringen og argumenterer for, at inverteringen ikke er en advarsel om en nært forestående recession.

Tidsvarierende volatilitet spiller en væsentlig rolle i både Kapitel 1 og 2 og således i forhold til at løse nogle af de vanskeligheder, som karakteriserer rentestrukturmodellering. Modellering af obligationsmarkedsvolatilitet viser sig dog, at være svært i traditionelle kontinuerttids affine rentestrukturmodeller med stokastisk volatilitet (Jacobs og Karoui, 2009). Det skyldes, at disse modeller beror på en linear sammenhæng mellem renteniveau og -volatilitet, hvilket viser sig at være i modstrid med data (Andersen og Benzoni, 2010, Collin-Dufresne, Goldstein og Jones, 2009).

I **Kapitel 3** ("A Joint Model for the Term Structures of Interest Rates and Realized Volatility") fremsætter jeg en ny metode til at modellere strukturen for obligationsmarkedsvolatilitet. Helt konkret præsenterer jeg en samlet model for rentestrukturens niveau og realiseret volatilitet, som konstrueres med højfrekvente data. Min fremgangsmåde beror på resultater der viser, at realiseret volatilitet kan fortolkes ved summen af betinget volatilitet og en fejl med middelværdi nul. Med udgangspunkt i denne identitet konstruerer jeg en måleligning, der forbinder den realiserede kovariansmatrix for rentestrukturen med en betinget kovariansmatrix, givet ved en multivariat GARCH-type model. Disse idéer integreres i en ingen-arbitrage rentestrukturmodel med en lavdimensionel latent tilstandsvektor. Jeg udleder løsninger i lukket form for ingen-arbitrage obligationsrenter, risikopræmier, betinget rentekurvevolatilitet og prognoser for både renter og realiseret volatilitet over flere perioder. Derudover udleder jeg en algoritme til filtrering af de latente faktorer baseret på basale principper fra det sædvanlige, lineære Kalman (1960) filter. Mit filter er eksakt modsat eksisterende ikke-lineære Kalman filtre. Empirisk viser jeg, at

min model beskriver både rentestrukturen og den tilhørende realiserede kovariansmatrix i amerikanske statsobligationer med høj præcision. Jeg præsenterer også lovende resultater for fremskrivning af renters realiserede varianser og kovarianser over flere perioder. Endelig bruger jeg min model til at vise, at de risikoneutrale dynamikker der udledes fra første- og andenordens momenter i rentekurven, ikke beskriver prissætning af rentederivater.

Samlet set bidrager denne afhandling med nye værktøjer til at løse nogle af de udfordringer, som litteraturen om rentestruktursmodellering står overfor. Disse værktøjer giver nye indblik i fortolkningen af bevægelser i rentestrukturen.

Chapter 1

Modeling Persistent Interest Rates with Volatility-Induced Stationarity

Anne Lundgaard Hansen
University of Copenhagen

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. I resolve this puzzle by proposing a term structure model with volatility-induced stationarity. The 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, I obtain term premia that are economically plausible and consistent with survey data. Compared with the Gaussian affine term structure model, I improve out-of-sample forecasting of the yield curve. The empirical evidence suggests that volatility-induced stationarity is unspanned by the yield curve.

Keywords: Yield curve, unit root, persistence problem, volatility-induced stationarity, level-dependent conditional volatility.

JEL classification: E43, E44, G12.

1.1 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, I bridge macro-finance term structure modeling with the double-autoregressive (DAR) model studied in Ling (2004) and Nielsen and Rahbek (2014).

I am 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, Hjalmarsson, and Österholm, 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, I 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):

$$x_t = \phi x_{t-1} + \left(\omega + \psi x_{t-1}^2\right)^{1/2} z_t \tag{1.1}$$

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 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

[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, Singleton, and Dai (2010). However, these models fail to capture the first two stylized facts that I emphasize.

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.

I 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. I therefore implement the DAR model with these features and benchmark the results against the cointegrated VAR model.[2] The 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, I show that the DAR model passes misspecification tests of the standardized residuals that the corresponding VAR model fails. Economically, I 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, I 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, Monfort, and Pegoraro (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. Thus, 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

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

[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, Anderson, and Granger (1992), and Shea (1992).

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, I 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, I 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. I propose a quadratic approximation to facilitate analytical computation of no-arbitrage bond yields. My 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 the 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, Hansen, Luttmer, and Scheinkman (1997) who consider Markov diffusion models with constant volatility elasticity as in the CKLS model in Chan, Karolyi, Longstaff, and Sanders (1992). Conley, Hansen, Luttmer, and Scheinkman (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 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, Piazzesi, and Wei, 2006, Diebold, Piazzesi, and Rudebusch, 2005, Duffee, 2006, Hördahl, Tristani, and Vestin, 2006, Joslin, Priebsch, and Singleton, 2014, Rudebusch and Wu, 2008). This paper is 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, Rudebusch, and Wu (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. Consequently, I show empirically that term premia implied by my model have different properties from those obtained by Bauer, Rudebusch, and Wu (2014). Abbritti, Gil-Alana, Lovcha, and Moreno (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, Monfort, and Pegoraro (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 1.2 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 1.3. In Section 1.4, I 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 1.5 concludes.

1.2 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, I specify the conditional variance, Ω_{t+1}.[4]

1.2.1 Conditional Variance Specification

I 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) I 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

$$
\begin{aligned}
X_{t+1} &= \mu + \Phi X_t + \Omega_{t+1}^{1/2}\varepsilon_{t+1}, \\
\Omega_{t+1} &= \Sigma_0\Sigma_0' + \Sigma_1 X_t X_t'\Sigma_1', \\
\varepsilon_{t+1} &\sim \text{i.i.d. } \mathcal{N}(0, \mathrm{I}_p).
\end{aligned}
\tag{1.2}
$$

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.

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, I 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, I present a macro-finance empirical application in Section 1.3.

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

[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).

1.2.2 Stationarity Condition and Time Series Properties

Nielsen and Rahbek (2014) show that X_t given by (1.2) is globally stationary and geometrically ergodic if the Lyapunov exponents are strictly negative, i.e.,

$$\gamma\left(\Phi, \Sigma_1\right) = \lim_{\xi \to \infty} \left[\mathbb{E}\left(\log \| \prod_{t=1}^{\xi} (\Phi + e_t) \| \right) \right] < 0, \tag{1.3}$$

where e_t is a $p \times p$ matrix that is i.i.d. normally distributed with mean zero and covariance matrix given by $\Sigma_1 \otimes \Sigma_1$. 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 (1.3) 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, let us look at these cases separately.

(i) Non-stationary 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 does not take on values that 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 $\gamma(\Phi, \Sigma_1) \geq 0$. This case is well-studied if $\Sigma_1 = 0_{p \times p}$. In particular, if the number of unit roots equals the dimension of the model, i.e., if $q = p$, then the model is I(1) and thus can be made 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, X_t is cointegrated with cointegrating vector $\beta \neq 0$ as defined by Johansen (1995). 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

[6] I abstract from cases where the characteristic polynomial corresponding to the conditional mean has solutions outside the unit circle although (1.3) can be satisfied under such cases.

due to a sufficiently large level effect in the conditional variance: $|\lambda_i| = 1$ for some $i = 1, \ldots, q \leq p$, $|\lambda_j| < 1$ for $j = q+1, \ldots, p$, 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, 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.

(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.

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

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

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

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.

1.2.3 Numerical Illustrations

I 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) Non-stationary I(1) model: $\mu = 0.01$, $\Phi = 1$, $\Sigma_0 = 0.1$, and $\Sigma_1 = 0$.

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

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

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

I 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 1.1. 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.

1.2.4 Likelihood

The process X_t in (1.2) is conditionally Gaussian given X_{t-1} with conditional mean and variance equal to

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

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

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

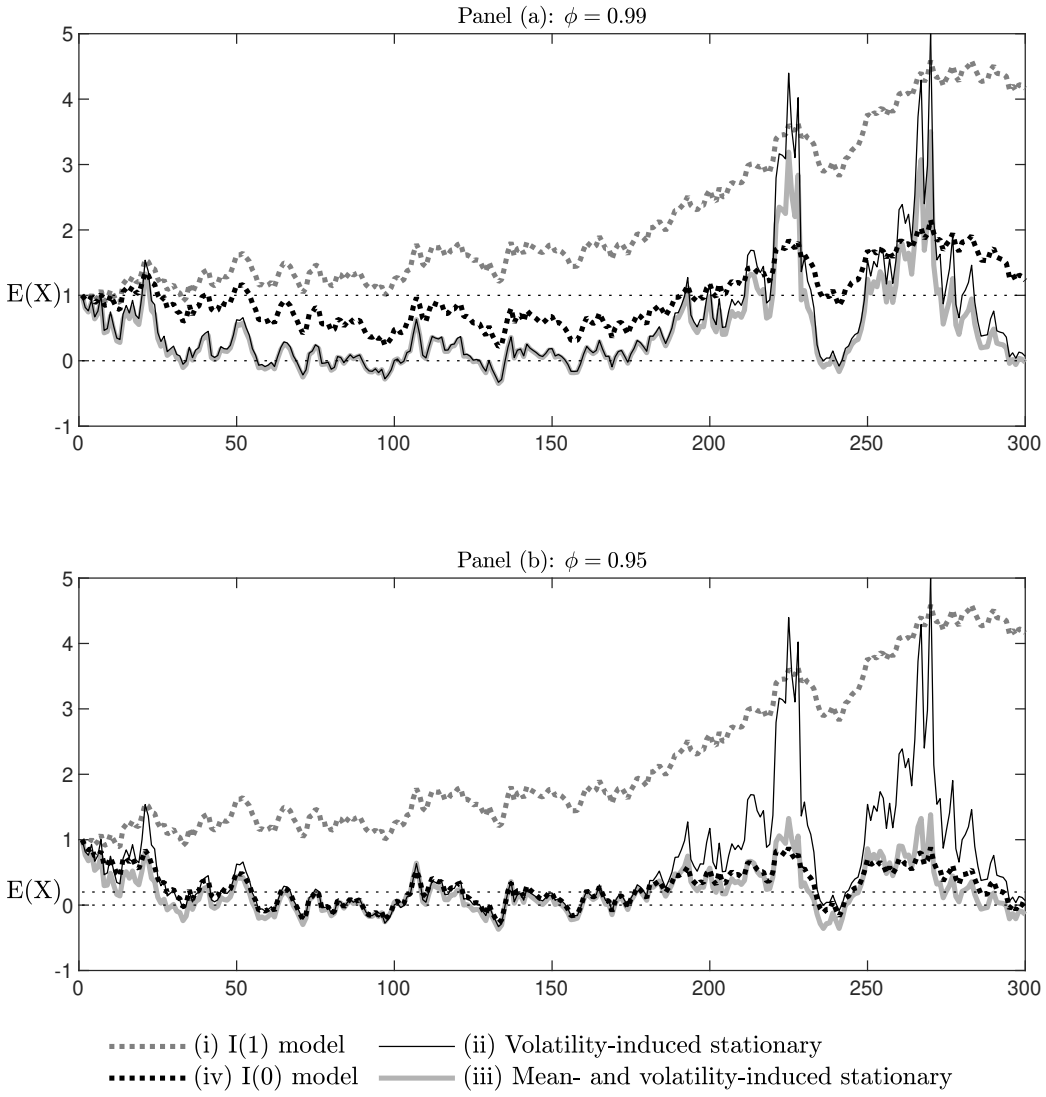
where I note that Ω_t is a function of Σ_0 and Σ_1 given in (1.2) and the parameters are given by $\Theta^{\mathbb{P}} = \{\mu, \Phi, \Sigma_0, \Sigma_1\}$. Consistency and asymptotic normality of the maximum likelihood estimator has been established in the univariate case by Ling (2004); in a bivariate model under certain parameter restrictions by Nielsen and Rahbek (2014); and in the multivariate setting but with a diagonal conditional covariance matrix in Zhu, Zhang, Liang, and Li (2017). Since there are no results available for the general multivariate specification, I confirm by simulations that the maximum likelihood estimators exhibit reasonable properties, i.e., are approximately Gaussian and centered around their true values.[9]

1.3 Empirical Analysis

For the empirical analysis, I 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. I consider a macro-finance setting for two reasons. Fist, macroeconomic variables are important predictors of term premia (Joslin, Priebsch, and Singleton, 2014, Wright, 2011). Second, the volatility specification in (1.2) has an economic

[9]The simulation results are available upon request.

Figure 1.1: Simulated Path of Univariate DAR Models



Notes: 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).

motivation involving inflation rates and unemployment as discussed in Section 1.2.1. In particular, I 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, $X_t = [r_t, R_t, \pi_t, g_t]'$.

1.3.1 Data

I 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. Define the long rate by the ten-year U.S. Treasury bond yield from Gürkaynak, Sack, and Wright (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 Reserve Economic Data; and the help-wanted-advertising-in-newspapers (HELP) index from Barnichon (2010).

Table 1.1 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 percent 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. The measure of real activity captures 66 percent of the variation in the underlying observables.

The data for $X_t = [r_t, R_t, \pi_t, g_t]'$ are exhibited in Figure 1.2. The series appear extremely persistent and use of conventional unit-root and stationarity tests indeed identify unit roots, see Table 1.2. 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, I propose the DAR model for these persistent and stationary data.

Table 1.1: Interpretation of the Macroeconomic Factors

	Explained (percent)	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

Notes: 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.

Table 1.2: 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]

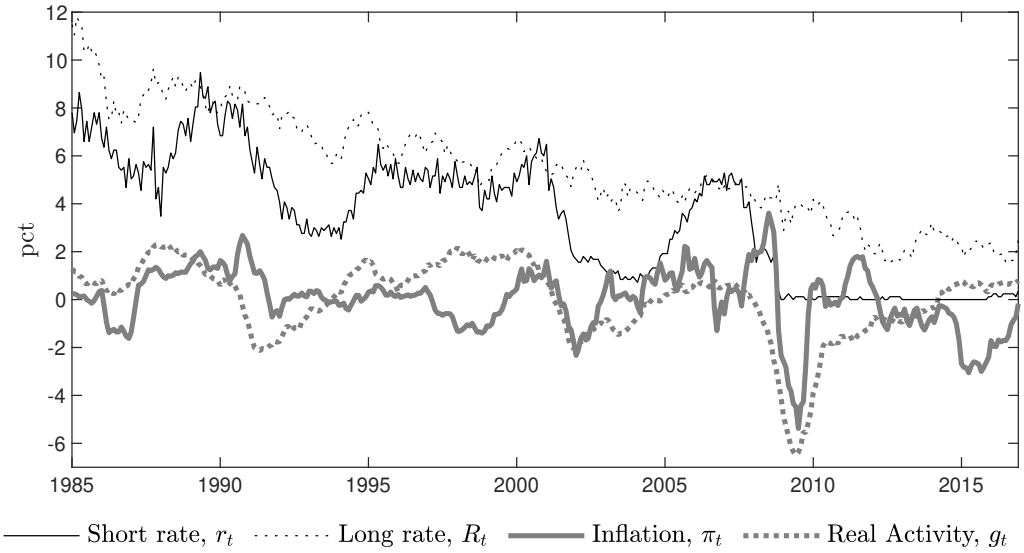
Notes: Augmented Dickey-Fuller and KPSS tests for respectively unit roots and stationarity. P-values in brackets.

1.3.2 Model Specification and Estimation

To achieve a well-specified model, I allow for an extensive lag structure in the conditional mean. I find that this generalization is sufficient to match the data and thus I leave the conditional variance as specified in (1.2). The resulting DAR model is given by:

$$\begin{aligned} X_{t+1} &= \mu + \Phi X_t + \sum_{k=1}^{K} \Gamma_k \Delta X_{t-k} + \Omega_{t+1}^{1/2} \varepsilon_{t+1}, \\ \Omega_{t+1} &= \Sigma_0 \Sigma_0' + \Sigma_1 X_t X_t' \Sigma_1', \\ \varepsilon_{t+1} &\sim \text{i.i.d. } \mathcal{N}_p(0, \mathrm{I}_p) \end{aligned} \tag{1.4}$$

Figure 1.2: Monthly Interest Rates and Macroeconomic Factors



Notes: 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.

for $K \geq 1$.[10] As I focus on the purely volatility-induced stationary case, I 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 $\Theta^{\mathbb{P}} = \{\mu, \alpha, \beta, \Gamma_1, \ldots, \Gamma_K, \Sigma_0, \Sigma_1\}$.

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 with the corresponding cointegrated VAR (CVAR) model that appears as the special case

[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

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

Table 1.3: 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 test	7.71 [0.10]	1.67 [0.80]	0.61 [0.96]	4.96 [0.29]	3.72 [0.44]	1.48 [0.83]	1.05 [0.90]	5.75 [0.22]
Engle's ARCH test	23.81 [0.00]	4.37 [0.04]	2.53 [0.11]	1.58 [0.21]	36.15 [0.00]	9.92 [0.00]	20.82 [0.00]	0.60 [0.44]
Kolmogorov-Smirnov test	0.07 [0.05]	0.06 [0.09]	0.06 [0.12]	0.06 [0.16]	0.09 [0.00]	0.06 [0.12]	0.10 [0.00]	0.05 [0.12]

Notes: 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.

when $\Sigma_1 = 0_{p\times p}$, the DAR model removes autocorrelation and improves normality tests of the standardized residuals, see Table 1.3. 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, 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.1.

Long-Run Equilibria

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

$$\hat{\beta}_1' X_t = r_t - 3.672\pi_t - 1.681 g_t,$$
$$\hat{\beta}_2' X_t = R_t - r_t,$$

see Table A.3 in Appendix. One relation is the spread between long and short rates as in Hall, Anderson, and Granger (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,

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

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

1.3.3 Conditional Volatilities

Model-implied conditional variances are exhibited in Figure 1.3. 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 1.3, the short rate stands out with a highly volatile conditional variance that ranges from high and rapidly changing to low and stable.

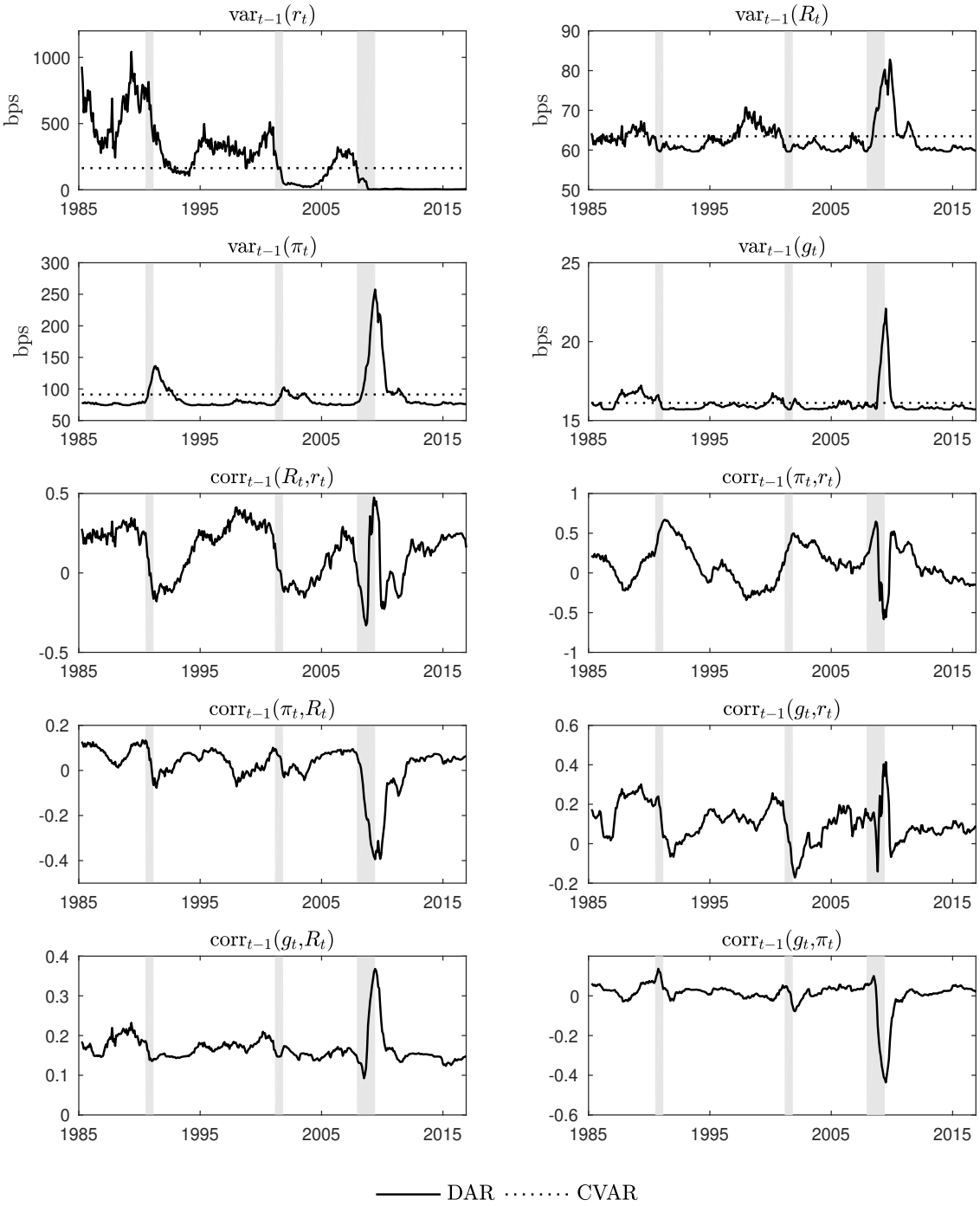
From the estimation results provided in Appendix A.1, the conditional variance of the short rate is given by

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

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 1.3 also shows the conditional correlations in the DAR model. I 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.

Figure 1.3: Estimated Conditional Variances and Correlations



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

1.3.4 Term Premia

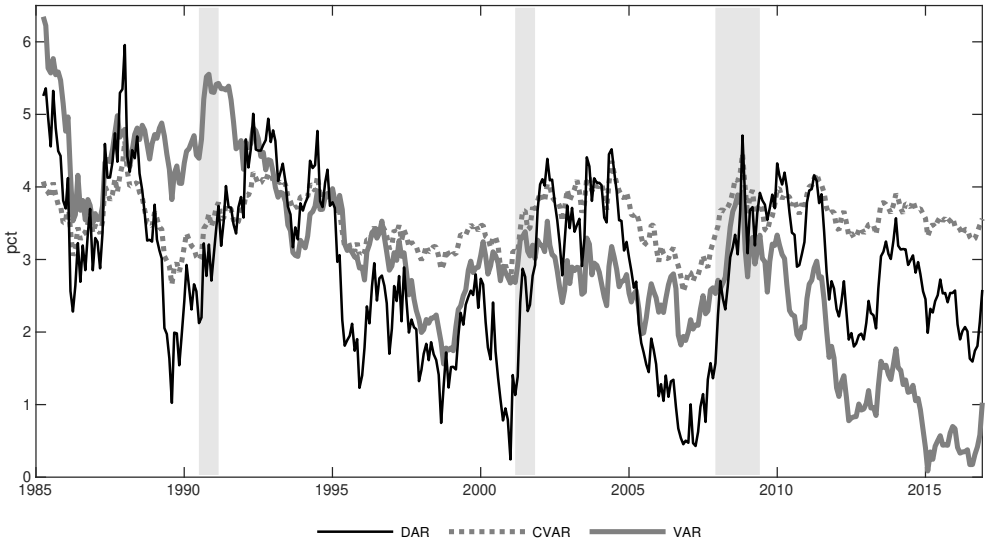
Next, I show that volatility-induced stationarity impacts model-implied term premia. Term premia are defined by an accounting identity that decomposes the bond yield, $Y_{t,n}$, into the yield that would prevail if investors were risk neutral, $\tilde{Y}_{t,n}$, and a residual, the term premium, $\text{TP}_{t,n}$: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

$$\tilde{Y}_{t,n} = -\frac{1}{n} \log \mathbb{E}_t \left(\exp \left[- \sum_{i=0}^{n-1} r_{t+i} \right] \right), \tag{1.5}$$

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 1.4. Besides comparing the DAR and CVAR models, I 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.

Figure 1.4: Ten-Year Term Premia



Notes: 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.

[12]The expectation in (1.5) can either be simulated or approximated by the method that will explained in Section 1.4.2, where the $\mathbb{Q}$-parameters are replaced by the corresponding parameters under the $\mathbb{P}$-measure. Here, I report term premia obtained by approximation.

I 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. Consequently, almost all of the variation in yields is attributed to term premia. This persistence problem has previously been recognized in the literature (Abbritti, Gil-Alana, Lovcha, and Moreno, 2016, Goliński and Zaffaroni, 2016, Jardet, Monfort, and Pegoraro, 2013, Kozicki and Tinsley, 2001, Shiller, 1979). Second, I note that the CVAR model predicts an almost constant ten-year term premium in the range of 3-4 percent throughout the entire sample even as the ten-year yield falls toward 2 percent 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 stationary VAR model. To numerically evaluate the DAR model's ability to decompose interest rates, I 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. I use median forecasts of the three-month Treasury Bill rate as a proxy for the short rate. I compare these data to expectations computed by respectively the DAR model and the cointegrated and stationary VAR models. Table 1.4 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. I interpret this result as an indication that the DAR model provides a more accurate term structure decomposition than the cointegrated and stationary VAR models.

Table 1.4: 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

Notes: Root mean squared errors in basis points of forecasts of the short rate in the DAR, CVAR, and stationary VAR models compared with 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.

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 1.5 compares model-implied five-by-five year forward term premia of the DAR, CVAR, and stationary VAR models compared with 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 1.5.

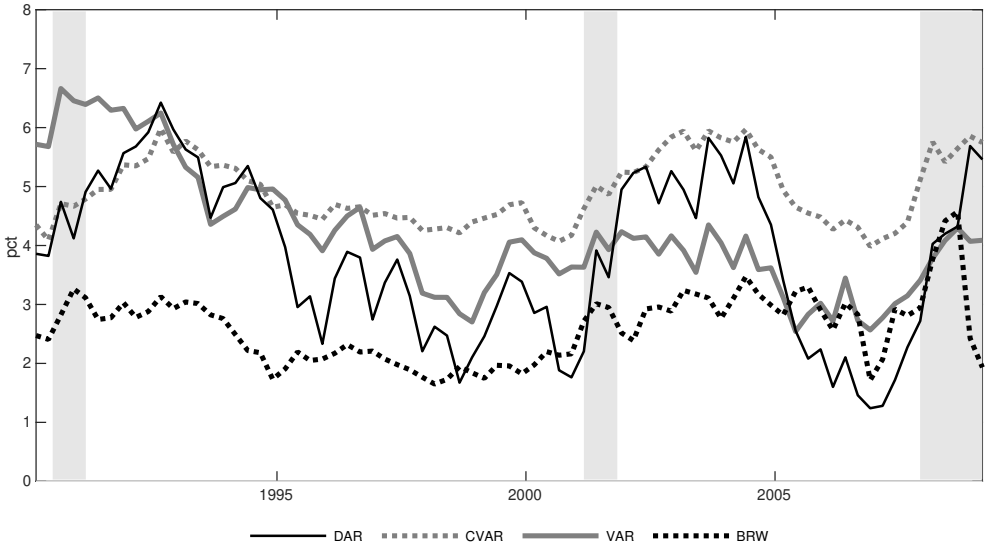
For the considered sample and at the quarterly frequency, the BRW forward term premia are as stable as those of the CVAR both with an empirical standard deviation of 0.6. Therefore, I expect that the BRW model encounters the same problem as the CVAR model if extrapolated into the zero-lower bound regime. 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, I would expect the correlation to be positive as predicted by the DAR and VAR models.

The persistence problem is also considered in Jardet, Monfort, and Pegoraro (2013), who suggest an averaging estimator that combines the parameter estimates of the stationary VAR and CVAR models. I adopt their weighting scheme to combine the estimated VAR models, which give term premia as depicted in Figure 1.6.[14] The term premia estimated by the averaging model are in-between those of the

[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.

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

Figure 1.5: Five-by-Five Year Forward Term Premia



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

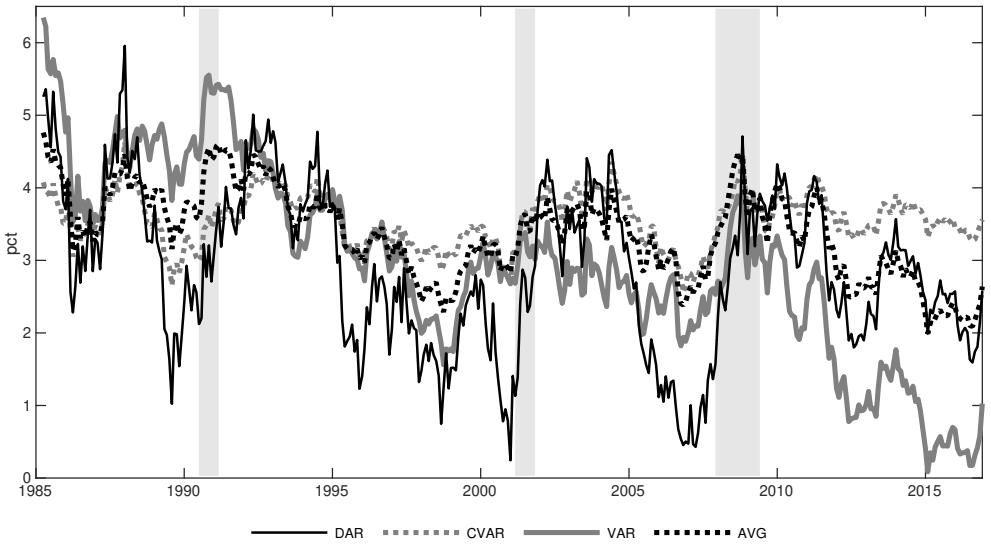
Table 1.5: Empirical Standard Deviations and Correlations of Five-by-Five Year Forward Term Premia

		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

Notes: 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, Rudebusch, and Wu (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.

stationary VAR and CVAR models and thus highly stable. Consequently, 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 zero-lower bound regime, where the term premia of the DAR model and averaging model coincide.

Figure 1.6: Ten-Year Term Premia



Notes: 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.

1.4 Volatility-Induced Stationary Term Structure Modeling

This section casts the DAR model analyzed thus far into a macro-finance term structure model. I 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 (1.4).

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

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

$$\mathcal{M}_{t+1} = \exp\left(-r_t - \frac{1}{2}\Lambda_t'\Omega_{t+1}\Lambda_t - \Lambda_t'\Omega_{t+1}^{1/2}\varepsilon_{t+1}\right), \tag{1.6}$$

where Λ_t is the market price of risk with risk measured by the conditional variance, Ω_{t+1}. I 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 I treat lagged variables as unspanned factors as in Joslin, Le, and Singleton (2013a). Thus, the lag structure determines the dynamics under the real-world measure but is not priced in the term structure cross-section. The market price of risk is defined by:

$$\Lambda_t = \Omega_{t+1}^{-1} \left[(\mu - \mu^{\mathbb{Q}}) + \left(\Phi - \Phi^{\mathbb{Q}}\right) X_t + \sum_{k=1}^{K} \Gamma_k \Delta X_{t-k} \right], \tag{1.7}$$

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

$$\begin{aligned} X_{t+1} &= \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t + \Omega_{t+1}^{1/2} \varepsilon_{t+1}^{\mathbb{Q}}, \\ \Omega_{t+1} &= \Sigma_0 \Sigma_0' + \Sigma_1 X_t X_t' \Sigma_1', \\ \varepsilon_{t+1}^{\mathbb{Q}} &\sim \text{i.i.d. } \mathcal{N}_p(0, \mathrm{I}_p). \end{aligned} \tag{1.8}$$

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, I use the acronym GATSM to describe the model based on the CVAR specification that is nested in my model.

1.4.2 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. My model does not admit a closed-form bond price expression that satisfies this equation. Instead, I 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, I make sure that for $\Sigma_1 = 0_{p \times p}$, bond yield computation must coincide with the solution of the GATSM. Appendix A.2 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[15]

$$P_{t,n} = \exp\left(A_n + B_n' X_t + C_n' \text{vec}\left(X_t X_t'\right)\right), \tag{1.9}$$

where

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

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

$$Y_{t,n} = -\frac{1}{n}\log\left(P_{t,n}\right) = -\frac{1}{n}A_n - \frac{1}{n}B_n' X_t - \frac{1}{n}C_n'\text{vec}\left(X_t X_t'\right). \tag{1.10}$$

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, Dittmar, and Gallant (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: Whereas the quadratic bond yield in my 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.

[15] To evaluate the approximation error, I 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 (1.8). Then, this simulated bond price is converted to yields. I repeat this procedure for all months of January in the sample using parameter values reported in Table 1.6. The approximation error is largest for the ten-year yield, for which the average absolute error is 38 basis pointsbasis points corresponding to 6.94 percent of the average ten-year yield level.

1.4.3 Estimation

I estimate $\mu^{\mathbb{Q}}$ and $\Phi^{\mathbb{Q}}$ by non-linear least squares given the parameters obtained for the factor dynamics in Section 1.3, $\hat{\Theta}^{\mathbb{P}} = \{\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\Sigma}_0, \hat{\Sigma}_1, \hat{\Gamma}_1, \hat{\Gamma}_2, \hat{\Gamma}_3\}$.[16] The data are U.S. Treasury bonds yields from Gürkaynak, Sack, and Wright (2007) with maturities $n = 1, 2, \ldots, 10$ years. The results are shown in Table 1.6. The DAR model remains stationary under the $\mathbb{Q}$-measure with a top-Lyapunov exponent of -0.017.

Table 1.6: Estimated $\mathbb{Q}$-Dynamics

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

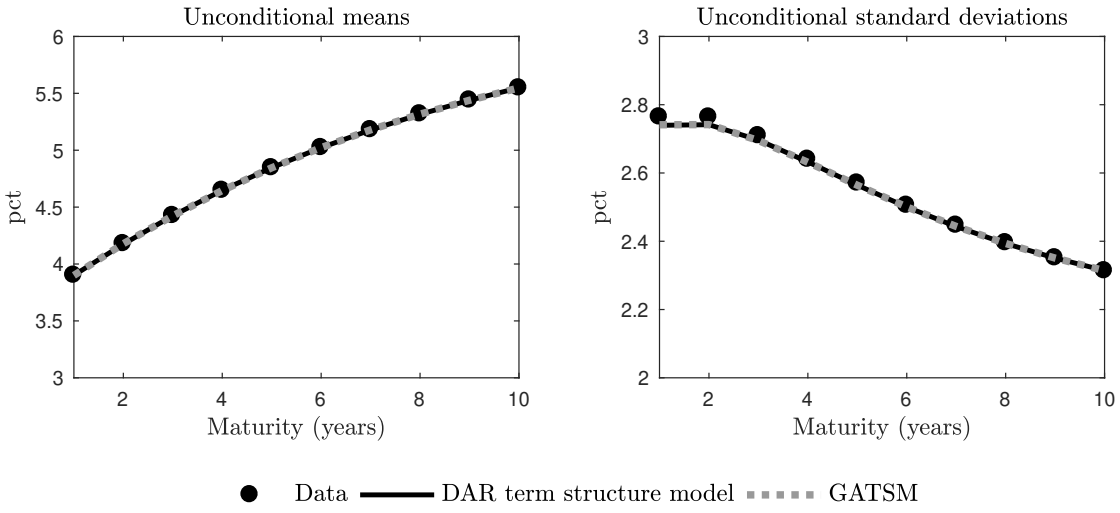
Notes: Estimated parameters related to the $\mathbb{Q}$-dynamics, $\mu^{\mathbb{Q}}$ and $\Phi^{\mathbb{Q}}$, in the DAR term structure model and the GATSM. The parameters are estimated given $\hat{\Theta}^{\mathbb{P}} = \{\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\Sigma}_0, \hat{\Sigma}_1, \hat{\Gamma}_1, \hat{\Gamma}_2, \hat{\Gamma}_3\}$ from Section 1.3. Standard errors are in paranthesis.

1.4.4 In-Sample Fit

Figure 1.7 shows how the DAR term structure model matches the unconditional first and second empirical moments of the yield curve cross-section compared with the GATSM. I observe that the models fit the yield curve and moreover, that the models are practically identical in this respect. To explain this similarity, I compare estimated factor loadings A_n and B_n in Figure 1.8. Indeed, the DAR term structure model does not imply different loadings from those obtained by the GATSM. In Figure 1.9, I plot the component that prices volatility-induced stationarity, $-n^{-1}C_n'$ vec $(X_t X_t')$ from (1.10). This quadratic component is very small and below 0.16 percent at all times. Therefore, the quadratic component generated by volatility-induced stationarity is not priced by the yield curve. Since the time-varying con-

[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, Piazzesi, and Wei (2006), and Wright (2011).

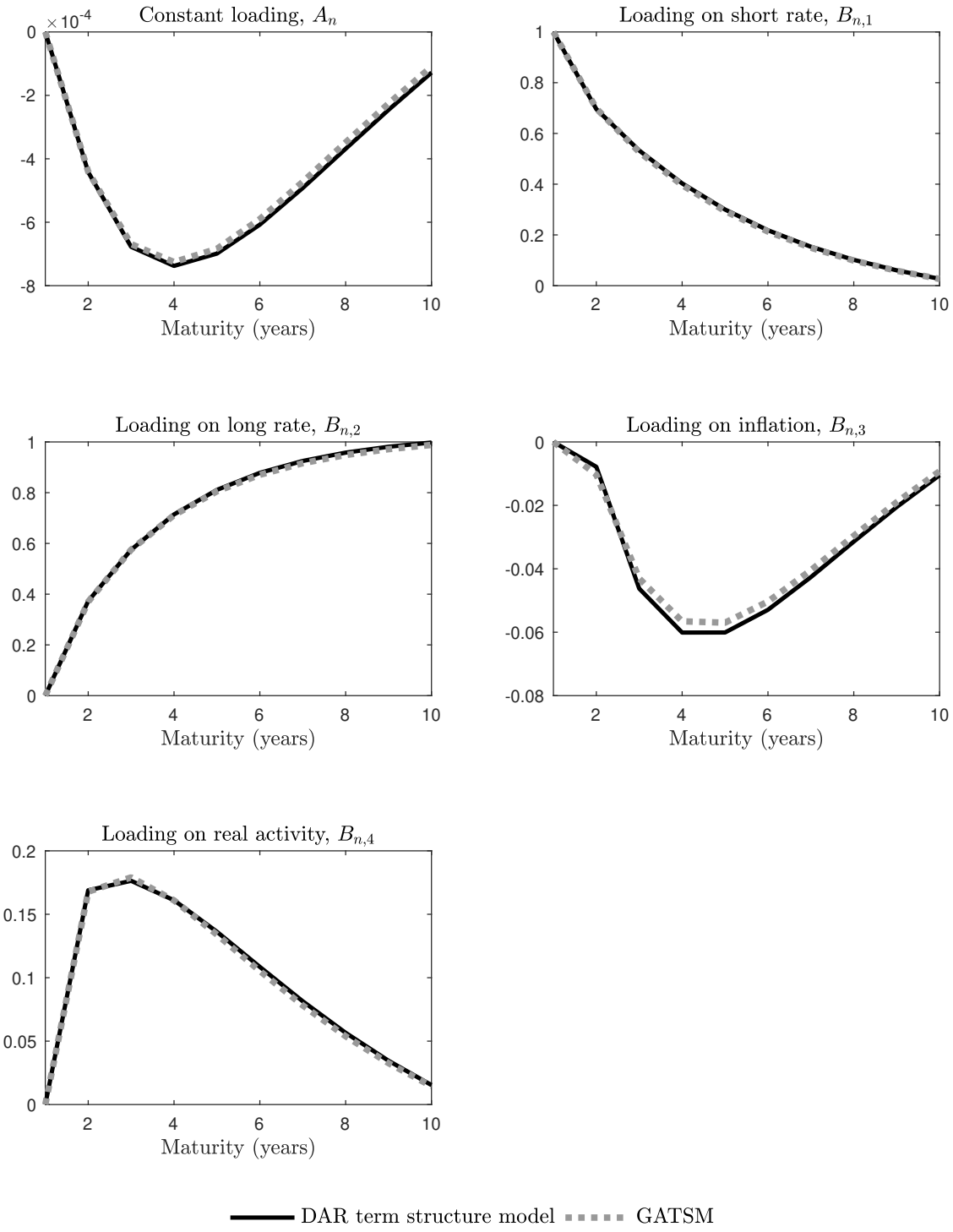
Figure 1.7: In-Sample Yield Curve Fit



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

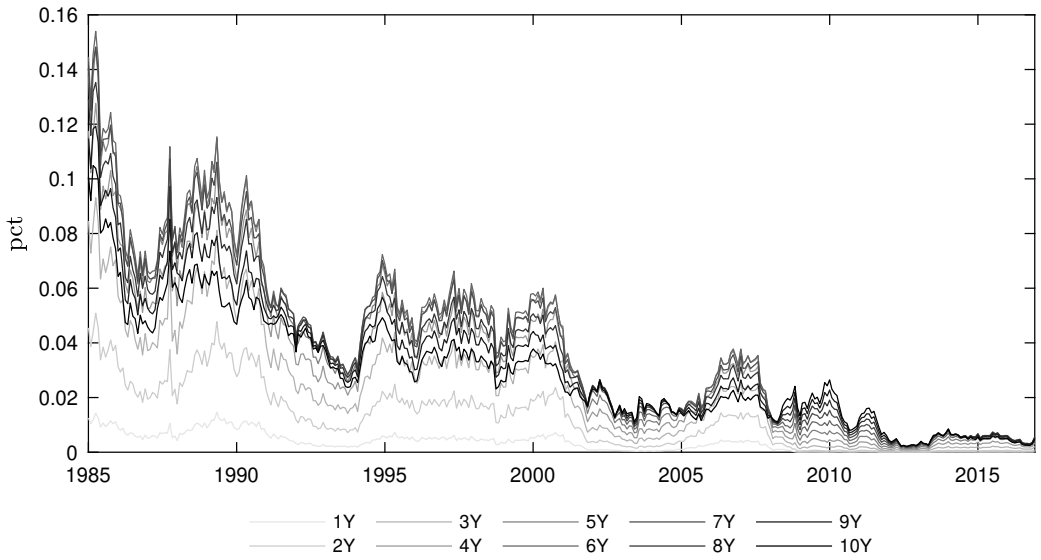
ditional variance of the DAR model only affects bond yields through a convexity effect, this finding is consistent with Joslin and Konchitchki (2018) who show that convexity effects under the $\mathbb{Q}$-measure are small.

Figure 1.8: Factor Loadings



Notes: Model-implied factor loadings, A_n and B_n, in the DAR term structure model and GATSM.

Figure 1.9: Quadratic Component of Model-Implied Bond Yield



Notes: The component $-n^{-1}C_n'\text{vec}\,(X_tX_t')$ from (1.10). Reported in percent.

1.4.5 Out-of-Sample Performance

I assess the out-of-sample performance through a rolling-window forecasting exercise. In particular, I estimate the models with one lag in the factor dynamics as in (1.2), 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. I 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 errors from this exercise are presented in Table 1.7 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.

Table 1.7: 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

Notes: 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 basis points per annum.

1.5 Conclusion

This paper presents 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 my 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. I show that this property is important for decomposing the term structure into expected future short rates and term premia. I embed the DAR model into a no-arbitrage term structure model and provide an approximation for computing model-implied bond yields analytically. Volatility-induced stationarity helps forecasting bond yields. However, compared with 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, my 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.

Acknowledgments

I thank Torben G. Andersen, Luca Benzoni, Laurs Randbøll Leth, Philipp Kless, 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.

Appendix

A.1 Specification and Estimation

I specify the model in (1.4) 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, Rahbek, and Taylor (2014). The test model is specified with a restricted constant, i.e., a constant term appears in the cointegrating relations only.[17] 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 1.3 in the body of the paper. The likelihood-ratio test, for which results are reported in Table A.1, suggests a reduced rank of $r = 2$. I interpret these findings as indications that the DAR model in (1.4) may be well-specified by $r = 2$ and lag length $K = 3$ as well. The left panel of Table 1.3 confirms this presumption.

[17] If a constant is unrestricted, the cointegrated VAR model implies that the data contains a linear trend (Johansen, 1995).

Table A.1: 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]

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

I estimate (1.4) with $r = 2$ and $K = 3$ by maximum likelihood under just-identifying restrictions. Then, I 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 A.3 and A.2.

Table A.2: Parameter Estimates of the Volatility Dynamics

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

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

Table A.3: 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)

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

A.2 Approximation of No-Arbitrage Bond Yields

Define $\epsilon_t^{\mathbb{Q}} = \Omega_t^{1/2}\varepsilon_t^{\mathbb{Q}}$. From the factor dynamics under the $\mathbb{Q}$-measure in (1.8), write

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

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

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

Next, I 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

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

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

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

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

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

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

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

Gathering terms result in factor loading recursions given by

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

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

Chapter 2

Yield Curve Volatility and Macroeconomic Risk

Anne Lundgaard Hansen
University of Copenhagen

I show that the relationship between the U.S. Treasury yield curve and macroeconomic risk fluctuates over time. I establish this result by introducing time-varying volatility and variance risk premia in a tractable term structure model. Based on my model, I characterize the joint behavior of the yield curve and macroeconomic risk captured by inflation and unemployment gap from 1971 to 2019. First, I find that the macroeconomic contribution to short-term yield volatility is high in the 1970s, low during the Great Moderation starting in the mid-1980s, and high again after the financial crisis. Second, investors are increasingly anchoring short-rate expectations to macroeconomic risk after the Great Moderation. Third, deflation fears increase term premia during the financial crisis. Finally, I show that macroeconomic shocks do not explain the yield curve inversion in 2019. My results suggest that the recent inversion is not a warning of an imminent recession and thus should not trigger monetary policy easing.

Keywords: Yield curve, macro-finance term structure model, bond market volatility, multivariate GARCH, variance risk premia.

JEL classification: C32, E43, E44, G12.

2.1 Introduction

Policy makers and investors pay close attention to the joint dynamics of the yield curve and macroeconomic variables. These dynamics can, for example, shed light upon how the yield curve is affected by the Federal Reserve's dual mandate of maximum sustainable employment and stable prices. The relationship between macroeconomic risk and the decomposition of yields into short-rate expectations and term premia is also useful for understanding the sources of movements in the yield curve. In light of the recent sharp decline in long-term yields, it is of particular interest to recognize if this development reflects low expectations on future economic activity and, consequently, is likely to anticipate a recession.

I focus on the question, how much variance in the yield curve, short-rate expectations, and term premia can be attributed to macroeconomic risk? I contribute to a large body of literature that uses dynamic term structure models with observed macroeconomic variables to decompose variances into macroeconomic and latent yield-curve-specific shares. Starting with Ang and Piazzesi (2003), the existing literature is centered around Gaussian term structure models (Bikbov and Chernov, 2010, Doshi, Jacobs, and Liu, 2018, Duffee, 2018). These models are celebrated for their tractability and ability to match the yield curve. However, they assume that yield curve volatility is constant, which contradicts the empirical evidence. My paper studies whether time-varying yield curve volatility has implications for the joint behavior of the yield curve and macroeconomic variables.

I build a novel term structure model that has time-varying volatility and variance risk premia, unlike the Gaussian term structure model. Specifically, I bridge the Baba-Engle-Kraft-Kroner (BEKK) multivariate GARCH model proposed by Engle and Kroner (1995) with the exponential-quadratic pricing kernel of Monfort and Pegoraro (2012). My model admits closed-form solutions for no-arbitrage bond yields and their decomposition into short-rate expectations and term premia. I match the U.S. Treasury yield curve at the monthly frequency from 1971 to 2019 using inflation and unemployment gap as macroeconomic variables and three latent yield-curve-specific factors. Model-implied conditional variances and covariances capture realized and rolling measures. In contrast, continuous-time term structure models with stochastic volatility struggle to match empirical volatility proxies with low and often negative correlation between predicted and realized volatilities (Christensen, Lopez, and Rudebusch, 2014, Collin-Dufresne, Goldstein, and Jones, 2009, Jacobs and Karoui, 2009). Finally, my model implies excess returns that explain the empirical failure of the expectations hypothesis. Thus, my model overcomes a trade-off in

the existing literature between modeling time-varying volatility and excess returns as documented by Dai and Singleton (2002).

By embracing the stylized fact that yield curve volatility is time-varying, I show that the relationship between the yield curve and macroeconomic variables fluctuates over time. Gaussian term structure models imply a constant relationship. I find that macroeconomic shocks explain more than half of the variation in yields in some periods, but that yield curve volatility is unrelated to macroeconomic risk in other periods. I argue that large month-to-month fluctuations are related to economic events, for example, the announcement of quantitative easing programs. Thus, macroeconomic news partly drives movements in the yield curve, which is consistent with Andersen, Bollerslev, Diebold, and Vega (2007), Feunou, Fontaine, and Roussellet (2019), and Piazzesi (2005).

I present a novel dynamic characterization of the historical joint behavior of the yield curve and macroeconomic variables. First, I show that the macroeconomic contribution to short-term yield variance has followed a U-shaped pattern since the 1970s. These macroeconomic shares were high in the 1970s, low during the Great Moderation starting from the mid-1980s, and high following the financial crisis. Second, I find an upward trend in the macroeconomic contribution to variance in 10-year short-rate expectations since the Great Moderation. Thus, investors increasingly form expectations on future short rates based on macroeconomic risk. This result possibly reflects a growing belief among investors that the Federal Reserve's commitment to the dual mandate is credible. Third, the macroeconomic share of variance in the 10-year term premium increases during the financial crisis due to inflation shocks. Thus, inflation risk premia are high despite a low-inflation environment. This result is generated by the exponential-quadratic pricing kernel as it allows investors to demand compensation for both positive and negative shocks to inflation as also shown by Roussellet (2018). As an implication, a symmetric inflation target enhances the efficiency of monetary policy during recessions.

In terms of methodology, I show that my model can be estimated by a step-wise approach without relying on filtering methods for estimating the latent yield-curve-specific factors. Specifically, my model is invariant to affine transformations such that latent factors can be rotated into portfolios of synthetic yields. The synthetic yields are constructed by residuals from yields explained with macroeconomic variables only. My estimation method is a generalization of Joslin, Singleton, and Zhu (2011), who provide similar arguments for the Gaussian term structure model. My method is most closely related to Ghysels, Le, Park, and Zhu (2014), who consider a model that differs from my model in two respects. First, my model is consistent

with the view that volatility is not priced in Treasury bonds (Andersen and Benzoni, 2010, Collin-Dufresne and Goldstein, 2002, Joslin, 2017). In contrast, Ghysels et al. let volatility affect yields in the spirit of the ARCH-in-mean model in Engle, Lilien, and Robins (1987). Second, I allow for both latent factors and observed macroeconomic variables, whereas Ghysels et al. use only latent factors. Other macro-finance term structure models with GARCH-type volatility in the literature do not allow for a rotation of the latent factors (Campbell, Sunderam, and Viceira, 2017, Haubrich, Pennacchi, and Ritchken, 2012, Koeda and Kato, 2015).

The introduction of time-varying volatility in the yield curve has two implications for model-implied risk compensation in fixed-income markets. First, my model decomposes long-term yields into persistent short-rate expectations and countercyclical term premia. The cyclicality is stronger than the term premia implied by the Gaussian term structure model. This difference arises because my model involves both a time-varying price and quantity of risk, whereas the Gaussian model only allows for variation in the price of risk. As an implication, the models disagree about the effectiveness of forward guidance. Specifically, my model is consistent with effective forward guidance as seen in Carlstrom, Fuerst, and Paustian (2015) and McKay, Nakamura, and Steinson (2016), but unlike Hagedorn, Luo, Manovskii, and Mitman (2019).

Second, I find that investors are willing to pay large variance risk premia to hedge macroeconomic uncertainty. My result indicates that macroeconomic uncertainty may increase trading activity in fixed-income derivative markets. To the extent that the policy rate is determined by the dual mandate, my results complement Cieslak and Povala (2016) who show that fixed-income variance risk premia are particularly related to uncertainty about future monetary policy. My term structure of estimated variance risk premia is downward-sloping in magnitudes, which is consistent with the literature (Choi, Mueller, and Vedolin, 2017, Trolle and Schwartz, 2009, Trolle and Schwartz, 2015). Gaussian term structure models abstract from the presence of variance risk premia by imposing exponential-affine pricing kernels that only allow investors to demand compensation for mean-based risk.

Finally, I show that macroeconomic variables did not drive the movements in the yield curve in the spring 2019. Specifically, I focus on the 50-basis-point decline in the 10-year yield from April to June. This decline led to an inversion of the yield curve, with the 10-year yield below the 3-month yield, in May. Yield curve inversions are sensational because history shows that inversions are strong predictors of recessions. The predictive power reflects that when investors expect a slowing economy, short-rate expectations and hence long-term yields decline. Thus, if the yield

curve inversion is likely to predict a recession, movements in short-rate expectations should be driven by macroeconomic variables. My model allows me to zoom in on the relationship between the yield curve and macroeconomic variables specifically during spring 2019. In contrast, the Gaussian term structure model can only draw conclusions that are valid on average across the full sample. I begin by showing that the yield curve inversion in August 2006 that preceded the financial crisis was related to macroeconomic risk through short-rate expectations. Following this, I show that the recent yield curve inversion is mainly driven by term premia and not short-rate expectations. Furthermore, I show that the relationship between short-rate expectations and macroeconomic risk weakened during the spring 2019. My results suggest that the recent yield curve inversion is not a warning of an imminent recession. Thus, the Federal Reserve should not launch a monetary policy easing cycle based on the inversion. This finding is consistent with the current strength in the U.S. economy. For example, in May 2019 the unemployment rate had been at or below 4 percent for over a year.

The remainder of this paper is organized as follows. First, I present my term structure model with time-varying volatility and variance risk premia in Section 2.2. Section 2.3 provides a tractable method for estimating my model. Section 2.4 details the data and discusses the empirical performance of my model. I analyze the relationship between yield curve volatility and macroeconomic risk in Section 2.5. Finally, I apply my model to study the 2019 yield curve inversion in Section 2.6. Conclusions follow in Section 2.7.

2.2 Term Structure Model

In this section, I present my term structure model. My model is distinct from the Gaussian term structure model in two dimensions: First, the dynamics under the physical probability measure exhibit multivariate GARCH volatility. Second, the pricing kernel is exponential-quadratic allowing investors to demand a compensation for exposure to conditional volatility. In the following, I describe these elements separately.

2.2.1 Physical Dynamics

The yield curve is driven by a n_X-dimensional state vector, X_t, consisting of n_x latent yield-curve-specific factors, x_t, and n_m observed macroeconomic variables, m_t. Thus, the state vector is $X_t = [x_t', m_t']'$. Assuming that x_t can be filtered from observed bond yields, let $\mathcal{F}_t$ be the filtration given by $(X_t, X_{t-1}, \ldots)$.

Conditional Mean

The physical dynamics of X_t are given by a vector autoregressive model allowing for some lag length, L, in the equation for the macroeconomic variables, m_t. Following the existing literature, latent factors have one lag.[18] The equations for x_t and m_t are given by

$$x_t = \mu_x + \Phi_x^{(1)} x_{t-1} + \Phi_{xm}^{(1)} m_{t-1} + \varepsilon_{x,t}, \tag{2.1}$$

$$m_t = \mu_m + \Phi_{mx}^{(1)} x_{t-1} + \Phi_m^{(1)} m_{t-1} + \Phi_m^{(2)} m_{t-2} + \ldots + \Phi_m^{(L)} m_{t-L} + \varepsilon_{m,t}, \tag{2.2}$$

where $\varepsilon_{x,t}$ and $\varepsilon_{m,t}$ are conditionally Gaussian given $\mathcal{F}_{t-1}$ with mean zero and conditional variance matrix, V_t.[19] Equivalently, I write

$$X_t = \mu_X + \Phi_X^{(1)} X_{t-1} + \ldots + \Phi_X^{(L)} X_{t-L} + \varepsilon_t \tag{2.3}$$

with appropriate zero restrictions in $\Phi_X^{(l)}$ for $l = 2, \ldots, L$ and $\varepsilon_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, V_t)$.

Conditional Variance Matrix

I model the full variance matrix, V_t, allowing for time-varying conditional covariances between the macroeconomic variables and latent factors. The volatility dynamics are given by the multivariate BEKK GARCH model of Engle and Kroner (1995):

$$V_t = \Sigma_X \Sigma_X' + \sum_{k=1}^{K} A_X^{(k)} \varepsilon_{t-1} \varepsilon_{t-1}' A_X^{(k)'} + \sum_{k=1}^{K} B_X^{(k)} V_{t-1} B_X^{(k)'}. \tag{2.4}$$

If Σ_X is lower triangular such that $\Sigma_X \Sigma_X'$ is positive definite, then V_t is also positive definite. Importantly, the model in (2.4) is invariant to affine transformations. This is a non-trivial property that many multivariate GARCH models do not satisfy.[20] I use this property to propose a tractable estimation method in Section 2.3.

[18] The model can easily be extended to a full vector autoregressive model of order L. However, allowing for a lag structure in the dynamics of x_t is costly in terms of parameters. Indeed, I show in Appendix B.2.2 that information criteria prefer models with multiple lags in the equation for m_t only.

[19] It should be noted that V_t denotes the conditional volatility of a process at time t that is adapted to $\mathcal{F}_{t-1}$. This choice follows the GARCH literature, but differs from the literature on stochastic volatility.

[20] For example, GARCH models with asymmetric components and the dynamic conditional correlations model of Engle (2002) are not invariant to affine transformations.

With multiple ARCH and GARCH components, $K > 1$, I allow for rich dynamics of conditional variances and covariances which are typically needed empirically.[21] When K is sufficiently large, the model can obtain a representation equivalent to the vech-GARCH model of Bollerslev, Engle, and Wooldridge (1988).

The full BEKK GARCH model, i.e., the model where $A_X^{(k)}$ and $B_X^{(k)}$ are fully parameterized matrices for some k, suffers from many econometric problems including the curse of dimensionality (Chang and McAleer, 2019). I therefore impose diagonal restrictions on $A_X^{(k)}$ and $B_X^{(k)}$ for all $k = 1, \ldots, K$.

2.2.2 Pricing Kernel

I relate the state vector to no-arbitrage bond yields by specifying the pricing kernel. Standard term structure models use exponential-affine pricing kernels, see, e.g., Le, Singleton, and Dai (2010). However, recent work suggests that higher-order pricing kernels are needed for capturing the conditional volatility of the yield curve (Creal and Wu, 2017, Ghysels, Le, Park, and Zhu, 2014, Joslin and Konchitchki, 2018). Based on these results, I apply the exponential-quadratic kernel in Monfort and Pegoraro (2012) given by

$$\mathcal{M}_{t+1} = \frac{\exp\left(-r_t + \xi_t' X_{t+1} + X_{t+1}' \Xi_t X_{t+1}\right)}{\mathbb{E}_t\left[\exp\left(\xi_t' X_{t+1} + X_{t+1}' \Xi_t X_{t+1}\right)\right]}, \tag{2.5}$$

where r_t is the short rate.

I show in Technical Appendix B.1.1 that exponential-quadratic pricing kernels can be structurally justified in preference-based models. Specifically, the long-run risk model of Bansal and Yaron (2004) implies an exponential-quadratic pricing kernel when solved by the second-order projection developed by Andreasen and Jørgensen (2019). In addition, Hansen and Heaton (2008) propose a long-run risk model based on vector autoregressive dynamics that implies an exponential-quadratic pricing kernel when allowing for a time-varying wealth-consumption ratio.

[21] To illustrate this notion, consider a bivariate model with $K = 1$ in which $A_X^{(1)}$ and $B_X^{(1)}$ are diagonal matrices with elements a_{ii} and b_{ii}, $i = 1, 2$. Then,

$$V_t = \Sigma_X \Sigma_X' + \begin{bmatrix} a_{11}^2 \varepsilon_{1,t-1}^2 & a_{11} a_{22} \varepsilon_{1,t-1} \varepsilon_{2,t-1} \\ a_{11} a_{22} \varepsilon_{1,t-1} \varepsilon_{2,t-1} & a_{22}^2 \varepsilon_{2,t-1}^2 \end{bmatrix} + \begin{bmatrix} b_{11}^2 V_{11,t-1} & b_{11} b_{22} V_{12,t-1} \\ b_{11} b_{22} V_{12,t-1} & b_{22}^2 V_{22,t-1} \end{bmatrix}.$$

The parameters a_{ii} and b_{ii} are the ARCH and GARCH effects related to variable i. However, these parameters also determine the conditional covariance between the variables. Thus, there is a tension between modeling conditional variances and covariances simultaneously. This problem is also present when $A_X^{(1)}$ and $B_X^{(1)}$ are full matrices.

The pricing kernel in (2.5) contains three components. The first component is a time discount, which is determined by the short rate. Following the literature, I impose a short-rate model that is affine in X_t:

$$r_t = \alpha_X + \beta_x' x_t + \beta_m' m_t = \alpha_X + \beta_X' X_t. \tag{2.6}$$

The remaining components represent compensation for mean-based risk, $\xi_t' X_{t+1}$, and variance-based risk, $X_{t+1}' \Xi_t X_{t+1}$. Thus, the exponential-quadratic pricing kernel allows investors to be averse towards both mean- and variance-based risk. The variables ξ_t and Ξ_t can be interpreted as market prices of risk associated with, respectively, the conditional mean and variance matrix of X_{t+1}. In this sense, the variances of X_{t+1} can be interpreted as quantities of risk. If $\Xi_t = 0$, the pricing kernel in (2.5) reduces to the standard exponential-affine pricing kernel. Thus, exponential-affine pricing kernels only allow investors to price mean-based risk.

Monfort and Pegoraro (2012) show that if the market prices of risk are chosen by

$$\xi_t = (V_{t+1}^{\mathbb{Q}})^{-1} \mathbb{E}_t^{\mathbb{Q}}(X_{t+1}) - V_{t+1}^{-1} \mathbb{E}_t(X_{t+1}), \tag{2.7}$$

$$\Xi_t = \frac{1}{2} \left(V_{t+1}^{-1} - (V_{t+1}^{\mathbb{Q}})^{-1} \right), \tag{2.8}$$

where $\mathbb{E}_t^{\mathbb{Q}}(X_{t+1})$ and $V_{t+1}^{\mathbb{Q}}$ denote the first and second conditional moments of X_{t+1} given $\mathcal{F}_t$ under the risk-neutral probability measure ($\mathbb{Q}$), then $X_{t+1}|\mathcal{F}_t \overset{\mathbb{Q}}{\sim} \mathcal{N}(\mathbb{E}_t^{\mathbb{Q}}(X_{t+1}), V_{t+1}^{\mathbb{Q}})$. Thus, my model allows the risk-neutral conditional variance matrix, $V_t^{\mathbb{Q}}$, to differ from the physical conditional variance matrix, V_t. The discrepancy between the conditional variance under the physical and risk-neutral probability measures is feasible in discrete-time models as they are not restricted by Girsanov's theorem. In contrast, Girsanov's theorem rules out distinct volatility specifications across probability measures in continuous-time models. Specifically, it follows from the market price of variance risk, Ξ_t, that it is the introduction of variance-based risk compensation that facilitates different variance specifications under the physical and risk-neutral probability measures.

The market price of mean-based risk in (2.7) is akin to that of the standard term structure model with exponential-linear pricing kernel. Specifically, it is given by the differences in the conditional means under physical and risk-neutral probability measures weighted by a conditional variance matrix. In contrast to the standard model, however, I weigh the physical and risk-neutral expectations by different variance matrices, accounting for the different variances across probability measures. The market price of variance-based risk in (2.8) is diminishing in the quantity of

risk under the physical measure, V_{t+1}. This property is consistent with the empirical fact that fixed-income variance risk premia are diminishing in magnitudes with the maturity of bonds (Choi, Mueller, and Vedolin, 2017, Trolle, 2009, Trolle and Schwartz, 2015).

In the following, I specify the risk-neutral moments which pin down the market prices of risk through (2.7)-(2.8). This approach contrasts with the standard literature in which the market prices of risk are specified to pin down the risk-neutral moments.

2.2.3 Risk-Neutral Moments

I specify moments under the risk-neutral probability measure such that my model is tractable and sufficiently flexible to match the yield curve in sample. An autoregressive model of order one with constant volatility satisfies both of these criteria (Joslin, Le, and Singleton, 2013b). I also impose independence of the macroeconomic variables and the latent yield-curve-specific factors under the risk-neutral measure. Intuitively, I allow for interdependent dynamics of macroeconomic variables and latent factors through the physical dynamics, but bond market investors do not price these interactions. Failure of this assumption affects the efficiency, but not the consistency, of my proposed estimator. Thus, the risk-neutral conditional mean and variance matrix are given by

$$\mathbb{E}_t^{\mathbb{Q}}\left(X_{t+1}\right) = \begin{bmatrix} \mu_x^{\mathbb{Q}} \\ \mu_m^{\mathbb{Q}} \end{bmatrix} + \begin{bmatrix} \Phi_x^{\mathbb{Q}} & 0 \\ 0 & \Phi_m^{\mathbb{Q}} \end{bmatrix} \begin{bmatrix} x_t \\ m_t \end{bmatrix} = \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t, \tag{2.9}$$

$$V_{t+1}^{\mathbb{Q}} = \begin{bmatrix} V_x^{\mathbb{Q}} & 0 \\ 0 & V_m^{\mathbb{Q}} \end{bmatrix} = V_X^{\mathbb{Q}}, \tag{2.10}$$

where $V_X^{\mathbb{Q}}$ is positive definite. The assumption of constant conditional variances and covariances under the risk-neutral measure is likely to be violated. However, recent literature indicates that a violation is unlikely to affect the pricing of bonds. First, Cieslak and Povala (2016) build a flexible term structure model with stochastic volatility under both the physical and risk-neutral probability measures. As a result, no-arbitrage bond yields depend directly on volatility. However, they empirically show that the loading on volatility is close to zero. This result is also supported by a large body of literature that argues that bond market volatility is not spanned by bond yields (Andersen and Benzoni, 2010, Collin-Dufresne and Goldstein, 2002, Joslin, 2017). This phenomenon is known as unspanned stochastic volatility. Second, Joslin and Konchitchki (2018) show that convexity effects un-

der the risk-neutral measure are small. Therefore, the conditional variance under the risk-neutral measure has only small effects on bond pricing.

2.2.4 No-Arbitrage Bond Pricing

The no-arbitrage bond yields implied by my model are equal to those of the Gaussian term structure model. These are given in the following theorem.

Theorem 2.1 *Let $P_{t,n}$ denote the no-arbitrage price of a n-period zero-coupon bond. Let $Y_{t,n} = -n^{-1}\log(P_{t,n})$ be the associated yield. Given* (2.6)-(2.10)*, $Y_{t,n}$ is given in closed form by*

$$Y_{t,n} = a_n + b'_{x,n}x_t + b'_{m,n}m_t = a_n + b'_n X_t, \tag{2.11}$$

where $a_n = -n^{-1}\tilde{a}_n$ and $b_{i,n} = -n^{-1}\tilde{b}_{i,n}$ for $i = \{x, m\}$ are given recursively by

$$\tilde{a}_{n+1} = -\alpha_X + \tilde{a}_n + \tilde{b}'_{x,n}\mu_x^{\mathbb{Q}} + \tilde{b}'_{m,n}\mu_m^{\mathbb{Q}} + \frac{1}{2}\tilde{b}'_{x,n}V_x^{\mathbb{Q}}b_{x,n} + \frac{1}{2}\tilde{b}'_{m,n}V_m^{\mathbb{Q}}b_{m,n}, \tag{2.12}$$

$$\tilde{b}'_{x,n+1} = -\beta'_x + \tilde{b}'_{x,n}\Phi_x^{\mathbb{Q}},, \tag{2.13}$$

$$\tilde{b}'_{m,n+1} = -\beta'_m + \tilde{b}'_{m,n}\Phi_m^{\mathbb{Q}}, \tag{2.14}$$

The recursions are initiated at $n = 0$ with $a_0 = 0$, $b_{x,0} = 0_{n_x\times 1}$, and $b_{m,0} = 0_{n_m\times 1}$.

Proof. The solution follows from the no-arbitrage recursion,

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

see Ang and Piazzesi (2003). □

Remark 2.1 *I note that the conditional variance matrix, V_t, does not appear in the bond pricing equation,* (2.11)*. Thus, my model exhibits a form of unspanned volatility, but without introducing restrictions as in the traditional literature on unspanned stochastic volatility.*[22]

Remark 2.2 *It follows from* (2.11) *that conditional yield variance is given by*

$$Var_t(Y_{t+1,n}) = b'_n V_{t+1} b_n. \tag{2.15}$$

[22]Joslin (2017) argues that traditional USV models restrict the cross-section of yield curve volatility in a way that is inconsistent with the data.

2.2.5 Risk Compensations

My model distinguishes between two concepts of risk compensation. First, term premia, $\Psi_{t,n}$, are defined by the difference between bond yields and expected future short rates, $\Upsilon_{t,n}$:

$$\Psi_{t,n} = Y_{t,n} - \Upsilon_{t,n}, \qquad \Upsilon_{t,n} = \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}_t\left(r_{t+i}\right). \tag{2.16}$$

The following theorem provides closed-form expressions of short-rate expectations and term premia.

Theorem 2.2 *Given the physical conditional mean of X_t in* (2.3) *and the assumptions of Theorem 2.1, short-rate expectations, $\Upsilon_{t,n}$, and term premia, $\Psi_{t,n}$, are given by*

$$\begin{aligned}
\Upsilon_{t,n} &= a_n^{EH} + (b_n^{EH})'\mathcal{X}_t, \\
\Psi_{t,n} &= \left(a_n - a_n^{EH}\right) + b_n' X_t - (b_n^{EH})'\mathcal{X}_t.
\end{aligned}$$

where $\mathcal{X}_t = [X_t,\ X_{t-1},\ \ldots, X_{t-L+1}]$ and

$$\begin{aligned}
a_n^{EH} &= \alpha_X + \frac{1}{n}\sum_{i=0}^{n-1}\sum_{j=0}^{i-1} \beta_{\mathcal{X}}'\Phi_{\mathcal{X}}^j \mu_{\mathcal{X}}, \\
b_n^{EH} &= \frac{1}{n}\sum_{i=0}^{n-1} \beta_{\mathcal{X}}'\Phi_{\mathcal{X}}^i.
\end{aligned}$$

Here, $\beta_{\mathcal{X}} = [\beta_x', \beta_m', 0_{1\times(p+m)(L-1)}]'$ and $\mu_{\mathcal{X}}$ and $\Phi_{\mathcal{X}}$ denote the parameters of the conditional mean associated with the companion form of (2.3).

Proof. Straightforward. □

Remark 2.3 *It follows that the conditional variances of short-rate expectations and term premia are given in closed form by*

$$\begin{aligned}
Var_{t-1}(\Upsilon_{t,n}) &= (b_n^{EH})' Var_{t-1}\left(\mathcal{X}_t\right) b_n^{EH} \\
Var_{t-1}(\Psi_{t,n}) &= \left(\check{b}_n - b_n^{EH}\right)' Var_{t-1}\left(\mathcal{X}_t\right)\left(\check{b}_n - b_n^{EH}\right),
\end{aligned}$$

where $\check{b}_n = (b_n', 0, 0, \ldots, 0)'$.

The second concept of risk compensation relates to variance risk. While V_t is not priced in bonds, see Theorem 2.1, investors may demand compensation for variance risk through fixed-income derivatives. This compensation is realized in the broader fixed-income market through the pricing of interest-rate derivatives. To analyze these premia and their relation to macroeconomic variables, I follow Bollerslev, Tauchen, and Zhou (2009) and consider the model-implied variance risk premium given by the difference between the expected variance under physical and risk-neutral probabilities. The next theorem provides a closed-form solution for the variance risk premium.

Theorem 2.3 *Variance risk premia defined by*

$$\Gamma_{t,n} = \mathbb{E}_t \left[Var_{t+1}(Y_{t+2,n}) \right] - \mathbb{E}_t^{\mathbb{Q}} \left[Var_{t+1}(Y_{t+2,n}) \right]$$

are given by

$$\Gamma_{t,n} = b_n' \left(\sum_{k=1}^{K} \left[A_X^{(k)} V_{t+1} A_X^{(k)'} - A_X^{(k)} V_X^{\mathbb{Q}} A_X^{(k)'} \right] \right) b_n.$$

Proof. Straightforward. □

Remark 2.4 *It follows that $\mathbb{E}_t^{\mathbb{Q}} \left[Var_{t+1}(Y_{t+2,n}) \right]$ is constant and the dynamics of my variance risk premium are determined solely from the physical dynamics.*

Due to the above remark, I abstract from analyzing time-variation in variance risk premia in my empirical application. Instead, I consider averages, which I believe are accurately estimated by my model.

2.3 Econometric Method

In this section, I present a simple approach to estimating my model. Specifically, I extend the ideas proposed by Joslin, Singleton, and Zhu (2011) for the Gaussian term structure model with latent yield-curve-specific factors only.

2.3.1 Estimation Problem

Consider bond yields of N different maturities given by $n_1, \ldots, n_N$ periods. Let $Y_t = (Y_{t,n_1}, \ldots, Y_{t,n_N})'$ denote a vector of these observations. I impose the following assumption on the accuracy of yield measurements:

Assumption 1 *Yields, Y_t, are measured with errors given by $\eta_t \sim$ i.i.d. $\mathcal{N}(0, \sigma^2 I_N)$.*

I identify latent yield-curve-specific factors, x_t, such that they capture yield curve variation that is not related to the macroeconomic variables. For this purpose, I decompose model-implied yields given in Theorem 2.1 by

$$Y_{t,n} = a_{m,n} + b'_{m,n} m_t + Y_{t,n}^{\perp m}, \tag{2.17}$$

where $Y_{t,n}^{\perp m} = Y_{t,n} - a_{m,n} - b'_{m,n} m_t$ and $a_{m,n} = n^{-1}\tilde{a}_{m,n}$ with

$$\tilde{a}_{m,n} = \tilde{a}_{m,n-1} + \tilde{b}'_{m,n}\mu_m^{\mathbb{Q}} + \frac{1}{2}\tilde{b}'_{m,n} V_m^{\mathbb{Q}} \tilde{b}_{m,n}.$$

The synthetic yields, $Y_{t,n}^{\perp m}$, capture the variation in the yield curve that cannot be explained by the macroeconomic variables. I therefore extract latent yield-curve-specific factors from synthetic yields.

I propose a step-wise estimation method to implement this strategy. A first step estimates parameters related to pricing bonds with macroeconomic variables, $\Theta_m^{\mathbb{Q}} = \{\beta_m, \mu_m^{\mathbb{Q}}, \Phi_m^{\mathbb{Q}}, V_m^{\mathbb{Q}}\}$. The marginal likelihood function to be maximized is given by

$$\log \ell_Y \left(Y_t | m_t; \Theta_m^{\mathbb{Q}}\right) = -\sum_{i=1}^{N} \left(Y_{t,n_i} - a_{m,n_i} - b_{m,n_i} m_t\right)^2.$$

I use these estimated parameters to construct estimates of the synthetic yields, $Y_{t,n}^{\perp m}$.

The remaining parameters can be estimated in a second step by the standard approach, combining maximum-likelihood estimation with a Kalman filter and parameter restrictions for statistical identification. This approach is, however, cumbersome because my model encounters non-linearities in both the conditional variance matrix, V_t, and in the loading recursions, a_n and $b_{x,n}$. In the remainder of this section, I provide a simpler approach that separates the estimation problem such that non-linearities in the physical dynamics are treated separately from the non-linear bond yield equation. Furthermore, my approach avoids Kalman filtering and hence the numerical stability issues that are likely to be associated with filtering methods.

2.3.2 Rotation and Identifying Restrictions

My model is invariant to affine transformations of the latent factors given by $\rho_t = c_x + C_x x_t$, see Lemma 2.1 in Technical Appendix B.1.2. Therefore, the latent factors can be rotated into linear combinations, or portfolios, of the synthetic yields, $Y_{t,n}^{\perp m}$. Let $\rho_t = W Y_t^{\perp m}$ denote a vector of these rotated factors. The matrix of portfolio weights, W, can be chosen freely, e.g., such that the portfolios are the first n_x principal components of the synthetic yield curve. The rotated model is unique given that W has full rank. Theorem 2.4 formalizes these ideas.

Theorem 2.4 *The model defined by* (2.3)-(2.10) *with synthetic yields defined in* (2.17) *is observationally equivalent to a unique model in which the latent factors are portfolios of synthetic yields given by* $\rho_t = WY_t^{\perp m}$*, where the weighting matrix* W *has full rank. Let* $\mathcal{P}_t = (\rho_t', m_t')'$ *denote the resulting observable state vector. Then, this unique model is given by*

$$\mathcal{P}_t = \mu_{\mathcal{P}} + \sum_{l=1}^{L} \Phi_{\mathcal{P}}^{(l)} \mathcal{P}_{t-l} + \varepsilon_{\mathcal{P},t}, \quad \varepsilon_{\mathcal{P},t} = V_t^{1/2} z_{\mathcal{P},t}, \tag{2.18}$$

$$V_t = \Sigma_{\mathcal{P}} \Sigma_{\mathcal{P}}' + \sum_{k=1}^{K} A_{\mathcal{P}}^{(k)} \varepsilon_{t-1} \varepsilon_{t-1}' A_{\mathcal{P}}^{(k)'} + \sum_{k=1}^{K} B_{\mathcal{P}}^{(k)} V_{t-1} B_{\mathcal{P}}^{(k)'}, \tag{2.19}$$

$$\mathcal{P}_t = \mu_{\mathcal{P}}^{\mathbb{Q}} + \Phi_{\mathcal{P}}^{\mathbb{Q}} \mathcal{P}_{t-1} + \varepsilon_{\mathcal{P},t}^{\mathbb{Q}}, \quad \varepsilon_{\mathcal{P},t}^{\mathbb{Q}} = (V_{\mathcal{P}}^{\mathbb{Q}})^{1/2} z_{\mathcal{P},t}^{\mathbb{Q}}, \tag{2.20}$$

$$r_t = \alpha_{\mathcal{P}} + \beta_m' m_t + \beta_\rho' \rho_t \tag{2.21}$$

with $z_{\mathcal{P},t} \sim$ *i.i.d.* $\mathcal{N}(0, I_{n_X})$ *and* $z_{\mathcal{P},t}^{\mathbb{Q}} \sim$ *i.i.d.* $\mathcal{N}(0, I_{n_X})$*.* $V_{\mathcal{P}}^{\mathbb{Q}}$ *is a positive definite matrix.*

Proof. See Technical Appendix B.1.2. □

Remark 2.5 *Given the rotation in Theorem 2.4, model-implied synthetic yields are given by*

$$Y_{t,n}^{\perp m} = a_{\rho,n} + b_{\rho,n}' \rho_t, \tag{2.22}$$

where $a_{\rho,n} = -n^{-1}\tilde{a}_{\rho,n}$ *and* $b_{\rho,n} = -n^{-1}\tilde{b}_{\rho,n}$ *with*

$$\tilde{a}_{\rho,n+1} = -\alpha_{\mathcal{P}} + \tilde{a}_{\rho,n} + \tilde{b}_{\rho,n}' \mu_\rho^{\mathbb{Q}} + \frac{1}{2} \tilde{b}_{\rho,n}' V_\rho^{\mathbb{Q}} b_{\rho,n}, \tag{2.23}$$

$$\tilde{b}_{\rho,n+1}' = -\beta_\rho' + \tilde{b}_{\rho,n}' \Phi_\rho^{\mathbb{Q}}. \tag{2.24}$$

The next step is to impose parameter restrictions such that my model is identified. In a model with a constant conditional variance matrix under the physical measure, these restrictions can be imposed on the risk-neutral dynamics only, see Joslin, Singleton, and Zhu (2011). I adopt these restrictions as formalized in part (i) of Theorem 2.5, which follows next. In my model, however, I need additional restrictions for identifying the physical dynamics. These are given by part (ii) of Theorem 2.5. To state the theorem, I define functions that map a set of parameters into a vector of bond yield loadings by $a(\alpha_{\mathcal{P}}, \beta_\rho, \mu_\rho^{\mathbb{Q}}, \Phi_\rho^{\mathbb{Q}}, V_\rho^{\mathbb{Q}}) = (a_{\rho,n_1}, \ldots, a_{\rho,n_N})'$ and $b(\beta_\rho, \Phi_\rho^{\mathbb{Q}}) = (b_{\rho,n_1}, \ldots, b_{\rho,n_N})'$.

Theorem 2.5 *The model defined by* (2.18)-(2.21) *in Theorem 2.4 with synthetic yields given by* (2.22) *is identified with the following parametrizations:*

(i) The risk-neutral dynamics and the short-rate equation are uniquely parametrized by $\Theta_\rho^{\mathbb{Q}} = \{k_\infty^{\mathbb{Q}}, \lambda^{\mathbb{Q}}, V_\rho^{\mathbb{Q}}\}$*, where* $\lambda^{\mathbb{Q}}$ *contains ordered eigenvalues of the transformed risk-neutral autoregressive coefficient matrix given by* $[Wb(\beta_x, \Phi_x^{\mathbb{Q}})]^{-1}\Phi_\rho^{\mathbb{Q}}[Wb(\beta_x, \Phi_x^{\mathbb{Q}})]$*, c.f., Lemma 2.1 in the Technical Appendix B.1.2. In particular,*

$$\begin{aligned}
\alpha_{\mathcal{P}} &= -\beta_\rho' a_W, \\
\beta_\rho &= (b_W^{-1})'\iota, \\
\mu_\rho^{\mathbb{Q}} &= k_\infty^{\mathbb{Q}} b_W e_1 + (I_{n_x} - \Phi_\rho^{\mathbb{Q}}) a_W, \\
\Phi_\rho^{\mathbb{Q}} &= b_W J(\lambda^{\mathbb{Q}}) b_W^{-1},
\end{aligned}$$

where $e_1 = (1, 0, \ldots, 0)'$ *and* $J(\lambda^{\mathbb{Q}}) = diag(J_1^{\mathbb{Q}}, J_2^{\mathbb{Q}}, \ldots, J_D^{\mathbb{Q}})$*, where* $J_d^{\mathbb{Q}}$ *for* $d = 1, 2, \ldots, D$ *are Jordan blocks. Also,* $a_W = Wa(0, \iota, k_\infty^{\mathbb{Q}} e_1, J(\lambda^{\mathbb{Q}}), b_W^{-1} V_\rho^{\mathbb{Q}} (b_W^{-1})')$ *and* $b_W = Wb(\iota, J(\lambda^{\mathbb{Q}}))$*.*

(ii) For identification of the conditional variance matrix under the physical measure, restrict the diagonal elements of $\Sigma_{\mathcal{P}}$ *to be strictly positive and the first entry of* $A_{\mathcal{P}}^{(k)}$ *and* $B_{\mathcal{P}}^{(k)}$ *for* $k = 1, \ldots, K$ *to be non-negative. Finally, a necessary condition for identification is that* $K \leq floor\left(\frac{1}{2}n_X + \frac{1}{2}\right)$*.*

Proof. See Technical Appendix B.1.3. □

2.3.3 Marginal Likelihood for Synthetic Yields

As in Joslin, Singleton, and Zhu (2011), I impose the following assumption to set up the likelihood function for synthetic yields.[23]

Assumption 2 *The portfolios of synthetic yields,* ρ_t*, balance out the measurement errors such that* ρ_t *is observed without errors.*

With ρ_t and hence $\mathcal{P}_t$ observed, the marginal log-likelihood of the synthetic yields can be separated by

$$\log \ell\left(Y_t^{\perp m} | Y_{t-1}^{\perp m}; \Theta_{\mathcal{P}}^{\mathbb{P}}, \Theta_\rho^{\mathbb{Q}}\right) = \log \ell_Y\left(Y_t^{\perp m} | \mathcal{P}_t; \Theta_\rho^{\mathbb{Q}}\right) + \log \ell_{\mathcal{P}}\left(\mathcal{P}_t | \mathcal{P}_{t-1}; \Theta_{\mathcal{P}}^{\mathbb{P}}\right).$$

[23] Relaxing this assumption requires a filtering technique such as Kalman filtering for estimating the model. Joslin, Singleton, and Zhu (2011) show in the Gaussian term structure model that the parameter estimates do not change significantly when this assumption is relaxed.

I note that the two terms of the log-likelihood contribution do not depend on the same parameters. In particular, the likelihood separation clarifies that $\Theta_\rho^{\mathbb{Q}}$ relates to the cross-section of the synthetic yield curve, while $\Theta_{\mathcal{P}}^{\mathbb{P}}$ governs the dynamics of the state vector. It follows that $\Theta_\rho^{\mathbb{Q}}$ and $\Theta_{\mathcal{P}}^{\mathbb{P}}$ can be estimated by solving two unrelated maximum likelihood problems. The marginal log-likelihood function related to the parameters of the physical dynamics is given up to constants by

$$\begin{aligned}&\log \ell_{\mathcal{P}}\left(\mathcal{P}_t|\mathcal{P}_{t-1};\Theta_{\mathcal{P}}^{\mathbb{P}}\right)\\&=-\frac{1}{2}\log|V_t|-\frac{1}{2}\left(\mathcal{P}_t-\mu_{\mathcal{P}}-\sum_{l=1}^{L}\Phi_{\mathcal{P}}^{(l)}\mathcal{P}_{t-l}\right)'V_t^{-1}\left(\mathcal{P}_t-\mu_{\mathcal{P}}-\sum_{l=1}^{L}\Phi_{\mathcal{P}}^{(l)}\mathcal{P}_{t-l}\right),\end{aligned}$$

where I recall that V_t depends on the parameters and data through (2.19). For the remaining parameters, the marginal log-likelihood function to be maximized is given by

$$\log \ell_Y\left(Y_t^{\perp m}|\mathcal{P}_t;\Theta_\rho^{\mathbb{Q}}\right)=\left(Y_t^{\perp m}-a-b\rho_t\right)'\left(Y_t^{\perp m}-a-b\rho_t\right), \tag{2.25}$$

where $a=a(\alpha_{\mathcal{P}},\beta_\rho,\mu_\rho^{\mathbb{Q}},\Phi_\rho^{\mathbb{Q}},V_\rho^{\mathbb{Q}})$ and $b=b(\beta_\rho,\Phi_\rho^{\mathbb{Q}})$ with $\alpha_{\mathcal{P}}$, β_ρ, $\mu_\rho^{\mathbb{Q}}$, and $\Phi_\rho^{\mathbb{Q}}$ given as explicit functions of $\Theta_\rho^{\mathbb{Q}}=\{k_\infty^{\mathbb{Q}},\lambda^{\mathbb{Q}},V_\rho^{\mathbb{Q}}\}$ in Theorem 2.5. It is noteworthy that the log-likelihood function in (2.25) is identical to that of the Gaussian term structure model. Thus, this part of the estimation is no more difficult to implement than Gaussian term structure models that are celebrated for their tractability.

2.4 Data and Empirical Performance

2.4.1 Data

I apply zero-coupon yields of U.S. Treasury bonds provided by Gürkaynak, Sack, and Wright (2007) at the monthly frequency with end-of-month observations. I consider the sample from September 1971 to June 2019, which contains both periods with highly volatile yields, e.g., in the beginning of the 1980s, and periods with low yield volatility, e.g., before the outbreak of the financial crisis and during the zero-lower bound regime. I include maturities of 1, 2, 3, 5, 7, and 10 years. To capture the short end of the yield curve, I also use the 3- and 6-month Treasury bill rates from Federal Reserve Economic Data.

Following the literature, I include two macroeconomic variables: inflation and economic activity. There is no consensus in the literature about which specific data to use for these variables.[24] Economic activity has been defined by both growth and

[24]For example, Ang and Piazzesi (2003) construct factors from principal components of a wide range of macroeconomic variables; Joslin, Priebsch, and Singleton (2014) use expected inflation measured

slack measures. Bauer and Rudebusch (2016) show that these series are widely different and uncorrelated over the business cycle. They argue that slack variables are appropriate as they are consistent with empirical monetary policy rules. Moreover, they show that the unemployment gap is related to the slope of the yield curve. In light of these results, I adopt the unemployment gap as a measure of economic activity. I construct the unemployment gap by the difference between the unemployment rate from the U.S. Bureau of Labor Statistics and the estimate of natural unemployment from the Congressional Budget Office.[25] I measure inflation by the 12-month change in headline consumer prices from the Bureau of Labor Statistics.

The data are exhibited in Figure 2.1 and the descriptive statistics are detailed in Table 2.1. The yield data are highly persistent for all maturities. They are upward-sloping on average with a decreasing term structure of standard deviations. Yield levels peak in the beginning of 1980 and approach the zero-lower bound in the aftermath of the financial crisis. The unemployment gap is weakly and negatively correlated with the yield curve, while the correlations between inflation and yields are higher and positive. The data are cyclical: Yields, particularly for short-term maturities, and inflation decrease in recessions, while the unemployment gap increases in recessions. The unemployment gap has decreased steadily during the current expansion.

The dual mandate of the Federal Reserve is related to inflation based on personal consumption expenditures (PCE) rather than CPI inflation. Moreover, it has been argued that core inflation, i.e., change in price indices that exclude food and energy prices, provides a better indicator of the path of future inflation compared to headline inflation (Blinder and Reis, 2005, Mishkin, 2007). I therefore provide results using PCE and core CPI inflation in Appendix B.3. I also apply my model using a growth measure of economic activity given by the Chicago Fed National Activity Index. The conclusions do not qualitatively change with these choices of macroeconomic variables.

2.4.2 Model Specification

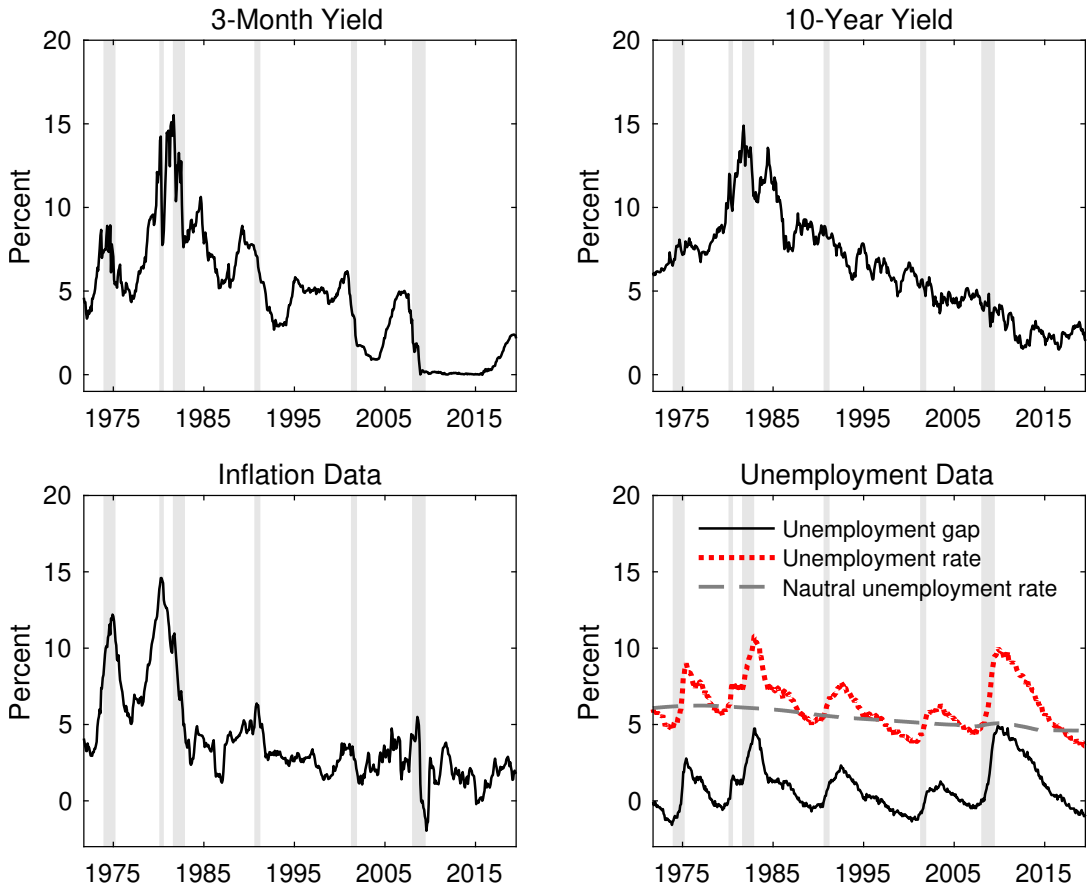
Number of Latent Factors

Since Litterman and Scheinkman (1991), the literature widely agrees that three latent factors are sufficient for modeling the yield curve. The macroeconomic vari-

from surveys along with the Chicago Fed National Activity Index, which is an estimate of overall economic growth; Bikbov and Chernov (2010) use CPI inflation and the Help Wanted Advertising in Newspapers index.

[25]The natural unemployment estimate is available at the quarterly frequency only. I assume that the natural unemployment rate is constant over the quarters.

Figure 2.1: Yield and Macroeconomic Data



Notes: The figure shows 3-month and 10-year bond yields (upper panel) and the macroeconomic data (lower panel). The yields are from U.S. Treasury bills and bonds provided by Federal Reserve Economic Data and Gürkaynak, Sack, and Wright (2007). Inflation is the 12-month change in consumer prices provided the Bureau of Labor Statistics. Unemployment gap is the unemployment rate from the Bureau of Labor Statistics minus the natural unemployment rate from the Congressional Budget Office. The data are from September 1971 to June 2019. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

ables can potentially reduce the number of latent factors necessary for capturing the residual variation in the synthetic yields. However, I find that three latent factors remain necessary.

As shown in Theorem 2.4, the latent factors can be rotated into observed portfolios of synthetic yields, $\rho_t = WY_t^{\perp m}$, where W has full rank. I choose W such that the yield portfolios correspond to the first three principal components of the synthetic yield curve. This choice ensures that the portfolios are orthogonal and thus span the entire three-dimensional subspace. As a result, the portfolios capture as much variation in the synthetic yields as possible with three factors.

Table 2.1: Descriptive Statistics of the Yield and Macroeconomic Data

	Yields with maturities in years								Inflation	Unemp. Gap
	0.25	0.5	1	2	3	5	7	10		
Average	4.63	4.76	5.13	5.37	5.56	5.88	6.12	6.39	3.98	0.82
Std.dev.	3.42	3.42	3.56	3.49	3.41	3.23	3.09	2.94	3.00	1.49
Skew-ness	0.58	0.52	0.46	0.38	0.37	0.39	0.41	0.44	1.49	0.93
Kurtosis	3.17	3.00	2.83	2.69	2.64	2.63	2.67	2.74	4.97	3.24
Autocorrelations:										
Lag 1	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Lag 2	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.98
Lag 3	0.96	0.97	0.97	0.97	0.97	0.98	0.98	0.98	0.96	0.97
Lag 6	0.93	0.94	0.94	0.95	0.95	0.96	0.96	0.96	0.90	0.90
Lag 12	0.87	0.88	0.89	0.9	0.91	0.92	0.92	0.91	0.75	0.73

Note: The table shows descriptive statistics of bond yield data and macroeconomic data in percent per annum. The yield data are U.S. Treasury bond yields from Gürkaynak, Sack, and Wright (2007) and Treasury bill rates from Federal Reserve Economic Data. The macroeconomic data are given by the 12-month change in consumer prices and the unemployment gap defined by the unemployment rate subtracted the natural unemployment rate estimated by the Congressional Budget Office. CPI inflation and unemployment rates are provided by the Bureau of Labor Statistics. The data are sampled monthly from September 1971 to June 2019.

Lag Length Under the Physical Measure

The order of the vector autoregression under the physical measure determines the lag length of the macroeconomic variables. I implement the model with an order of $L = 12$ corresponding to an annual lag structure, which is a common choice in the macro-finance term structure literature (Ang and Piazzesi, 2003, Jardet, Monfort, and Pegoraro, 2013, Joslin, Le, and Singleton, 2013b). The lag length is chosen based on general-to-specific tests and information criteria as documented in Appendix B.2.1.

Generality of the Conditional Variance Matrix Under the Physical Measure

With three latent factors and two macroeconomic variables, the state vector has dimension $N_X = 5$, which identifies up to $K = 3$ components in the conditional variance matrix. In fact, for $K = 3$, the BEKK GARCH model is equivalent to the diagonal vech-GARCH model with restrictions for positive definiteness. As already

discussed, $K = 1$ does not give the model sufficient freedom to capture both conditional variances and covariances. In contrast, I find that the model can match realized and rolling variances of bond yields closely with $K = 2$, see Section 2.4.3. In Appendix B.2.3, I formally compare models specified with $K = \{1, 2, 3\}$ and show that the data is best modeled with $K = 2$.

2.4.3 Estimation Results

Parameter estimates are reported in Appendix B.2.4. The model is highly persistent under both the physical and risk-neutral measures with maximum eigenvalues of the autoregressive coefficients of respectively 0.995 and 0.997. The conditional variance matrix of the physical dynamics is persistent but remains stationary because

$$\max \left[\sum_{k=1}^{K} (A_{\mathcal{P}}^{(k)} \odot A_{\mathcal{P}}^{(k)} + B_{\mathcal{P}}^{(k)} \odot B_{\mathcal{P}}^{(k)}) \right] = 0.981 < 1,$$

where $\odot$ denotes element-wise Hadamard product.[26]

The estimated latent factors are exhibited in Figure 2.2 along with the factor loadings $b_{x,n}$ across maturities n. The loadings show that the latent factors capture level, slope, and curvature risk of the synthetic yields, $Y_t^{\perp m}$. The figure shows that these factors are different from the level, slope, and curvature of observed yields, Y_t. In particular, I observe that the macroeconomic variables relate to variations in the level and slope of the yield curve. This is consistent with Rudebusch and Wu (2008), who show that the level and slope of the yield curve have macroeconomic underpinnings, and Bauer and Rudebusch (2016), who show that the unemployment gap is related to the slope.

2.4.4 Empirical Performance

Goodness of In-Sample Yield Curve Fit

The ability of my model to match the yield curve is summarized by root mean squared errors in Panel A of Table 2.2. My model fits the data well with an average error across maturities of 8 basis points. This result is not surprising since the risk-neutral dynamics of my model are similar to those of the Gaussian term structure model, which is known to have good in-sample properties in terms of matching the yield curve.

[26] The stationarity condition for the BEKK GARCH model is given in Engle and Kroner (1995).

Table 2.2: In-Sample Performance

Panel A. Yield Curve Fit

	Maturities in years								
	0.25	0.5	1	2	3	5	7	10	Average
RMSE	8.74	10.90	12.42	2.97	5.64	7.86	3.64	8.35	8.18

Panel B.1. Fit to Realized Variances and Covariances

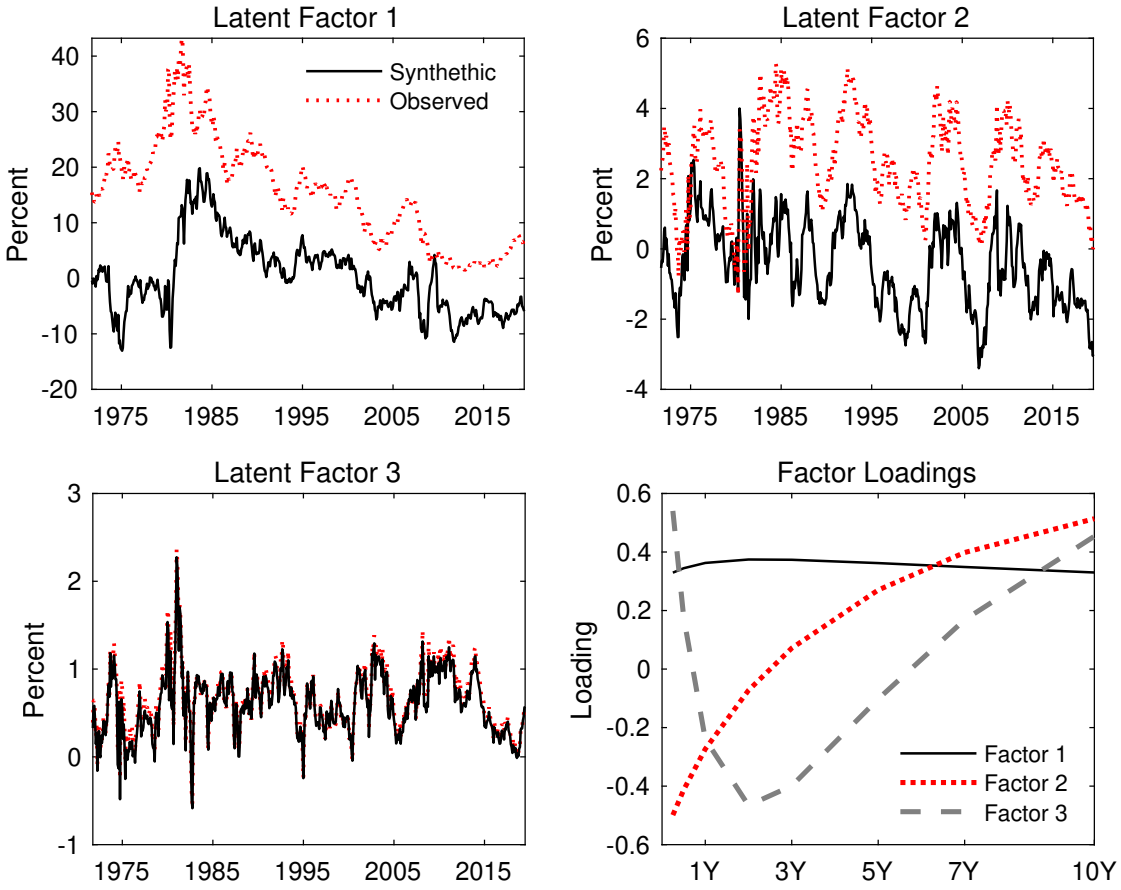
	Maturities in years								
	0.25	1	2	5	10	(0.25,5)	(0.25,10)	(5,10)	Average
RMSE	4.83	4.53	3.77	2.04	1.23	3.37	1.60	1.47	2.85
Constant	0.03 (0.02)	0.05 (0.02)	0.06 (0.02)	0.04 (0.02)	-0.01 (0.02)	0.03 (0.01)	0.02 (0.01)	0.02 (0.02)	0.03
Slope	0.77 (0.06)	0.50 (0.05)	0.53 (0.06)	0.63 (0.05)	0.80 (0.06)	0.59 (0.05)	0.69 (0.06)	0.69 (0.06)	0.65
$\bar{R}^2$	0.63	0.56	0.52	0.49	0.50	0.63	0.63	0.48	0.56

Panel B.2. Fit to Rolling Variances and Covariances

	Maturities in years								
	0.25	1	2	5	10	(0.25,5)	(0.25,10)	(5,10)	Average
RMSE	3.91	3.13	2.16	1.68	1.38	2.78	1.82	1.58	2.31
Constant	-0.01 (0.02)	-0.01 (0.03)	-0.01 (0.03)	-0.00 (0.02)	-0.04 (0.02)	0.02 (0.02)	0.02 (0.01)	-0.03 (0.02)	-0.01
Slope	0.67 (0.09)	0.63 (0.09)	0.62 (0.09)	0.65 (0.08)	0.70 (0.08)	0.64 (0.08)	0.71 (0.09)	0.67 (0.08)	0.66
$\bar{R}^2$	0.53	0.49	0.44	0.42	0.41	0.50	0.48	0.43	0.46

Notes: The table shows the in-sample performance of my model. Panel A shows root mean squared errors (RMSEs) of model-implied against observed yields in basis points per annum. Panels B.1 and B.2 show RMSEs of model-implied against realized and rolling variances and covariances. The table also shows constants and slope coefficients with standard errors in parenthesis and adjusted R-squared values from Mincer and Zarnowitz (1969) regressions of realized and rolling variances and covariances onto a constant and model-implied variances and covariances.

Figure 2.2: Estimated Latent Factors and Factor Loadings



Notes: The figure shows the estimated latent factors in the upper panel and in the lower left chart (solid lines). These are estimated by the first three principal components of the synthetic yields given by the residuals from projecting yields on the macroeconomic variables. The figure compares these with the principal components of the observed yield curve (dotted lines). The lower-right chart shows the estimated factor loadings across maturities as given in Theorem 2.1.

Matching Conditional Yield Curve Volatility

I compare model-implied conditional variances and covariances of the yield curve to two empirical proxies. First, I construct a realized variance matrix using daily yield data.[27] This measure is likely to be a good proxy of conditional variance, because realized variance is a consistent estimator of integrated variance (Barndorff-Nielsen

[27]I construct the realized variance matrix as follows. Let $y_{t,s,n}$ denote the n-maturity yield on day s in month t and let S denote the total number of daily observations in month t. Following this, I define the realized covariance between yields with maturities n_1 and n_2 by

$$\text{RV}_{t,n_1,n_2} = \sum_{s=1}^{S} \left(y_{t,s,n_1} - y_{t,s-1,n_1}\right)\left(y_{t,s,n_2} - y_{t,s-1,n_2}\right).$$

and Shephard, 2004) and conditional variance can be viewed as a noisy measure of integrated variance.[28] As a second proxy, I use a rolling conditional variance matrix constructed with daily data using a 3-month look-back.

Figure 2.3 shows model-implied conditional variances and covariances along with the corresponding realized and rolling measures for the 3-month and 10-year maturities. My model captures low-frequency variation of empirical measures closely in both the highly volatile period in the beginning of the 1980s and in the periods with low volatility. Short-lived bursts observed in the empirical proxies that are not reproduced by my model may be due to measurement errors (Andersen, Bollerslev, and Meddahi, 2005) and outliers that should not be predictable ex ante (Cieslak and Povala, 2016).

To provide a formal assessment, I show root mean squared errors in Panels B.1 and B.2 of Table 2.2. Errors are generally lower for long-term yields, which may reflect that volatility of short-term bond yields is distorted by noise, institutional effects, and factors specific to Treasury bills, see Cieslak and Povala (2016) and references therein. I also run Mincer and Zarnowitz (1969) regressions of the empirical measures on a constant and the model-implied conditional variances and covariances. Table 2.2 shows the constants, slope coefficients, and adjusted R-squared values from these regressions. I find that a majority of the constants are estimated insignificantly. The estimated slope coefficients are between 0.5 and 0.8 with adjusted R-squared values between 0.4 and 0.6.

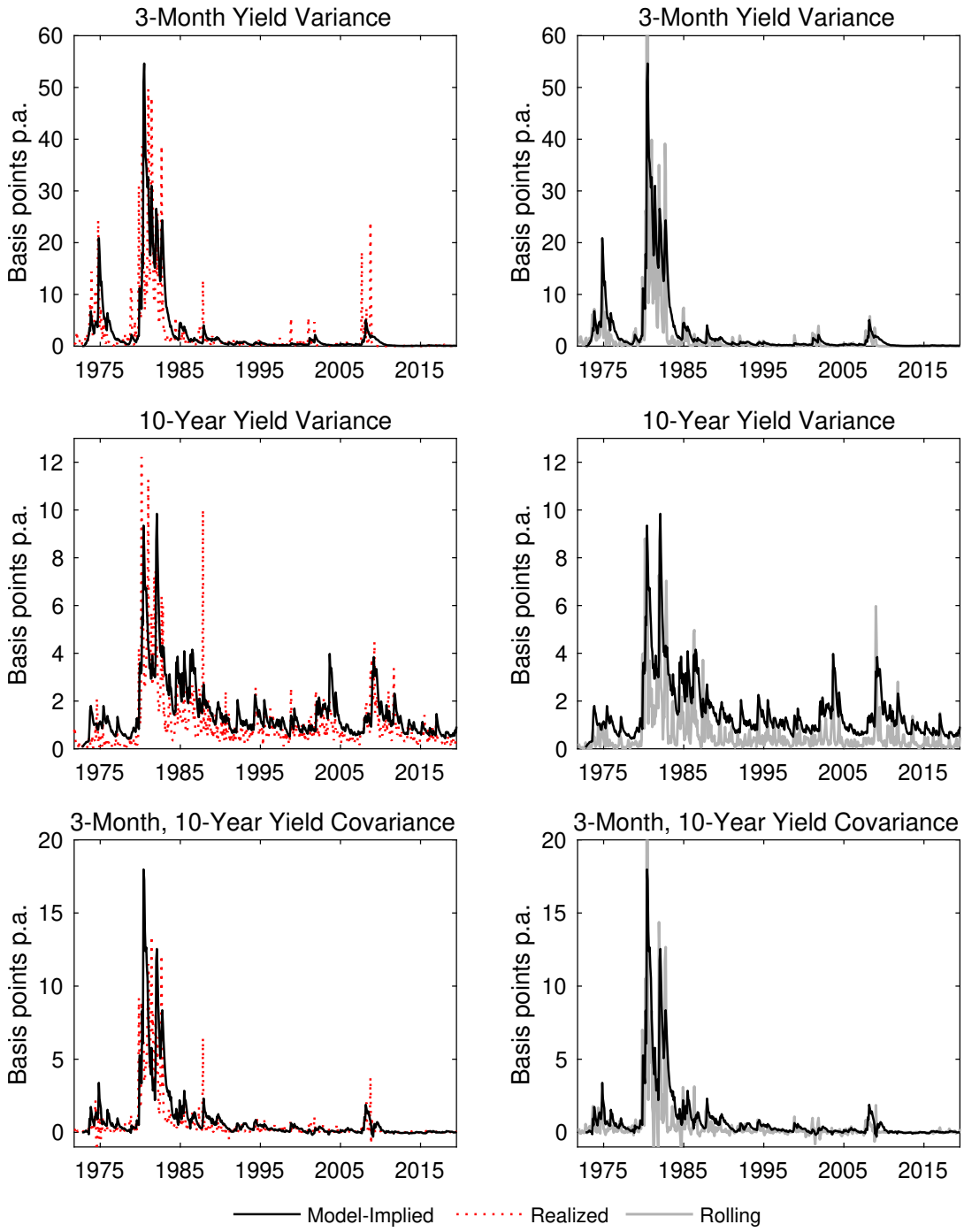
Thus, I conclude that my model can match empirical measures of yield curve variances and covariances. This finding contrasts with continuous-time term structure models with stochastic volatility, which result in low and often negative correlation between predicted and realized variances (Christensen, Lopez, and Rudebusch, 2014, Collin-Dufresne, Goldstein, and Jones, 2009, Jacobs and Karoui, 2009). The success of my model can be attributed to the exponential-quadratic pricing kernel. Specifically, the distinction between conditional variance matrices under the physical and risk-neutral probability measures gives my model freedom to capture empirical volatility as seen in Creal and Wu (2017), Ghysels, Le, Park, and Zhu (2014), and Joslin and Konchitchki (2018).

Term Structure Decomposition

Next, I show how my model decomposes yields into short-rate expectations and term premia, see (2.16). The estimated decomposition is shown for the 10-year yield in

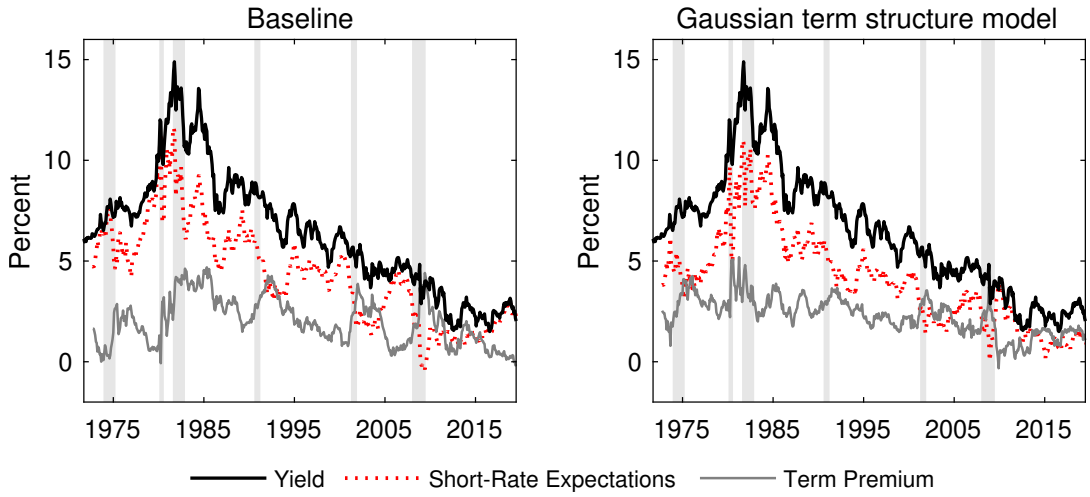
[28] See Andersen, Bollerslev, Diebold, and Labys, (2001, 2003), Koopman and Scharth (2013), and Hansen, Huang, and Shek (2011).

Figure 2.3: Model-Implied, Realized, and Rolling Variances and Covariances



Notes: The figure shows model-implied conditional variances and covariances along with empirical measures given by: (i) Realized variances constructed from daily data (left panel). (ii) Rolling variances of daily data with a 3-month look-back (right panel). Results are shown results for the 3-month and 10-year bond yields.

Figure 2.4: Decomposition of the 10-Year Yield

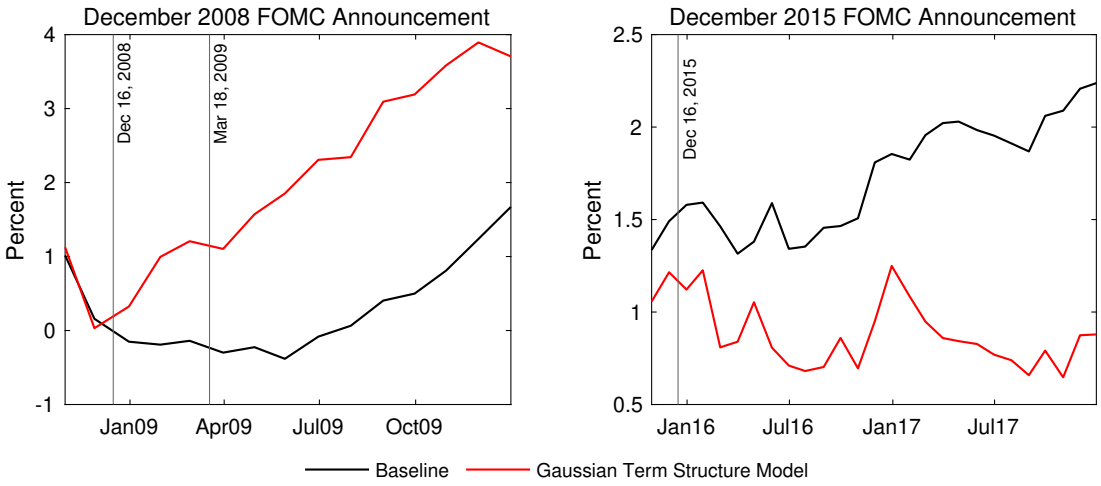


Notes: The figure shows the model-implied decomposition of the 10-year yield into short-rate expectations and a term premium. The left chart shows the decomposition implied by my model, whereas the right chart shows the decomposition implied by the Gaussian term structure model. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

the left chart of Figure 2.4. The expected short rates reflect the dynamics of long-term yields due to the high degree of persistence in yield data. As a result, model-implied term premia are relatively stable. However, the term premia are counter-cyclical as investors demand a higher compensation for risk in times of crisis.

Time-varying volatility can be interpreted as a time-varying quantity of risk. Thus, my model involves time-variation in both the quantity and price of risk, see (2.7)-(2.8). In contrast, Gaussian term structure models only allow for a time-varying price of risk. Thus, comparing term premia estimates of my model with those of the Gaussian term structure model provides an insight into the impact of a time-varying quantity of risk on term premia. The right chart in Figure 2.4 shows the term structure decomposition of the Gaussian term structure model. The time-varying quantity of risk in my model causes stronger counter-cyclicality in term premia compared to the constant quantity of risk in the Gaussian model. Specifically, I observe large differences between the models after the financial crisis. My model implies that term premia decreased slowly and became negative in May 2019. This result signals an increasing optimism among investors, which is consistent with the economic expansion. In contrast, the Gaussian term structure model implies term premia that became negative immediately after the recession. This estimate points to investors regaining optimism shortly after the recession. It is likely that this result

Figure 2.5: Short-Rate Expectations Following Forward Guidance

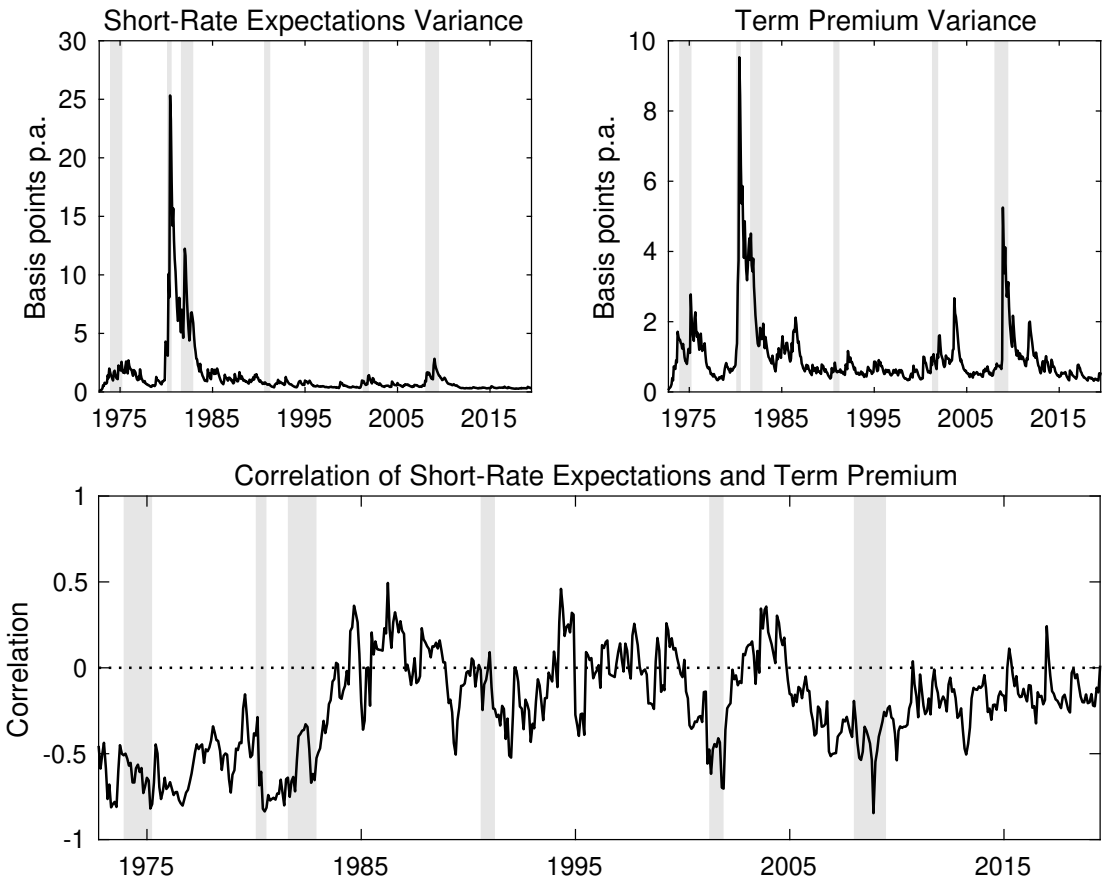


Notes: The figure shows model-implied 10-year short-rate expectations following the Federal Open Market Committee announcements containing forward guidance on December 16, 2008 (left chart) and December 15, 2015 (right chart). The charts compare expectations from my model to those of the Gaussian term structure model.

arises because the Gaussian model fails to capture the decrease in the quantity of risk following the recession.

Furthermore, my model disagrees with the Gaussian term structure model in regard to the effectiveness of forward guidance. If forward guidance works as intended, short-rate expectations should follow central bank communications. In my model, this is indeed the case for many examples of statements from the Federal Open Market Committee. For example, as shown in the right chart of Figure 2.5, my model implies increased short-rate expectations as the committee signaled an upward path of the federal funds rate during the lift-off from the zero-lower bound in December 2015. In contrast, the Gaussian term structure model estimates decreasing short-rate expectations during this period, which implies that the forward guidance was not effective. Another example is in December 2008, where the Federal Open Market Committee indicated that the federal funds rate was likely to remain low in the future. My model is consistent with markets reacting to this statement as the short-rate expectations remain low. However, short-rate expectations in the Gaussian term structure model increase over this period, as observed in the left chart on Figure 2.5. My conclusion that forward guidance is effective is supported by Carlstrom, Fuerst, and Paustian (2015) and McKay, Nakamura, and Steinson (2016), while Hagedorn, Luo, Manovskii, and Mitman (2019) present counter-arguments.

Figure 2.6: Conditional Variances of and Correlation Between 10-Year Short-Rate Expectations and Term Premium



Notes: The figure shows model-implied conditional variances of the 10-year short-rate expectations and term premium (upper panel). The lower panel shows their conditional correlation implied by the model. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

My model also generates time-varying conditional variances of the expected short rates and term premia. These variances are shown at the 10-year maturity in Figure 2.6. The variance of short-rate expectations is counter-cyclical and mimics the behavior of the conditional yield variances. Specifically, the large burst of volatility in short-term yields in the beginning of the 1980s is transmitted to the volatility of short-rate expectations. The term premium is more stable. Figure 2.6 also shows the model-implied conditional correlation between the 10-year short-rate expectations and term premium. The correlation is negative throughout the majority of the sample. This is intuitively appealing: Whereas a high term premium signals that bond market investors fear risk, i.e., pessimism, high expected short rates reflect an optimism about future economic conditions. The sign, however, tends to

reverse during expansions, which may reflect the fact that mid-expansions are typically characterized by a flattening yield curve and low volatility.[29]

Specification Test

The ability of the model to decompose the term structure can be tested by projecting holding period returns, $Y_{t+1,n-1} - Y_{t,n}$, on a constant and the slope of the yield curve, $(Y_{t,n} - r_t)/(n-1)$. Under the expectations hypothesis, such regressions give coefficients equal to one. However, Campbell and Shiller (1991) show that empirically estimated coefficients are negative and exhibit a downward pattern across maturities. I test the ability of the model to generate coefficients that match the empirical failure of the expectations hypothesis, which has been named the LPY-I test by Dai and Singleton (2002).[30] The test results are depicted in the upper-left chart of Figure 2.7. With the exception of the short maturities, my model produces coefficients that are well within the 95 percent confidence intervals for the empirically estimated coefficients. Thus, the model passes the LPY-I test.

Dai and Singleton (2002) also propose a LPY-II test which projects the risk-adjusted holding period returns onto a constant and the slope of the yield curve.[31] If the model is well-specified, the resulting coefficients should be equal to one. The lower-left chart on Figure 2.7 shows that the model-implied 95 percent confidence interval contains one for all maturities. Hence, my model also passes the LPY-II test.

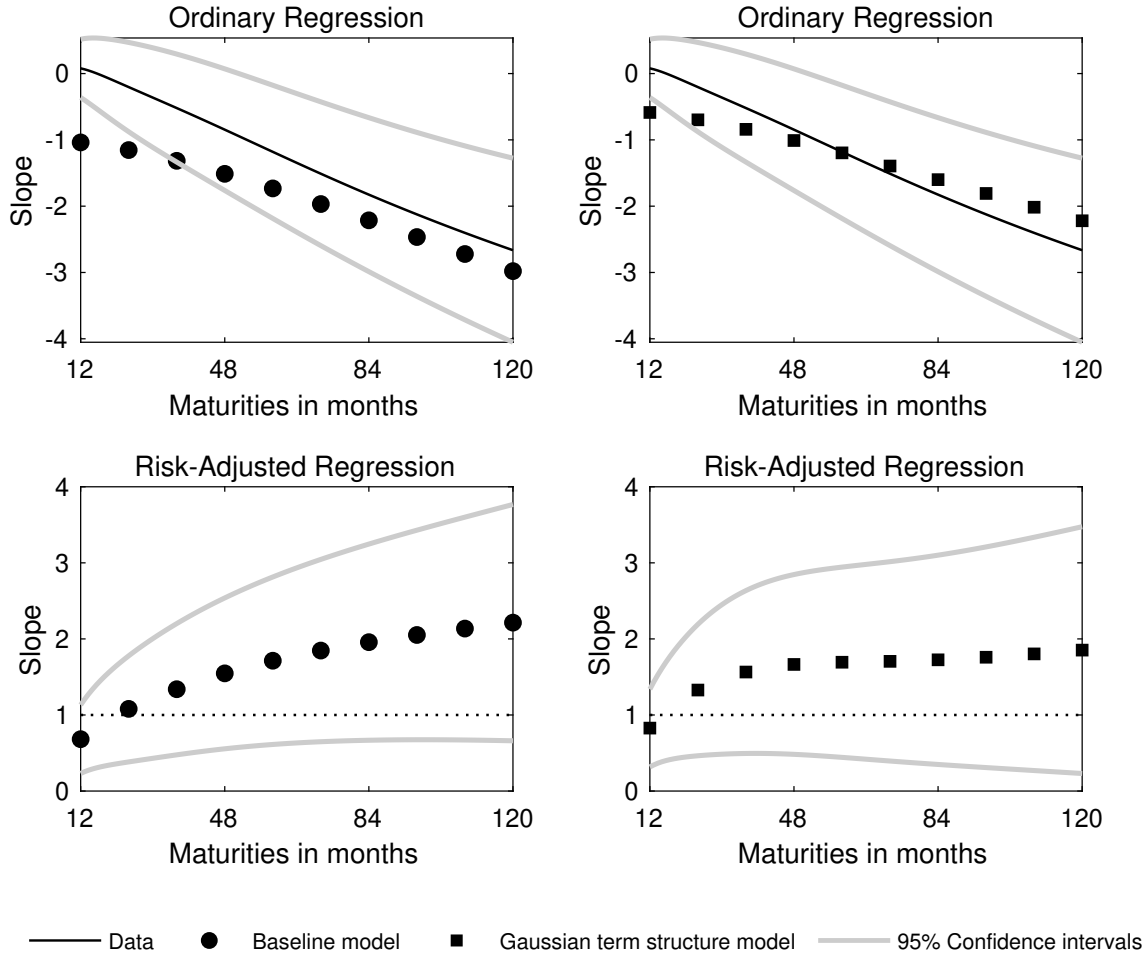
Traditional affine term structure models with stochastic volatility (Dai and Singleton, 2000) fail both the LPY-I and LPY-II tests. In contrast, the Gaussian term structure model passes both tests (Dai and Singleton, 2002), which I confirm in the right panel of Figure 2.7, but fails to capture time-varying volatility. Thus, the existing literature poses a trade-off between modeling time-varying volatility and passing the LPY tests. My model overcomes this trade-off as I simultaneously match empirical proxies of the conditional yield curve variance matrix and pass both LPY tests.

[29] This relationship has been established by Norland (2018) using the VIX index as a measure of volatility. While the VIX index relates to equity markets, it is positively correlated with volatility measures related to bond markets, e.g., the Merill-Lynch Option Implied Volatility Estimate (MOVE).

[30] In particular, I construct yield curve data with maturities $1, 2, \ldots, 120$ months using the method of Gürkaynak, Sack, and Wright (2007). Then, I regress $Y_{t+1,n-1} - Y_{t,n} = c + \phi_{T,n}(Y_{t,n} - r_t)/(n-1) + e_t$, which results in the empirically estimated coefficients $\hat{\phi}_{T,n}$. I compare these with model-implied population coefficients obtained by simulating 100,000 observations from my model.

[31] The risk-adjusted returns are given by $Y_{t+1,n-1} - Y_{t,n} - (\Psi_{t+1,n-1} - \Psi_{t,n-1}) + \Psi^F_{t,n-1}/(n-1)$, where $Y_{t,n}$ are observed yields, $\Psi_{t,n}$ is the model-implied term premium defined in (2.16), and Ψ^F is the model-implied forward term premium defined by $\Psi^F_{t,n} = F_{t,n} - \mathbb{E}_t(r_{t+n})$, where $F_{t,n}$ is the forward rate.

Figure 2.7: Campbell-Shiller Regression Slope Coefficients



Notes: The figure shows slope coefficients from ordinary and risk-adjusted Campbell and Shiller (1991) regressions. For the ordinary regressions, model-implied slope coefficients are obtained by simulating 100,000 observations in my model (left panel) and the Gaussian term structure model (right panel). Confidence intervals are estimated with Newey-West robust standard errors with automatically selected lag length.

2.5 Yield Curve Volatility and Macroeconomic Risk

In this section, I show that the relationship between the yield curve and macroeconomic risk fluctuates over time. My model provides insights into the patterns and sources of these fluctuations. Specifically, I examine the macroeconomic contribution to the variances of yields, short-rate expectations, and term premia. Finally, I relate the model-implied price of variance risk to macroeconomic variables.

2.5.1 Macroeconomic Contribution to Yield Curve Variances

I analyze the relationship between the yield curve and macroeconomic risk by decomposing model-implied variances into contributions from macroeconomic variables and latent yield-curve-specific factors. For this purpose, I apply a recursive identification scheme assuming that macroeconomic variables are slow-moving and do not react to contemporary yield-curve-specific shocks. Furthermore, I assume that the unemployment gap responds to inflation shocks with a one-month lag. These assumptions are standard in the structural vector autoregressive literature with Cholesky identification schemes (Bernanke, Boivin, and Eliasz, 2005, Christiano, Eichenbaum, and Evans, 1999, Stock and Watson, 2001). My approach is also standard in the macro-finance term structure literature in which more sophisticated identification methods are not yet widespread. My results are robust to different orderings of the variables, see Appendix B.3.2.

Given these assumptions, conditional yield variances can be decomposed into variation generated by, respectively, macroeconomic variables and latent yield-curve-specific factors. Moreover, the macroeconomic contribution can be decomposed into an inflation and an unemployment-gap component:

$$\text{Var}_{t-1}\left(Y_{t,n}\right) = \text{Var}_{t-1}\left(\epsilon_t^{\text{inf}}\right) + \text{Var}_{t-1}\left(\epsilon_t^{\text{ugap}}\right) + \text{Var}_{t-1}\left(\epsilon_t^{\text{yield}}\right),$$

where ϵ_t^{inf}, ϵ_t^{ugap}, and $\epsilon_t^{\text{yield}}$ are structural inflation, unemployment-gap, and yield-curve-specific shocks. I define the macroeconomic share of variance in $Y_{t,n}$ by

$$\text{Var}_{t-1}\left(\epsilon_t^{\text{inflation}} + \epsilon_t^{\text{unemployment gap}}\right) / \text{Var}_{t-1}\left(Y_{t,n}\right).$$

Analogously, I construct macroeconomic shares of variance in short-rate expectations and term premia. The novelty of my analysis compared to the previous literature is that the macroeconomic shares are time-varying. I characterize the time-series properties of the term structure of macroeconomic shares in Table 2.3. I emphasize two results from these descriptive statistics.

First, I observe that the macroeconomic shares vary over wide ranges. For example, the macroeconomic contribution to variance in the 3-month yield varies between 0.02 and 55.56 percent. The ranges are similar for yields of bonds with longer maturities. Thus, in some periods, latent yield-curve-specific factors fully account for movements in the yield curve, whereas macroeconomic shocks explain more than half of the variation in other periods. On average, macroeconomic shocks explain 7.92 percent of the 3-month yield variance. This average is consistent with the constant macroeconomic share obtained with a Gaussian term structure model, which equals 7.72 percent for the 3-month maturity. The models agree on average for all

Table 2.3: Descriptive Statistics of Macroeconomic Shares of Yield Variances

	Maturities in years							
	0.25	0.5	1	2	3	5	7	10
Average	7.92	8.51	7.61	7.14	7.33	7.91	8.14	7.91
Std. dev.	7.94	7.95	7.27	6.87	7.00	7.51	7.67	7.59
Skewness	2.07	2.01	2.58	2.33	2.06	1.88	1.80	1.84
Kurtosis	9.07	8.84	12.79	10.41	8.44	7.48	7.11	7.43
Min	0.02	0.05	0.03	0.03	0.02	0.00	0.01	0.00
Max	55.56	56.24	54.03	45.05	45.91	46.39	50.84	53.86
Median	5.23	6.40	5.78	5.16	5.27	5.48	5.66	5.49
Autocorrelation	0.71	0.71	0.68	0.64	0.65	0.67	0.69	0.69
Quantiles:								
10 percent	1.06	1.20	1.18	0.99	0.97	1.03	1.05	0.95
25 percent	2.53	2.97	3.05	2.82	2.73	2.9	2.84	2.74
75 percent	10.85	11.74	9.76	9.11	9.25	10.57	11.17	10.87
90 percent	17.68	19.34	16.02	15.96	16.39	17.89	18.26	17.73
Gaussian term structure model	7.72	8.29	8.39	8.03	8.01	8.17	7.93	6.92

Note: The table shows descriptive statistics of the estimated shares of yield variances determined by the macroeconomic variables. The statistics are reported in percent. The table also shows the constant macroeconomic share of variances estimated using the Gaussian term structure model.

considered maturities. Since the Gaussian term structure model captures the average level of yield curve variance, this result can be interpreted as a validation of my model specification.

Second, the sample kurtosis along with the quantiles suggest that there are occasional large month-to-month fluctuations. Some of these occasions can be tied to meaningful economic events. For example, when the first round of quantitative easing was announced in November 2008, the macroeconomic share of variance in the 10-year yield increased from 5 to 28 percent. The intention of quantitative easing is that central banks impact long-term yields directly by large-scale asset purchases. Thus, successful quantitative easing programs result in a tighter link between long-term yields and the dual mandate. This intuition is indeed consistent with an increased macroeconomic share of variance in the 10-year yield. Another example is when Donald Trump was elected president in the end of 2016, where it was gen-

erally believed that Trump's stimulus plans would lead to higher interest rates.[32] During the month following the election, the macroeconomic share of variance in the 10-year yield indeed increased from 5 to 25 percent. More recently, the Federal Reserve announced in January 2019 that the monetary policy tightening cycle would be put on hold due to factors unrelated to the domestic economy.[33] During this month, the macroeconomic share of the 3-month yield variance fell from 16 to 3 percent. This anecdotal evidence is consistent with the widespread idea that macroeconomic announcement effects impact financial markets (Andersen, Bollerslev, Diebold, and Vega, 2007, Balduzzi, Elton, and Green, 2001, Bollerslev, Cai, and Song, 2000, Feunou, Fontaine, and Roussellet, 2019, Fleming and Remolona, 1999, Johannes, 2004, Piazzesi, 2005).

The link between variance in long-term yields and macroeconomic risk can be generated through two channels, namely short-rate expectations and term premia, as shown by the term structure decomposition in (2.16). Table 2.4 shows descriptive statistics of the macroeconomic shares of variances in 5- and 10-year short-rate expectations and term premia.

Macroeconomic variables account for 0.32 to 75.66 percent of the variance in 10-year short-rate expectations and 41.88 percent on average. The macroeconomic share of variance in the 10-year term premium ranges between 0.36 and 68.58 percent, with an average of 21.30 percent. These averages do not align with the Gaussian term structure model that implies macroeconomic shares of 26.36 and 26.16 percent for the 10-year short-rate expectations and term premium. These differences arise because the Gaussian term structure model obtains a different decomposition of yields by ignoring the time-varying quantity of risk, as discussed in Section 2.4.4. Thus, even if our interest lies in explaining variation across the full sampling period, it is important to account for time-varying volatility. The results also show that macroeconomic shocks explain more variation in, respectively, short-rate expectations and term premia compared to the yield curve. Hence, investors pay closer attention to macroeconomic risk when forming expectations about future monetary policy compared to when they price bonds. This is consistent with the idea of unspanned macroeconomic risk proposed by Joslin, Priebsch, and Singleton (2014).

[32] See, for example, "How high can US rates go in the Trump era?" Financial Times, December 13, 2016, https://www.ft.com/content/b5551882-c057-11e6-81c2-f57d90f6741a.

[33] "Federal Reserve puts rate rises on hold as global economy slows", Financial Times, January 30, 2019, https://www.ft.com/content/2565e154-24b7-11e9-b329-c7e6ceb5ffdf.

Table 2.4: Descriptive Statistics of Macroeconomic Shares of Variances in the Decomposition of 5- and 10-Year Yields

	Short-rate expectations		Term premia	
Maturity	5-year	10-year	5-year	10-year
Average	40.24	41.88	34.08	21.30
Std. dev.	18.12	18.28	14.91	12.44
Skewness	-0.12	-0.15	0.38	0.88
Kurtosis	1.94	1.96	2.85	3.78
Min	0.23	0.32	0.73	0.36
Max	75.8	75.66	75.84	68.59
Median	40.29	42.04	32.33	19.06
Autocorrelation	0.89	0.90	0.67	0.69
Quantiles:				
10 percent	15.49	16.67	15.23	6.55
25 percent	26.25	27.64	23.60	12.11
75 percent	56.56	58.60	43.63	28.80
90 percent	64.51	65.57	54.52	38.32
Gaussian term structure model	33.36	26.36	48.36	26.16

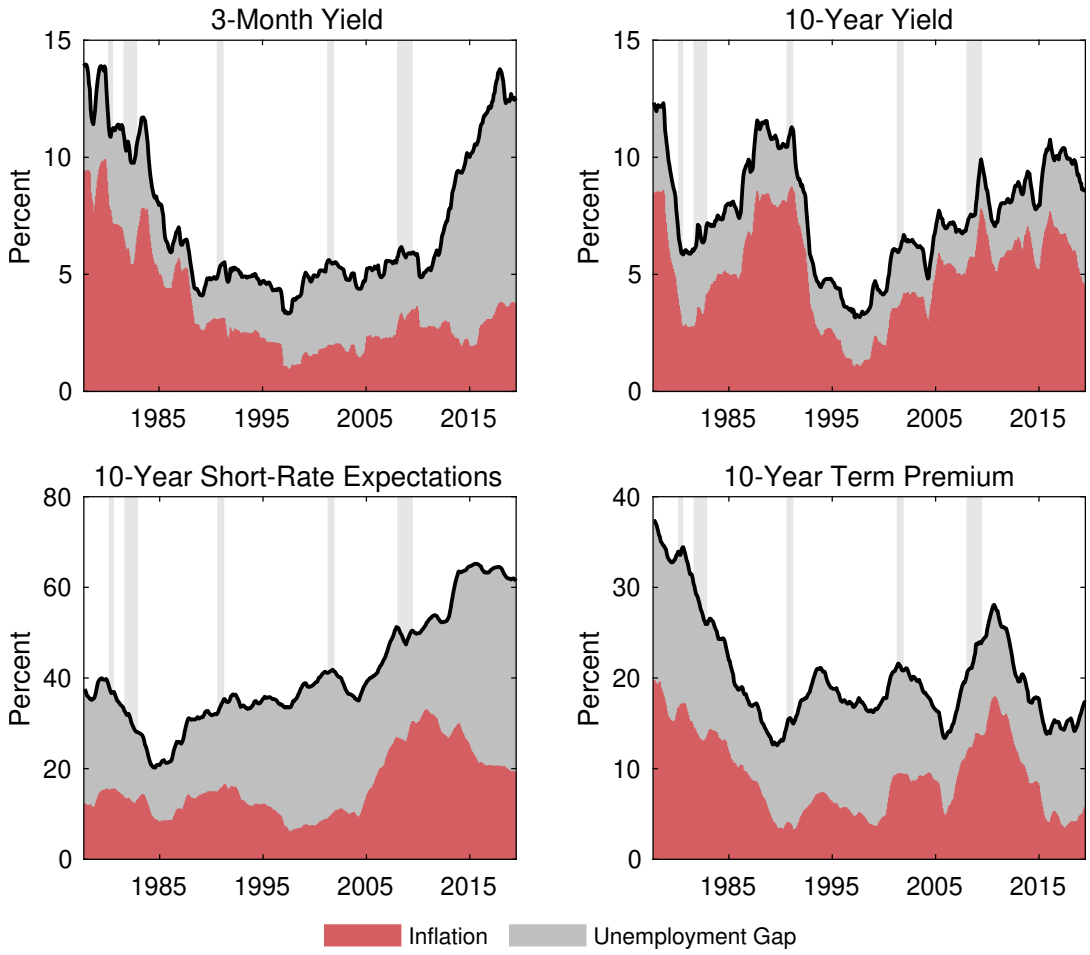
Note: The table shows descriptive statistics of the estimated shares of variances in short-rate expectations and term premia determined by the macroeconomic variables. The statistics are reported in percent. The table also shows the constant macroeconomic share of variances estimated using the Gaussian term structure model.

2.5.2 Historical Joint Behavior of the Yield Curve and Macroeconomic Risk

Next, I characterize the historical relationship between movements in the yield curve and macroeconomic variables throughout my sample. Since investors and policy makers typically care about horizons longer than one month, I consider fluctuations in macroeconomic shares of variances over 5-year periods. Figure 2.8 shows 5-year moving-average macroeconomic shares of the 3-month and 10-year yield variances along with variances of the 10-year short-rate expectations and term premium.

In the upper-left chart on Figure 2.8, I observe a U-shaped pattern in the macroeconomic share of the 3-month yield variance. Specifically, macroeconomic variables are important short-term yield factors in the 1970s and 1980s, but cease to be relevant during the Great Moderation starting from the mid-1980s. This result is intuitively appealing as the stabilization of macroeconomic variables reduces the extent

Figure 2.8: 5-Year Moving-Average Macroeconomic Shares of Variances



Notes: The figure shows 5-year moving averages of the macroeconomic contribution to variances in the 3-month and 10-year yields, 10-year short-rate expectations, and the 10-year term premium. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

to which these shocks can explain volatility in financial markets. In the aftermath of the financial crisis, however, the link between the macroeconomic variables and short-term yields has strengthened. Figure 2.8 also decomposes the macroeconomic contribution into inflation and the unemployment gap. The role of inflation attenuates over time and only accounts for a small fraction of the short-term yield variance after the Great Moderation. The unemployment gap, on the other hand, accounts for the increasing macroeconomic share after the financial crisis, which coincides with the recovery of the unemployment rate.

The upper-right chart on Figure 2.8 shows the trend in the macroeconomic share of the 10-year yield variance. This trend can be described by a W-shape with low values during the Volcker period and in the aftermath of the early 1990s recession.

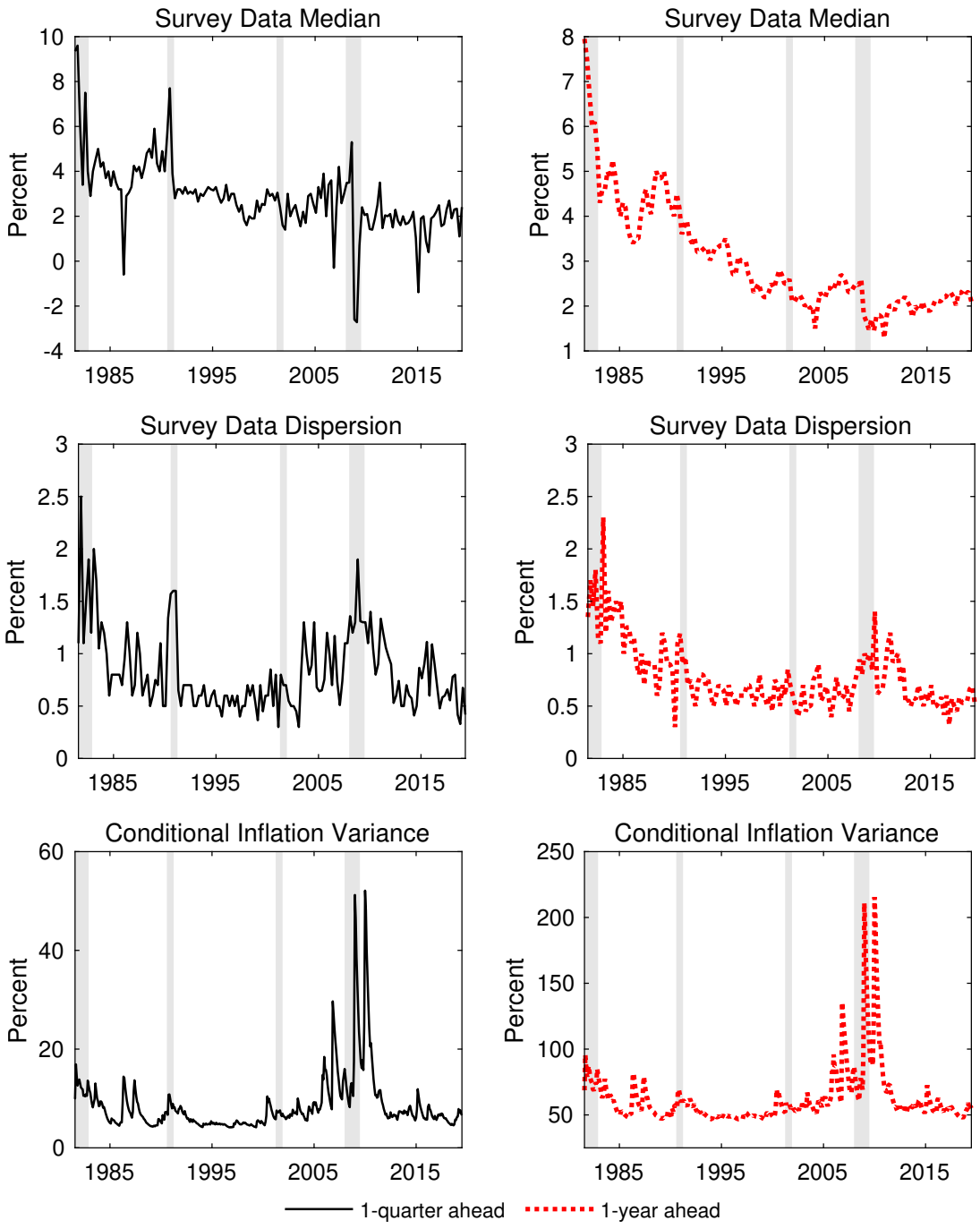
The W-shape is driven by inflation, while the contribution from the unemployment gap is stable. The macroeconomic shares of variances in the decomposition of the 10-year yield provides additional insights, see the lower charts of Figure 2.8.

First, the macroeconomic share of variance in the 10-year short-rate expectations exhibits an upward trend since the Great Moderation. Thus, the expectations hypothesis is an increasingly important channel for how macroeconomic variables shape the yield curve. One interpretation is that there is a growing belief among investors that the Federal Reserve's commitment to the dual mandate is credible. The upward trend is primarily attributed to the unemployment gap with the exception of the financial crisis and its aftermath, where inflation accounts for more variation. This result is consistent with survey data on CPI inflation from the Survey of Professional Forecasters. To show this, I plot the median one-quarter and one-year ahead forecasts of CPI inflation along with the dispersion of these forecasts defined by the difference between the 75th and 25th quantiles in Figure 2.9. I observe that the median forecasts vary more during recessions and in particular during the financial crisis. Moreover, the dispersions are larger during recessions. I capture this pattern in the model-implied conditional variance of inflation, which is plotted in the lower panel of Figure 2.9 for forecasting horizons of one quarter and one year.

Second, the macroeconomic share of variance in the 10-year term premium decreases rapidly in the 1980s, but evolves around a constant mean in the remainder of the sample with counter-cyclical fluctuations caused primarily by inflation. Thus, my model is consistent with a counter-cyclical inflation risk premium, as estimated by Buraschi and Jiltsov (2005), Chernov and Mueller (2012), and D'Amigo, Kim, and Wei (2018). The contribution of the inflation risk premium is particularly high during the financial crisis. In my model, a counter-cyclical inflation risk premium is generated by the exponential-quadratic pricing kernel, which allows investors to price the risk associated with both negative and positive inflation shocks. I illustrate this mechanism by simulating 100,000 paths of my model under the risk-neutral ($\mathbb{Q}$) and physical ($\mathbb{P}$) probability measures and estimating kernel density functions of these simulated series. I plot the conditional densities of the inflation variable along with their ratio ($\mathbb{Q}/\mathbb{P}$) for the one-month and ten-year horizons in Figure 2.10.

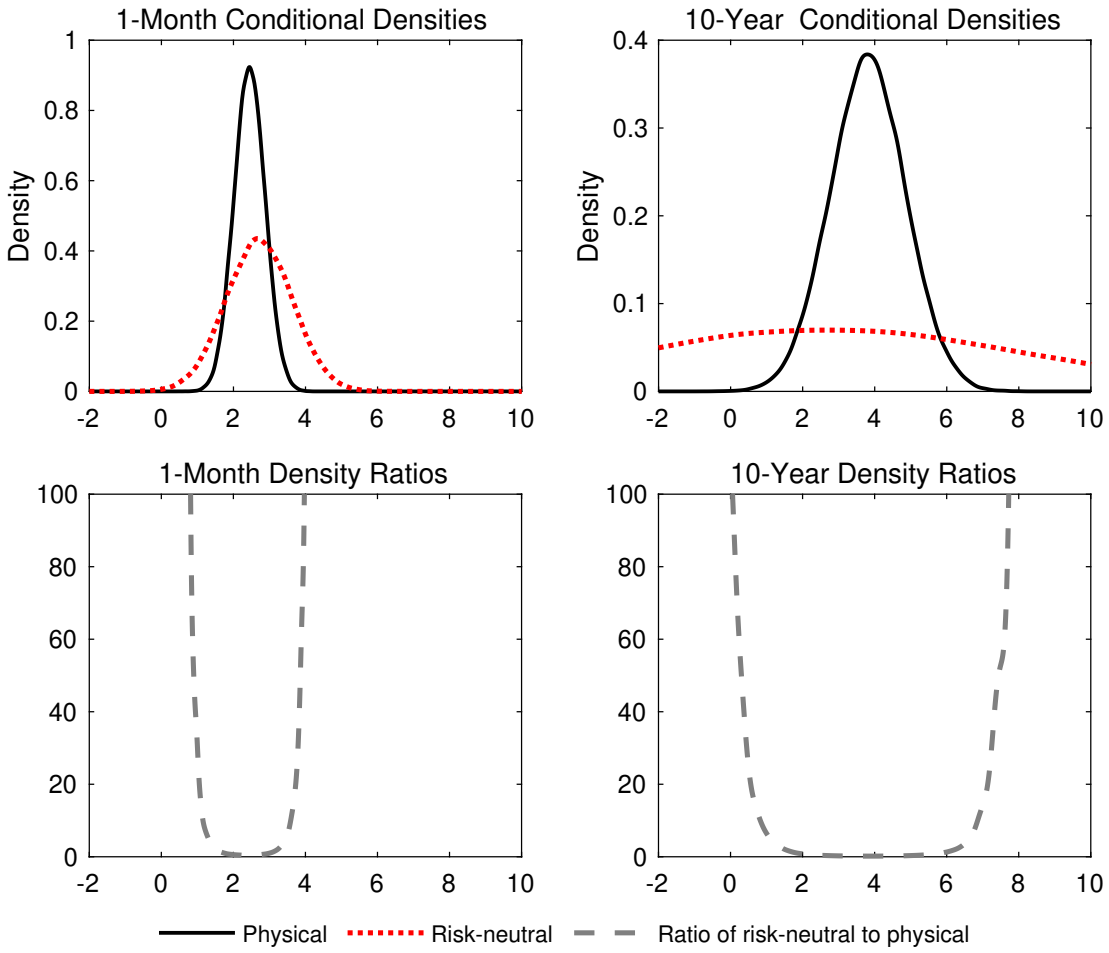
The risk-neutral densities are wider than the physical ones, which reflects the presence of variance risk premia, or, equivalently, different conditional variances under the probability measures. This difference is more pronounced for the long horizon, which involves more uncertainty and thus a larger risk compensation. Differences in the conditional variances result in U-shaped density ratios such that in-

Figure 2.9: Inflation Survey Expectations and Model-Implied Conditional Inflation Variance



Notes: The figure shows the median of inflation survey forecasts from the Survey of Professional Forecasters in the upper panel. The middle panel shows forecast dispersions defined by the difference between the 75th and 25th quantiles. These data are available at quarterly frequency from the third quarter of 1981. The lower panel shows the model-implied conditional inflation variance from July 1981. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

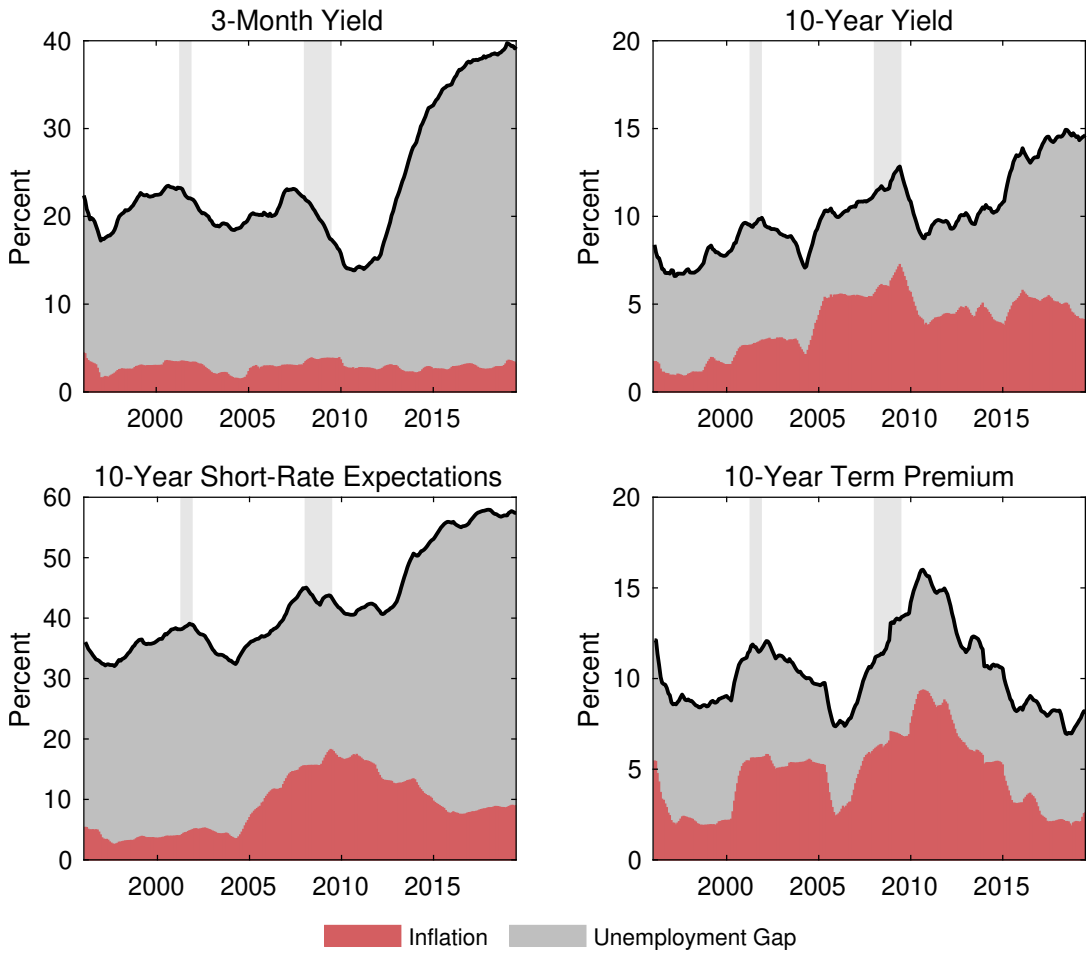
Figure 2.10: Conditional Physical and Risk-Neutral Inflation Densities



Notes: The figure shows model-implied conditional densities under the physical and risk-neutral probability measures (upper panel). The lower panel shows ratios of the risk-neutral to physical densities. The densities are estimated by the Epanechnikov kernel function applied to 100,000 simulated paths with horizon of one month and 10 years. The simulations are initiated at the sample average from January 1990. The bandwidth is selected by the optimal choice for estimating normal densities.

vestors are averse to both high and low states of the distribution of inflation. Thus, my model allows investors to demand a premium for deflation fears, which can generate a counter-cyclical inflation risk premium. In this respect my model is similar to Roussellet (2018). An implication of this result is that a symmetric inflation target enhances the efficiency of monetary policy during recessions.

Figure 2.11: 5-Year Moving-Average Macroeconomic Shares of Variances from 1990-2019



Notes: The figure shows 5-year moving averages of the macroeconomic contribution to variances in the 3-month and 10-year yields, 10-year short-rate expectations, and the 10-year term premium. The model is estimated using data from January 1990 to June 2019. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

Subsample Analysis: January 1990 to June 2019

Applications in term structure modeling often start the sample in 1990 to avoid issues related to the Volcker period. As a robustness check, I show in Figure 2.11 that the time-varying patterns of macroeconomic shares of yield curve volatility identified after the Great Moderation for the full sample continue to hold when the model is estimated using the shorter sample.

For the 3-month yield, I reproduce half of the U-shaped trend in the 5-year moving-average macroeconomic share of variance with the increase driven by the unemployment gap only. The levels of the ratios are higher than observed from

Figure 2.8, reflecting that the short-maturity loading on unemployment gap is overestimated compared to the full sample. For the 10-year yield, I find that the macroeconomic share of variance is increasing after 1995 due to inflation, which is consistent with the results obtained using the full sample. I also generate macroeconomic shares of the variances in the 10-year short-rate expectations and term premium that are similar to the full-sample results.

2.5.3 Variance Risk Premia and Macroeconomic Risk

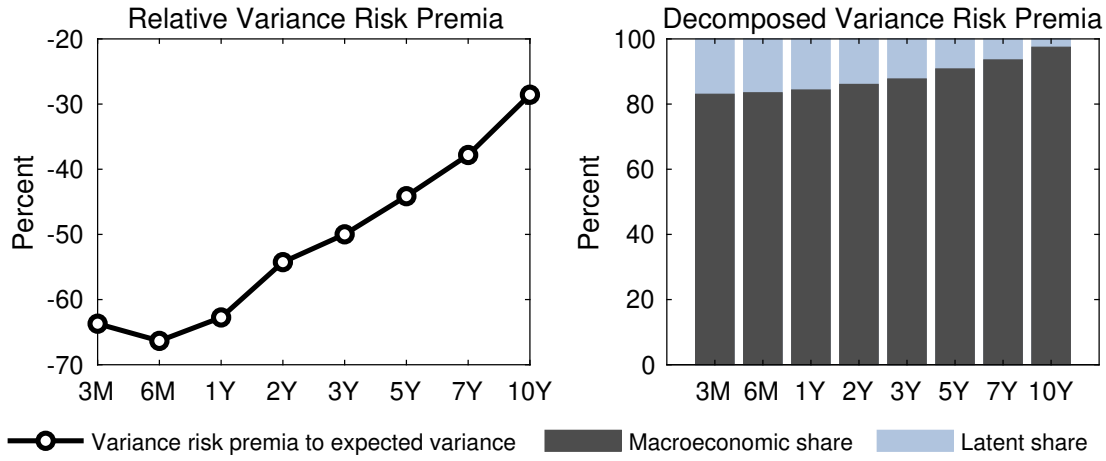
A central distinction of my model from the Gaussian term structure model is the fact that I allow investors to be compensated for variance risk, as defined by Theorem 2.3. This compensation is not earned by investments in Treasury bonds, as shown in the bond yield equation (2.11). Rather, the variance risk premium is demanded by investors in the broader fixed-income market through interest-rate derivatives.

Although my model results in time-varying variance risk premia, I abstract from analyzing these dynamics because they – by the assumption of constant risk-neutral volatility – are solely given by the specification of volatility under the physical measure. Instead, I focus on the average variance risk premia.

Following Carr and Wu (2009), I report variance risk premia as ratios of expected variance under the physical measure. Thus, I show the sample average of $\Gamma_{t,n}$ as a fraction of the sample average of $\mathbb{E}_t(\text{Var}_{t+1}(Y_{t+2,n}))$ in the left chart of Figure 2.12. As expected, the variance risk premia are negative across all maturities, which is consistent with the stylized fact that implied volatilities exceed realized volatilities. The sizes of the variance risk premia are between 30 to 70 percent of the expected variance. This ratio decreases in magnitude across maturities with a small hump at the short end of the yield curve. Thus, fixed-income investors pay large premia for protection against variance risk, particularly in the short end of the yield curve. These results are aligned with existing estimates of variance risk premia in Treasury bond markets (Choi, Mueller, and Vedolin, 2017, Trolle, 2009, Trolle and Schwartz, 2015).

As a novel contribution to the literature, I analyze the link between variance risk premia and the macroeconomic variables. The right chart in Figure 2.12 decomposes the variance risk premia into contributions from the macroeconomic variables and the latent yield-curve-specific factors. A majority of the variance risk premia, between 80 and near 100 percent, are related to macroeconomic uncertainty. This share is increasing with maturity. Thus, fixed-income investors demand large compensations for exposure to macroeconomic uncertainty, whereas uncertainty related to latent yield-curve-specific factors is less important for variance risk premia. It fol-

Figure 2.12: Variance Risk Premia



Notes: The figure shows model-implied variance risk premia. The left panel presents the ratios of average variance risk premia to the average expected variance under the physical measure. The right panel decomposes the average variance risk premia into macroeconomic and latent yield-curve-specific contributions. Results are shown for different maturities.

lows that macroeconomic uncertainty can increase trading activity in fixed-income derivative markets. To the extent that monetary policy is determined by inflation and unemployment gap through the dual mandate, my model is consistent with Cieslak and Povala (2016).

2.6 The 2019 Yield Curve Inversion and Macroeconomic Risk

The 10-year U.S. Treasury bond yield declined by 50 basis points between April and June 2019. The yield curve inverted with the 10-year yield below the 3-month yield during May. Yield curve inversions draw attention from policy makers and investors because history shows that inversions are strong predictors of recessions. Indeed, the yield curve has inverted before every recession since 1971. However, the time between inversion and recession varies and history shows examples of yield curve inversions that are not followed by recessions. A yield curve inversion is therefore subject to debate on whether a recession will occur and when this is likely to happen.

Long-term yields may fall below short-term yields because investors expect a slow-down in future economic growth. Such expectations are reflected in the yield curve by low short-rate expectations, and thus low long-term yields for a given term premium. Inversions that are credible warnings of economic slowdowns and potential recessions should therefore be driven by declining short-rate expectations and,

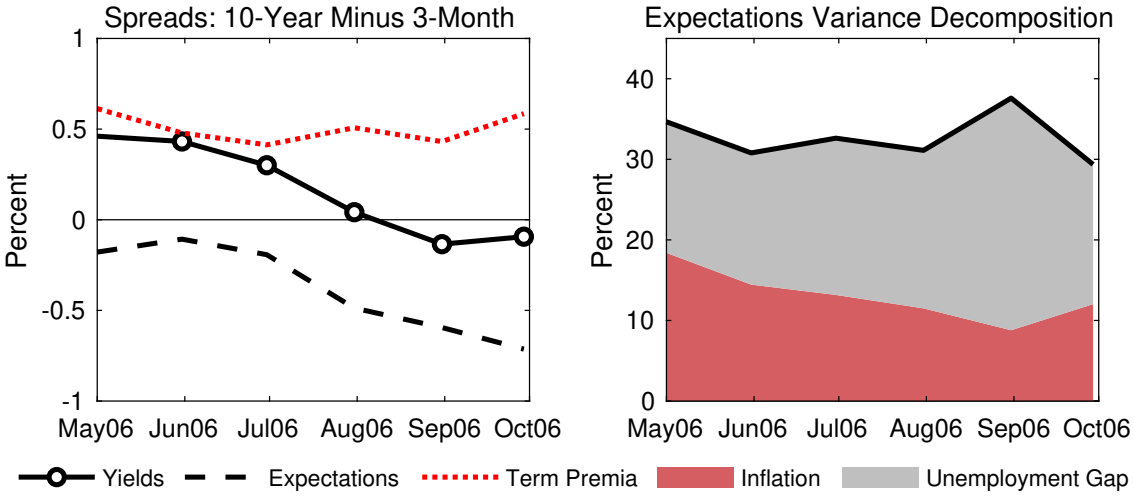
furthermore, this decline should be related to macroeconomic variables. I therefore use my model to show whether the recent yield curve inversion is driven by macroeconomic variables through low short-rate expectations, and thus is likely to pose a warning of an economic slowdown. My model is promising for this purpose because it can analyze time-varying variances conditional on current economic conditions. In contrast, the Gaussian term structure model can only draw conclusions that are valid on average across the full sample.

To support my arguments, I begin by analyzing the yield curve inversion in August 2006 that preceded the financial crisis. Panel (a) in Figure 2.13 analyzes the yield curve from May to October 2006. The left-hand chart shows the spread between the 10-year and 3-month yields along with the corresponding spreads in short-rate expectations and term premia. I observe that the decline in the yield spread can be attributed almost exclusively to the short-rate expectations. The right-hand chart shows that the macroeconomic contribution to the variation in short-rate-expectations spread was consistently around 35 percent over the period. Thus, the inversion in 2006 is related to macroeconomic risk through short-rate expectations.

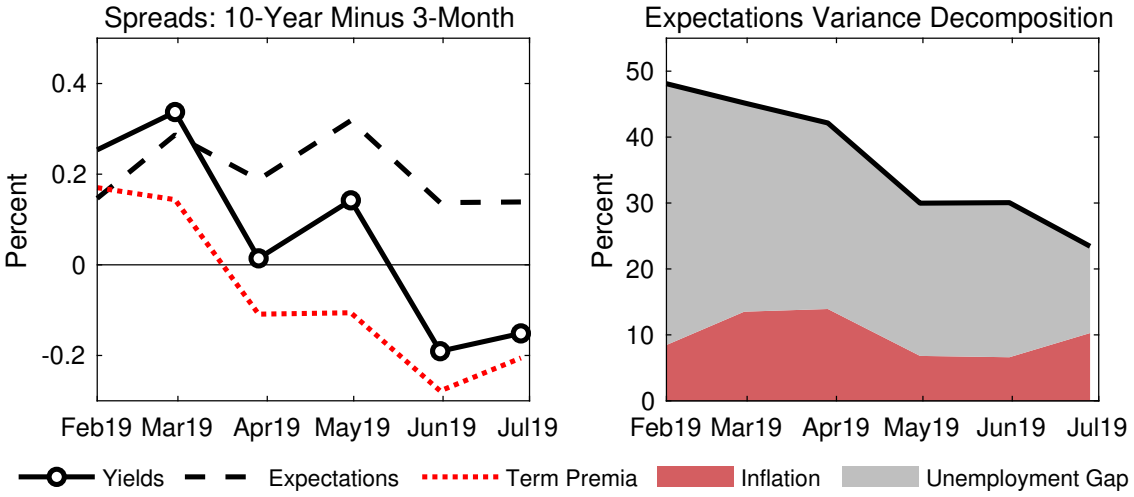
Given these insights, I return to the yield curve inversion in 2019. Panel (b) in Figure 2.13 shows the analysis of the yield curve between February and July 2019. I observe that the decline in the yield spread is mainly driven by a decline in the term-premium spread, unlike the yield curve inversion in 2006. In fact, the short-rate-expectations spread was stable over the period and its contribution to the yield spread becoming negative in May is small. Furthermore, the macroeconomic contribution to variation in the short-rate-expectations spread declined by 25 percentage points over the period. Thus, I find that the relationship between short-rate expectations and the real economy weakened over the period in which the yield curve inverted. Rather, investors increasingly formed short-rate expectations based on information not related to inflation and unemployment gap.

Figure 2.13: Analysis of Yield Curve Inversions

(a) The 2006 Yield Curve Inversion



(b) The 2019 Yield Curve Inversion



Notes: The figure shows yield curve movements from May 2006 to October 2006 in panel (a) and from February 2019 and July 2019 in panel (b). Left-hand charts shows the 10-year-minus-3-month spread in yields, short-rate expectations, and term premia. The yield curve inverted in August 2006 and May 2019. Right-hand charts shows the macroeconomic contribution to variances in the short-rate-expectations spread.

These results indicate that the yield curve inversion in May 2019 is only weakly related to macroeconomic risk through short-rate expectations. I therefore conclude that the recent inversion is not likely to signal an imminent recession. The strength of the current U.S. economy supports this conclusion. For instance, in May 2019 the unemployment rate had been at or below 4 percent for more than a year. Moreover, it has been argued that the relationship between yields and macroeconomic risk is currently weak due to distress in bond markets caused by low levels of interest rates and quantitative easing programs that are yet to be reversed.[34]

What does the yield curve inversion, then, reflect? Global factors and in particular the trade war between the U.S. and China are likely to play a role. Obviously, these circumstances can impact the domestic economy and contribute to a recession in the future. My results therefore do not reject all signs of a recession. However, I emphasize that the yield curve inversion is not currently related to macroeconomic risk and therefore does not signal that the risk of a recession is imminent. Thus, the Federal Reserve should not begin a cycle of monetary policy easing based on the yield curve inversion.

2.7 Conclusion

Using a novel macro-finance term structure model with time-varying volatility and variance risk premia, I document a time-varying relationship between the yield curve and macroeconomic risk. In particular, I estimate macroeconomic shares of yield curve volatility that fluctuate over time both at the monthly frequency and over longer periods. I relate large month-to-month fluctuations to meaningful economic events. I characterize the historical joint dynamics of yields and macroeconomic risk. First, I show a U-shaped pattern in the link between the short end of the yield curve and macroeconomic risk. Macroeconomic variables are important short-term yield factors in the 1970s-1980s, become less relevant during the Great Moderation, and regain importance after the financial crisis. Second, investors increasingly form expectations on future short rates based on macroeconomic risk. Third, macroeconomic shares of variation in 10-year term premia increase during the financial crisis due to deflation fears.

In conclusion, my results show that the assumption of constant volatility in standard Gaussian term structure models shield information about: (i) Fluctuations

[34]Research from Pictet Wealth Management concludes that “the distortions created by extraordinary post-crisis monetary policies have led to the breakdown in the relationship between interest rate expectations and economic growth" (“Why the yield curve is not the economic guide it once was," April 2, 2019, `https://www.ft.com/content/15d4048e-552f-11e9-91f9-b6515a54c5b1`).

in the relationship between the yield curve and macroeconomic risk. (ii) Countercyclicality of term premia. (iii) The demand of compensation for macroeconomic uncertainty.

My model provides new insights about the predictive power of the 2019 yield curve inversion for a future recession. I find a weak link between the recent inversion and current conditions of the economy. Specifically, the large decline in long-term yield is not driven by inflation and the unemployment gap. I conclude that the yield curve inversion does not reflect low expectations on future economic growth and, consequently, is not likely to predict a recession in the near future.

Acknowledgments

I thank Torben G. Andersen, Martin M. Andreasen, Simon Lund Hetland, Erik Hjalmarsson, Mike Joyce, Iryna Kaminska, Tamas Kiss, Philipp Kless, Heino Bohn Nielsen, Søren Hove Ravn, Rasmus Søndergaard Pedersen, Paolo Pesenti, Anders Rahbek, Guillaume Roussellet, Claudia Sahm, and Viktor Todorov. I also thank seminar and conference participants at the Adam Smith Business School, Deutsche Bundesbank, Federal Reserve Bank of Richmond, University of California San Diego, Anderson School of Management, University of North Carolina at Chapel Hill, Tilburg University, Bank of Spain, Bank of England, Danmarks Nationalbank, Kellogg School of Management, University of Chicago, University of Copenhagen, University of Gothenburg, the Quantitative Economics Doctorate Annual Meeting 2019, the 3rd International Conference on Econometrics and Statistics, the 13th Nordic Summer Symposium in Macroeconomics, and the 3rd Annual Workshop on Financial Econometrics at Örebro University.

Appendix

B.1 Technical Appendix

B.1.1 Structural Justification of Exponential-Quadratic Pricing Kernels

I provide two examples from the equilibrium asset pricing literature that result in exponential-quadratic pricing kernels. In both examples, the pricing kernel is determined by the marginal rate of intertemporal substitution of a representative household with Epstein and Zin (1989), Kreps and Porteus (1978), and Weil (1990) recursive preferences.

Long-Run Risk Model with Second-Order Projection Solution

Consider an indirect utility function given by

$$\mathcal{U}_t = \frac{1}{1-\frac{1}{\psi}} C_t^{1-\frac{1}{\psi}} + \beta \left(\mathbb{E}_t \left[\mathcal{U}_{t+1}^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$$

and a pricing kernel given by

$$\mathcal{M}_{t+1} = \beta \exp(-r_t) \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left(\frac{\mathcal{U}_{t+1}}{\mathbb{E}_t \left(\mathcal{U}_{t+1}^{1-\alpha} \right)} \right)^{-\alpha},$$

where C_t is consumption. Bansal and Yaron (2004) specify a model with a long-run growth factor, g_t, and a volatility factor, σ_t. The state vector also includes log consumption, $c_t = \log C_t$. The state dynamics are given by

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_c \sigma_t z_{c,t+1}, \\
g_{t+1} &= \rho_g g_t + \sigma_g \sigma_t z_{g,t+1}, \\
\sigma_{t+1}^2 &= 1 - \rho_\sigma + \rho_\sigma \sigma_t^2 + \sigma_\sigma \varepsilon_{\sigma,t+1},
\end{aligned}$$

where $z_{c,t+1} \sim$ i.i.d. $\mathcal{N}(0,1)$, $z_{g,t+1} \sim$ i.i.d. $\mathcal{N}(0,1)$, and $z_{\sigma,t+1} \sim$ i.i.d. $\mathcal{N}(0,1)$.

Bansal and Yaron (2004) propose an analytical solution using a log-linearization. This approximation results in an exponential-affine pricing kernel. However, Pohl, Schmedders, and Wilms (2018) argue that a first-order approximation is likely to generate large numerical errors. In response, Andreasen and Jørgensen (2019) solve the model using a second-order projection in which

$$\begin{aligned}
u_t &= \gamma_0^u + \gamma_X^u X_t + X_t' \gamma_{XX}^u X_t, \\
\tilde{u}_t &= \gamma_0^{\tilde{u}} + \gamma_X^{\tilde{u}} X_t + X_t' \gamma_{XX}^{\tilde{u}} X_t,
\end{aligned}$$

where $u_t = \log \mathcal{U}_t$, $\tilde{u}_t = \log \mathbb{E}_t \left(\mathcal{U}_{t+1}^{1-\alpha} \right)$, and $X_t = [c_t', g_t', \sigma_t^{2'}]'$. The implied log pricing kernel is then given by

$$\begin{aligned}
\log \mathcal{M}_{t+1} &= \log \beta - r_t - \frac{1}{\psi} \Delta c_{t+1} - \alpha u_{t+1} + \alpha \tilde{u}_{t+1} \\
&\propto \xi' X_{t+1} + X_{t+1}' \Xi X_{t+1}
\end{aligned}$$

with $\xi = -\frac{1}{\psi} - \alpha \gamma_X^u + \alpha \gamma_X^{\tilde{u}}$ and $\Xi = \gamma_{XX}^u + \gamma_{XX}^{\tilde{u}}$. This pricing kernel is a special case of (2.5) with constant prices of risk.

Long-Run Risk Model with Time-Varying Wealth-Consumption Ratio

Monfort and Pegoraro (2012) note that the exponential-quadratic kernel can be justified by a long-run risk model with a time-varying wealth-consumption ratio as considered in Hansen and Heaton (2008) and Hansen, Heaton, Roussanov, and Lee (2007). Here, I derive the details to clarify this justification. Indirect utility is specified by

$$\mathcal{U}_t = \left\{(1-\beta)C_t^{1-\rho} + \beta\left[\mathbb{E}_t\left(\mathcal{U}_{t+1}^{1-\gamma}\right)\right]^{\frac{1-\rho}{1-\gamma}}\right\}^{\frac{1}{1-\rho}}.$$

The log-consumption dynamics are defined by

$$\Delta c_{t+1} = \mu_c + \rho_{cx} x_t + \sigma_c z_{t+1},$$

where x_t is the state vector that follows first-order vector autoregressive dynamics given by

$$x_{t+1} = \Phi x_t + \Sigma z_{t+1}$$

with $z_t \sim$ i.i.d. $\mathcal{N}(0, \mathrm{I})$. Indirect utility is related to the ratio of wealth to consumption by

$$\frac{W_t}{C_t} = \frac{1}{1-\beta}\left(\frac{\mathcal{U}_t}{C_t}\right)^{1-\rho}.$$

When the wealth-consumption ratio is constant, i.e., when $\rho = 1$, the model implies a linear-exponential pricing kernel given by

$$\log \mathcal{M}_{t+1,1} = \mu_{\mathcal{M}} + \rho_{\mathcal{M}x} x_t + \sigma_{\mathcal{M}} z_{t+1}.$$

In the more realistic case of $\rho \neq 1$, Hansen and Heaton (2008) derives an approximate solution by expanding around the case of $\rho = 1$:

$$\begin{aligned}&\log \mathcal{M}_{t+1,1}\\ &\quad\approx \log \mathcal{M}_{t+1,1}|_{\rho=1} + (\rho-1)\left[\frac{1}{2} z_{t+1}' \Gamma_1 z_{t+1} + z_{t+1}' \Gamma_1 x_t + \varrho_0 + \varrho_{\mathcal{M}x} x_t + \varrho_{\mathcal{M}z} z_{t+1}\right].\end{aligned}$$

Thus, by allowing for a time-varying wealth-consumption ratio, the pricing kernel becomes quadratic as in (2.5) with prices of risk given by $\xi_t = \varrho_{\mathcal{M}z} + \Gamma_1 x_t$ and $\Xi = \frac{1}{2}\Gamma_1$.

B.1.2 Proof of Theorem 2.4

The proof follows the steps in Joslin, Singleton, and Zhu (2011) closely. I rely on invariant affine transformations of the state vector, X_t, given by $\mathcal{P}_t = c + CX_t$, where transformation of the observed macroeconomic variables can be prevented by restricting

$$c = \begin{bmatrix} c_x \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_x & 0 \\ 0 & \mathrm{I}_{n_m} \end{bmatrix}. \tag{B.1}$$

The following lemma gives the model parameters resulting from this rotation.

Lemma 2.1 *Consider the affine transformation $\mathcal{P}_t = c + CX_t$ restricted by* (B.1)*. Applying this transformation to the model in* (2.3)-(2.10) *gives an observationally equivalent model with state vector $\mathcal{P}_t = (\rho_t', m_t)'$ and parameters given by*

$$\begin{aligned}
\mu_{\mathcal{P}} &= c + C\mu_X - \sum_{l=1}^{L} C\Phi_X^{(l)} C^{-1} c, \\
\Phi_{\mathcal{P}}^{(l)} &= C\Phi_X^{(l)} C^{-1}, \quad l = 1, \ldots, L \\
\Sigma_{\mathcal{P}} &= C\Sigma_X, \\
A_{\mathcal{P}}^{(k)} &= CA_X^{(k)} C^{-1}, \quad k = 1, \ldots, K \\
B_{\mathcal{P}}^{(k)} &= CB_X^{(k)} C^{-1}, \quad k = 1, \ldots, K \\
\alpha_{\mathcal{P}} &= \alpha_X - \beta_x' C_x^{-1} c_x, \\
\beta_{\mathcal{P}}' &= \beta_X' C^{-1}, \\
\mu_{\mathcal{P}}^{\mathbb{Q}} &= c + C\mu_X^{\mathbb{Q}} - C\Phi_X^{\mathbb{Q}} C^{-1} c, \\
\Phi_{\mathcal{P}}^{\mathbb{Q}} &= C\Phi_X^{\mathbb{Q}} C^{-1}, \\
V_{\mathcal{P}}^{\mathbb{Q}} &= CV_X^{\mathbb{Q}} C'.
\end{aligned}$$

Proof. Straightforward. □

To prove uniqueness, I apply the following lemma, which is identical to Proposition 1 in JSZ. Although my model differs from the Gaussian term structure model considered in JSZ, the lemma remains valid because the model involves Gaussian-affine risk neutral dynamics, which is the only part of the model invoked in the proof of the result.

Lemma 2.2 *The model defined by* (2.3)-(2.10) *with synthetic yields defined in* (2.17) *is observationally equivalent to a model in real ordered Jordan form with* $r_t^{\perp m} \equiv r_t - \alpha_m - \beta_m' m_t = \iota' x_t$*, where* ι *is a vector of ones. The parameters determining the* $\mathbb{Q}$*-dynamics of this model are given by* $\mu_x^{\mathbb{Q}} = (k_\infty^{\mathbb{Q}}, 0, \ldots, 0)'$*, a positive definite matrix* $V_x^{\mathbb{Q}}$*, and* $\Phi_x^{\mathbb{Q}} = diag(J_1^{\mathbb{Q}}, J_2^{\mathbb{Q}}, \ldots, J_D^{\mathbb{Q}})$*, where for* $d = 1, \ldots, D$*,*

$$J_d^{\mathbb{Q}} = \begin{bmatrix} \lambda_d^{\mathbb{Q}} & 1 & \ldots & 0 \\ 0 & \lambda_d^{\mathbb{Q}} & \ldots & 0 \\ \vdots & \vdots & \ddots & 1 \\ 0 & \ldots & 0 & \lambda_d^{\mathbb{Q}} \end{bmatrix}.$$

Proof. See JSZ. □

Given these lemmas, I prove Theorem 2.4. The portfolios of synthetic yields given by $\rho_t = W Y_t^{\perp m}$ are affine transformation of the latent factors because

$$\rho_t = W Y_t^{\perp m} = W a_x + W b_x x_t = a_W + b_W x_t.$$

Thus, $\mathcal{P}_t = c + C X_t$ with

$$c = \begin{bmatrix} a_W \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} b_W & 0 \\ 0 & \mathrm{I}_{n_m} \end{bmatrix}. \tag{B.2}$$

By an application of Lemma 2.1, the model can be rotated into an equivalent model with state vector given by $\mathcal{P}_t$. I prove by contradiction that this model is unique. Suppose that two models with state vector $\mathcal{P}_t$ exist. Assume that these models are parametrized by, respectively, Θ_1 and Θ_2. By Lemma 2.2, each model is observationally equivalent to the model in real ordered Jordan form with $r_t^{\perp m} = \iota' x_t$, whose parametrization I denote by Θ_J. It follows that $\Theta_1 = \Theta_J = \Theta_2$.

B.1.3 Proof of Theorem 2.5

By Lemma 2.2, the model defined by (2.3)-(2.10) with synthetic yields defined in (2.17) can be rewritten in real ordered Jordan form and the rotation only affects parameters of the $\mathbb{Q}$-dynamics and the short-rate equation. These rotated parameters are given by

$$\left\{\alpha_X, \quad \beta_x, \quad \mu_x^{\mathbb{Q}}, \quad \Phi_x^{\mathbb{Q}}, \quad V_x^{\mathbb{Q}}\right\} = \left\{0, \quad \iota, \quad k_\infty^{\mathbb{Q}} e_1, \quad J(\lambda^{\mathbb{Q}}), \quad V_x^{\mathbb{Q}}\right\},$$

where e_1 is a vector of zeros except with the first entry equal to one and $V_x^{\mathbb{Q}}$ is positive definite. Given the model in Jordan form from Lemma 2.2, I apply the invariant transformation $\mathcal{P}_t = c + CX_t$ with c and C given by (B.2). By an application of Lemma 2.1, the parameters of the rotated model are given by $\Theta_\rho^{\mathbb{Q}} = \{k_\infty^{\mathbb{Q}}, \lambda^{\mathbb{Q}}, V_\rho^{\mathbb{Q}}\}$ and

$$\begin{aligned}\Theta_{\mathcal{P}}^{\mathbb{P}} &\equiv \left\{\mu_{\mathcal{P}}, \quad \{\Phi_{\mathcal{P}}^{(l)}\}_{l=1}^{L}, \quad \Sigma_{\mathcal{P}}, \quad \{A_{\mathcal{P}}^{(k)}, B_{\mathcal{P}}^{(k)}\}_{k=1}^{K}\right\} \\ &= \left\{c + C\mu_X - \textstyle\sum_{l=1}^{L} C\Phi_{\mathcal{P}}^{(l)}C^{-1}c, \quad \{C\Phi_X^{(l)}C^{-1}\}_{l=1}^{L}, \quad C\Sigma_X, \quad \{CA_{\mathcal{P}}^{(k)}C^{-1}, CB_{\mathcal{P}}^{(k)}C^{-1}\}_{k=1}^{K}\right\}.\end{aligned}$$

Since $\Theta_X^{\mathbb{P}} = \left\{\mu_X, \{\Phi_X^{(l)}\}_{l=1}^{L}, \Sigma_X, \{A_X^{(k)}, B_X^{(k)}\}_{k=1}^{K}\right\}$ is not involved in the rotation into the Jordan form, restricting $\Theta_{\mathcal{P}}^{\mathbb{P}}$ is not necessary to preclude rotations of the state vector given $\Theta_\rho^{\mathbb{Q}} = \{k_\infty^{\mathbb{Q}}, \lambda^{\mathbb{Q}}, V_\rho^{\mathbb{Q}}\}$.

Restrictions are, however, necessary to identify parameters of the conditional mean under the $\mathbb{P}$-measure. To ensure a unique Cholesky decomposition, $\Sigma_{\mathcal{P}}\Sigma_{\mathcal{P}}'$, I restrict the diagonal of the lower triangular matrix $\Sigma_{\mathcal{P}}$ to be strictly positive. Moreover, since $A_{\mathcal{P}}^{(k)}$ and $B_{\mathcal{P}}^{(k)}$ enter the model in quadratic form, I require the first elements in these matrices to be non-negative. The following lemma derives an upper bound for K that is necessary for identification.

Lemma 2.3 *Consider the variance specification given by*

$$V_t = \Sigma_{\mathcal{P}}\Sigma_{\mathcal{P}}' + \sum_{k=1}^{K} A_{\mathcal{P}}^{(k)}\varepsilon_{t-1}\varepsilon_{t-1}'A_{\mathcal{P}}^{(k)'} + \sum_{k=1}^{K} B_{\mathcal{P}}^{(k)}V_{t-1}B_{\mathcal{P}}^{(k)'}, \tag{B.3}$$

where $\Sigma_{\mathcal{P}}$ is lower triangular with strictly positive elements on the diagonal. Let $A_{\mathcal{P}}^{(k)}$ and $B_{\mathcal{P}}^{(k)}$ be diagonal matrices for $k = 1, \ldots, K$. Then, a necessary condition for the parameters to be identified is $K \leq$ floor $\left(\frac{1}{2}n_X + \frac{1}{2}\right)$, where n_X is the dimension of V_t.

Proof. I give the proof for a trivariate model, but the arguments can be extended to higher dimensions. I assume that $K = \text{floor}\left(\frac{1}{2}n_X + \frac{1}{2}\right) = 2$ and show that this degree of generalization gives the maximum number of identified parameters. The ARCH and GARCH terms of the BEKK GARCH model can be treated separately. Furthermore, the arguments for the two terms are identical. I focus on the ARCH term below. Let

$$A_{\mathcal{P}}^{(1)} = \begin{bmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{1,2} & 0 \\ 0 & 0 & a_{1,3} \end{bmatrix} \quad \text{and} \quad A_{\mathcal{P}}^{(2)} = \begin{bmatrix} a_{2,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{2,3} \end{bmatrix}.$$

Then,

$$A_{\mathcal{P}}^{(1)}\varepsilon_t\varepsilon_t' A_{\mathcal{P}}^{(1)'} + A_{\mathcal{P}}^{(2)}\varepsilon_t\varepsilon_t' A_{\mathcal{P}}^{(2)'} =$$

$$\begin{bmatrix} \left(a_{1,1}^2 + a_{2,1}^2\right)\varepsilon_{1,t}^2 & \left(a_{1,1}a_{1,2} + a_{2,1}a_{2,2}\right)\varepsilon_{1,t}\varepsilon_{2,t} & \left(a_{1,1}a_{1,3} + a_{2,1}a_{2,3}\right)\varepsilon_{1,t}\varepsilon_{3,t} \\ \left(a_{1,1}a_{1,2} + a_{2,1}a_{2,2}\right)\varepsilon_{1,t}\varepsilon_{2,t} & \left(a_{1,2}^2 + a_{2,2}^2\right)\varepsilon_{2,t}^2 & \left(a_{1,2}a_{1,3} + a_{2,2}a_{2,3}\right)\varepsilon_{2,t}\varepsilon_{3,t} \\ \left(a_{1,1}a_{1,3} + a_{2,1}a_{2,3}\right)\varepsilon_{1,t}\varepsilon_{3,t} & \left(a_{1,2}a_{1,3} + a_{2,2}a_{2,3}\right)\varepsilon_{2,t}\varepsilon_{3,t} & \left(a_{1,3}^2 + a_{2,3}^2\right)\varepsilon_{3,t}^2 \end{bmatrix}$$

Suppose that $a_{2,1}$, $a_{2,2}$, and $a_{2,3}$ are identified. Then, $a_{1,1}$, $a_{1,2}$, and $a_{1,3}$ follows from the coefficients on respectively $\varepsilon_{1,t}^2$, $\varepsilon_{2,t}^2$, and $\varepsilon_{3,t}^2$. This leaves the following three equations to show that $a_{2,1}$, $a_{2,2}$, and $a_{2,3}$ are indeed identified:

$$a_{2,1}a_{2,2} = c_1$$

$$a_{2,1}a_{2,3} = c_2$$

$$a_{2,2}a_{2,3} = c_3.$$

This system is obviously identified. Thus, the model is identified with $K = \text{floor}(\frac{1}{2}n_X + \frac{1}{2})$. Naturally, it follows that the model is also identified with $K < \text{floor}\left(\frac{1}{2}n_X + \frac{1}{2}\right)$. Now, suppose that K exceeds floor $\left(\frac{1}{2}n_X + \frac{1}{2}\right) = 2$, say $K = 3$. For the ARCH equation, this leaves $n_X(n_X+1)/2 = 6$ equations to identify $Kn_X = 9$ parameters, which is not feasible. □

B.2 Model Specification

In this appendix, I show that the specified model in terms of lag length under the physical measure and generalization of the conditional variance matrix is preferred by the data. I also provide parameter estimates and results from misspecification testing which indicate that the model is well specified.

B.2.1 Lag Length of Macroeconomic Variables

I test for the lag length of the macroeconomic variables under the physical measure given the choice of generalization of the conditional variance matrix of $K = 2$. Model selection with respect to the choice of K is considered in Section B.2.3. Table B.1 shows information criteria and likelihood-ratio tests for general-to-specific model selection. The maximum number of lags considered is an annual lag length, $L = 12$. I only report results for lag lengths of one months and for each quarter, i.e., $L = \{1, 3, 6, 9, 12\}$. Choices of L in-between these values do not give further insights. The table unambiguously shows that an extensive lag structure is needed to match the data. The maximal lag length, $L = 12$, is preferred by all information criteria, and the likelihood-ratio tests reject all specifications with lower lag lengths.

Table B.1: Criteria for Lag-Length Selection of Macroeconomic Variables

	Log-likelihood	AIC	BIC	HQC	LR-test
L=12	-335.10	**868.21**	**942.43**	**756.75**	-
L=9	-371.11	936.22	1009.17	784.87	72.02 [0.00]
L=6	-391.72	953.44	1017.55	820.84	41.22 [0.00]
L=3	-421.78	989.56	1044.80	875.71	60.12 [0.00]
L=1	-445.50	1020.99	1070.27	919.63	47.44 [0.00]

Note: The table shows criteria for lag-length selection of the macroeconomic variables under the physical measure. Log-likelihood values relate to the estimation of the physical dynamics. Information criteria are the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQC). Minimum values of these criteria across model specifications are in bold. Likelihood-ratio-(LR-)test statistics are reported with p-values in brackets. The test is performed under the null that the log-likelihood value for a given L is equivalent to that of the model with L+1. The models are estimated with the K=2 components in the conditional variance matrix.

B.2.2 Lag Structure in Latent Factors

By allowing for a lag structure not only in the macroeconomic variables, but also in the latent factors, I estimate a maximum log-likelihood value of -226.30. However, this model involves $(L-1)\cdot N_x^2 = 99$ additional parameters, which result in information criteria given by 868.26 (AIC), 1024.55 (BIC), and 544.97 (HQC). Thus, with the exception of the HQC for which the penalty is extremely small, I find that the additional number of parameters involved in allowing for a lag structure in the latent factors is not justified by a sufficiently large improvement in the in-sample fit. In addition, I find that the many additional parameters result in the estimation problem in large estimation inaccuracy.

B.2.3 Generalization of the Conditional Variance Matrix

With three latent factors and two macroeconomic variables, the model is identified with up to $K = 3$ components in the conditional variance matrix, see part (ii) of Theorem 2.5 in the main paper. In this appendix, I test the model with choices of $K = \{1, 2, 3\}$. First, I consider information criteria and likelihood-ratio tests in Table B.2. I observe a large increase in the log-likelihood value from choosing $K = 2$ compared to the simplest model with $K = 1$. This difference persists when punishing for the additional parameters as reflected by all information criteria. Moreover, the difference is statistically significant. Thus, as explained in the main paper, it is

Table B.2: Criteria for Generalization of the Conditional Variance Matrix

	Log-likelihood	AIC	BIC	HQC	LR-test
K=3	-332.69	873.38	951.35	**756.75**	-
K=2	-335.10	**868.21**	**942.43**	757.19	1.30 [0.99]
K=1	-363.75	915.50	985.98	810.09	57.30 [0.00]

Note: The table shows criteria for selection of the generalization of the conditional variance matrix under the physical measure. Log-likelihood values relate to the estimation of the physical dynamics. Information criteria are the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQC). Minimum values of these criteria across model specifications are in bold. Likelihood-ratio-(LR-)test statistics are reported with p-values in brackets. The test is performed under the null that the log-likelihood value for a given K is equivalent to the model with K+1. The models are estimated with the L=12 lags in the macroeconomic variables.

necessary to include multiple components in the conditional variance matrix to fit the data. In contrast, the improvement in fit obtained by increasing K to $K = 3$ is modest and the AIC and BIC are lower for $K = 2$. The HQC is practically the same for the models with $K = 2$ and $K = 3$. Finally, the difference between these specifications is not statistically significant. As a result, $K = 2$ seems to be sufficient to match the data.

B.2.4 Parameter Estimates

Parameter estimates of the selected model are shown in Tables B.3, B.4, and B.5. Almost all parameters are highly significant by conventional standard normal critical values.

B.2.5 Misspecification Testing

Figure B.1 shows the standardized residuals, their histograms superimposed with the Gaussian density, and autocorrelation functions. The figure also shows autocorrelation functions of the absolute values of standardized residuals, which serve as a test of the ability of the conditional variance matrix to eliminate ARCH effects. The standardized residuals appear to be Gaussian and the autocorrelation functions of both levels and absolute values are mostly insignificant. There is, however, some autocorrelation in the residual of the equation for inflation.

Table B.3: Parameter Estimates Related to the Physical Conditional Mean

$\mu_{\mathcal{P}}$				
-0.103 (0.180)	-0.253 (0.048)	0.069 (0.027)	-0.014 (0.060)	-0.058 (0.023)
$\Phi_{\mathcal{P}}^{(1)}$				
0.990 (0.007)	0.026 (0.049)	-0.035 (0.175)	-0.032 (0.035)	-0.050 (0.036)
-0.001 (0.003)	0.935 (0.017)	0.226 (0.057)	0.021 (0.009)	0.000 (0.013)
0.002 (0.001)	-0.005 (0.007)	0.851 (0.025)	-0.002 (0.005)	0.018 (0.006)
0.001 (0.003)	-0.017 (0.018)	0.025 (0.061)	1.057 (0.030)	-0.124 (0.056)
-0.001 (0.001)	0.002 (0.007)	0.052 (0.027)	-0.002 (0.020)	0.889 (0.065)
$\Phi_m^{(2)}$		$\Phi_m^{(3)}$		
-0.071 (0.039)	0.030 (0.061)	0.005 (0.033)	0.003 (0.051)	
0.012 (0.036)	0.175 (0.061)	-0.014 (0.037)	0.010 (0.056)	
$\Phi_m^{(4)}$		$\Phi_m^{(5)}$		
0.052 (0.033)	0.017 (0.048)	-0.057 (0.037)	-0.009 (0.054)	
0.034 (0.038)	0.007 (0.058)	-0.058 (0.040)	0.019 (0.061)	
$\Phi_m^{(6)}$		$\Phi_m^{(7)}$		
-0.010 (0.037)	0.003 (0.052)	0.047 (0.033)	0.002 (0.049)	
0.054 (0.035)	-0.018 (0.059)	-0.024 (0.030)	-0.064 (0.059)	
$\Phi_m^{(8)}$		$\Phi_m^{(9)}$		
-0.048 (0.039)	0.049 (0.048)	0.062 (0.041)	0.027 (0.058)	
0.030 (0.031)	-0.051 (0.056)	0.052 (0.034)	-0.008 (0.054)	
$\Phi_m^{(10)}$		$\Phi_m^{(11)}$		
-0.029 (0.038)	0.021 (0.055)	0.003 (0.037)	-0.002 (0.053)	
0.052 (0.034)	-0.008 (0.054)	-0.035 (0.032)	0.115 (0.056)	
$\Phi_m^{(12)}$				
-0.013 (0.022)	-0.042 (0.041)			
0.021 (0.021)	-0.087 (0.044)			

Note: The table shows parameter estimates related to the conditional mean of the state vector under the physical measure. Robust standard errors computed by the Huber sandwich estimator are reported in parentheses.

Table B.4: Parameter Estimates Related to the Physical Conditional Variance

$\Sigma_{\mathcal{P}}$				
0.247 (0.037)	-	-	-	-
-0.007 (0.011)	0.084 (0.011)	-	-	-
-0.025 (0.008)	-0.019 (0.007)	0.030 (0.008)	-	-
-0.081 (0.013)	0.046 (0.008)	-0.025 (0.010)	0.000 (0.000)	-
0.003 (0.015)	-0.025 (0.018)	-0.002 (0.018)	-0.054 (0.016)	0.000 (0.000)

diag($A_{\mathcal{P}}^{(1)}$)	diag($A_{\mathcal{P}}^{(2)}$)	diag($B_{\mathcal{P}}^{(1)}$)	diag($B_{\mathcal{P}}^{(2)}$)
0.026 (0.106)	0.448 (0.050)	0.893 (0.009)	0.087 (0.107)
0.133 (0.081)	0.408 (0.046)	0.892 (0.011)	0.072 (0.119)
0.290 (0.075)	0.282 (0.085)	0.883 (0.034)	-0.096 (0.177)
0.013 (0.111)	0.409 (0.047)	0.901 (0.010)	0.080 (0.110)
0.267 (0.164)	0.246 (0.196)	0.418 (0.260)	0.745 (0.155)

Note: The table shows parameter estimates of the conditional variance-matrix of the state vector under the physical measure. Robust standard errors computed by the Huber sandwich Estimator are reported in parentheses.

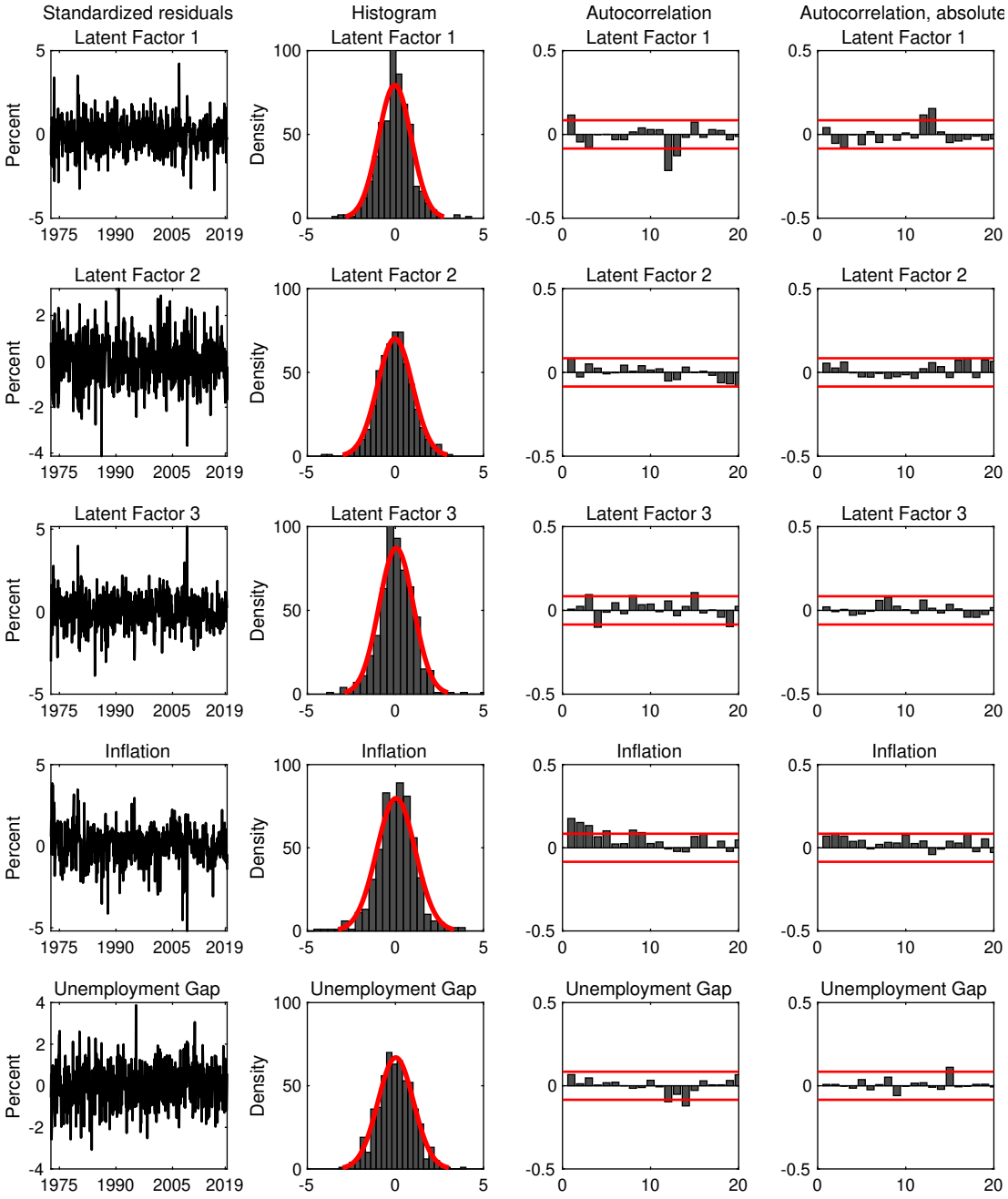
Table B.5: Parameter Estimates Related to the Risk-Neutral Measure

β_m	$\mu_m^{\mathbb{Q}}$	$\Phi_m^{\mathbb{Q}} - I_{n_m}$			$(V_m^{\mathbb{Q}})^{1/2}$	
0.943 (2.5×10^{-4})	4.087 (0.009)	-0.023 (0.001)	-0.042 (1.9×10^{-4})		0.925 (0.436)	-
-0.196 (0.3×10^{-4})	17.091 (0.048)	-0.051 (0.005)	-0.191 (0.001)		0.076 (0.230)	0.6×10^{-5} (0.012)

$k_\infty^{\mathbb{Q}}$	$\lambda^{\mathbb{Q}} - 1$	$(V_\rho^{\mathbb{Q}})^{1/2}$		
0.020 (0.069)	-0.003 (1.6×10^{-4})	1.251 (0.3×10^{-5})	-	-
	-0.024 (0.6×10^{-4})	-0.213 (0.8×10^{-7})	0.393 (0.018)	-
	-0.138 (0.004)	-0.027 (0.083)	-0.086 (0.070)	0.165 (0.062)

Note: The table shows parameter estimates related to the risk-neutral measure. Robust standard errors computed by the Huber sandwich estimator are reported in parentheses.

Figure B.1: Misspecification Testing



Notes: The figure shows misspecification tests of the standardized residuals shown in the first column. Charts in the second column exhibit histograms of the standardized residuals with a Gaussian density function superimposed. The third and fourth columns show autocorrelation functions of, respectively, levels and absolute values of standardized residuals for lags between 1 and 20 with 95% confidence intervals computed under the assumption that the standardized residuals are Gaussian white noise series.

B.3 Robustness Checks

B.3.1 Measures of Macroeconomic Variables

In this appendix, I apply the model to alternative choices of macroeconomic variables. In particular, I substitute CPI inflation with, respectively, core CPI and personal consumption expenditures (PCE) inflation, and the unemployment gap with the Chicago Fed National Activity Index. The empirical performance of the model does not change by the choice of macroeconomic variables. Instead, I focus on the broader validity of the main results related to the relationship between macroeconomic variables and the yield curve through time.

Core CPI Inflation

The literature has suggested that core inflation measures, i.e., inflation that excludes food and energy prices, are better indicators of the path of future inflation compared to headline inflation (Blinder and Reis, 2005, Mishkin, 2007). Core CPI inflation is available from the U.S. Bureau of Labor Statistics. As in the main paper, I use the unemployment gap as a measure of economic activity. The macroeconomic share of variance in bond yields, 10-year short-rate expectations, and the 10-year term premium are summarized in Table B.6.

By excluding food and energy prices from the CPI inflation series, inflation shocks explain more variation of yields on average, especially at the short end of the yield curve. This is consistent with the findings in Ajello, Benzoni, and Chyruk (2018), who show that while the impact of shocks to food and energy inflation on the yield curve is small, core-inflation shocks have a large impact on yields, in particular at short maturities. The macroeconomic shares of variances to the decomposition of the 10-year yield into short-rate expectations and term premium are smaller than those obtained with CPI inflation in the main paper. However, I maintain the conclusion that macroeconomic shocks primarily explain yield variation through the expectations channel. Using core CPI inflation also results in highly time-varying macroeconomic shares of variances with wide ranges and large month-to-month changes.

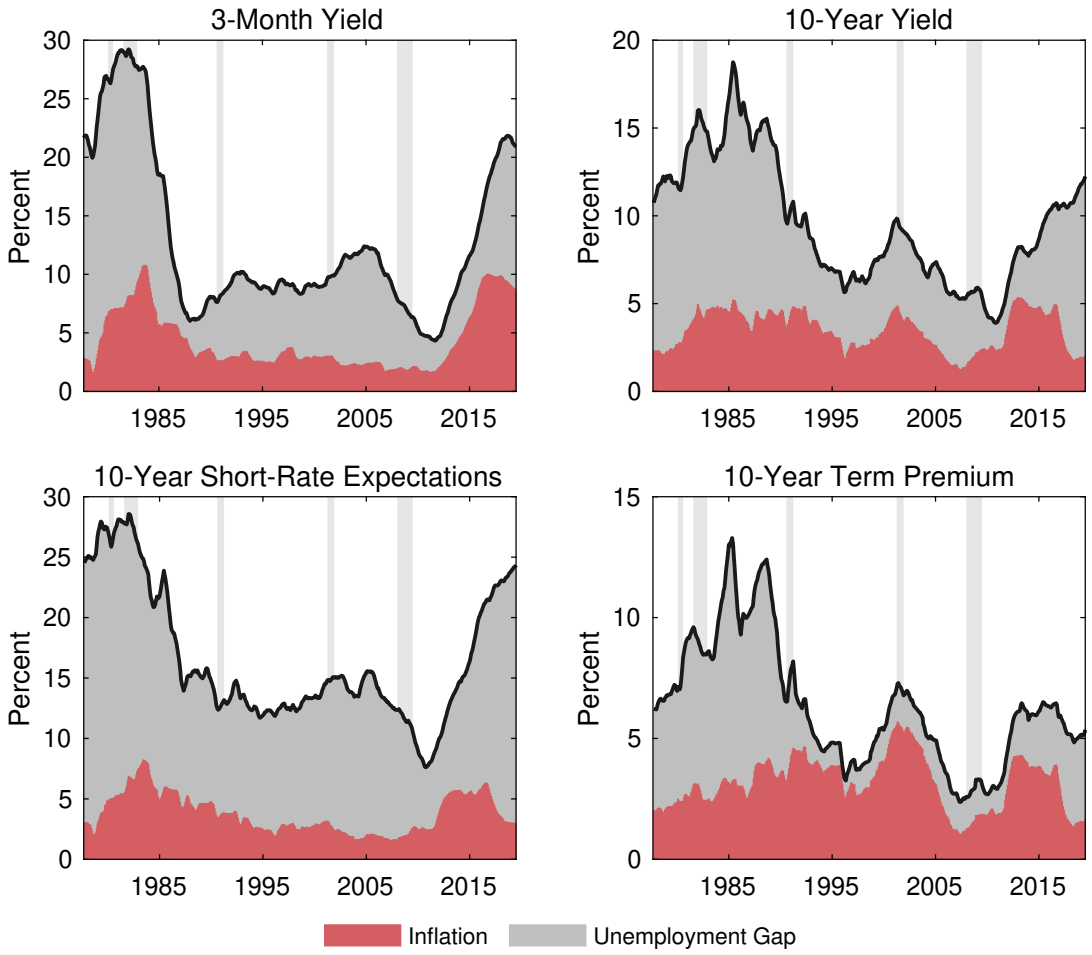
The 5-year moving averages of the estimated macroeconomic shares of variances are depicted Figure B.2. These exhibit similar patterns as the trends estimated using headline inflation. In particular, the U-shape is maintained at the short end of the yield curve. For long-term yields, I observe increased macroeconomic shares of variances after the Volcker period and in the most recent part of the sample.

Table B.6: Descriptives of Macroeconomic Shares of Variances Using Core CPI Inflation

	Yields								Short-rate expectations	Term premia
Maturities in years	0.25	0.5	1	2	3	5	7	10	10	10
Average	13.71	13.76	11.81	10.27	10.1	10.28	10.21	9.80	17.30	6.20
Std dev	10.47	10.09	8.98	8.10	7.86	7.70	7.48	7.12	9.78	6.22
Skewness	1.24	1.14	1.40	1.53	1.45	1.25	1.13	1.04	0.67	2.00
Kurtosis	4.85	4.82	6.07	6.33	5.88	4.89	4.32	3.94	3.40	7.64
Min	0.02	0.06	0.02	0.01	0.08	0.23	0.17	0.02	0.04	0.02
Max	60.32	60.27	54.31	52.36	51.81	48.93	44.95	38.98	54.61	35.48
Median	11.96	12.61	10.41	8.42	8.23	8.4	8.49	8.37	16.00	4.40
Quantiles:										
10 percent	2.61	2.49	2.13	2.03	2.15	2.24	2.37	1.95	5.78	0.71
25 percent	5.73	5.70	4.80	4.24	4.36	4.36	4.56	4.46	9.61	2.04
75 percent	18.57	18.81	16.64	14.53	14.15	14.23	14.14	13.65	23.53	8.09
90 percent	27.95	27.26	22.18	20.24	19.60	20.78	20.95	19.80	30.08	13.48

Note: The table shows descriptive statistics of the estimates shares of variances in yields, 10-year short-rate expectations, and 10-year term premia determined by the macroeconomic variables (core CPI inflation and unemployment gap). The statistics are reported in percent. The table also shows the constant macroeconomic share of variances estimated using the Gaussian term structure model.

Figure B.2: 5-Year Moving-Average Macroeconomic Shares of Variances with Core CPI Inflation



Notes: The figure shows 5-year moving averages of the macroeconomic contribution to variances in the 3-month and 10-year yields, 10-year short-rate expectations, and the 10-year term premium. The macroeconomic variables are core CPI inflation and the unemployment gap. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

PCE Inflation

Since PCE inflation is an important indicator for U.S. policy makers, I also report results using this measure of inflation. I obtain PCE inflation data from the U.S. Bureau of Economic Analysis and continue to use the unemployment gap as a measure of economic activity. Macroeconomic shares of variances in yields, 10-year short-rate expectations, and 10-year term premia are summarized in Table B.7. These are quantitatively similar to the results reported in the main paper for CPI inflation when considering variation in the yield curve. The macroeconomic shares are, however, larger on average for the 10-year short-rate expectations and term premium compared to the main paper. This difference may reflect that the Federal Reserve does indeed pay attention to PCE inflation, which results in a tighter link between investors' beliefs about future monetary policy and the macroeconomic variables.

The 5-year moving averages of the macroeconomic shares are reported in Figure B.3. For the 3-month and 10-year yields, the patterns in these series are similar to those obtained with CPI inflation. For the 10-year short-rate expectations, the upward trend is only observed between 1985 and 1990, whereafter the macroeconomic shares are constant. For the macroeconomic shares of variance in the 10-year term premium, I find an even stronger link to inflation than reported in the main paper. This result suggests that the inflation risk premium involving PCE inflation is higher than that of CPI inflation.

Chicago Fed National Activity Index

The unemployment gap is a measure of economic slack. Another interpretation of economic activity, which has been applied in macro-finance, is growth in real variables. In the following, I apply the model to CPI inflation and the 3-month moving average of growth in the Chicago Fed National Activity Index (CFNAI) index.[35]

Table B.8 summarizes the macroeconomic shares of variances in yields, 10-year short-rate expectations, and 10-year term premia with these data. Compared to the results obtained with the unemployment gap in the main paper, I find slightly lower average macroeconomic shares of yield curve variances, but similar ranges and kurtosis. This difference may reflect that the unemployment gap is more closely related to the yield curve through monetary policy rules and the Federal Reserve's dual mandate compared to growth in real variables.

[35]The Chicago Fed National Activity Index is constructed by a weighted average of 85 economic indicators within four categories: production and income (23 indicators); employment, unemployment, and hours (24 indicators); personal consumption and housing (15 indicators); and sales, orders, and inventories (23 indicators). The data series are adjusted for inflation.

Table B.7: Descriptives of Macroeconomic Shares of Variances Using PCE Inflation

	Yields								Short-rate expectations	Term premia
Maturities in years	0.25	0.5	1	2	3	5	7	10	10	10
Average	6.47	6.07	6.29	5.48	4.79	5.05	5.23	5.94	73.18	52.83
Std dev	7.13	6.68	6.35	5.59	5.35	5.84	5.58	6.06	11.60	17.11
Skewness	2.29	2.61	2.53	2.40	3.01	3.30	2.57	1.96	-1.34	-0.28
Kurtosis	10.89	13.2	12.29	10.07	15.68	19.48	12.57	7.35	5.42	2.37
Min	0.03	0.03	0.07	0.02	0.00	0.01	0.03	0.03	24.29	10.33
Max	55.29	53.2	46.51	35.63	41.9	49.2	41.33	38.06	92.34	88.17
Median	4.00	3.78	4.83	4.00	3.31	3.20	3.29	3.84	75.28	53.95
Quantiles;										
10 percent	0.67	0.76	0.80	0.86	0.69	0.65	0.76	0.96	60.11	27.00
25 percent	1.62	1.81	2.02	2.03	1.66	1.57	1.67	1.81	67.48	41.47
75 percent	9.02	8.04	8.18	6.67	5.78	6.53	7.21	7.59	81.73	65.17
90 percent	16.5	14.65	13.23	12.28	10.03	11.31	11.88	14.01	84.88	74.67

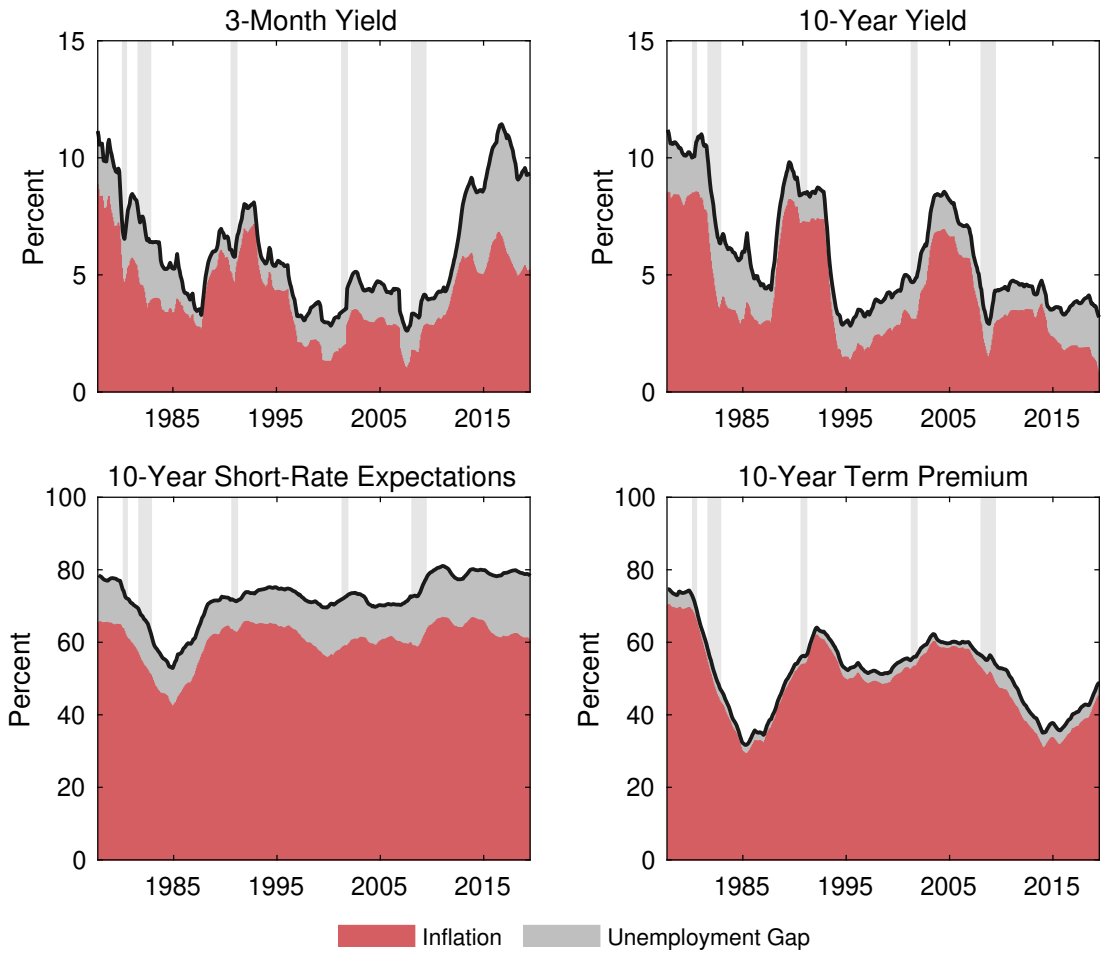
Note: The table shows descriptive statistics of the estimated shares of variances in yields, 10-year short-rate expectations, and 10-year term premia determined by the macroeconomic variables (PCE inflation and unemployment gap). The statistics are reported in percent. The table also shows the constant macroeconomic share of variances estimated using the Gaussian term structure model.

Table B.8: Descriptives of Macroeconomic Shares of Variances Using the CFNAI Series

	Yields								Short-rate expectations	Term premia
Maturities in years	0.25	0.5	1	2	3	5	7	10	10	10
Average	5.85	6.07	6.09	5.80	5.86	6.29	6.40	6.09	93.88	76.03
Std dev	6.36	6.64	6.73	6.61	6.73	7.21	7.47	7.34	5.04	9.60
Skewness	2.07	2.06	2.14	2.04	2.04	2.13	2.35	2.67	-2.23	-0.71
Kurtosis	8.89	8.69	9.18	7.95	8.08	8.77	10.31	12.73	10.16	3.35
Min	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	57.92	39.65
Max	47.38	48.02	46.25	40.08	42.56	45.37	46.9	50.21	98.83	94.11
Median	3.65	3.78	4.02	3.83	3.40	3.68	3.82	3.44	96.01	77.29
Quantiles:										
10 percent	0.57	0.62	0.52	0.39	0.36	0.41	0.59	0.56	86.97	62.04
25 percent	1.41	1.48	1.32	1.09	1.09	1.31	1.19	1.16	92.24	70.51
75 percent	7.78	8.24	8.39	7.92	8.34	9.14	9.41	9.09	97.19	82.85
90 percent	14.77	15.01	14.47	13.88	14.5	15.61	15.21	14.48	97.78	87.30

Note: The table shows descriptive statistics of the estimated shares of variances in yields, 10-year short-rate expectations, and 10-year term premia determined by the macroeconomic variables (CPI inflation and the CFNAI series). The statistics are reported in percent. The table also shows the constant macroeconomic share of variances estimated using the Gaussian term structure model.

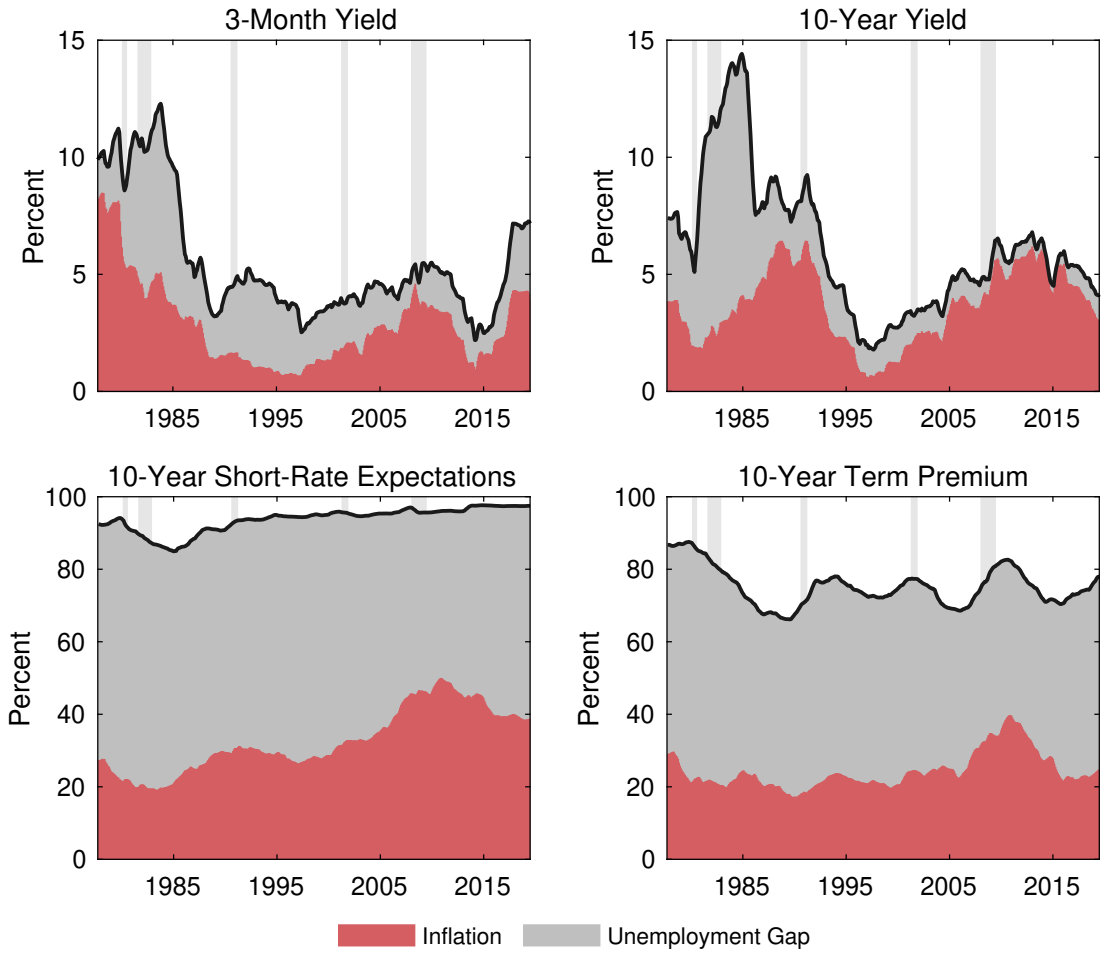
Figure B.3: 5-Year Moving-Average Macroeconomic Shares of Variances with PCE Inflation



Notes: The figure shows 5-year moving averages of the macroeconomic contribution to variances in the 3-month and 10-year yields, 10-year short-rate expectations, and the 10-year term premium. The macroeconomic variables are PCE inflation and the unemployment gap. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

The 5-year moving averages of the macroeconomic shares are reported in Figure B.4. The conclusions obtained on macroeconomic shares of yield curve variances with the unemployment gap in the main paper are maintained with the CFNAI series. However, the estimated role of macroeconomic variables for the term structure decomposition is different from the main paper. In particular, the CFNAI series and CPI inflation accounts for almost 100 percent of the variation in short-rate expectations throughout the sample. This difference is due to the role of the CFNAI series only, as the results for CPI inflation are similar to those reported in the main paper. Thus, whereas shocks to the unemployment gap have become increasingly important for

Figure B.4: 5-Year Moving-Average Macroeconomic Shares of Variances with the CFNAI Series



Notes: The figure shows 5-year moving averages of the macroeconomic contribution to variances in the 3-month and 10-year yields, 10-year short-rate expectations, and the 10-year term premium. The macroeconomic variables are CPI inflation and the CFNAI series. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

explaining movements in short-rate expectations, the CFNAI series has consistently been a very important factor that investors consider when forming these expectations. I also estimate higher macroeconomic shares of variance in the 10-year term premium with the CFNAI series.

B.3.2 Cholesky Ordering

The main results of the paper are based on the standard assumptions that (i) macroeconomic variables do not respond to contemporaneous shocks to the latent factors, and (ii) the unemployment gap does not respond to contemporaneous inflation shocks. I show that my results are not sensitive to these assumptions by evaluating the model under alternative orderings of the variables.

First, I consider assumption (i) by reversing the ordering of the macroeconomic variables and latent factors. Thus, I assume that the three latent factors respond to macroeconomic shocks with a delay of one month. I maintain assumption (ii). The resulting macroeconomic shares of variances in yields, 10-year short-rate expectations, and 10-year term premia are detailed in Table B.9. By ordering latent factors before the macroeconomic variables, I obtain larger macroeconomic shares on average compared to those reported in the main paper in Table 2.3. However, the macroeconomic shares remain highly time-varying with wide ranges and occasional large month-to-month changes.

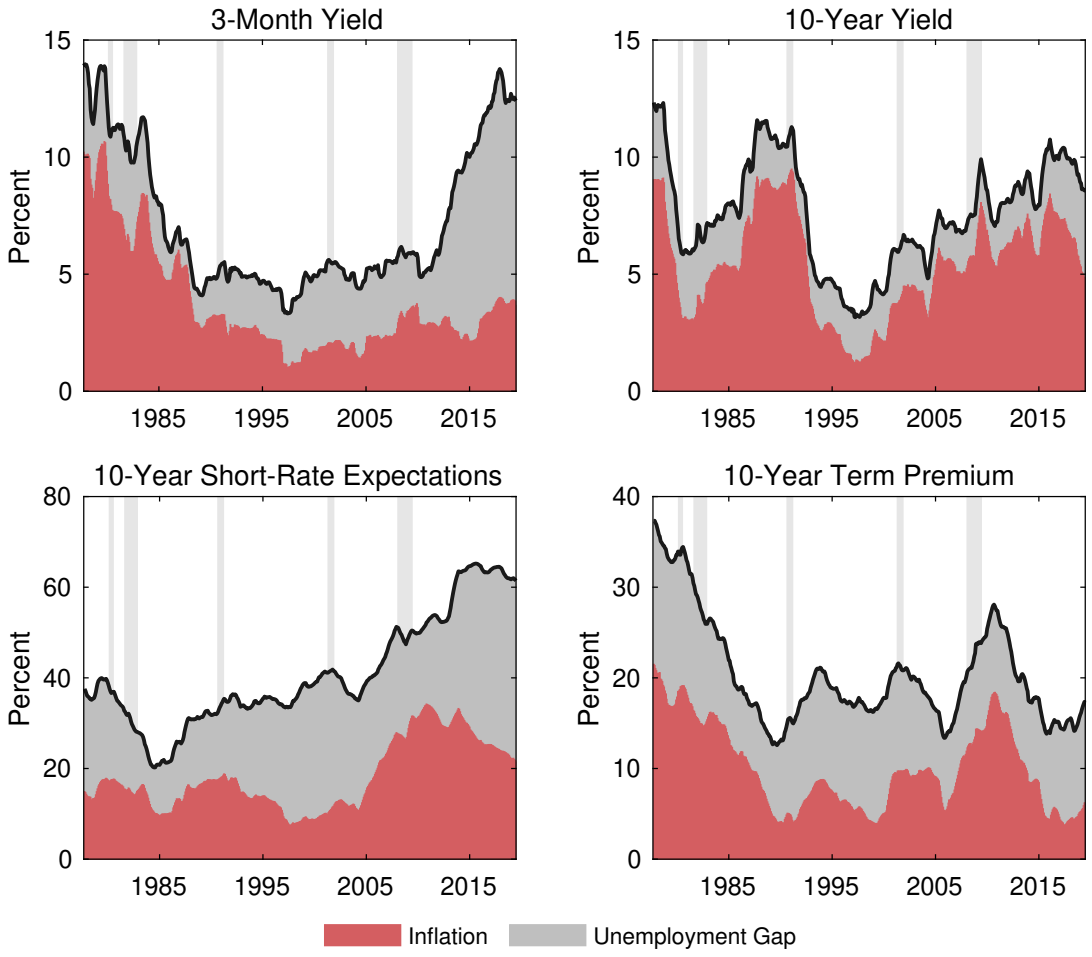
Next, I assume that inflation is slower moving than the unemployment gap, violating assumption (ii), but maintaining assumption (i). By construction, this alteration only affects the decomposition of macroeconomic shares into inflation and the unemployment gap, but not the total macroeconomic shares. I show the decomposition of the 5-year moving average of the macroeconomic shares in Figure B.5. When comparing with Figure 2.8 in the main paper, I see that reversing the order of the macroeconomic variables has no implications for my conclusions. In fact, even quantitatively, there are very few differences between the results reported in Figure B.5 and those reported for the chosen order in the main paper.

Table B.9: Descriptives of Macroeconomic Shares of Variances with Alternative Cholesky Ordering

	Yields								Short-rate expectations	Term premia
Maturities in years	0.25	0.5	1	2	3	5	7	10	10	10
Average	61.33	55.63	34.95	23.32	20.03	17.27	15.22	12.27	51.74	17.43
Std dev	26.54	23.15	15.26	12.55	11.92	11.9	11.75	10.37	13.96	6.90
Skewness	-0.57	-0.42	0.89	1.34	1.31	1.43	1.63	1.75	-0.08	0.75
Kurtosis	2.18	2.25	4.00	5.10	4.91	5.36	6.13	6.51	4.03	4.33
Min	2.72	2.70	2.72	2.92	3.22	2.92	1.98	1.12	5.12	2.44
Max	99.93	94.87	91.39	74.62	71.83	70.81	69.69	63.10	97.05	44.99
Median	67.63	59.1	31.07	20.51	16.95	14.13	11.97	9.50	50.43	17.24
Quantiles:										
10 percent	20.17	21.92	19.55	10.76	8.03	5.62	4.15	2.76	36.62	8.86
25 percent	41.9	38.06	25.33	13.81	10.79	8.36	6.53	4.88	44.31	12.68
75 percent	83.71	73.44	43.61	29.89	25.88	22.57	20.79	16.05	59.50	20.99
90 percent	92.26	84.27	55.69	40.00	36.21	33.77	31.76	26.76	69.76	25.32

Note: The table shows descriptive statistics of the estimated shares of variances in yields, 10-year short-rate expectations, and 10-year term premia determined by the macroeconomic variables. The macroeconomic shares are identified by assuming that latent factors do not respond to contemporaneous macroeconomic shocks. The statistics are reported in percent. The table also shows the constant macroeconomic share of variances estimated using the Gaussian term structure model.

Figure B.5: 5-Year Moving-Average Macroeconomic Shares of Variances with Alternative Cholesky Ordering



Notes: The figure shows 5-year moving averages of the macroeconomic contribution to variances in the 3-month and 10-year yields, 10-year short-rate expectations, and the 10-year term premium. These are identified under the assumption that inflation responds to unemployment-gap shocks with a delay of one month. Shaded areas represent recessions, as defined by the National Bureau of Economic Research.

Chapter 3

A Joint Model for the Term Structures of Interest Rates and Realized Volatility

Anne Lundgaard Hansen
University of Copenhagen

I develop a term structure model that describes the yield curve and realized bond market volatility jointly. My model includes a measurement equation for the realized yield covariance matrix expressed by the sum of a conditional covariance matrix and a mean-zero error. The conditional covariance matrix is described by a multivariate GARCH-type model. My model admits closed-form solutions for no-arbitrage bond yields, term premia, conditional yield curve volatility, and multi-step ahead forecasts of both yields and realized volatility. I derive an exact algorithm for filtering latent state variables of my model from the basic principles of the Kalman filter. An empirical application to U.S. Treasury data shows that my model matches the realized yield curve covariance matrix both in- and out-of-sample. Finally, I show that the risk-neutral dynamics extracted from first and second moments of the yield curve do not describe the pricing of interest-rate derivatives.

Keywords: Discrete-time term structure model, realized yield curve covariance matrix, multivariate GARCH, non-linear Kalman filter.

JEL classification: C13, C32, G12.

3.1 Introduction

This paper introduces a novel method for modeling the term structure of volatility in no-arbitrage bond yields. Bond market volatility is an important input for risk management and the pricing of fixed-income derivatives. In addition, Cieslak and Povala (2016) argue that the decomposition of bond yield volatility into volatilities of short-rate expectations and term premia contains important economic information, e.g., on the uncertainty about the expected future path of monetary policy.

Modeling bond market volatility has proven difficult in traditional continuous-time affine term structure models with stochastic volatility (Jacobs and Karoui, 2009). These models imply bond yields that are linear functions of the quadratic variation in yields, i.e., that bond yields span bond market volatility (Dai and Singleton, 2000, Duffie and Kan, 1996). This prediction has, however, been rejected by an extensive body of literature (Andersen and Benzoni, 2010, Collin-Dufresne, Goldstein, and Jones, 2009). Collin-Dufresne and Goldstein (2002) propose models with so-called "unspanned stochastic volatility" (USV) restrictions to relieve the tension between bond yields and volatility. But, Joslin (2017) shows that these restrictions are rejected by the data.

In this paper, I propose a discrete-time model that describes the term structures of interest rates and their quadratic variation jointly. Specifically, I use realized volatility to approximate quadratic variation (Andersen and Bollerslev, 1998, Andersen, Bollerslev, Diebold, and Labys, 2001). My model explores the result that a realized covariance matrix can be expressed by the sum of a conditional covariance matrix and an error with mean zero and variance given by multivariate realized quarticity (Barndorff-Nielsen and Shephard, 2002, 2004). I use a multivariate GARCH-type model to describe the conditional covariance matrix in a parsimonious way. These ideas are integrated into a no-arbitrage term structure model with a low-dimensional latent state vector.

My model is tractable: I derive closed-form solutions for bond yield levels, term premia, conditional yield variances and covariances, and their multi-step ahead forecasts. In addition, I set up my model as a state space model with measurement equations for both yield levels and realized variances and covariances. Then, I show that this non-linear state space model can be estimated by a filtering algorithm derived from the basic principles of the standard, linear Kalman (1960) filter. The filter is exact, unlike existing algorithms for non-linear state space models such as the extended (see Jazwinski, 1970) and unscented Kalman filters (Julier, Uhlmann, and Durrant-Whyte, 1995).

This work is most closely related to Cieslak and Povala (2016), who propose a continuous-time term structure model with both yield and volatility factors. My approach offers three advantages over their important contribution. First, my model is complete in the sense that there is a well-founded measurement equation for the realized covariance matrix. Therefore, my model admits computation of multi-step ahead forecasts of the realized yield curve covariance matrix. Second, I avoid introducing additional factors in my model by using a GARCH-type framework to model the conditional covariance matrix. Thus, my model is consistent with the stylized fact that bond markets can be characterized by a few, e.g., three, factors Litterman and Scheinkman (1991). And finally, my model can be estimated by an exact filtering algorithm, whereas Cieslak and Povala (2016) rely on an approximation.

My approach is also related to the Realized GARCH framework proposed by Hansen, Huang, and Shek (2011), the MEM model by Engle and Gallo (2006), and the HEAVY model by Shephard and Sheppard (2010). Similarly to my approach, these models describe returns and realized measures of volatilities jointly. I show how these ideas can be embedded into term structure modeling with no-arbitrage restrictions and multivariate dynamics. Finally, the proposed term structure model is a generalization of the models in Heston and Nandi (2003) and Koeda and Kato (2015). Specifically, my generalizations allow for (i) time-varying conditional correlations in the state vector, hence genuinely multivariate dynamics, and (ii) the inclusion of realized variances and covariances in the measurement equation.

The paper proceeds as follows: I present my model in Section 3.2. Section 3.3 summarizes the model as a non-linear state space model and derives a Kalman-type filtering algorithm. In Section 3.4, I present an empirical analysis on U.S. Treasury data. I show that my model captures yield curve levels and the associated realized covariance matrix with a high precision. I also present encouraging out-of-sample results for forecasting the realized covariance matrix with multi-step ahead horizons. These promising results are obtained despite evidence that the linear relationship between bond yields and volatility is weak. My model therefore contrasts with the traditional continuous-time affine term structure models with stochastic volatility for which USV restrictions are needed to capture bond market volatility. Finally, I compare model-implied risk-neutral volatility to implied volatility from options on Treasury bonds. I demonstrate that there are differences between these series, which can be interpreted as an indication that the risk-neutral dynamics extracted from first and second moments of bond yields alone do not generalize to those determining prices of interest-rate derivatives. Section 3.5 concludes and discusses promising paths for future research.

Notation: I introduce the following notation for the manipulation of matrices. For a symmetric matrix A, let vec (A) be a function that performs a vectorization by stacking the columns of A to a vector. Likewise, vech (A) performs a half-vectorization by stacking the columns of the lower triangular part of A. Define the elimination matrix, $\mathcal{E}$, such that $\mathcal{E}\text{vec}(A) = \text{vech}\,(A)$ and the duplication matrix, $\mathcal{D}$, by $\mathcal{D}\text{vech}\,(A) = \text{vec}(A)$. Let diag$(A)$ denote the vector with elements given by the diagonal of the square, but not necessarily symmetric, matrix A. Finally, I define a matrix square root $A^{1/2}$ such that $A^{1/2}(A^{1/2})' = A$.

3.2 Model

Consider N zero-coupon bonds with maturities $n_1, n_2, \ldots, n_N$ and denote a yield curve by $Y_t = (Y_{t,n_1}, Y_{t,n_2}, \ldots, Y_{t,n_N})'$. The yield curve is modeled at the monthly frequency Δt. Yields are, however, also observed at a higher frequency between each time interval $[t, t+1[$, e.g., daily. Denote these intra-period yields by $y_{t,s} = (y_{t,s,n_1}, y_{t,s,n_2}, \ldots, y_{t,s,n_N})'$ for $s = 0, 1, \ldots, S$ where $y_{t,0} = Y_t$. The realized covariance matrix is constructed by cumulating the outer products of $\Delta y_{t,s} = y_{t,s} - y_{t,s-1}$ over $s = 1, 2 \ldots, S$:

$$\text{RV}_t = \sum_{s=1}^{S} \Delta y_{t,s} \Delta y_{t,s}'.$$

In this section, I build a model for the yield curve, Y_t, and the realized yield covariance matrix, RV$_t$. I start by summarizing results from the literature that relate realized variances and covariances to conditional variances and covariances. Then, I set up a term structure model for the yield curve and its conditional covariance matrix.

3.2.1 Relating the Realized and Conditional Covariance Matrices

As an outset for my model, consider the following well-established results. Andersen, Bollerslev, Diebold, and Labys (2001, 2003) argue that for semi-martingale processes:

(i) Quadratic variation is related to conditional variance up to a mean-zero error.

(ii) Quadratic variation can be approximated by realized variance.

Furthermore, Barndorff-Nielsen and Shephard (2004) show that:

(iii) For a general class of continuous stochastic volatility semi-martingales with spot covariance matrix Ψ_t, as $S \to \infty$,

$$\sqrt{S}\left\{\text{vech}\left(\text{RV}_t\right) - \text{vech}\left(\int_t^{t+1} \Psi(u)du\right)\right\} \xrightarrow{L} \mathcal{N}(0, \text{IQ}_t),$$

where IQ_t denotes integrated quarticity as defined in Barndorff-Nielsen and Shephard (2004).

(iv) Integrated quarticity can be consistently estimated by realized quarticity. In particular, define $\xi_{t,s} = \text{vech}\left(\Delta y_{t,s}\Delta y_{t,s}'\right)$ and

$$\text{RQ}_t = \sum_{s=1}^{S} \xi_{t,s}\xi_{t,s}' - \frac{1}{2}\sum_{s=1}^{S-1}\left(\xi_{t,s+1}\xi_{t,s} + \xi_{t,s+1}\xi_{t,s}'\right). \tag{3.1}$$

Then $S\,\text{RQ}_t$ converges in probability to IQ_t.

The observations in (i)-(iv) imply that when intra-period sampling is frequent, i.e., for S large, the realized covariance matrix can be written by the sum of a conditional covariance matrix and a mean-zero error with variance given by realized quarticity.

3.2.2 Term Structure Model

Define a filtration $\mathcal{F}_t = \{Y_i, \text{RV}_i, \text{RQ}_i\}_{i=1}^t$.[36] My model describes yields, Y_t, and the realized covariance matrix, RV_t, but treats realized quarticity, RQ_t as exogenous. I build a no-arbitrage term structure model based on the notion that the yield curve can be described by a low-dimensional factor structure (Duffie and Kan, 1996, Duffie, Pan, and Singleton, 2000, Litterman and Scheinkman, 1991). Therefore, let X_t denote a latent p-dimensional state vector. No-arbitrage bond yields and their covariance matrix are derived given $\{X_t, X_{t-1}, \ldots\}$. Section 3.3.3 shows how the latent state vector X_t can be filtered from $\mathcal{F}_t$.

Measurement Equations

Observed bond yields are given by the sum of the no-arbitrage equilibrium yield, $\tilde{Y}_t$, given by

$$\tilde{Y}_{t,n} = -\frac{1}{n}\log \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{i=0}^{n-1}\tilde{Y}_{t+i,1}\right)\,\middle|\, X_t, X_{t-1}, \ldots\right], \tag{3.2}$$

[36] Since RV_t and RQ_t are constructed using intra-period data between t and $t+1$ (not included), it is technically more correct to denote the filtration by $\mathcal{F}_{t+\frac{s-1}{S}}$. However, since time is discrete with intervals Δt in the considered model, I define $\mathcal{F}_t = \{Y_t, \text{RV}_t, \text{RQ}_t, Y_{t-1}, \text{RV}_{t-1}, \text{RQ}_{t-1}, \ldots, Y_1, \text{RV}_1, \text{RQ}_1\}$.

and a measurement error, $\varepsilon_t^Y \sim$ i.i.d. $(0, R^Y)$, where R^Y is a matrix of free parameters:

$$Y_t = \tilde{Y}_t + \varepsilon_t^Y, \tag{3.3}$$

Based on results from Section 3.2.1, I also write the realized covariance matrix as a sum of the conditional covariance matrix of $\tilde{Y}_{t+1}$ given $\{X_t, X_{t-1}, \ldots\}$, Var$(\tilde{Y}_{t+1}|X_t, X_{t-1}, \ldots)$, and a mean-zero error. This error reflects both a Gaussian noise with covariance matrix estimated by RQ$_t$, an error introduced by approximating quadratic variation by realized volatility, and potential measurement errors in the high-frequency yield observations $y_{t,s}$. I accommodate these errors by introducing $\varepsilon_t^{\text{RV}} \sim \left(0, R_t^{\text{RV}}\right)$ with $R_t^{\text{RV}} = \text{RQ}_t + R^{\text{RV}}$, where R^{RV} is a matrix of free parameters.[37] Furthermore, let $(R_t^{\text{RV}})^{-1/2}\varepsilon_t^{\text{RV}}$ be i.i.d. Thus,

$$\text{vech}\left(\text{RV}_t\right) = \text{vech}\left(\text{Var}(\tilde{Y}_{t+1}|X_t, X_{t-1}, \ldots)\right) + \varepsilon_t^{\text{RV}}. \tag{3.4}$$

This equation is a multivariate extension of the measurement equation for realized volatility in the realized GARCH model in Hansen, Huang, and Shek (2011).

Modeling No-Arbitrage Equilibrium Bond Yields, $\tilde{Y}_t$

Let the one-period spot interest rate $\tilde{Y}_{t,1}$ be given by an affine function of the state vector, X_t:

$$\tilde{Y}_{t,1} = \delta_0 + \delta_1' X_t. \tag{3.5}$$

The state vector follows a vector autoregressive process with a conditional covariance matrix $\Omega_t = \text{Var}(X_{t+1}|X_t, X_{t-1}, \ldots)$:[38]

$$X_{t+1} = \mu + \Phi X_t + \Omega_t^{1/2} Z_{t+1} \tag{3.6}$$

where $Z_t \sim$ i.i.d. $\mathcal{N}(0, \text{I}_p)$ and $\Omega_t^{1/2}$ is a $p \times p$ matrix. The conditional covariance matrix is modeled by

$$\Omega_t = CC' + AZ_tZ_t'A' + B\Omega_{t-1}B' \tag{3.7}$$

with A and B diagonal matrices. This specification adopts the structure of the BEKK model from Engle and Kroner (1995), but is distinct from a standard multivariate GARCH framewrok by letting Z_t rather than $\Omega_{t-1}^{1/2} Z_t$ determine Ω_t. The volatility model can be viewed as a multivariate generalization of that in Heston and Nandi (2003).

[37] The measurement error covariance matrix R_t^{RV} can also be defined by an appropriate multiplication of RQ$_t$ and R^{RV}. The empirical results in Section 3.4 are robust to this choice of specification.

[38] Note that this notation follows the stochastic volatility literature but contrasts with the GARCH literature.

Including standardized errors Z_t in the conditional covariance matrix can be intuitively appealing. In the standard GARCH framework, a large shock results in an increased variance in the following period. In the model in (3.7), however, the degree to which a shock generates more variance in the following period depends inversely on the current level of conditional volatility. That means that if the economy is in a state of low uncertainty when a large shock arrives, then there will be a larger impact on future volatility than there would have been if the shock had arrived in an already highly uncertain period.

Finally, under a standard exponential-linear pricing kernel, see, e.g., Ang and Piazzesi (2003), the risk-neutral dynamics of X_t is given by:

$$X_{t+1} = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t + \Omega_t^{1/2} Z_{t+1}^{\mathbb{Q}} \tag{3.8}$$

with $Z_t^{\mathbb{Q}} \sim$ i.i.d. $\mathcal{N}(0, \mathrm{I}_p)$ and

$$\Omega_t = CC' + AZ_t^{\mathbb{Q}} Z_t^{\mathbb{Q}'} A' + B\Omega_{t-1}B'. \tag{3.9}$$

Model Solution

The model implies a closed-form solution for the risk-neutral expectation in (3.2), hence for no-arbitrage bond yields. This result is given in the next theorem.

Theorem 3.1 *Define $\check{c}_n$ and c_n for $n = 0, 1, \ldots$ such that vec $(\check{c}_n) = (A \otimes A)'\mathcal{E}'c_n$ and assume that $(I_p - 2\check{c}_n)$ is positive definite. Then, the no-arbitrage yield of a n-period zero-coupon bond under the model in* (3.5)-(3.9) *is given by*

$$\tilde{Y}_{t,n} = -\frac{1}{n}\left\{a_n + b_n' X_t + c_n' \textit{vech}\left(\Omega_t\right)\right\}, \tag{3.10}$$

where the loadings are given recursively by

$$\begin{aligned}
a_{n+1} &= -\delta_0 + a_n + b_n'\mu^{\mathbb{Q}} + c_n' \textit{vech}\left(CC'\right) - \frac{1}{2}\ln\left[\det(I_p - 2\check{c}_n)\right], \\
b_{n+1}' &= -\delta_1' + b_n'\Phi^{\mathbb{Q}}, \\
c_{n+1}' &= c_n'\mathcal{E}(B \otimes B)\mathcal{D} + \frac{1}{2}(b_n \otimes b_n)'\mathcal{D}.
\end{aligned}$$

The recursions are initiated at $n = 0$ with $a_0 = 0$, $b_0 = \mathbf{0}_{p\times 1}$, and $c_0 = \mathbf{0}_{p(p+1)/2\times 1}$.

Proof. See Appendix C.1. □

Corollary 3.1 *It follows that the conditional covariance between the n- and m-period bond yields, $Y_{t+1,n}$ and $Y_{t+1,m}$, given $\{X_t, X_{t-1}, \ldots\}$ is given by*

$$\textit{Cov}(\tilde{Y}_{t+1,n}, \tilde{Y}_{t+1,m} | X_t, X_{t-1}, \ldots) = \frac{1}{nm}\left\{\alpha_{n,m} + \beta_{n,m}' \textit{vech}\left(\Omega_t\right)\right\} \tag{3.11}$$

with

$$\alpha_{n,m} = c_n' \mathcal{E}(A \otimes A)\left[I_{p^2} + \textit{diag}(\textit{vec}\,(I_p))\right](A \otimes A)'\mathcal{E}' c_m$$
$$\beta_{n,m}' = (b_m \otimes b_n)'\mathcal{D}.$$

These conditional covariances define the elements in the conditional covariance matrix $\text{Var}(\tilde{Y}_{t+1}|X_t, X_{t-1}, \ldots)$ in (3.4).

Spanning of Conditional Volatility

According to (3.11), conditional variances and covariances of yields span Ω_t and, furthermore, bond yields span Ω_t through (3.10). Therefore, my model implies that bond yields span conditional yield variances and covariances. In the terminology of Engle, Lilien, and Robins (1987), the model-implied bond yields exhibit ARCH-in-mean effects.

The spanning may be weak if the loading on vech (Ω_t), c_n, is estimated to be small. Weak spanning will, however, not deteriorate the ability of the model to estimate conditional variances and covariances because (i) Ω_t is determined deterministically from Z_t and Ω_{t-1}, and (ii) the observed realized covariance matrix helps recovering both Ω_t and the loadings $\alpha_{n,m}$ and $\beta_{n,m}$. This contrasts with the traditional continuous-time affine term structure model with stochastic volatility when relying on yield curve level data to identify yield volatilities (Collin-Dufresne, Goldstein, and Jones, 2009, Jacobs and Karoui, 2009).

Statistical Properties

From the recursive structure of both X_t and Ω_t, it is straightforward to show that the h-step conditional first moments of the state vector and the associated conditional covariance matrix are given for $h > 0$ by

$$\mathbb{E}\left(X_{t+h}|\mathcal{F}_t\right) = \sum_{i=0}^{h-1} \Phi^i \mu + \Phi^h \mathbb{E}(X_t|\mathcal{F}_t),$$
$$\text{vech}\left(\mathbb{E}\left(\Omega_{t+h}|\mathcal{F}_t\right)\right) = \sum_{i=0}^{h-1} [\mathcal{E}(B \otimes B)\mathcal{D}]^i \text{vech}\left(CC' + AA'\right) + [\mathcal{E}(B \otimes B)\mathcal{D}]^h \text{vech}\left(\mathbb{E}(\Omega_t|\mathcal{F}_t)\right)$$

assuming Z_{t+h} independent from $\mathcal{F}_t$ for $h > 0$. Section 3.3.3 presents an algorithm for filtering the model, hence obtaining $\mathbb{E}(X_t|\mathcal{F}_t)$ and $\mathbb{E}(\Omega_t|\mathcal{F}_t)$. It follows that if the eigenvalues of Φ and $(B \otimes B)$ are inside the unit circle, then unconditional first moments are

$$\mathbb{E}\left(X_t\right) = \left(\mathrm{I}_p - \Phi\right)^{-1}\mu, \tag{3.12}$$
$$\text{vech}\left(\mathbb{E}\left(\Omega_t\right)\right) = \mathcal{E}\left(\mathrm{I}_{p^2} - B \otimes B\right)^{-1} \text{vec}\left(CC' + AA'\right). \tag{3.13}$$

The expression for $\mathbb{E}(\Omega_t)$ is an input in the Kalman-type algorithm derived in Section 3.3.3. Given the moments of the state vector process, the h-step conditional first moments of yields and the realized covariance matrix are given by

$$\begin{aligned}\mathbb{E}(Y_{t+h,n}|\mathcal{F}_t) &= -\frac{1}{n}\left\{a_n + b_n'\mathbb{E}(X_{t+h}|\mathcal{F}_t) + c_n'\text{vech}\left(\mathbb{E}(\Omega_{t+h}|\mathcal{F}_t)\right)\right\},\\ \mathbb{E}(\text{RV}_{t+h,n,m}|\mathcal{F}_t) &= \frac{1}{nm}\left\{\alpha_{n,m} + \beta_{n,m}'\text{vech}\left(\mathbb{E}(\Omega_{t+h}|\mathcal{F}_t)\right)\right\},\end{aligned}$$

where $\text{RV}_{t,n,m}$ is the realized covariance between $Y_{t,n}$ and $Y_{t,m}$. Therefore, the model facilitates closed-form expressions for multi-step ahead forecasting of both yields and their realized variances and covariances. The unconditional moments of Y_t and RV_t are given analogously.

Term Premia

Term structure models can be used to study term premia, i.e., risk premia in bond markets. There are multiple definitions in the literature. I work with the following:

$$\text{TP}_{t,n} = \tilde{Y}_{t,n} - \tilde{Y}_{t,n}^{\mathbb{Q}=\mathbb{P}}, \tag{3.14}$$

where $\tilde{Y}_{t,n}^{\mathbb{Q}=\mathbb{P}}$ is computed by (3.2) under the assumption that investors are risk-neutral, i.e.,

$$\tilde{Y}_{t,n}^{\mathbb{Q}=\mathbb{P}} = -\frac{1}{n}\log\mathbb{E}\left[\exp\left(-\sum_{i=0}^{n-1}\tilde{Y}_{t+i,1}\right)\,\middle|\, X_t, X_{t-1},\ldots\right]. \tag{3.15}$$

It follows that if investors actually *are* risk-neutral, then $\tilde{Y}_{t,n} = \tilde{Y}_{t,n}^{\mathbb{Q}=\mathbb{P}}$ and $\text{TP}_{t,n} = 0$. The following theorem gives a closed-form solution for the risk-neutral yield and hence for the term premium.

Theorem 3.2 *Define $\check{c}_n^{\mathbb{Q}=\mathbb{P}}$ and $c_n^{\mathbb{Q}=\mathbb{P}}$ for $n = 0, 1, \ldots$ such that vec$\left(\check{c}_n^{\mathbb{Q}=\mathbb{P}}\right) = (A\otimes A)'\mathcal{E}'c_n^{\mathbb{Q}=\mathbb{P}}$ and assume that $(I_p - 2\check{c}_n^{\mathbb{Q}=\mathbb{P}})$ is positive definite. Then, the risk-neutral yield of a n-period bond under the model in* (3.5)-(3.9) *is given by*

$$\tilde{Y}_{t,n}^{\mathbb{Q}=\mathbb{P}} = -\frac{1}{n}\left\{a_n^{\mathbb{Q}=\mathbb{P}} + b_n^{\mathbb{Q}=\mathbb{P}'}X_t + c_n^{\mathbb{Q}=\mathbb{P}'}\textit{vech}\left(\Omega_t\right)\right\}, \tag{3.16}$$

where

$$\begin{aligned}a_{n+1}^{\mathbb{Q}=\mathbb{P}} &= -\delta_0 + a_n^{\mathbb{Q}=\mathbb{P}} + b_n^{\mathbb{Q}=\mathbb{P}'}\mu + c_n^{\mathbb{Q}=\mathbb{P}'}\textit{vech}\left(CC'\right) - \frac{1}{2}\ln\left[\det(I_p - 2\check{c}_n^{\mathbb{Q}=\mathbb{P}})\right],\\ b_{n+1}^{\mathbb{Q}=\mathbb{P}'} &= -\delta_1' + b_n^{\mathbb{Q}=\mathbb{P}'}\Phi,\\ c_{n+1}^{\mathbb{Q}=\mathbb{P}'} &= c_n^{\mathbb{Q}=\mathbb{P}'}\mathcal{E}(B\otimes B)\mathcal{D} + \frac{1}{2}(b_n^{\mathbb{Q}=\mathbb{P}}\otimes b_n^{\mathbb{Q}=\mathbb{P}})'\mathcal{D}.\end{aligned}$$

The recursions are initiated at $n = 0$ with $a_0^{\mathbb{Q}=\mathbb{P}} = 0$, $b_0^{\mathbb{Q}=\mathbb{P}} = \mathbf{0}_{p\times 1}$, and $c_0^{\mathbb{Q}=\mathbb{P}} = \mathbf{0}_{p(p+1)/2\times 1}$.

Proof. Follows straightforwardly by an application of Theorem 3.1. □

Corollary 3.2 *It follows that under the risk-neutral measure, the conditional covariance matrix between the n- and m-period bond yields, $Y_{t+1,n}$ and $Y_{t+1,m}$, given $\{X_t, X_{t-1}, \ldots\}$ is given by*

$$Cov^{\mathbb{Q}=\mathbb{P}}(\tilde{Y}_{t+1,n}, \tilde{Y}_{t+1,m}|X_t, X_{t-1}, \ldots) = \frac{1}{nm}\left\{\alpha_{n,m}^{\mathbb{Q}=\mathbb{P}} + \beta_{n,m}^{\mathbb{Q}=\mathbb{P}'} vech\left(\Omega_t\right)\right\} \tag{3.17}$$

where

$$\begin{aligned}\alpha_{n,m}^{\mathbb{Q}=\mathbb{P}} &= c_n^{\mathbb{Q}=\mathbb{P}'}\mathcal{E}(A\otimes A)\left[I_{p^2} + diag(vec\left(I_p\right))\right](A\otimes A)'\mathcal{E}'c_m \\ \beta_{n,m}^{\mathbb{Q}=\mathbb{P}'} &= (b_m^{\mathbb{Q}=\mathbb{P}}\otimes b_n^{\mathbb{Q}=\mathbb{P}})'\mathcal{D}.\end{aligned}$$

Corollary 3.3 *The term premium of a n-period bond is given by*

$$TP_{t,n} = -\frac{1}{n}\left\{(a_n - a_n^{\mathbb{Q}=\mathbb{P}}) + (b_n - b_n^{\mathbb{Q}=\mathbb{P}})'X_t + (c_n - c_n^{\mathbb{Q}=\mathbb{P}})'vech\left(\Omega_t\right)\right\} \tag{3.18}$$

with loadings given in Theorems 3.1 and 3.2.

3.3 Econometric Method

3.3.1 State Space Model

The model described in Section 3.2 can be written as a state space model. The measurement equations follow from (3.3) and (3.4) along with the model solutions in (3.10) and (3.11). Specifically, by stacking the loadings appropriately and scaling by periods to maturity, the measurement equations are given by

$$Y_t = a + bX_t + c\,\text{vech}\left(\Omega_t\right) + \varepsilon_t^Y, \tag{3.19}$$

$$\text{vech}\left(\text{RV}_t\right) = \alpha + \beta\,\text{vech}\left(\Omega_t\right) + \varepsilon_t^{\text{RV}}. \tag{3.20}$$

The loadings (a, b, c, α, and β) depend on the parameters of the model although this dependency has been suppressed to ease notation. The transition equations are given by (3.6)-(3.7). Define $\varepsilon_t = [\varepsilon_t^{Y'}, \varepsilon_t^{\text{RV}'}]'$ with

$$R_t = \begin{pmatrix} R^Y & 0 \\ 0 & R_t^{\text{RV}} \end{pmatrix}$$

such that $R_t^{-1/2}\varepsilon_t$ is i.i.d. I assume that the errors in the measurement and transition equations (ε_t and Z_t) are independent.

3.3.2 Parameters and Econometric Identification

The parameters of the model are $\Theta = \{\delta_0, \delta_1, \mu, \Phi, C, A, B, \mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}, R^Y, R^{\text{RV}}\}$. I introduce the following parameter restrictions to ensure econometric identification, i.e., to prevent affine transformations of X_t given by $\check{X}_t = v + VX_t$ along with $\check{\Omega}_t = V\Omega_t V'$. These restrictions are similar to the identification scheme in Cieslak and Povala (2016). First, setting $\mu = 0_{p\times 1}$ ensures identification of δ_0. Second, as noted in Dai and Singleton (2000), both Φ and Ω_t determine the interdependencies in X_{t+1}. I therefore restrict Φ to be diagonal. Third, restricting δ_1 to a vector of ones identifies the diagonal elements of A and B. Fourth, C lower triangular with positive elements on the diagonal ensures that Ω_t is positive definite. Finally, to ensure identification of the parameters in Ω_t, let the first entries of A and B be positive (Engle and Kroner, 1995).

3.3.3 Kalman-Type Filtering Algorithm

This section presents an algorithm for filtering the latent states in the non-linear state space model described in Section 3.3.1. The algorithm is derived from the basic principles of the standard, linear Kalman (1960) filter, hence I refer to the method as a Kalman-type filter. The Kalman-type filter is exact for the model considered here, unlike existing non-linear filters, namely the extended and unscented Kalman filters (see Jazwinski, 1970 and Julier, Uhlmann, and Durrant-Whyte, 1995). The following theorem gives the algorithm.

Theorem 3.3 *Consider the state space model defined by the transition equations in* (3.6)-(3.7) *and the measurement equations in* (3.19)-(3.20). *Let* $\mathcal{M}_t = [Y_t', \textit{vech}\,(RV_t)']'$ *and define the constants* $\Psi = \mathcal{E}(B \otimes B)\mathcal{D}$ *and*

$$\kappa = \mathbb{E}\left[\textit{vech}\left(A(Z_tZ_t' - I_p)A'\right)\textit{vech}\left(A(Z_tZ_t' - I_p)A'\right)'\right].$$

Assume that $E(\Omega_t) < \infty$. *Then, for well-chosen values of* $\hat{X}_{0|0}$, $\hat{\Omega}_{0|0}$, $V^X_{0|0}$, $V^{\Omega}_{0|0}$, *and* $V^{X,\Omega}_{0|0}$, *best linear mean-squared prediction can be obtained by the following recursion.*

State prediction

$$\hat{X}_{t|t-1} = \mathbb{E}\left[X_t|\mathcal{F}_{t-1}\right] = \mu + \Phi\hat{X}_{t-1|t-1},$$

$$\text{vech}\left(\hat{\Omega}_{t|t-1}\right) = \mathbb{E}\left[\text{vech}\left(\Omega_t\right)|\mathcal{F}_{t-1}\right] = \text{vech}\left(CC' + AA'\right) + \Psi\text{vech}\left(\hat{\Omega}_{t-1|t-1}\right),$$

$$V^X_{t|t-1} = \mathbb{E}\left[(X_t - \hat{X}_{t|t-1})(X_t - \hat{X}_{t|t-1})'\right] = \Phi V^X_{t-1|t-1}\Phi' + \mathbb{E}(\Omega_{t-1}),$$

$$V^{\Omega}_{t|t-1} = \mathbb{E}\left[\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t-1}\right)\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t-1}\right)'\right] = \Psi V^{\Omega}_{t-1|t-1}\Psi' + \kappa,$$

$$V_{t|t-1}^{X,\Omega} = \mathbb{E}\left[\text{vech}\left(X_t - \hat{X}_{t|t-1}\right)\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t-1}\right)'\right] = \Phi V_{t-1|t-1}^{X,\Omega}\Psi'.$$

Measurement prediction

$$\hat{Y}_{t|t-1} = \mathbb{E}\left[Y_t|\mathcal{F}_{t-1}\right] = a + b\hat{X}_{t|t-1} + c\,\text{vech}\left(\hat{\Omega}_{t|t-1}\right),$$

$$\text{vech}\left(\hat{\text{RV}}_{t|t-1}\right) = \mathbb{E}\left[\text{vech}\left(\text{RV}_t\right)|\mathcal{F}_{t-1}\right] = \alpha + \beta\text{vech}\left(\hat{\Omega}_{t|t-1}\right)$$

with mean-squared error $S_{t|t-1} = \begin{pmatrix} S_{t|t-1}^{Y} & S_{t|t-1}^{Y,\text{RV}} \\ S_{t|t-1}^{Y,\text{RV}'} & S_{t|t-1}^{\text{RV}} \end{pmatrix}$, where

$$S_{t|t-1}^{Y} = \mathbb{E}\left[(Y_t - \hat{Y}_{t|t-1})(Y_t - \hat{Y}_{t|t-1})'\right] = bV_{t|t-1}^{X}b' + cV_{t|t-1}^{\Omega}c' + bV_{t|t-1}^{X,\Omega}c' + cV_{t|t-1}^{X,\Omega'}b' + R^{Y},$$

$$S_{t|t-1}^{\text{RV}} = \mathbb{E}\left[\text{vech}\left(\text{RV}_t - \hat{\text{RV}}_{t|t-1}\right)\text{vech}\left(\text{RV}_t - \hat{\text{RV}}_{t|t-1}\right)'\right] = \beta V_{t|t-1}^{\Omega}\beta' + R_t^{\text{RV}},$$

$$S_{t|t-1}^{Y,\text{RV}} = \mathbb{E}\left[(Y_t - \hat{Y}_{t|t-1})\text{vech}\left(\text{RV}_t - \hat{\text{RV}}_{t|t-1}\right)'\right] = cV_{t|t-1}^{\Omega}\beta' + bV_{t|t-1}^{X,\Omega}\beta'.$$

Kalman gains

$$K_t^{X} = V_{t|t-1}^{X}b'S_{t|t-1}^{-1} + V_{t|t-1}^{X,\Omega}c'S_{t|t-1}^{-1},$$

$$K_t^{\Omega} = V_{t|t-1}^{X,\Omega'}b'S_{t|t-1}^{-1} + V_{t|t-1}^{\Omega}c'S_{t|t-1}^{-1}.$$

Filtering

$$\hat{X}_{t|t} = \mathbb{E}\left[X_t|\mathcal{F}_t\right] = \hat{X}_{t|t-1} + K_t^{X}\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right),$$

$$\text{vech}\left(\hat{\Omega}_{t|t}\right) = \mathbb{E}\left[\text{vech}\left(\Omega_t\right)|\mathcal{F}_t\right] = \text{vech}\left(\hat{\Omega}_{t|t-1}\right) + K_t^{\Omega}\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right),$$

$$V_{t|t}^{X} = \mathbb{E}\left[(X_t - \hat{X}_{t|t})(X_t - \hat{X}_{t|t})'\right] = (\text{I}_p - K_t^{X}b)V_{t|t-1}^{X} - K_t^{X}cV_{t|t-1}^{X,\Omega'},$$

$$V_{t|t}^{\Omega} = \mathbb{E}\left[\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t}\right)\right)\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t}\right)'\right] = (\text{I}_{p(p+1)/2} - K_t^{\Omega}c)V_{t|t-1}^{\Omega} - K_t^{\Omega}bV_{t|t-1}^{X,\Omega},$$

$$V_{t|t}^{X,\Omega} = \mathbb{E}\left[\left(X_t - \hat{X}_{t|t}\right)\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t}\right)'\right] = V_{t|t-1}^{X,\Omega}(\text{I}_{p(p+1)/2} - K_t^{\Omega}c)' - V_{t|t-1}^{X}b'K_t^{\Omega'}.$$

Proof. See Appendix C.1 □

The algorithm can be initialized at the unconditional first moments of X_t and Ω_t as given in (3.12) and (3.13). These priors are diffuse when the initial mean squared errors, $V_{0|0}^{X}$, $V_{0|0}^{\Omega}$, and $V_{0|0}^{X,\Omega}$, have high values. Given the prediction errors $e_t = [(Y_t - \hat{Y}_{t|t-1})', \text{vech}(\text{RV}_t - \hat{\text{RV}}_{t|t-1})']'$ and associated mean-squared error matrix $S_{t|t-1}$, my model can be estimated by quasi-maximum likelihood.

3.4 Empirical Analysis

3.4.1 Data

The empirical analysis focuses on end-of-month yields of U.S. Treasury bonds from January 1990 to August 2019. I construct monthly realized variances and covariances along with realized quarticity based on daily data. Zero-coupon bond yields for maturities between one and ten years are available from Gürkaynak, Sack, and Wright (2007) at the daily frequency. I use yields of bonds with maturities of 1, 2, 3, 5, 7, and 10 years and realized variances of 2-, 5-, and 10-year bond yields along with their covariances.[39]

Figure 3.1 shows the data for the 2- and 10-year maturities. The 2-year yield is impacted by jumps in the Federal Funds rate, whereas the 10-year yield exhibits a more persistent downward slope. Realized variances are time-varying with many short-lived bursts. Since the realized quarticities determine the measurement error variance of the measurement equation for the realized covariance matrix, these data indicate the extent that one can expect a conditional variance model to match the variance bursts. Note that the realized quarticities display large jumps during the dot-com bubble and the financial crisis, indicating that these periods are particularly characterized by ex-ante unpredictable events.

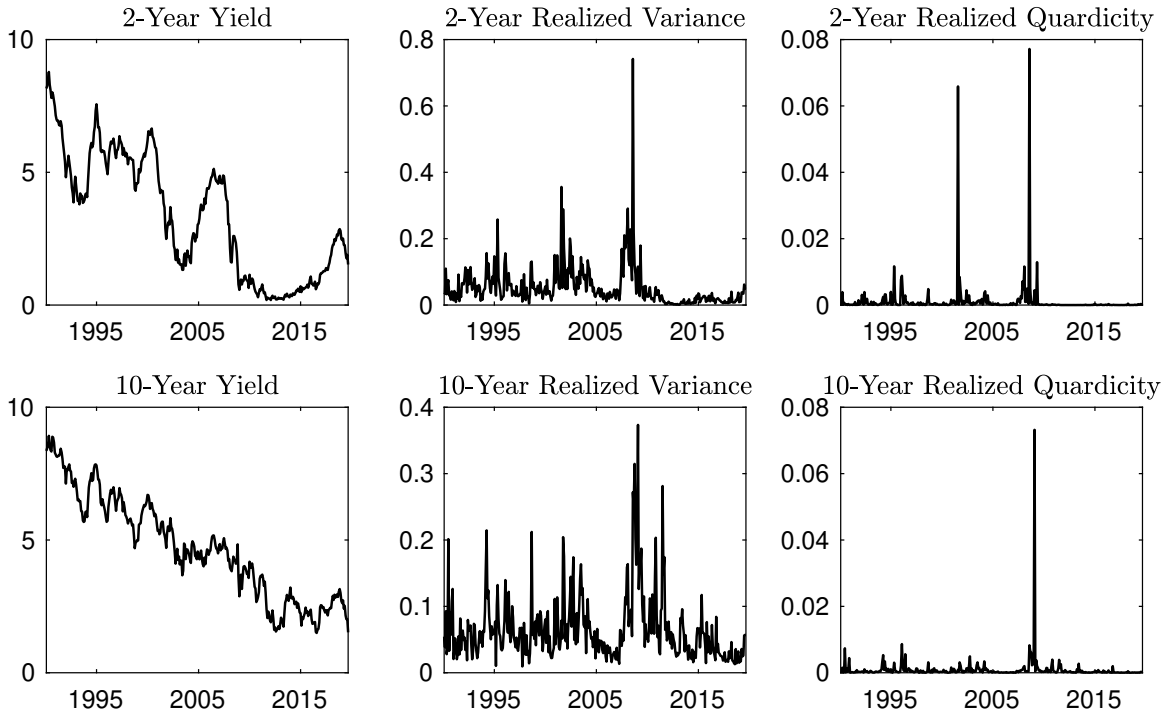
3.4.2 Estimation

Model Specification and Additional Parameter Restrictions

The yield curve can be summarized by a three-dimensional state vector, i.e., $p = 3$ (Litterman and Scheinkman, 1991). With a three-dimensional state vector, the model allows for six latent factors driving the conditional yield curve covariance matrix. However, I only include data on six variances and covariances given by the realized covariance matrix of the 2-, 5-, and 10-year bond yields. A principal component analysis shows that three factors explain 99.56 percent of the variation in this realized covariance matrix. I therefore restrict the model such that there are two heteroskedastic state variables and one state variable with constant variance. This is easily implemented by imposing one zero-row and -column in the matrices A and B in (3.7). Note that the element in δ_1 corresponding to the homoskedastic state variable must still be normalized, e.g., to one, to ensure identification of the constant variance.

[39] Cieslak and Povala (2016) argue that realized variances of maturities below two years are distorted by money market noise and institutional effects.

Figure 3.1: Data



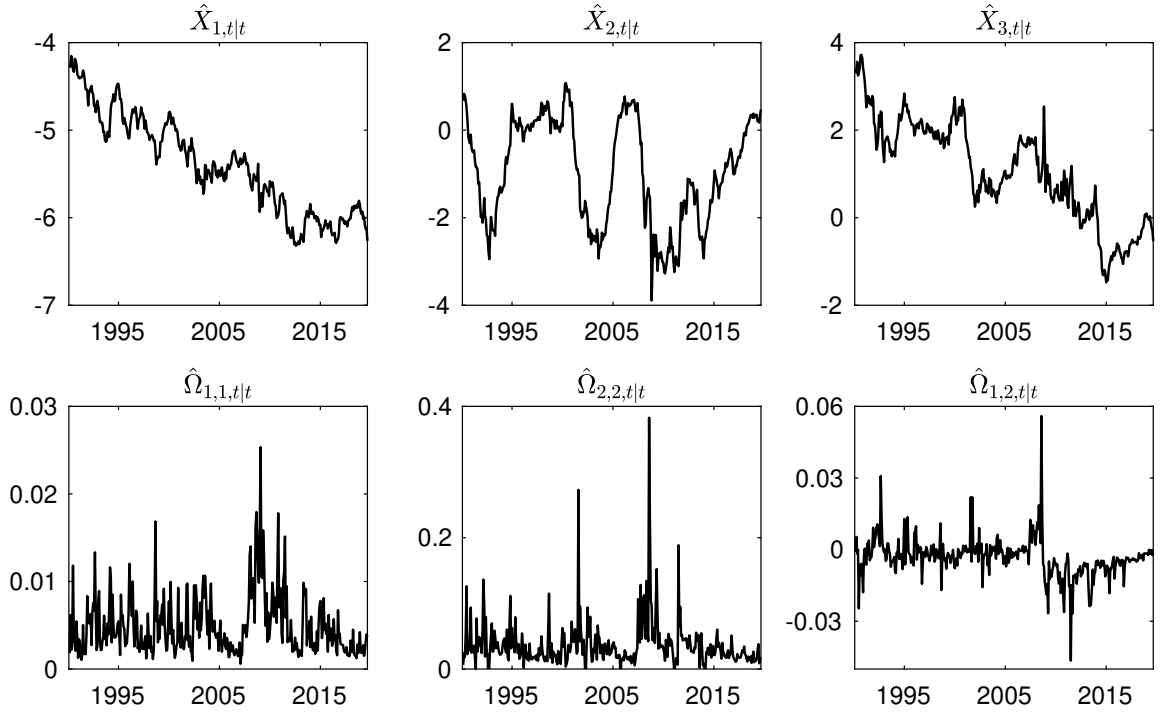
Notes: The figure shows monthly 2- and 10-year yields along with their realized variances and quarticities constructed from daily data. The data is sampled from January 1990 to August 2019.

I also restrict the measurement error variance parameters R^Y and R^{RV} to be diagonal matrices. The measurement errors of the realized covariance matrix are still allowed to be correlated through the realized quarticity, RQ_t.

Estimation Results

Parameter estimates are shown in Appendix C.2 and Figure 3.2 shows the filtered state variables and the time-varying part of the filtered conditional state covariance matrix, $\hat{\Omega}_{t|t}$. The third element in the state vector, $\hat{X}_{3,t|t}$, is homoskedastic such that the time-varying part of $\hat{\Omega}_{t|t}$ relates to the first two elements of $\hat{X}_{t|t}$ only. I observed from the figure that the conditional variances of $\hat{X}_{1,t|t}$ and $\hat{X}_{2,t|t}$ are clearly time-varying with particular sharp bursts during the 1998 Russian crisis, the early 2000's dot-com bubble, and the financial crisis. The filtered covariance, $\hat{\Omega}_{1,2,t|t}$, exhibits interesting variation around and after the financial crisis: The covariance increase during the spring 2008 as the Federal Reserve begins to lower the Federal Funds rate target and take action to prevent a housing bust. As the Lehman Brothers defaults in September 2008 and global markets begin to panic, the covariance falls sharply.

Figure 3.2: Filtered State Vector and Conditional Covariance Matrix

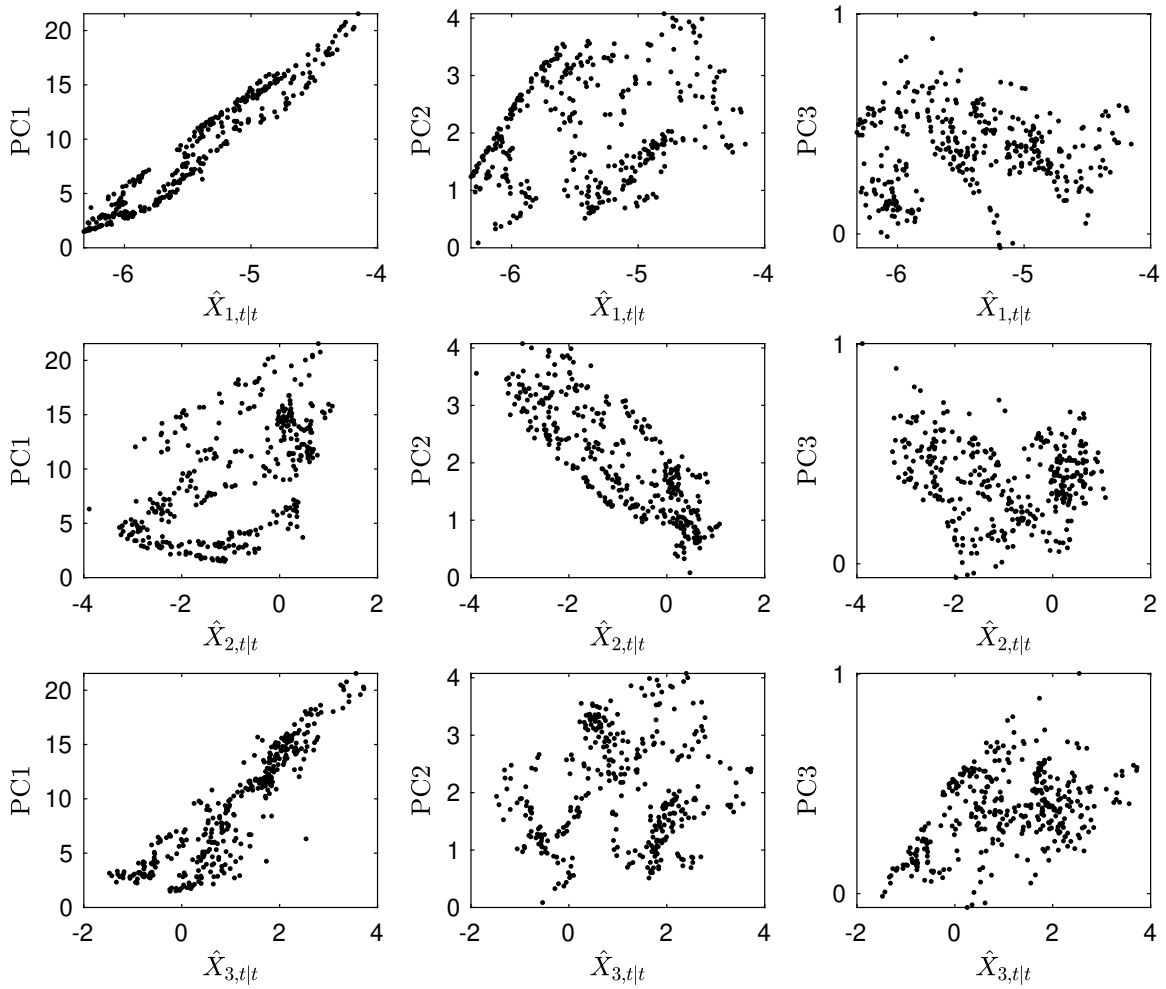


Notes: The figure shows the filtered state vector $\hat{X}_{t|t}$ and the time-varying variances and covariances in $\hat{\Omega}_{t|t}$.

The covariance falls abruptly again in June 2011 as markets react to the fear of a Greek default.

The filtered state vector is related to, but different from, the principal components of the yield curve. Specifically, Figure 3.3 shows that the first and third elements in $\hat{X}_{t|t}$ are related to the first principal component, often interpreted as the level of the yield curve. The second element in $\hat{X}_{t|t}$ is related to the second principal component, hence the yield curve slope. It is well-known that the first three principal components of the yield curve can approximate the latent factors filtered from affine term structure models with a constant covariance matrix. Joslin, Singleton, and Zhu (2011) even show that such models can be identified by rotating the latent state vector into principal components of the yield curve. It is therefore a novel result that time-varying conditional variances and covariances, when identified through realized measurements, change the interpretation of the yield curve factors. This result is also an indication that conditional variances and covariances are priced in the yield curve.

Figure 3.3: Filtered State Variables Versus Principal Components



Notes: The figures shows scatter plots of the filtered state vector $\hat{X}_{t|t}$ against the first three principal components of the yield data.

3.4.3 In-Sample Performance

Table 3.1 shows the in-sample fit in terms of yield levels and the realized covariance matrix. Starting with yield levels, the average root mean squared error is about four basis points per annum, which is highly precise. Considering individual maturities reveals that the state vector and its conditional covariance matrix are filtered such that some maturities (2, 5, and 7 years) are priced almost perfectly while the remaining maturities are priced with an error.

The realized covariance matrix is matched with a precision of 0.2 to 0.3 basis points per annum. For reference, Cieslak and Povala (2016) match realized variances and covariances with root mean squared errors of 0.4 to 0.9 basis points per

Table 3.1: In-Sample Performance

	Yield Levels						
Maturity	1-year	2-year	3-year	5-year	7-year	10-year	Average
Model	8.992	2.4×10^{-6}	1.426	9.8×10^{-6}	6.3×10^{-6}	3.825	4.032

	Realized Yield Variances			Realized Yield Covariances		
Maturity	2-year	5-year	10-year	(2,5)-year	(2,10)-year	(5,10)-year
Model	0.285	0.205	0.284	0.247	0.264	0.240
Naive	0.751	0.664	0.593	0.657	0.489	0.568

Notes: The upper panel shows root mean squared errors of model-implied against observed yield levels. The lower panel shows root mean squared errors of model-implied conditional variances and covariances against realized variances and covariances. These are compared to a naive model, where errors are constructed by the realized measures subtracted their sample mean. All numbers are reported in basis points per annum.

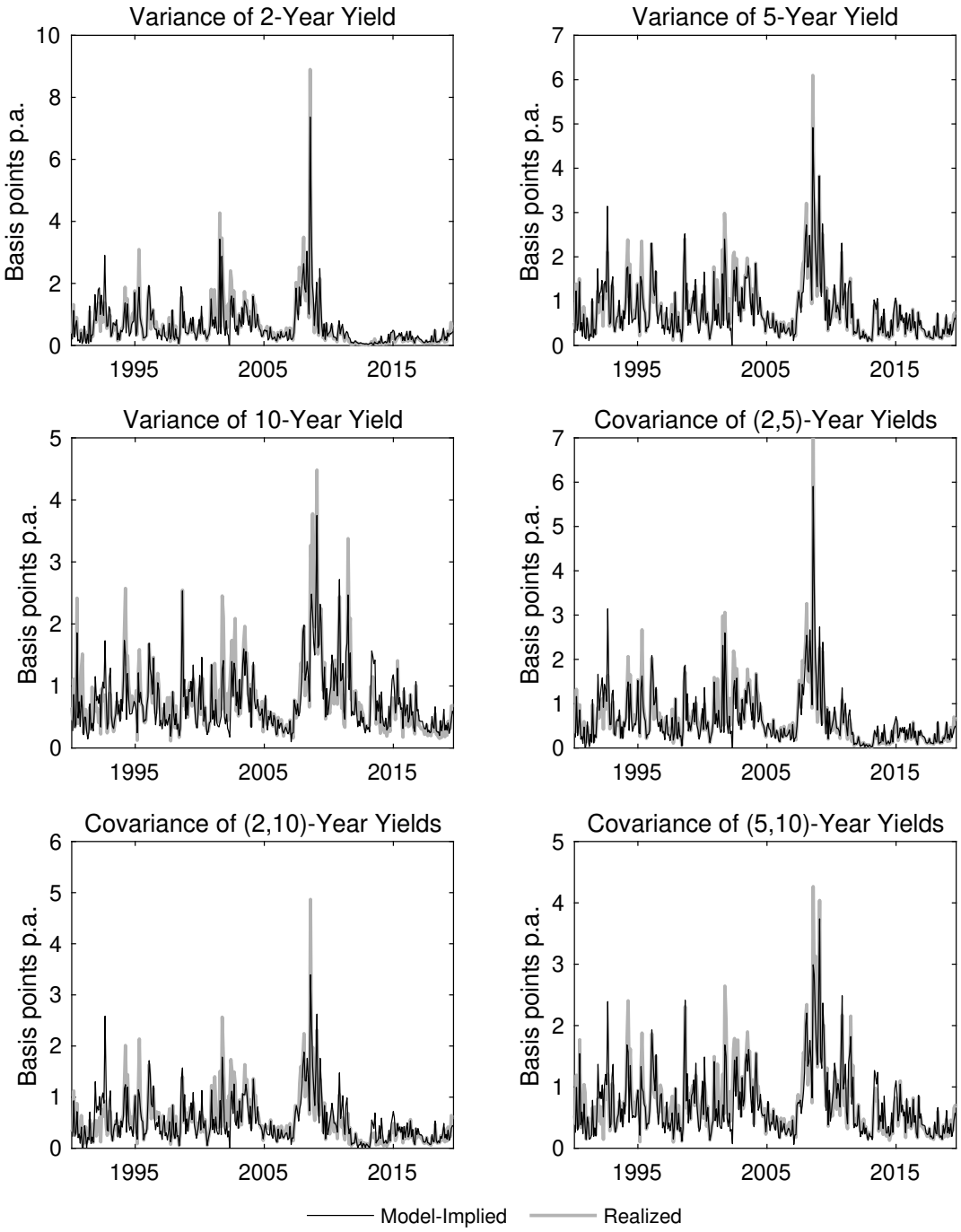
annum, although different samples and sampling frequencies impede a direct comparison. As another benchmark, I compare these measures with a naive model, where realized variances and covariances are given by their sample mean. I observe that the naive model results in root mean squared errors that are two to three times larger compared with those from my model.

Figure 3.4 shows the time series of model-implied conditional against realized variances and covariances. As expected from gauging the realized quarticity data in Figure 3.1, the pricing errors are particularly large during the dot-com bubble and the financial crisis. The model also does not predict the responses of realized variances to the Federal Reserve's mid-course correction in 1995-96. In conclusion, prediction errors are small on average and, for the second moments, higher in periods with high realized quarticity.

Cross-Validation

The estimation sample involves six maturities of yields, but only the realized covariance matrix of three of the yields. This choice allows me to cross-validate my model by assessing the fit to variances and covariances of the yields for which realized measures are not included in the estimation. Table 3.2 shows the resulting root mean squared errors and compares them to the naive model. I observe that my model outperforms the naive benchmark for all but (1,3)- and (1,7)-year covariances for which the models performance equally well.

Figure 3.4: Realized and Model-Implied Covariance Matrix



Notes: The figure shows model-implied conditional against realized variances and covariances.

Table 3.2: Cross Validation

	Realized Yield Variances			Realized Yield Covariances		
Maturity	1-year	3-year	7-year	(1,3)-year	(1,7)-year	(3,7)-year
Model	0.374	0.401	0.617	0.442	0.370	0.441
Naive	0.509	0.633	0.751	0.437	0.374	0.547

Notes: The table shows root mean squared errors in basis points per annum of model-implied conditional variances and covariances against realized variances and covariances for maturities that are not included in the estimation. These are compared to a naive model, where errors are constructed by the realized measures subtracted the sample mean.

Model-Implied Term Premium

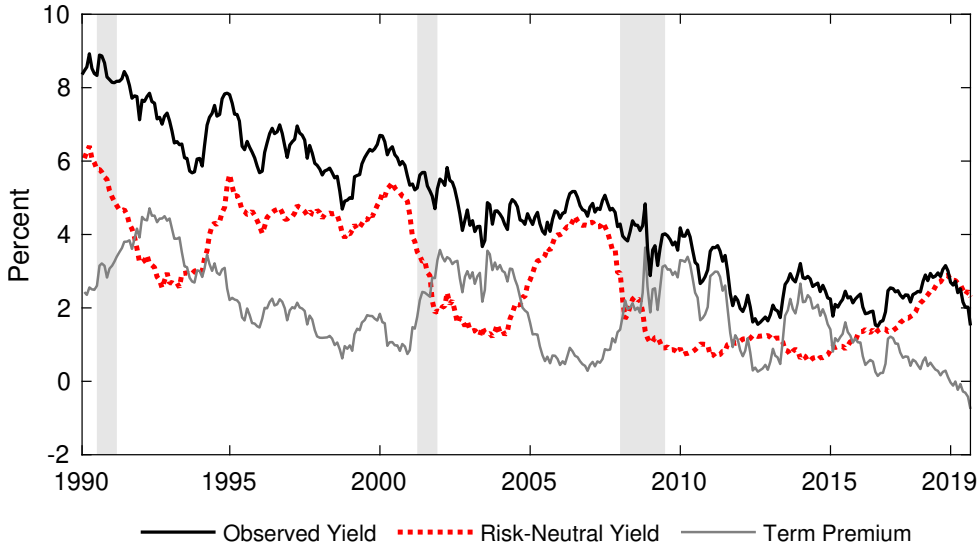
Figure 3.5 shows the model-implied decomposition of the 10-year yield into a risk-neutral yield, $\tilde{Y}_{t,n}^{\mathbb{Q}=\mathbb{P}}$, and a term premium. The term premium is counter-cyclical, which is consistent with existing literature (Adrian, Crump, and Moench, 2013, Kim and Orphanides, 2007, 2012). The term premium dynamics are, however, different from estimates obtained in affine term structure models with constant conditional variances and covariances. Specifically, I estimate term premia whose dynamics are largely unrelated to the downward trend in the 10-year yield, whereas affine models generally result in premia whose dynamics mimic those of the yields (Jardet, Monfort, and Pegoraro, 2013, Kim and Wright, 2005, Kozicki and Tinsley, 2001). This result is similar to Hansen (2019), who argues that the introduction of time-varying volatility can break the strong relationship between yields and term premia.

The Role of Conditional Volatility

My model allows the conditional covariance matrix Ω_t to directly impact model-implied bond yields, risk-neutral yields, and term premia through (3.10), (3.16), and (3.18). While it is well-established that the yield curve does not span stochastic volatility, it is largely unexplored whether the predictable component of volatility is spanned by the yield curve. To investigate the role of conditional volatility in the yield curve, consider the following model-implied bond yield, risk-neutral yield, and term premium under *unspanned conditional volatility*:

$$\begin{aligned}
\tilde{Y}_{t,n}^{UCV} &= -\frac{1}{n}\left\{a_n + b_n' X_t\right\}, \\
Y_{t,n}^{\mathbb{Q}=\mathbb{P},UCV} &= -\frac{1}{n}\left\{a_n^{\mathbb{Q}=\mathbb{P}} + b_n^{\mathbb{Q}=\mathbb{P}'} X_t\right\}, \\
\text{TP}_{t,n}^{UCV} &= -\frac{1}{n}\left\{\left(a_n - a_n^{\mathbb{Q}=\mathbb{P}}\right) + \left(b_n - b_n^{\mathbb{Q}=\mathbb{P}}\right)' X_t\right\}.
\end{aligned}$$

Figure 3.5: Decomposition of the 10-Year Yield



Notes: The figure shows how the model decomposes the 10-year yield into a risk-neutral yield and a term premium. The decomposition is computed from the solution for risk-neutral yields given in Theorem 3.2. Shaded areas show periods of recession as defined by the NBER.

Table 3.3: Root Mean Squared Errors Under Unspanned Conditional Volatility

	In-Sample Performance: Yield Levels						
Maturity	1-year	2-year	3-year	5-year	7-year	10-year	Average
Yield	0.043	0.129	0.239	0.464	0.643	0.854	0.489
Risk-Neutral Yield	0.036	0.076	0.112	0.170	0.214	0.262	0.165
Term Premium	0.012	0.062	0.147	0.339	0.501	0.698	0.383

Notes: The table shows root mean squared errors of $\tilde{Y}_{t,n}^{UCV}$ against $\tilde{Y}_{t,n}$ (yield), $Y_{t,n}^{\mathbb{Q}=\mathbb{P},UCV}$ against $Y_{t,n}^{\mathbb{Q}=\mathbb{P}}$ (risk-neutral yield), and $\text{TP}_{t,n}^{UCV}$ against $\text{TP}_{t,n}$ (term premium). The numbers are reported in basis points per annum.

Table 3.3 reports the root mean squared errors that incur under unspanned conditional volatility assuming that my model is correctly specified. These errors are small with an average of 0.5 basis points for yields and even smaller for risk-neutral yields and term premia. Thus, the linear impact of conditional volatility on the yield curve is small.

Table 3.4: Rolling Forecasting Performance

	Realized Yield Variances			Realized Yield Covariances		
Maturity	1-year	5-year	10-year	(1,5)-year	(1,10)-year	(5,10)-year
Horizons:						
1-month	0.658	0.731	0.775	0.729	0.859	0.785
6-month	0.666	0.794	0.889	0.752	0.891	0.878
12-month	0.677	0.881	0.925	0.786	0.933	0.976

Notes: The table shows the out-of-sample root mean squared errors of my model relative to the naive model, where errors are constructed by the realized measures subtracted their sample mean. Forecasts are constructed by re-estimating the model over a rolling window. The window length is 20 years resulting in the test sample from January 2010 to August 2019.

3.4.4 Forecasting the Realized Covariance Matrix

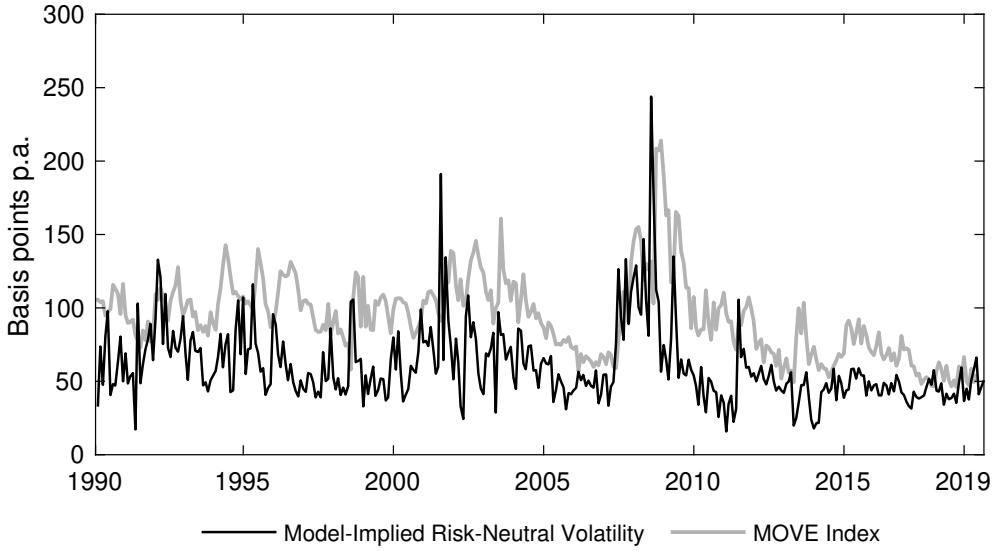
One advantage of a model for both levels and the realized covariance matrix is that it can generate multi-step ahead forecasts of the realized covariance matrix. I evaluate the ability of my model to produce reliable forecasts by a rolling forecasting exercise. Thus, I re-estimate my model over a rolling window with a length of 20 years, hence, the out-of-sample forecasting performance is tested on the sample from January 2010 to August 2019. I compare the performance of the model to the naive model that estimates realized variances and covariances by the sample average over the rolling sample. Root mean squared errors relative to those of the naive model are shown in Table 3.4 for forecasting horizons of 1, 6, and 12 months. Numbers lower than one indicate that my model forecasts are more precise than the rolling sample average. The table shows that my model uniformly outperforms the naive forecasts with root mean squared errors that are up to 35 percent lower.

3.4.5 Risk-Neutral versus Option-Implied Volatility

My model implies a risk-neutral covariance matrix given by (3.17). In this section, I compare this model-implication with implied volatilities from prices of options on Treasury bonds. This exercise can show whether the risk-neutral dynamics extracted from the Treasury bond market generalize to the broader fixed-income market.

Specifically, I compare risk-neutral volatilities (the square root of risk-neutral variances) to the Merrill Lynch option volatility estimate (MOVE) index, which is a weighted average of implied volatilities from one-month options on 2-year, 5-year, 10-year and 30-year Treasury bonds with weights 20, 20, 40, and 20 percent. For a

Figure 3.6: Model-Implied Risk Neutral Volatility and the MOVE Index



Notes: The figure shows a linear combination of model-implied risk-neutral volatilities of maturities 2, 5, and 10-year with weights 20, 20, and 40 percent along with the MOVE index.

fair comparison, I construct a corresponding weighted average of the model-implied risk-neutral volatilities of the 2, 5, and 10-year bonds with weights 20, 20 and 60 percent.

Figure 3.6 shows that there are large differences in the variation of the series. Specifically, model-implied risk-neutral volatilities are generally lower than the MOVE index. The average difference between the series is 33 basis points. These findings are consistent with a negative variance risk premium as documented in Treasury bond markets by Choi, Mueller, and Vedolin (2017) and Trolle and Schwartz (2009).

Thus, the risk-neutral dynamics extracted from the cross-section of yield curve levels and the realized yield covariance matrix do not describe the risk-neutral second moments implied by derivatives. It follows that the estimated risk-neutral dynamics may be improved by including implied volatilities from Treasury options in the measurement equation along with a variance risk premium. I leave this extension for further research.

3.5 Conclusion

I present a new framework that jointly describes the term structures of interest rates and realized volatility. My approach relies on well-established results showing that a realized covariance matrix can be expressed by the sum of a conditional covariance matrix and a mean-zero error. From this result, I set up a measurement equation that relates the conditional yield covariance matrix implied by a parsimonious model to an observed realized covariance matrix. I cast these ideas into a tractable term structure model with a conditional covariance matrix given by a GARCH-type model.

My approach offers numerous advantages over existing methods. First, unlike traditional continuous-time affine term structure models with stochastic volatility, a weak relationship between bond yield levels and volatility does not obstruct the ability of my model to match realized volatility. Second, my model can capture both yield levels, variances, and covariances while maintaining a low-dimensional factor structure. Third, the measurement equation for realized volatility enables computation of multi-step ahead volatility forecasts. Furthermore, my model facilitates computation of these forecasts in closed form. Finally, despite non-linearities in the state-space model, I show that my model can be estimated by an exact filtering algorithm derived from the basic principles of the standard, linear Kalman filter.

I evaluate the empirical performance of my model using U.S. Treasury bond yield data at the monthly frequency with realized volatility constructed from daily data. The data is considered from January 1990 to August 2019. Both in- and out-of-sample results are promising. However, I show that risk-neutral volatility dynamics may be improved by using volatility implied from prices of options on Treasury bonds perhaps along with a variance risk premium.

Another promising direction for future research relates to understanding the relationship between bond yields and bond market volatility. While I find evidence that realized volatility affects the filtered state vector, I show that the linear dependence between bond yields and bond market volatility is weak. Therefore, it is likely that volatility is priced in the yield curve in a non-linear way. An understanding of the specification and modeling of such non-linearity is warranted.

Acknowledgments

I thank Torben G. Andersen, Martin Møller Andreasen, Simon Hetland, Philipp Kless, Heino Bohn Nielsen, Viktor Todorov, Cynthia Wu, and participants at the Quanti-

tative Finance Seminar at Kellogg School of Management and the annual DAEINA meeting 2018.

Appendix

C.1 Proofs

C.1.1 Proof of Theorem 3.1

Lemma 3.1 *Consider a vector $Z_t \sim \mathcal{N}(0, I_p)$. Define a $p \times 1$ vector ξ and a $p \times p$ symmetric matrix Ξ. Then, given that $I_p - 2\Xi$ is positive definite,*

$$\mathbb{E}\left[\exp\left(\xi' Z_t + Z_t' \Xi Z_t\right)\right] = \det\left(I_p - 2\Xi\right)^{-1/2} \exp\left(\frac{1}{2}\xi'\xi\right).$$

Proof of Lemma 3.1.

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\xi' Z_t + Z_t' \Xi Z_t\right)\right] &= \int_{-\infty}^{\infty} \exp\left(\xi' z + z' \Xi z\right) (2\pi)^{-p/2} \exp\left(-\frac{1}{2} z' z\right) \mathrm{d}z \\
&= \int_{-\infty}^{\infty} \exp\left(\xi' z\right) (2\pi)^{-p/2} \exp\left(-\frac{1}{2} z' \left(\mathrm{I}_p - 2\Xi\right) z\right) \mathrm{d}z \\
&= \mathbb{E}\left[\exp\left(\xi' (\mathrm{I}_p - 2\Xi)^{1/2} \mathcal{Z}_t\right)\right] \det\left(\mathrm{I}_p - 2\Xi\right)^{-1/2}
\end{aligned}$$

with $\mathcal{Z}_t \sim \mathcal{N}(0, (\mathrm{I}_p - 2\Xi)^{-1})$. Finally,

$$\mathbb{E}\left[\exp\left(\xi' (\mathrm{I}_p - 2\Xi)^{1/2} \mathcal{Z}_t\right)\right] = \exp\left(\frac{1}{2} \xi' (\mathrm{I}_p - 2\Xi)^{1/2} \left(\mathrm{I}_p - 2\Xi\right)^{-1} (\mathrm{I}_p - 2\Xi)^{1/2} \xi\right) = \exp\left(\frac{1}{2}\xi'\xi\right).$$

□

Proof of Theorem 3.1. Let $\tilde{P}_{t,n}$ denote the no-arbitrage price of a n-period bond. Given a risk-neutral probability measure $\mathbb{Q}$,

$$\begin{aligned}
\tilde{P}_{t,n+1} &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{i=0}^{n} \tilde{Y}_{t+i,1}\right) \;\middle|\; X_t, X_{t-1}, \ldots\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\tilde{Y}_{t,1}\right) \tilde{P}_{t+1,n} \;\middle|\; X_t, X_{t-1}, \ldots\right], \qquad \text{(C.1)}
\end{aligned}$$

where $\tilde{Y}_{t,1}$ is the yield of a one-period bond. Suppose the solution to (C.1) is given by

$$\tilde{P}_{t,n} = \exp\left(a_n + b_n' X_t + c_n' \mathsf{vech}\left(\Omega_t\right)\right) = \exp\left(a_n + b_n' X_t + \tilde{c}_n' \mathsf{vec}\left(\Omega_t\right)\right), \qquad \text{(C.2)}$$

where $c_n' = \tilde{c}_n' \mathcal{D}$. Then,

$$\begin{aligned}\tilde{P}_{t,n+1} =& \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\tilde{Y}_{t,1} + a_n + b_n' X_{t+1} + \tilde{c}_n' \text{vec}\left(\Omega_{t+1}\right)\right) \;\middle|\; X_t, X_{t-1}, \ldots\right] \\ =& \exp\left(-\tilde{Y}_{t,1} + a_n + b_n'(\mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t) + \tilde{c}_n' \text{vec}\left(AA' + B\Omega_t B'\right)\right) \\ & \times \mathbb{E}^{\mathbb{Q}}\left[\exp\left(b_n' \Omega_t^{1/2} Z_{t+1}^{\mathbb{Q}} + \tilde{c}_n' \text{vec}\left(A Z_{t+1}^{\mathbb{Q}} Z_{t+1}^{\mathbb{Q}\prime} A'\right)\right) \;\middle|\; X_t, X_{t-1}, \ldots\right].\end{aligned}$$

Next, note that $\tilde{c}_n' \text{vec}\left(A Z_{t+1}^{\mathbb{Q}} Z_{t+1}^{\mathbb{Q}\prime} A'\right) = \tilde{c}_n'(A \otimes A) \text{vec}\left(Z_{t+1}^{\mathbb{Q}} Z_{t+1}^{\mathbb{Q}\prime}\right) = \tilde{c}_n'(A \otimes A)(Z_{t+1}^{\mathbb{Q}} \otimes Z_{t+1}^{\mathbb{Q}}) = (Z_{t+1}^{\mathbb{Q}} \otimes Z_{t+1}^{\mathbb{Q}})'(A \otimes A)' \tilde{c}_n = Z_{t+1}^{\mathbb{Q}\prime} \text{unvec}((A \otimes A)' \tilde{c}_n) Z_{t+1}^{\mathbb{Q}}$, where the function unvec is defined such that $\text{unvec}(\text{vec}\,(A)) = A$ for some square matrix A. Define $\check{c}_n = \text{unvec}[(A \otimes A)' \tilde{c}_n] = \text{unvec}[(A \otimes A)' \mathcal{E}' c_n]$. Then, if $\check{c}_n$ is symmetric with $\mathrm{I}_p - 2\check{c}_n$ positive definite, it follows from an application of Lemma 3.1 that

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}\left[\exp\left(b_n' \Omega_t^{1/2} Z_{t+1}^{\mathbb{Q}} + \tilde{c}_n' \text{vec}\left(A Z_{t+1}^{\mathbb{Q}} Z_{t+1}^{\mathbb{Q}\prime} A'\right)\right) \;\middle|\; X_t, X_{t-1}, \ldots\right] \\ &\quad = \det\left(\mathrm{I}_p - 2\check{c}_n\right)^{-1/2} \exp\left(\frac{1}{2} b_n' \Omega_t b_n\right) \\ &\quad = \det\left(\mathrm{I}_p - 2\check{c}_n\right)^{-1/2} \exp\left(\frac{1}{2}(b_n \otimes b_n)' \text{vec}\left(\Omega_t\right)\right).\end{aligned}$$

Thus, it follows that $\tilde{P}_{t,n+1}$ can be written as in (C.2) with loadings given by the recursions in Theorem 3.1. Finally, note that $\check{c}_n$ is indeed symmetric given the recursion for c_n.

C.1.2 Proof of Theorem 3.3

Lemma 3.2 *Consider the filtration $\mathcal{F}_t = \{\mathcal{Z}_t, \mathcal{Z}_{t-1}, \ldots, \mathcal{Z}_1\}$ and define $e_t = \mathcal{Z}_t - \mathbb{E}[\mathcal{Z}_t|\mathcal{F}_{t-1}]$. Then, for any non-degenerate random variable $\mathcal{X}_t$,*

$$\mathbb{E}\left[\mathcal{X}_t|\mathcal{F}_t\right] = \mathbb{E}\left[\mathcal{X}_t|\mathcal{F}_{t-1}\right] + \mathbb{E}\left[\mathcal{X}_t|e_t\right]. \tag{C.3}$$

Proof of Lemma 3.2. Since $\mathcal{F}_t = \{\mathcal{F}_{t-1}, e_t\}$, $\mathbb{E}\left[\mathcal{X}_t|\mathcal{F}_t\right] = \mathbb{E}\left[\mathcal{X}_t|\mathcal{F}_{t-1}, e_t\right]$. Next, the prediction error, e_t is mean zero and for $s = 1, \ldots, t-1$,

$$\mathbb{E}\left[\mathcal{Z}_s e_t'\right] = \mathbb{E}\left[\mathcal{Z}_s\left(\mathcal{Z}_t - \mathbb{E}[\mathcal{Z}_t|\mathcal{F}_{t-1}]\right)'\right] = \mathbb{E}\left[\mathcal{Z}_s \mathcal{Z}_t'\right] - \mathbb{E}\left[\mathbb{E}\left[\mathcal{Z}_s \mathcal{Z}_t'|\mathcal{F}_{t-1}\right]\right] = 0$$

because $\mathcal{Z}_s \in \mathcal{F}_s \subseteq \mathcal{F}_{t-1}$ and by the law of iterated expectations. Thus, $\mathcal{F}_{t-1}$-measurable processes are uncorrelated with e_t. It follows that $\mathbb{E}\left[\mathcal{X}_t|\mathcal{F}_{t-1}, e_t\right] = \mathbb{E}\left[\mathcal{X}_t|\mathcal{F}_{t-1}\right] + \mathbb{E}\left[\mathcal{X}_t|e_t\right]$.

□

Proof of Theorem 3.3. Let $\mathcal{M}_t$ denote a vector of measurements observed at time t: $\mathcal{M}_t = [Y_t', \text{vech}\,(RV_t)']'$ and denote by m the dimension of $\mathcal{M}_t$, i.e., $m = N + (N + 1)/2$. Then, the measurement equations can be written as

$$\begin{aligned}\mathcal{M}_t &= \begin{pmatrix} Y_t \\ \text{vech}\,(\text{RV}_t)\end{pmatrix} = \begin{pmatrix} a \\ \alpha\end{pmatrix} + \begin{pmatrix} b \\ \mathbf{0}_{N_{\text{RV}}\times p}\end{pmatrix} X_t + \begin{pmatrix} c \\ \beta\end{pmatrix} \text{vech}\,(\Omega_t) + \begin{pmatrix} \varepsilon_t^Y \\ \varepsilon_t^{\text{RV}}\end{pmatrix}, \\ &= \mathcal{A} + \mathcal{B}X_t + \mathcal{C}\text{vech}\,(\Omega_t) + \varepsilon_t,\end{aligned}$$

where $\mathcal{A} = (a', \alpha')'$, $\mathcal{B} = (b', \mathbf{0}_{N_{\text{RV}}\times p}')'$, $\mathcal{C} = (c', \beta')'$, and N_{RV} is the number of measurements in the measurement equation for the realized covariance matrix. Note that $\varepsilon_t \sim (0, R_t)$ with $R_t^{-1/2}\varepsilon_t$ i.i.d.

The best mean-squared prediction at time t, $\hat{\mathcal{M}}_{t|t-1}$, is given by

$$\begin{aligned}\hat{\mathcal{M}}_{t|t-1} &= \mathcal{A} + \mathcal{B}\mathbb{E}\left[X_t|\mathcal{F}_{t-1}\right] + \mathcal{C}\mathbb{E}\left[\text{vech}\,(\Omega_t)\,|\mathcal{F}_{t-1}\right] \\ &= \mathcal{A} + \mathcal{B}\hat{X}_{t|t-1} + \mathcal{C}\text{vech}\left(\hat{\Omega}_{t|t-1}\right). \qquad \text{(C.4)}\end{aligned}$$

Given the autoregressive structure of X_t and Ω_t, the best predictions of these variables are easily derived by

$$\begin{aligned}\hat{X}_{t|t-1} &= \mathbb{E}[X_t|\mathcal{F}_{t-1}] = \mu + \Phi\mathbb{E}[X_{t-1}|\mathcal{F}_{t-1}] = \mu + \Phi\hat{X}_{t-1|t-1}, \qquad \text{(C.5)} \\ \text{vech}\left(\hat{\Omega}_{t|t-1}\right) &= \mathbb{E}[\text{vech}\,(\Omega_t)\,|\mathcal{F}_{t-1}] = \text{vech}\,(CC' + AA') + \Psi\text{vech}\left(\hat{\Omega}_{t-1|t-1}\right),\end{aligned}$$

where $\Psi = \mathcal{E}(B \otimes B)\mathcal{D}$. I use Lemma 3.2 to derive the filtered state vector, $\hat{X}_{t|t}$:

$$\hat{X}_{t|t} = \mathbb{E}[X_t|\mathcal{F}_t] = \mathbb{E}[X_t|\mathcal{F}_{t-1}] + \mathbb{E}[X_t|\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}] = \hat{X}_{t|t-1} + \mathbb{E}[X_t|\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}]. \qquad \text{(C.6)}$$

It follows from (C.5) that $\hat{X}_{t|t-1}$ is linear if $\hat{X}_{t-1|t-1}$ is linear. The second term in (C.6) is a best linear predictor if there exist a $\Gamma \in \mathbb{R}^{p\times m}$ such that

$$\mathbb{E}[X_t|\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}] = \Gamma\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right) \qquad \text{(C.7)}$$

and

$$\left[X_t - \Gamma\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)\right] \perp \left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right).$$

Then,

$$\begin{aligned}&\mathbb{E}\left\{\left[X_t - \Gamma\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)\right]\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)'\right\} = \mathbf{0}_{p\times m} \\ \Leftrightarrow\quad &\Gamma = \mathbb{E}\left[X_t\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)'\right] S_{t|t-1}^{-1}, \qquad \text{(C.8)}\end{aligned}$$

where $S_{t|t-1} = \mathbb{E}\left[\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)'\right]$. From (C.6), (C.7), and (C.8), the filtered state vector is given by

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + \mathbb{E}\left[X_t\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)'\right] S_{t|t-1}^{-1}\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right). \tag{C.9}$$

Using (C.4) and $Z_t \perp \varepsilon_t$, it follows that

$$\begin{aligned}
&\mathbb{E}\left[X_t\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)'\right] \\
&\quad = \mathbb{E}\left[X_t\left(X_t - \hat{X}_{t|t-1}\right)'\right]\mathcal{B}' + \mathbb{E}\left[X_t \text{vech}\left(\Omega_t - \hat{\Omega}_{t|t-1}\right)'\right]\mathcal{C}' \\
&\quad = \mathbb{E}\left[\left(X_t - \hat{X}_{t|t-1}\right)\left(X_t - \hat{X}_{t|t-1}\right)'\right]\mathcal{B}' + \mathbb{E}\left[\left(X_t - \hat{X}_{t|t-1}\right)\text{vech}\left(\Omega_t - \hat{\Omega}_{t|t-1}\right)'\right]\mathcal{C}' \\
&\quad = V_{t|t-1}^{X}\mathcal{B}' + V_{t|t-1}^{X,\Omega}\mathcal{C}'
\end{aligned}$$

with $V_{t|t-1}^{X} = \mathbb{E}[(X_t - \hat{X}_{t|t-1})(X_t - \hat{X}_{t|t-1})']$ and $V_{t|t-1}^{x,\Omega} = \mathbb{E}[(X_t - \hat{X}_{t|t-1})\text{vech}(\Omega_t - \hat{\Omega}_{t|t-1})']$. Defining the Kalman gain for X_t, K_t^X, as in Theorem 3.3 completes the derivation of the filtered states $\hat{X}_{t|t}$. Analogous steps are taken to show the filtered conditional state covariance matrix:

$$\begin{aligned}
\text{vech}\left(\hat{\Omega}_{t|t}\right) &= \mathbb{E}[\text{vech}\left(\Omega_t\right)|\mathcal{F}_t] \\
&= \mathbb{E}[\text{vech}\left(\Omega_t\right)|\mathcal{F}_{t-1}] + \mathbb{E}[\text{vech}\left(\Omega_t\right)|\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}] \\
&= \text{vech}\left(\hat{\Omega}_{t|t-1}\right) + K_t^{\Omega}\left(\mathcal{M}_t - \hat{\mathcal{M}}_{t|t-1}\right)
\end{aligned}$$

with $K_t^{\Omega} = (V_{t|t-1}^{X,\Omega'}\mathcal{B}' + V_{t|t-1}^{\Omega}\mathcal{C}')S_{t|t-1}^{-1}$ and $V_{t|t-1}^{\Omega} = \mathbb{E}[\text{vech}(\Omega_t - \hat{\Omega}_{t|t-1})\text{vech}(\Omega_t - \hat{\Omega}_{t|t-1})']$. Finally, the mean-squared errors follow straightforwardly.

C.2 Parameter Estimates

Table C.1: Parameter Estimates

	Estimate	Standard Deviation		Estimate	Standard Deviation
δ_0	8.380	(0.065)	$\Phi^{\mathbb{Q}}_{1,3}$	0.109	(0.038)
$\Phi_{1,1}$	0.999	(0.001)	$\Phi^{\mathbb{Q}}_{2,1}$	-0.004	(0.003)
$\Phi_{2,2}$	0.990	(0.009)	$\Phi^{\mathbb{Q}}_{2,2}$	0.984	(0.006)
$\Phi_{3,3}$	0.979	(0.009)	$\Phi^{\mathbb{Q}}_{2,3}$	0.003	(0.007)
$C_{1,1}$	0.018	(0.017)	$\Phi^{\mathbb{Q}}_{3,1}$	-0.002	(0.008)
$C_{2,1}$	-0.066	(0.038)	$\Phi^{\mathbb{Q}}_{3,2}$	-0.066	(0.013)
$C_{2,2}$	0.179	(0.024)	$\Phi^{\mathbb{Q}}_{3,3}$	0.949	(0.009)
$C_{3,1}$	8.20×10^{-14}	(4.43×10^{-7})	R^{Y}_{1}	0.008	(0.001)
$C_{3,2}$	9.59×10^{-7}	(1.81×10^{-4})	R^{Y}_{2}	2.31×10^{-21}	(3.34×10^{-18})
$C_{3,3}$	1.44×10^{-4}	(0.018)	R^{Y}_{3}	2.03×10^{-4}	(1.33×10^{-5})
$A_{1,1}$	0.044	(0.019)	R^{Y}_{4}	2.52×10^{-19}	(1.19×10^{-16})
$A_{2,2}$	0.154	(0.031)	R^{Y}_{5}	6.42×10^{-21}	(7.54×10^{-19})
$B_{1,1}$	0.762	(0.036)	R^{Y}_{6}	0.001	(1.29×10^{-4})
$B_{2,2}$	0.615	(0.073)	R^{RV}_{1}	2.22×10^{-5}	(4.29×10^{-6})
$\mu^{\mathbb{Q}}_{1}$	0.047	(0.086)	R^{RV}_{2}	6.31×10^{-22}	(5.42×10^{-20})
$\mu^{\mathbb{Q}}_{2}$	0.871	(0.442)	R^{RV}_{3}	8.74×10^{-20}	(5.16×10^{-18})
$\mu^{\mathbb{Q}}_{3}$	0.650	(0.168)	R^{RV}_{4}	1.05×10^{-6}	(8.19×10^{-7})
$\Phi^{\mathbb{Q}}_{1,1}$	1.007	(0.012)	R^{RV}_{5}	8.15×10^{-23}	(7.99×10^{-21})
$\Phi^{\mathbb{Q}}_{1,2}$	0.143	(0.090)	R^{RV}_{6}	1.84×10^{-5}	(4.86×10^{-6})

Notes: The table shows parameter estimates and standard deviations computed by the Huber sandwich estimator.

Bibliography

Abbritti, M., L. A. Gil-Alana, Y. Lovcha, and A. Moreno (2016): "Term Structure Persistence," *Journal of Financial Econometrics*, 14(2), 331–352.

Adrian, T., R. K. Crump, and E. Moench (2013): "Pricing the Term Structure with Linear Regressions," *Federal Reserve Bank of New York Staff Reports*, 340.

Ahn, D.-H., R. F. Dittmar, and A. R. Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15(1), 243–288.

Ajello, A., L. Benzoni, and O. Chyruk (2018): "Core and 'Crust': Consumer Prices and the Term Structure of Interest Rates," *Working Paper*.

Andersen, T. G., and L. Benzoni (2010): "Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models," *Journal of Finance*, 65(2), 603–653.

Andersen, T. G., and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts," *International Economic Review*, 39(4), 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96(453), 42–55.

——— (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2007): "Real-Time Price Discovery in Global Stock, Bond and Foreign Exchange Markets," *Journal of International Economics*, 73, 251–277.

Andersen, T. G., T. Bollerslev, and N. Meddahi (2005): "Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities," *Econometrica*, 73(1), 279–296.

Andreasen, M. M., and K. Jørgensen (2019): "The Importance of Timing Attitudes in Consumption-Based Asset Pricing Models," *Journal of Monetary Economics*, forthcoming.

Ang, A., and M. Piazzesi (2003): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50(4), 745–787.

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

Balduzzi, P., E. J. Elton, and T. C. Green (2001): "Economic News and Bond Prices: Evidence from the U.S. Treasury Market," *Journal of Financial and Quantitative Analysis*, 36(4), 523–543.

Ball, L. (1992): "Why Does High Inflation Raise Inflation Uncertainty?," *Journal of Monetary Economics*, 29, 371–388.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society B*, 64, 253–280.

——— (2004): "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression and Correlation in Financial Economics," *Econometrica*, 72(3), 885–925.

Barnichon, R. (2010): "Building a Composite Help-Wanted Index," *Economics Letters*, 109(3), 175–178.

Bauer, M. D., and G. D. Rudebusch (2016): "Resolving the Spanning Puzzle in Macro-Finance Term Structure Models," *Review of Finance*, 21(2), 511–553.

Bauer, M. D., G. D. Rudebusch, and J. C. Wu (2012): "Correcting Estimation Bias in Dynamic Term Structure Models," *Journal of Business and Economic Statistics*, 30(3), 454–467.

——— (2014): "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset: Comment," *American Economic Reivew*, 101(4), 1514–1534.

Beechey, M., E. Hjalmarsson, and P. Österholm (2009): "Testing the Expectations Hypothesis When Interest Rates Are Near Integrated," *Journal of Banking and Finance*, 33(953), 934–943.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005): "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR)," *Quarterly Journal of Economics*, 120(1), 387–422.

Bikbov, R., and M. Chernov (2010): "No-Arbitrage Macroeconomic Determinants of the Yield Curve," *Journal of Econometrics*, 159, 166–182.

Blinder, A. S., and R. Reis (2005): "Understanding the Greenspan Standard," in *The Greenspan Era: Lessons for the Future*, ed. by A. Greenspan, pp. 11–96. Federal Reserve Bank of Kansas City, Wyoming.

Bollerslev, T., J. Cai, and F. M. Song (2000): "Intraday Periodicity, Long Memory Volatility, and Macroeconomic Announcement Effects in the US Treasury Bond Market," *Journal of Empirical Finance*, 7, 37–55.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988): "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96(1), 116–131.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22(11), 4463–4492.

Buraschi, A., and A. Jiltsov (2005): "Inflation Risk Premia and the Expectations Hypothesis," *Journal of Financial Economics*, 75, 429–490.

Campbell, J. Y., and R. J. Shiller (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95(5), 1062–1088.

——— (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58(3), 495–514.

Campbell, J. Y., A. Sunderam, and L. M. Viceira (2017): "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds," *Critical Finance Review*, 6(2), 263–301.

Carlstrom, C. T., T. S. Fuerst, and M. Paustian (2015): "Inflation and Output in New Keynesian Models with a Transient Interest Rate Peg," *Journal of Monetary Economics*, 76, 230–243.

Carr, P., and L. Wu (2009): "Variance Risk Premiums," *Review of Financial Studies*, 22(3), 1311–1341.

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

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders (1992): "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47(3), 1209–1227.

Chang, C.-L., and M. McAleer (2019): "The Fiction of Full BEKK: Pricing Fossil Fuels and Carbon Emissions," *Finance Research Letters*, 28, 11–19.

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., and D. Creal (2019): "The PPP View of Multihorizon Currency Risk Premiums," *Working Paper*.

Chernov, M., and P. Mueller (2012): "The Term Structure of Inflation Forecasts," *Journal of Financial Economics*, 106, 367–394.

Choi, H., P. Mueller, and A. Vedolin (2017): "Bond Variance Risk Premiums," *Review of Finance*, 987-1022.

Christensen, J. H. E., J. A. Lopez, and G. D. Rudebusch (2014): "Can Spanned Term Structure Factors Drive Stochastic Yield Volatility?," *Federal Reserve Bank of San Fransisco Working Paper Series*.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary Policy Shocks: What Have We Learned and to What End?," in *Handbook of Macroeconomics*, pp. 65–148. Elsevier, Amsterdam.

Cieslak, A., and P. Povala (2016): "Information in the Term Structure of Yield Curve Volatility," *Journal of Finance*, 71, 1393–1436.

Collin-Dufresne, P., and R. S. Goldstein (2002): "Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility," *Journal of Finance*, 57(4), 1685–1730.

Collin-Dufresne, P., R. S. Goldstein, and C. S. Jones (2009): "Can Interest Rate Volatility be Extracted from the Cross-section of Bond Yields?," *Journal of Financial Economics*, 94, 47–66.

Conley, T. G., L. P. Hansen, E. G. J. Luttmer, and J. A. Scheinkman (1997): "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10(3), 525–577.

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

——— (2017): "Monetary Policy Uncertainty and Economic Fluctuations," *International Economic Review*, 58(4), 1317–1354.

Dai, Q., and K. J. Singleton (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55(5), 1943–1978.

——— (2002): "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure," *Journal of Financial Economics*, 63, 415–441.

D'Amigo, S., D. Kim, and M. Wei (2018): "Tips from TIPS: the Information Content of Treasury Inflation-Protected Security Prices," *Journal of Financial and Quantitative Analysis*, 53(1), 395–436.

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

Doshi, H., K. Jacobs, and R. Liu (2018): "Macroeconomic Determinants of the Term Structure: Long-Run and Short-Run Dynamics," *Journal of Empirical Finance*, 48, 99–122.

Duffee, G. R. (2002): "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57(1), 405–443.

——— (2006): "Term Structure Estimation Without Using Latent Factors," *Journal of Financial Economics*, 79, 507–536.

——— (2018): "Expected Inflation and Other Determinants of Treasury Yields," *Journal of Finance*, 73(5), 2139–2180.

Duffie, D., and R. Kan (1996): "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.

Duffie, D., J. Pan, and K. J. Singleton (2000): "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68(6), 1343–1376.

Engle, R. (2002): "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business & Economic Statistics*, 20(3), 339–350.

Engle, R. F., and G. M. Gallo (2006): "A Multiple Indicators Model for Volatility Using Intra-Daily Data," *Journal of Econometrics*, 131, 3–27.

Engle, R. F., and K. F. Kroner (1995): "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11(1), 122–150.

Engle, R. F., D. M. Lilien, and R. P. Robins (1987): "Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model," *Econometrica*, 55(2), 391–407.

Epstein, L. G., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–969.

Feunou, B., J.-S. Fontaine, and G. Roussellet (2019): "Macro-Finance Term Structure Models with External Macro Shocks," *Working Paper*.

Fischer, S., and F. Modigliani (1978): "Towards an Understanding of the Real Effects and Costs of Inflation," *Weltwirtschaftliches Archiv*, 114, 810–833.

Fleming, M. J., and E. M. Remolona (1999): "Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information," *Journal of Finance*, 54, 1901–1915.

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(3), 451–472.

Ghysels, E., A. Le, S. Park, and H. Zhu (2014): "Risk and Return Trade-off in the U.S. Treasury Market," *Working Paper*.

Goliński, A., and P. Zaffaroni (2016): "Long Memory Affine Term Structure Models," *Journal of Econometrics*, 191, 33–56.

Golob, J. E. (1994): "Does Inflation Uncertainty Increase with Inflation?," *Economic Review, Federal Reserve Bank of Kansas City Economic Review*, Third Quarter, 27–38.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54, 2291–2304.

Hagedorn, M., J. Luo, I. Manovskii, and K. Mitman (2019): "Forward Guidance," *Journal of Monetary Economics*, 102, 1–23.

Hall, A., H. Anderson, and C. Granger (1992): "A Cointegration Analysis of Treasury Bill Yields," *Review of Economics and Statistics*, 74(1), 116–126.

Hansen, A. L. (2019): "Modeling Persistent Interest Rates with Volatility-Induced Stationarity," *Danmarks Nationalbank Working Paper*, 142.

Hansen, L. P., J. Heaton, N. Roussanov, and J. Lee (2007): "Intertemporal Substitution and Risk Aversion," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. Leamer, vol. 6A, chap. 61, pp. 3967–4056. Elsevier: North Holland.

Hansen, L. P., and J. C. Heaton (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116(2), 260–302.

Hansen, P. R., Z. Huang, and H. H. Shek (2011): "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, 27(6), 877–906.

Hartmann, M., and H. Herwartz (2012): "Causal Relations Between Inflation and Inflation Uncertainty Cross-Sectional Evidence in Favour of the Friedman-Ball Hypothesis," *Economic Letters*, 115, 144–147.

Haubrich, J., G. Pennacchi, and P. Ritchken (2012): "Inflation Expectations, Real Rates, and Risk Premia: Evidence from Inflation Swaps," *Review of Financial Studies*, 25(5), 1588–1629.

Hayford, M. D. (2000): "Inflation Uncertainty, Unemployment Uncertainty and Economic Activity," *Journal of Macroeconomics*, 22(2), 315–329.

Heston, S., and S. Nandi (2003): "A Two-Factor Term Structure Model under GARCH Volatility," *Journal of Fixed Income*, 13, 87–95.

Hilton, S. (2005): "Trends in Federal Funds Rate Volatility," *Federal Reserve Bank of New York, Current Issues in Economics and Finance*, 11(7), 1–7.

Hördahl, P., O. Tristani, and D. Vestin (2006): "A Joint Econometric Model of Macroeconomic and Term-Structure Dynamics," *Journal of Econometrics*, 131, 405–444.

Jacobs, K., and L. Karoui (2009): "Conditional Volatility in Affine Term-Structure Models: Evidence from Treasury and Swap Markets," *Journal of Financial Economics*, 91(3), 288–318.

Jardet, C., A. Monfort, and F. Pegoraro (2013): "No-Arbitrage Near-Cointegrated VAR(p) Term Structure Models, Term Premia and GDP Growth," *Journal of Banking & Finance*, 37, 389–402.

Jazwinski, A. (1970): *Stochastic Processes and Filtering Theory*, vol. 64. Academic Press, 1 edn.

Johannes, M. (2004): "The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models," *Journal of Finance*, 59(1), 227–260.

Johansen, S. (1991): "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59(6), 1551–1580.

——— (1995): *Likelihood-based Inference in Cointegated Vector Autoregressive Models*. Oxford University Press, New York.

Joslin, S. (2017): "Can Unspanned Stochastic Volatility Models Explain the Cross Section of Bond Volatilities?," *Management Science*, 64(4), 1707–1726.

Joslin, S., and Y. Konchitchki (2018): "Interest Rate Volatility, the Yield Curve, and the Macroeconomy," *Journal of Financial Economics*, 128, 344–362.

Joslin, S., A. Le, and K. J. Singleton (2013a): "Gaussian Macro-Finance Term Structure Models with Lags," *Journal of Financial Econometrics*, 11(4), 581–609.

——— (2013b): "Why Gaussian Macro-Finance Term Structure Models are (Nearly) Unconstrained Factor-VARs," *Journal of Financial Economics*, 109(3), 604–622.

Joslin, S., M. Priebsch, and K. J. Singleton (2014): "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," *Journal of Finance*, 69(3), 1197–1233.

Joslin, S., K. J. Singleton, and H. Zhu (2011): "A New Perspective on Gaussian Dynamic Term Structure Models," *Review of Financial Studies*, 24(3), 926–970.

Julier, S. J., J. K. Uhlmann, and H. Durrant-Whyte (1995): "A New Approach for Filtering Nonlinear Systems," *Proceedings of American Control Conference*, 3, 1628–1632.

Kalman, R. E. (1960): "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, 83, 35–45.

Kim, D.-H., and S.-C. Lin (2012): "Inflation and Inflation Volatility Revisited," *International Finance*, 15(3), 327–345.

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

——— (2012): "Term Structure Estimation with Survey Data on Interest Rate Forecasts," *Journal of Financial and Quantitative Analysis*, 47(1), 241–272.

Kim, D. H., and J. H. Wright (2005): "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates," *Federal Reserve Board: Finance and Economics Discussion Series*, (33).

Koeda, J., and R. Kato (2015): "The Role of Uncertainty in the Term Structure of Interest Rates: A GARCH-ATSM Approach," *Applied Economics*, 47(34-35), 3710–3722.

Koopman, S. J., and M. Scharth (2013): "The Analysis of Stochastic Volatility in the Presence of Daily Realized Measures," *Journal of Financial Econometrics*, 11(1), 76–115.

Kozicki, S., and P. A. Tinsley (2001): "Shifting Endpoints in the Term Structure of Interest Rates," *Journal of Monetary Economics*, 47(3), 613–652.

Kreps, D. M., and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice," *Econometrica*, 46(1), 185–200.

Le, A., K. J. Singleton, and Q. Dai (2010): "Discrete-Time Affine Term Structure Models with Generalized Market Prices of Risk," *Review of Financial Studies*, 23(5), 2184–2227.

Leippold, M., and L. Wu (2002): "Asset Pricing under the Quadratic Class," *Journal of Financial and Quantitative Analysis*, 37(2), 271–295.

Ling, S. (2004): "Estimation and Testing of Stationarity for Double-Autoregressive Models," *Journal of Statistics*, 66, 63–78.

Litterman, R., and J. Scheinkman (1991): "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, 1(1), 54–61.

Logue, D. E., and T. D. Willett (1976): "A Note on the Relation Between the Rate and Variability of Inflation," *Economica*, 43(170), 151–158.

McKay, A., E. Nakamura, and J. Steinson (2016): "The Power of Forward Guidance Revisited," *American Economic Review*, 106(10), 3133–3158.

Mincer, J., and V. Zarnowitz (1969): "The Evaluation of Economic Forecasts," in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, ed. by J. Mincer, pp. 3–46. Columbia University Press, New York.

Mishkin, F. S. (2007): "Headline versus Core Inflation in the Conduct of Monetary Policy," Speech at the Business Cycles, International Transmission and Macroeconomic Policies Conference, HEC Montreal, Montreal, Canada.

Monfort, A., and F. Pegoraro (2012): "Asset Pricing with Second-Order Esscher Transforms," *Journal of Banking & Finance*, 36, 1678–1687.

Nicolau, J. (2005): "Processes With Volatility-Induced Stationarity: An Application for Interest Rates," *Statistica Neerlandica*, 59(4), 376–396.

Nielsen, H. B., and A. Rahbek (2014): "Unit Root Vector Autoregression With Volatility Induced Stationarity," *Journal of Empirical Finance*, 29(C), 144–167.

Norland, E. (2018): "VIX-Yield Curve: At the Door of High Volatility?," CME Group Reports.

Piazzesi, M. (2005): "Bond Yields and the Federal Reserve," *Journal of Political Economy*, 113(2), 311–344.

Pohl, W., K. Schmedders, and O. Wilms (2018): "Higher Order Effects in Asset Pricing Models," *Journal of Finance*, 73(3), 1061–1111.

Realdon, M. (2006): "Quadratic Term Structure Models in Discrete Time," *Finance Research Letters*, 3, 277–289.

Roussellet, G. (2018): "The Term Structure of Macroeconomic Risks at the Zero Lower Bound," *Working Paper*.

Rudebusch, G., and T. Wu (2008): "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy," *Economic Journal*, 118(530), 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(3), 347–366.

Shephard, N., and K. Sheppard (2010): "Realising the Future: Forecasting with High Frequency Based Volatility (HEAVY) Models," *Journal of Applied Econometrics*, 25, 197–231.

Shiller, R. (1979): "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," *Journal of Political Economy*, 87(6), 1190–1219.

Stock, J. H., and M. W. Watson (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15, 101–115.

Trolle, A. B. (2009): "The Price of Interest Rate Variance Risk and Optimal Investments in Interest Rate Derivatives," *EFA 2009 Bergen Meetings Paper*.

Trolle, A. B., and E. S. Schwartz (2009): "A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives," *Review of Financial Studies*, 22(5), 2007–2057.

——— (2015): "The Swaption Cube," *Review of Financial Studies*, 27(8), 2307–2353.

Weil, P. (1990): "Nonexpected Utility in Macroeconomics," *Quarterly Journal of Economics*, 105(1), 29–42.

Wright, J. H. (2011): "Term Premiums and Inflation Uncertainty: Empirical Evidence from an International Dataset," *American Economic Reivew*, 101(4), 1514–1534.

Zhu, H., X. Zhang, X. Liang, and Y. Li (2017): "On a Vector Double Autoregressive Model," *Statistics & Probability Letters*, 129, 85–96.