1-1-2014

Essays on Risk Measurement and Modeling in Macroeconomics and Finance

Dongho Song
University of Pennsylvania, donghos@sas.upenn.edu

Essays on Risk Measurement and Modeling in Macroeconomics and Finance

Abstract
This dissertation consists of four essays that focus on the measurement and economic analysis of key risk factors behind macroeconomic and financial variables using state-space models. Chapters 2, 3, 4, and 5 develop and implement estimation approaches that can handle nonlinear linkages of economic forces and tackle issues when data are missing or contaminated by errors. Chapter 2 estimates an equilibrium term structure model that includes real and nominal uncertainty in particular that allows for changes in the responsiveness of the Federal Reserve to inflation fluctuations. These uncertainty, particularly those concerning monetary policy action are considered potential sources of risk variations that can explain several features in the U.S. government bond market including the upward sloping yield curve. Chapter 3, co-authored with Frank Schorfheide and Amir Yaron, develops a nonlinear state-space model to estimate predictable mean and volatility components in monthly consumption growth using a mixed-frequency data and accounting for serially-correlated measurement errors. We provide a methodological contribution that allow to maximize the span of the estimation sample to recover the predictable component and at the same time use high-frequency data to efficiently identify the volatility processes. The estimation provides strong evidence for predictable mean and volatility components in consumption growth. We show that the model can go a long way in explaining several well known asset pricing facts of the data. Chapter 4, co-authored with Boragan Aruoba, Francis Diebold, Jeremy Nalewaik, and Frank Schorfheide, considers the fundamental question of GDP estimation, focusing on the U.S., and provides estimates superior to the ubiquitous expenditure-side series by applying optimal signal-extraction techniques to the noisy expenditure-side and income-side GDP estimates. The quarter-by-quarter values of the new measure often differ noticeably from those of the traditional measures, and dynamic properties differ as well, indicating that the persistence of aggregate output dynamics is stronger than previously thought. Chapter 5, co-authored with Frank Schorfheide, develops the idea of using mixed-frequency data in state-space form. We show that adding monthly observations to a quarterly VAR, which then is estimated with Bayesian methods under a Minnesota-style prior, substantially improves its forecasting performance.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Francis X. Diebold

Second Advisor
Frank Schorfheide

ESSAYS ON RISK MEASUREMENT AND MODELING IN MACROECONOMICS AND FINANCE

Dongho Song

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

Francis X. Diebold

Professor of Economics

Co-Supervisor of Dissertation

Frank Schorfheide

Professor of Economics

Graduate Group Chairperson

George J. Mailath

Professor of Economics

Dissertation Committee

Amir Yaron, Professor of Finance

Ivan Shaliastovich, Assistant Professor of Finance

ESSAYS ON RISK MEASUREMENT AND MODELING

IN MACROECONOMICS AND FINANCE

Acknowledgments

I am immensely grateful to my advisors Francis X. Diebold, Frank Schorfheide, Ivan Shaliastovich, and Amir Yaron for their invaluable guidance and encouragement at every stage. Without their help, this dissertation could not be completed. I would like to express my special appreciation to Frank Schorfheide and Amir Yaron, you have been tremendous mentors for me. Thank you for your patience, understanding, guidance, and most of all the encouragement you have given me during my graduate studies. I thank Francesco Bianchi, Xu Cheng, Francis DiTraglia, Jesus Fernandez-Villaverde, Christian Opp, Nikolai Roussanov, and Xun Tang for the feedback, direction, and assistance when I needed it.

Thanks are also extended to all the members of our Doctoral cohort. In particular, I would like to thank my officemates, Luigi Bocola, Yousuk Kim, and Minchul Shin. This is a special group of people that I consider members of my extended family.

Finally, special recognition goes out to my parents, for their support, encouragement and patience during my pursuit of the Doctorate in Economics. To my wife Ami Ko, who inspired me and provided constant encouragement during the entire process.

ABSTRACT

ESSAYS ON RISK MEASUREMENT AND MODELING IN MACROECONOMICS AND FINANCE

Dongho Song

Francis X. Diebold

Frank Schorfheide

This dissertation consists of four essays that focus on the measurement and economic analysis of key risk factors behind macroeconomic and financial variables using state-space models. Chapters 2, 3, 4, and 5 develop and implement estimation approaches that can handle nonlinear linkages of economic forces and tackle issues when data are missing or contaminated by errors. Chapter 2 estimates an equilibrium term structure model that includes real and nominal uncertainty in particular that allows for changes in the responsiveness of the Federal Reserve to inflation fluctuations. These uncertainty, particularly those concerning monetary policy action are considered potential sources of risk variations that can explain several features in the U.S. government bond market including the upward sloping yield curve. Chapter 3, co-authored with Frank Schorfheide and Amir Yaron, develops a nonlinear state-space model to estimate predictable mean and volatility components in monthly consumption growth using a mixed-frequency data and accounting for serially-correlated measurement errors. We provide a methodological contribution that allow to maximize the span of the estimation sample to recover the predictable component and at the same time use high-frequency data to efficiently identify the volatility processes. The estimation provides strong evidence for predictable mean and volatility components in consumption growth. We show that the model can go a long way in explaining several well

known asset pricing facts of the data. Chapter 4, co-authored with Boragan Aruoba, Francis Diebold, Jeremy Nalewaik, and Frank Schorfheide, considers the fundamental question of GDP estimation, focusing on the U.S., and provides estimates superior to the ubiquitous expenditure-side series by applying optimal signal-extraction techniques to the noisy expenditure-side and income-side GDP estimates. The quarter-by-quarter values of the new measure often differ noticeably from those of the traditional measures, and dynamic properties differ as well, indicating that the persistence of aggregate output dynamics is stronger than previously thought. Chapter 5, co-authored with Frank Schorfheide, develops the idea of using mixed-frequency data in state-space form. We show that adding monthly observations to a quarterly VAR, which then is estimated with Bayesian methods under a Minnesota-style prior, substantially improves its forecasting performance.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation consists of four essays that focus on the measurement and economic analysis of key risk factors behind macroeconomic and financial variables using state-space models. Chapters 2, 3, 4, and 5 develop and implement estimation approaches that can handle nonlinear linkages of economic forces and tackle issues when data are missing or contaminated by errors. Chapter 2 provides strong empirical evidence of structural changes in the United States government bond markets and examines whether the changes are brought about by external shocks, monetary policy, or by both. To explore this, Chapter 2 estimates an equilibrium term structure model that includes real and nominal uncertainty in particular that allows for changes in the responsiveness of the Federal Reserve to inflation fluctuations. These uncertainty, particularly those concerning monetary policy action are considered potential sources of risk variations that can explain several features in the bond market including the upward sloping yield curve. Chapter 3, co-authored with Frank Schorfheide and Amir Yaron, develops a nonlinear state-space model to estimate predictable mean and volatility components in monthly consumption growth using a mixed-frequency data and accounting for serially-correlated measurement errors. We provide a methodological contribution that allows to maximize the span of the estimation sample to recover the predictable component and at the same time uses high-frequency data to efficiently identify

the volatility processes. The estimation provides strong evidence for predictable mean and volatility components in consumption growth. We show that the model can go a long way in explaining several well known asset pricing facts of the data. Chapter 4, co-authored with Boragan Aruoba, Francis Diebold, Jeremy Nalewaik, and Frank Schorfheide, considers the fundamental question of GDP estimation, focusing on the U.S., and provides estimates superior to the ubiquitous expenditure-side series by applying optimal signal-extraction techniques to the noisy expenditure-side and income-side GDP estimates. The quarter-by-quarter values of the new measure often differ noticeably from those of the traditional measures, and dynamic properties differ as well, indicating that the persistence of aggregate output dynamics is stronger than previously thought. Chapter 5, co-authored with Frank Schorfheide, develops the idea of using mixed-frequency data in state-space form. We show that adding monthly observations to a quarterly VAR, which then is estimated with Bayesian methods under a Minnesota-style prior, substantially improves its forecasting performance.

1.1 "Bond Market Exposures to Macroeconomic and Monetary Policy Risks"

Chapter 2 provides empirical evidence of changes in the U.S. Treasury yield curve and related macroeconomic factors, and investigates whether the changes are brought about by external shocks, monetary policy, or by both. To explore this, I characterize bond market exposures to macroeconomic and monetary policy risks, using an equilibrium term structure model with recursive preferences in which inflation dynamics are endogenously determined. In my model, the key risks that affect bond market prices are changes in the correlation between growth and inflation and changes in the conduct of monetary policy. Using a novel estimation technique, I find that the changes in monetary policy affect the volatility of yield spreads, while the changes in the correlation between growth and inflation affect both the

level as well as the volatility of yield spreads. Consequently, the changes in the correlation structure are the main contributor to bond risk premia and to bond market volatility. The time variations within a regime and risks associated with moving across regimes lead to the failure of the Expectations Hypothesis and to the excess bond return predictability regression of Cochrane and Piazzesi (2005), as in the data.

1.2 "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach"

Chapter 3, co-authored with Frank Schorfheide and Amir Yaron, develops a nonlinear state-space model that captures the joint dynamics of consumption, dividend growth, and asset returns. Building on Bansal and Yaron (2004), our model consists of an economy containing a common predictable component for consumption and dividend growth and multiple stochastic volatility processes. The estimation is based on annual consumption data from 1929 to 1959, monthly consumption data after 1959, and monthly asset return data throughout. We maximize the span of the sample to recover the predictable component and use high-frequency data, whenever available, to efficiently identify the volatility processes. Our Bayesian estimation provides strong evidence for a small predictable component in consumption growth (even if asset return data are omitted from the estimation). Three independent volatility processes capture different frequency dynamics; our measurement error specification implies that consumption is measured much more precisely at an annual than monthly frequency; and the estimated model is able to capture key asset-pricing facts of the data.

1.3 "Improving GDP Measurement: A Measurement-Error Perspective"

Chapter 4, co-authored with Boragan Aruoba, Francis Diebold, Jeremy Nalewaik, and Frank Schorfheide, provides a new measure of U.S. GDP growth, obtained by applying optimal

signal-extraction techniques to the noisy expenditure-side and income-side GDP estimates. The quarter-by-quarter values of our new measure often differ noticeably from those of the traditional measures. Its dynamic properties differ as well, indicating that the persistence of aggregate output dynamics is stronger than previously thought.

1.4 "Real-Time Forecasting with a Mixed-Frequency VAR"

Chapter 5, co-authored with Frank Schorfheide, develops a vector autoregression (VAR) for time series which are observed at mixed frequencies – quarterly and monthly. The model is cast in state-space form and estimated with Bayesian methods under a Minnesota-style prior. We show how to evaluate the marginal data density to implement a data-driven hyperparameter selection. Using a real-time data set, we evaluate forecasts from the mixed-frequency VAR and compare them to standard quarterly-frequency VAR and to forecasts from MIDAS regressions. We document the extent to which information that becomes available within the quarter improves the forecasts in real time.

Chapter 2

Bond Market Exposures to Macroeconomic and Monetary Policy Risks

2.1 Introduction

There is mounting evidence that the U.S. Treasury yield curve and relevant macroeconomic factors have undergone structural changes over the past decade. For example, recent empirical studies have come to understand that U.S. Treasury bonds have served as a hedge to stock market risks in the last decade.[1] In sharp contrast to the 1980s, during which both bond and stock returns were low and tended to co-move positively, the bond-stock return correlation has turned strongly negative in the 2000s. Several other aspects of bond markets have changed over the years between 1998 and 2011. Among them are a flattening of the yield curve and a substantial drop in the degree of time variation in excess bond returns. The striking feature is that the correlation between the macroeconomic factors, that is, consumption growth and inflation, have also changed from negative to positive in the same period.[2] In this paper, I study the role of structural changes in the macroeconomic factors

[1] See Baele, Bekaert, and Inghelbrecht (2010); Campbell, Pflueger, and Viceira (2013); Campbell, Sunderam, and Viceira (2013); and David and Veronesi (2013).

[2] See Table 2.1 for descriptive statistics.

as well as in the conduct of monetary policy in explaining the bond market changes over the last decade. The central contributions of this paper are to investigate whether the bond market changes are brought about by external shocks, by monetary policy, or by both, and to quantify and characterize bond market price exposures to macroeconomic and monetary policy risks.

I develop a state-space model to capture the joint dynamics of consumption growth, inflation, and asset returns. The real side of the model builds on the work of Bansal and Yaron (2004) and assume that consumption growth contains a small predictable component (i.e., long-run growth), which in conjunction with investor's preference for early resolution of uncertainty determine the price of real assets. The nominal side of the model extends Gallmeyer, Hollifield, Palomino, and Zin (2007) in that inflation dynamics are derived endogenously from the monetary policy rule, and the nominal assets inherit the properties of monetary policy. My model distinguishes itself from the existing literature in two important dimensions. First, it allows for changes in the monetary policy rule, both in the inflation target and in the stabilization rule (i.e., the central bank's response to deviations of actual inflation from the inflation target and to fluctuations in consumption). The regime-switches in stabilization policy coefficients are modeled through a Markov process. Second, I allow for a channel that breaks the long-run dichotomy between the nominal and real sides of the economy. I assume that the fluctuations in the long-run growth component are not just driven by its own innovation process but also by the innovation to the inflation target of the central bank. I add flexibility to this channel by allowing for both positive-negative fluctuations. In essence, there is a regime-switching Markov process that captures the sign-switching behavior of conditional covariance between long-run growth and the inflation target.

As a consequence of my model features, the asset prices and macroeconomic aggregates

are affected by two distinct channels: (1), changes in the conditional covariance between the inflation target and long-run growth, and (2), changes in the stabilization policy rule. This leads to endogenous inflation dynamics and resulting nominal bond market prices are differentially affected by both channels. In order to empirically assess the relative strength of the two channels, I apply a novel Bayesian approach to the estimation of the model parameters and to the nonlinear filtering problem, which arises due to hidden Markov states (i.e., regimes) and stochastic volatilities.

The estimation of the model delivers several important empirical findings. First, the estimation results suggest that the economic environment involves two regimes with different conditional covariance dynamics: one with a negative covariance between the inflation target and long-run growth (*countercyclical* inflation) and one with a positive covariance (*procyclical* inflation). The relative magnitude of the conditional heteroscedasticity present is larger in the countercyclical inflation regime. In each inflation regime the central bank either increases interest rates more than one-for-one with inflation (*active* monetary policy) or does not (*passive* monetary policy). Overall, there are a total of 4 different regimes that affect comovement of bond prices and macroeconomic aggregates. Second, the changes in the conditional covariance between the inflation target and long-run growth alter the dynamics of long-run components and have a persistent effect on bond markets. On the other hand, the changes in the conduct of monetary policy are more targeted toward affecting the short-run dynamics of inflation and therefore their effect on bond markets is short-lived. I find the changes in the conditional covariance dynamics to be the main driver of structural changes in bond markets, such as sign changes in the stock-bond return correlation and the drop in time variation in excess bond returns.

Third, each regime carries distinctly different risk prices, and uncertainty concerning moving across regimes poses additional risks to bond markets. The risks channels can

be broadly classified into two types: "*within*-regime" and "*across*-regime" risks. For the purpose of explanation, I decompose the bond yields into the expected sum of future short rates (the expectations component) and the term premium (risk compensation for long-term bonds). Risks associated with the countercyclical inflation regime raise both the expectations component and the term premium.[3] Risks for the procyclical inflation work in the opposite direction. With regard to monetary policy risks, the effect is mostly on the expectations component, but its directional influence depends on the inflation regime. When the policy stance is active, monetary policy works toward lowering the inflation expectation and produces a downward shift in the level of the term structure (i.e., lowers the expectations component). With passive monetary policy and a countercyclical inflation regime, agents understand that the central bank is less effective in stabilizing the economy (raising the expectations component) and demand a greater inflation premium, leading to the steepest term structure. With passive monetary policy and a procyclical inflation regime, the inherent instability associated with the passive monetary policy will amplify the "procyclicality" (lower the expectations component). The across-regime risks imply that the risks properties of alternative regimes are incorporated as agents are aware of the possibility of moving across regimes. This is a prominent feature of the model that generates an upward-sloping yield curve even when the economy is in the procyclical inflation regime. As long as the switching probability is sufficiently high, agents will always demand an inflation premium as compensation for the countercyclical inflation risks.

Fourth, the time variations within a regime and risks associated with moving across regimes give rise to time variations in risk premia, which provide testable implications for the Expectations Hypothesis (EH). The estimated model as a whole overwhelmingly

[3]Note that this is how Piazzesi and Schneider (2006) and Bansal and Shaliastovich (2013) generate the inflation premium.

rejects the EH and provides strong empirical evidence of time variations in expected excess bond returns. The evidence is supported by the model-implied term spread regression of Campbell and Shiller (1991) and the excess bond return predictability regression of Cochrane and Piazzesi (2005). However, I find that the degree of violation of the EH is least apparent with a procyclical inflation regime and passive monetary policy. The increase in the term premium will be minimal in the procyclical inflation regime and the relative importance of the expectations component on the long-term rate movements will be large in the passive monetary policy stance, which together bring the bond market closer to what the EH predicts. I believe I am the first to show that this interesting feature of the model is also documented in the data once I partition them based on the identified regimes.

Related Literature. This paper is related to several strands of literature. My work is related to a number of recent papers that study the changes in bond-stock return correlation. Baele, Bekaert, and Inghelbrecht (2010) utilize a dynamic factor model in which stock and bond returns depend on a number of economic state variables, e.g., macroeconomic, volatility, and liquidity factors, and attribute the cause of changes in bond-stock return correlation to liquidity factors. Campbell, Sunderam, and Viceira (2013) embed time-varying bond-stock return covariance in a quadratic term-structure model and argue that the root cause is due to changes in nominal risks in bond markets. What distinguishes my work from these reduced-form studies is that it builds on a consumption-based equilibrium model to understand the macroeconomic driving forces behind the yield curve changes. In this regard, the approach of Campbell, Pflueger, and Viceira (2013) and David and Veronesi (2013) are more relevant to my study. Campbell, Pflueger, and Viceira (2013) examine the role of monetary policy using a New Keynesian model and David and Veronesi (2013) explore the time-varying signaling role of inflation in a consumption-based model. My work complements these two studies because it studies the role of structural changes

in the macroeconomic factors as well as in the conduct of monetary policy in a unified framework, and investigates their role in explaining the bond market fluctuations.

By investigating time variation of the stance of monetary policy, my work also contributes to the monetary policy literature, e.g., Clarida, Gali, and Gertler (2000), Coibon and Gorodnichenko (2011), Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010), Lubik and Schorfheide (2004), Schorfheide (2005), and Sims and Zha (2006).[4] While most of these papers study the impact of changes in monetary policy on macroeconomic aggregates, Ang, Boivin, Dong, and Loo-Kung (2011) and Bikbov and Chernov (2013) focus on their bond market implications (using reduced-form modeling frameworks). My work distinguishes itself from these last two papers as I focus on a fully specified economic model and characterize time-varying bond market exposures to monetary policy risks.

In terms of modeling term structure with recursive preferences, this paper is closed related to those of Bansal and Shaliastovich (2013), Doh (2012), and Piazzesi and Schneider (2006), who work in an endowment economy setting, and, van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez (2012) who study in a production-based economy. My work generalizes the first three by endogenizing inflation dynamics from monetary policy rule. While van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez (2012) allow for endogenous capital and labor supply and analyze their interaction with the yield curve, which are ignored in my analysis, they do not allow for time variations in volatilities and in monetary policy stance, both of which are key risk factors in my analysis.

There is a growing and voluminous literature in macro and finance that highlights the importance of volatility for understanding the macroeconomy and financial markets (see Bansal, Kiku, and Yaron (2012a); Bansal, Kiku, Shaliastovich, and Yaron (2013); Bloom (2009); and Fernández-Villaverde and Rubio-Ramírez (2011)). This paper further con-

[4]Note that I am including those that explicitly account for changes in monetary policy.

tributes to the literature by incorporating time-varying covolatility specifications. Finally, the estimation algorithm builds on Schorfheide, Song, and Yaron (2013), yet further develops to accommodate Markov-switching processes (see Kim and Nelson (1999) for a comprehensive overview of estimation methods for the Markov switching models) and efficiently implements Bayesian inference using particle filtering in combination with a Markov chain Monte Carlo (MCMC) algorithm.

The remainder of the paper is organized as follows. Section 3.2 introduces the model environment and describes the model solution. Section 3.3 presents the empirical state-space model and describes the estimation procedure. Section 3.4 discusses the empirical findings, and Section 5.5 provides concluding remarks.

2.2 The Long-Run Risks (LRR) Model with Monetary Policy

2.2.1 Preferences and Cash-flow Dynamics

I consider an endowment economy with a representative agent who maximizes her lifetime utility,

$$V_t = \max_{C_t} \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \big(\mathbb{E}_t [V_{t+1}^{1-\gamma}] \big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}},$$

subject to budget constraint

$$W_{t+1} = (W_t - C_t) R_{c,t+1},$$

where W_t is the wealth of the agent, $R_{c,t+1}$ is the return on all invested wealth, γ is risk aversion, $\theta = \frac{1-\gamma}{1-1/\psi}$, and ψ is intertemporal elasticity of substitution (IES).

Following Bansal and Yaron (2004), consumption growth, $g_{c,t+1}$, is decomposed into a (persistent) long-run growth component, $x_{c,t+1}$, and a (transitory) short-run component, $\bar{\sigma}_c \eta_{c,t+1}$. The persistent long-run growth component is modeled as an AR(1) process with two fundamental shocks: shock to growth, $\sigma_{c,t} e_{c,t+1}$, and shock to the inflation target,

$\sigma_{\pi,t}e_{\pi,t+1}$ (both with stochastic volatilities). The inflation target is modeled by an AR(1) process with its own stochastic volatilities and the persistence is allowed to switch regimes. The persistence of the long-run growth, $\rho_c(S_{t+1})$, and its exposure to inflation target shock, which is captured by $\chi_{c,\pi}(S_{t+1})$, are subject to regime changes, where S_{t+1} denotes the regime indicator variable. The value of $\chi_{c,\pi}(S_{t+1})$ can be either negative or positive. The economic reasoning behind this follows the view that there are periods in which the inflation target is above the so-called desirable rate of inflation,[5] and that any positive shock to the inflation target during those periods creates distortions and hampers long-run growth. The negative $\chi_{c,\pi}(S_{t+1})$ values correspond to these periods. The periods with positive $\chi_{c,\pi}(S_{t+1})$ values depict periods during which the inflation target is assumed to be lower than the desirable one, and a positive shock to the inflation target removes distortions and facilitates long-run growth. Dividend streams, $g_{d,t+1}$, have levered exposures to both $x_{c,t+1}$ and $\bar{\sigma}_c\eta_{c,t+1}$, whose magnitudes are governed by the parameters ϕ_x and ϕ_η, respectively. I allow $\bar{\sigma}_d\eta_{d,t+1}$ to capture idiosyncratic movements in dividend streams. Overall, the joint dynamics for the cash-flows are

$$\begin{aligned}
\begin{bmatrix} g_{c,t+1} \\ g_{d,t+1} \end{bmatrix} &= \begin{bmatrix} \mu_c \\ \mu_d \end{bmatrix} + \begin{bmatrix} 1 \\ \phi_x \end{bmatrix} x_{c,t+1} + \begin{bmatrix} 1 & 0 \\ \phi_\eta & 1 \end{bmatrix} \begin{bmatrix} \bar{\sigma}_c\eta_{c,t+1} \\ \bar{\sigma}_d\eta_{d,t+1} \end{bmatrix} \\
x_{c,t+1} &= \rho_c(S_{t+1})x_{c,t} + \sigma_{c,t}e_{c,t+1} + \chi_{c,\pi}(S_{t+1})\sigma_{\pi,t}e_{\pi,t+1}, \\
x_{\pi,t+1} &= \rho_\pi(S_{t+1})x_{\pi,t} + \sigma_{\pi,t}e_{\pi,t+1}
\end{aligned} \tag{2.1}$$

where the stochastic volatilities evolve according to

$$\sigma_{j,t} = \varphi_j\bar{\sigma}_c\exp(h_{j,t}), \quad h_{j,t+1} = \nu_j h_{j,t} + \sigma_{h_j}\sqrt{1-\nu_j^2}w_{j,t+1}, \quad j = \{c,\pi\}, \tag{2.2}$$

and the shocks are assumed to be

$$\eta_{i,t+1}, e_{j,t+1} \sim N(0,1), \quad i \in \{c,d\}.$$

[5] In a New Keynesian model, the desirable rate of inflation would be the rate at which prices can be changed without costs. See Aruoba and Schorfheide (2011) for a more detailed discussion.

Following Schorfheide, Song, and Yaron (2013), the logarithm of the volatility process is assumed to be normal, which ensures that the standard deviation of the shocks remains positive at every point in time.

2.2.2 Monetary Policy

Monetary policy consists of two components: stabilization and a time-varying inflation target. Stabilization policy is "active" or "passive" depending on its responsiveness to the consumption gap and inflation fluctuations relative to the target. The monetary policy shock, $x_{m,t}$, is also modeled as an AR(1) process. In sum, monetary policy follows a regime-switching Taylor rule,

$$\begin{aligned} i_t &= \mu_i^{MP}(S_t) + \underbrace{\tau_c(S_t)(g_{c,t} - \mu_c)}_{\text{consumption gap}} + \underbrace{\tau_\pi(S_t)(\pi_t - x_{\pi,t})}_{\text{short-run inflation}} + x_{\pi,t} + x_{m,t}, \qquad (2.3) \\ &= \mu_i^{MP}(S_t) + \big[\tau_c(S_t), 1 - \tau_\pi(S_t), 1, \tau_c(S_t)\big] X_t^B + \tau_\pi(S_t)\pi_t, \quad X_t^B = [x_{c,t}, x_{\pi,t}, x_{m,t}, \eta_{c,t}]', \end{aligned}$$

where $\tau_c(S_t)$ and $\tau_\pi(S_t)$ capture central bank's reaction to the consumption gap and to short-run inflation variation, respectively. To recap, the dynamics of the inflation target and monetary policy shocks are

$$\begin{aligned} x_{\pi,t+1} &= \rho_\pi(S_{t+1}) x_{\pi,t} + \sigma_{\pi,t} e_{\pi,t+1} \\ x_{m,t+1} &= \rho_m x_{m,t} + \sigma_m e_{m,t+1}. \end{aligned}$$

Observe that several important modifications have been made in (2.3). To begin with, the role of interest rate smoothing is assumed absent. While (2.3) may look quite restrictive in its form, it yields much a simpler expression in that the current short-rate is affine with respect to the "current" state variables, X_t^B, and "realized" inflation, π_t, without any "lagged" term. Moreover, given the argument posited in Rudebusch (2002), it seems sensible to consider the monetary policy rule without interest rate smoothing motive in

order to study the term structure.[6] More importantly, however, (2.3) assumes that the central bank makes informed decisions with respect to inflation fluctuations at different frequencies. While the central bank attempts to steer actual inflation towards the inflation target at low frequencies, it aims to stabilize inflation fluctuations relative to its target at high frequencies. Furthermore, in the context of the term structure models, it is very important to consider an explicit role for the target inflation since it behaves similarly to a level factor of the nominal term structure. The specification of (2.3) resembles specifications in which the level factor of the term structure directly enters into the monetary policy rule (see Rudebusch and Wu (2008) for example).[7] Finally, (2.3) assumes that the strength with which the central bank tries to pursue its goal—a stabilization policy—changes over time along the lines explored in Clarida, Gali, and Gertler (2000).

2.2.3 Endogenous Inflation Dynamics

Inflation dynamics can be determined endogenously from the monetary policy rule (2.3) and a Fisher-type asset-pricing equation which is given below,

$$\begin{aligned} i_t &= -\mathbb{E}_t\left[m_{t+1} - \pi_{t+1}\right] - \frac{1}{2}\mathbb{V}_t\left[m_{t+1} - \pi_{t+1}\right] \\ &\approx \mu_i^{AP}(S_t) + \big[\frac{1}{\psi}\mathbb{E}_t[\rho_c(S_{t+1})], 0, 0, 0\big] X_t^B + E_t\left[\pi_{t+1}\right], \quad X_t^B = [x_{c,t}, x_{\pi,t}, x_{m,t}, \eta_{c,t}]'. \end{aligned} \tag{2.4}$$

(see Cochrane (2011) and Backus, Chernov, and Zin (2013) for a similar discussion.) The approximation is exact if the short-rate contains no risk premium.[8] Substituting the asset-pricing equation (2.4) into the monetary policy rule (2.3), the system reduces to a single

[6] Based on the term structure evidence, Rudebusch (2002) shows that monetary policy inertia is not due to the smoothing motive but is due to persistent shocks.

[7] Note also that incorporating a time-varying inflation target is quite common in the monetary policy literature (see Ascari and Sbordone (2013); Coibon and Gorodnichenko (2011); and Aruoba and Schorfheide (2011)).

[8] This assumption is not unreasonable given the results of the variance decomposition of the short-rate in the subsequent section, see Table 2.5. Also, Campbell, Pflueger, and Viceira (2013) apply similar assumption.

regime-dependent equation

$$\tau_\pi(S_t)\pi_t \quad = \quad \mathbb{E}_t\left[\pi_{t+1}\right] + \Lambda(S_t)X_t^B, \tag{2.5}$$

where $\Lambda(S_t) = \left[\frac{1}{\psi}\mathbb{E}_t[\rho_c(S_{t+1})], 0, 0, 0\right] - \left[\tau_c(S_t), 1-\tau_\pi(S_t), 1, \tau_c(S_t)\right]$.[9] In the appendix, I show that the equilibrium inflation dynamics can be expressed as

$$\pi_t \quad = \quad \Gamma(S_t)X_t^B, \quad \text{where} \quad \Gamma(S_t) = [\underbrace{\Gamma_{x,c}(S_t), \Gamma_{x,\pi}(S_t), \Gamma_{x,m}(S_t)}_{\Gamma_x(S_t)}, \Gamma_\eta(S_t)]. \tag{2.6}$$

2.2.4 Markov-Chain

In order to achieve flexibility while maintaining parsimony,[10] I assume that the model parameters evolve according to a four-state Markov-chain $S_t = (S_t^X, S_t^M)$ (i.e., that the regime-switching is not synchronized). It can be further decomposed into two independent two-state Markov-chains, S_t^X, S_t^M,

$$\mathbb{P}_X = \begin{bmatrix} p_{X_1} & 1-p_{X_1} \\ 1-p_{X_2} & p_{X_2} \end{bmatrix}, \quad \mathbb{P}_M = \begin{bmatrix} p_{M_1} & 1-p_{M_1} \\ 1-p_{M_2} & p_{M_2} \end{bmatrix}$$

where X_i and M_i are indicator variables for correlation and monetary policy regimes, $i = 1, 2$. Define

$$S_t = \begin{cases} 1 & \text{if} \quad S_t^X = X_1 \quad \text{and} \quad S_t^M = M_1 \\ 2 & \text{if} \quad S_t^X = X_1 \quad \text{and} \quad S_t^M = M_2 \\ 3 & \text{if} \quad S_t^X = X_2 \quad \text{and} \quad S_t^M = M_1 \\ 4 & \text{if} \quad S_t^X = X_2 \quad \text{and} \quad S_t^M = M_2, \end{cases}$$

from which I construct the transition probability $\mathbb{P} = \mathbb{P}_X \otimes \mathbb{P}_M$.

[9] Equation (2.5) holds true if $\mu_i^{MP}(S_t) = \mu_i^{AP}(S_t)$.

[10] There is no reason to assume *a priori* that the coefficient, $\chi_{c,\pi}$, and the monetary policy parameters, τ_c, τ_π, switch simultaneously.

2.2.5 Solution

The first-order condition of the agent's expected utility maximization problem yields the Euler equations

$$\mathbb{E}_t\left[\exp\left(m_{t+1}+r_{k,t+1}\right)\right]=1, \quad k\in\{c,m\}, \quad \text{(Real Assets)} \tag{2.7}$$

$$p_{n,t}=\log\mathbb{E}_t[\exp(m_{t+1}-\pi_{t+1}+p_{n-1,t+1})], \quad \text{(Nominal Assets)} \tag{2.8}$$

where $m_{t+1}=\theta\log\delta-\frac{\theta}{\psi}g_{c,t+1}+(\theta-1)r_{c,t+1}$ is the log of the real stochastic discount factor (SDF), $r_{c,t+1}$ is the log return on the consumption claim, $r_{m,t+1}$ is the log market return, and $p_{n,t}$ is the log price of an n-month zero-coupon bond.

The solutions to (2.7) and (2.8) depend on the joint dynamics of consumption, dividend growth, and inflation, which can be conveniently broken up into three parts and be re-written as:

1. Fundamental Dynamics

$$\begin{bmatrix} g_{c,t+1}\\ g_{d,t+1}\\ \pi_{t+1}\end{bmatrix} = \begin{bmatrix}\mu_c\\ \mu_d\\ \mu_\pi\end{bmatrix}+\begin{bmatrix} e_1\\ \phi_x e_1\\ \Gamma_x(S^X_{t+1},S^M_{t+1})\end{bmatrix}X_{t+1}+\begin{bmatrix}1 & 0 & 0\\ \phi_\eta & 1 & 0\\ \Gamma_\eta(S^X_{t+1},S^M_{t+1}) & 0 & 1\end{bmatrix}\begin{bmatrix}\bar\sigma_c\eta_{c,t+1}\\ \bar\sigma_d\eta_{d,t+1}\\ \bar\sigma_\pi\eta_{\pi,t+1}\end{bmatrix} \tag{2.9}$$

2. The Conditional Mean Dynamics

$$\underbrace{\begin{bmatrix}x_{c,t+1}\\ x_{\pi,t+1}\\ x_{m,t+1}\end{bmatrix}}_{X_{t+1}} = \underbrace{\begin{bmatrix}\rho_c(S^X_{t+1}) & 0 & 0\\ 0 & \rho_\pi(S^X_{t+1}) & 0\\ 0 & 0 & \rho_m\end{bmatrix}}_{\Upsilon(S^X_{t+1})}\underbrace{\begin{bmatrix}x_{c,t}\\ x_{\pi,t}\\ x_{m,t}\end{bmatrix}}_{X_t}+\underbrace{\begin{bmatrix}1 & \chi_{c,\pi}(S^X_{t+1}) & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}}_{\Omega(S^X_{t+1})}\underbrace{\begin{bmatrix}\sigma_{c,t}e_{c,t+1}\\ \sigma_{\pi,t}e_{\pi,t+1}\\ \sigma_m e_{m,t+1}\end{bmatrix}}_{E_{t+1}} \tag{2.10}$$

3. The Conditional Volatility Dynamics

$$\underbrace{\begin{bmatrix}\sigma^2_{c,t+1}\\ \sigma^2_{\pi,t+1}\end{bmatrix}}_{\Sigma_{t+1}} = \underbrace{\begin{bmatrix}(1-\nu_c)(\varphi_c\bar\sigma)^2\\ (1-\nu_\pi)(\varphi_\pi\bar\sigma)^2\end{bmatrix}}_{\Phi_\mu}+\underbrace{\begin{bmatrix}\nu_c & 0\\ 0 & \nu_\pi\end{bmatrix}}_{\Phi_\nu}\underbrace{\begin{bmatrix}\sigma^2_{c,t}\\ \sigma^2_{\pi,t}\end{bmatrix}}_{\Sigma_t}+\underbrace{\begin{bmatrix}\sigma_{w_c}w_{c,t+1}\\ \sigma_{w_\pi}w_{\pi,t+1}\end{bmatrix}}_{W_{t+1}}, \quad W_{t+1}\sim N(0,\Phi_w). \tag{2.11}$$

In the above, derivations of $\Gamma_x(S^X_{t+1},S^M_{t+1}),\Gamma_\eta(S^X_{t+1},S^M_{t+1})$ are provided in (2.6), $e_1=[1,0,0]$, and the shocks follow $\eta_{j,t+1},e_{k,t+1},w_{l,t+1}\sim N(0,1)$ for $j\in\{c,d,\pi\}$, $k\in\{c,\pi,m\}$, and $l\in\{c,\pi\}$. I approximate the exponential Gaussian volatility process in (3.2) by linear Gaussian processes (2.11) such that the standard analytical solution techniques that have been widely used in the LRR literature can be applied. The approximation of the exponential volatility process is used only to derive the solution coefficients in the law of

motion of the asset prices. $\{S_{t+1}, X_{t+1}, \Sigma_{t+1}\}$ are sufficient statistics for the evolution of the fundamental macroeconomic aggregates.

Real Equity Asset Solutions

Real asset prices are determined from the approximate analytical solution described in Bansal and Zhou (2002) and Schorfheide, Song, and Yaron (2013). Let I_t denote the current information set $\left\{S_t^X, X_t, \Sigma_t\right\}$ and define $I_{t+1}{=}I_t \cup \{S_{t+1}^X\}$ that includes information regarding S_{t+1}^X in addition to I_t.[11] Suppose $S_t^X = i$ for i=1,2. Derivation of (2.7) follows Bansal and Zhou (2002), who make repeated use of the law of iterated expectations and log-linearization, and Schorfheide, Song, and Yaron (2013) who utilize log-linear approximation for returns and for volatilities

$$\begin{aligned}
1 &= \mathbb{E}\Big(\mathbb{E}\left[\exp\left(m_{t+1} + r_{m,t+1}\right) \mid I_{t+1}\right] \mid I_t\Big) \\
&= \sum_{j=1}^{2} \mathbb{P}_{X_{ij}} \mathbb{E}\Big(\exp\left(m_{t+1} + r_{m,t+1}\right) \mid S_{t+1}^X = j, X_t, \Sigma_t\Big) \\
0 &= \sum_{j=1}^{2} \mathbb{P}_{X_{ij}} \underbrace{\left(\mathbb{E}\left[m_{t+1} + r_{m,t+1} \mid S_{t+1}^X = j, X_t, \Sigma_t\right] + \frac{1}{2}\mathbb{V}\left[m_{t+1} + r_{m,t+1} \mid S_{t+1}^X = j, X_t, \Sigma_t\right]\right)}_{B}.
\end{aligned}$$

The first line uses the law of iterated expectations, second line uses the definition of Markov-chain; and the third line applies log-linearization (i.e., $\exp(B) - 1 \approx B$), log-normality assumption, and log-linearization for returns and for volatilities.

The state-contingent solution to the log price to consumption ratio follows

$$z_t(i) = A_0(i) + A_1(i)X_t + A_2(i)\Sigma_t,$$

[11] Note that regime information on S_t^M is irrelevant for real equity asset solutions.

where

$$\begin{aligned}
\left[\begin{array}{cc} A_1(1) & A_1(2) \end{array}\right] &= (1-\frac{1}{\psi})e_1 \left[\begin{array}{cc} p_{X_1}\Upsilon(1)+(1-p_{X_1})\Upsilon(2) & (1-p_{X_2})\Upsilon(1)+p_{X_2}\Upsilon(2) \end{array}\right] \\
&\quad \times \left[\begin{array}{cc} \mathbb{I}_2 - p_{X_1}\kappa_{1,c}\Upsilon(1) & -(1-p_{X_2})\kappa_{1,c}\Upsilon(1) \\ -(1-p_{X_1})\kappa_{1,c}\Upsilon(2) & \mathbb{I}_2 - p_{X_2}\kappa_{1,c}\Upsilon(2) \end{array}\right]^{-1} \\
\left[\begin{array}{c} A_{2,c}(1) \\ A_{2,c}(2) \end{array}\right] &= \frac{\theta}{2}\left[\begin{array}{c} \mathbb{I}_2 - \kappa_{1,c}\nu_c\mathbb{P}_X \end{array}\right]^{-1} \times \mathbb{P}_X \times \left[\begin{array}{c} \left\{\left((1-\frac{1}{\psi})e_1+\kappa_{1,c}A_1(1)\right)\cdot\Omega(1)e_1'\right\}^2 \\ \left\{\left((1-\frac{1}{\psi})e_1+\kappa_{1,c}A_1(2)\right)\cdot\Omega(2)e_1'\right\}^2 \end{array}\right] \\
\left[\begin{array}{c} A_{2,\pi}(1) \\ A_{2,\pi}(2) \end{array}\right] &= \frac{\theta}{2}\left[\begin{array}{c} \mathbb{I}_2 - \kappa_{1,c}\nu_\pi\mathbb{P}_X \end{array}\right]^{-1} \times \mathbb{P}_X \times \left[\begin{array}{c} \left\{\left((1-\frac{1}{\psi})e_1+\kappa_{1,c}A_1(1)\right)\cdot\Omega(1)e_2'\right\}^2 \\ \left\{\left((1-\frac{1}{\psi})e_1+\kappa_{1,c}A_1(2)\right)\cdot\Omega(2)e_2'\right\}^2 \end{array}\right].
\end{aligned}$$

The log price to consumption ratio loading with respect to long-run growth, $A_{1,c}(i)$, will be positive whenever the IES, ψ, is greater than 1. The loadings on the inflation target, $A_{1,\pi}(i)$, and on the monetary policy shock, $A_{1,m}(i)$, are zero. The sign of the responses of the log price to consumption ratio to long-run growth and inflation target volatilities, $A_{2,c}(i)$ and $A_{2,\pi}(i)$, will be negative if $\theta < 0$ (i.e., $\gamma > 1$ and $\psi > 1$).

Nominal Bond Asset Solutions

Similar to the previous case, the approximate analytical expressions for the state-contingent log bond price coefficients $p_{n,t} = C_{n,0}(i) + C_{n,1}(i)X_t + C_{n,2}(i)\Sigma_t$ are derived by exploiting the law of iterated expectations and log-linearization,

$$p_{n,t} \approx \sum_{j=1}^{4}\mathbb{P}_{ij}\log\Big(\mathbb{E}[\exp(m_{t+1}-\pi_{t+1}+p_{n-1,t+1})|S_{t+1}=j,S_t=i]\Big),$$

where

$$\begin{aligned}
C_{n,1}(i) &= \sum_{j=1}^{4}\mathbb{P}_{ij}\Big(C_{n-1,1}(j)-\frac{1}{\psi}e_1-\Gamma_x(j)\Big)\Upsilon(j) \\
C_{n,2}(i) &= \sum_{j=1}^{4}\mathbb{P}_{ij}\Big(C_{n-1,2}(j)\Phi_\nu+(\theta-1)\left\{\kappa_{1,c}A_2(j)\Phi_\nu-A_2(i)\right\} \\
&\quad + \frac{1}{2}\left[\begin{array}{c} \{(C_{n-1,1}(j)-\gamma e_1-\Gamma_x(j)+(\theta-1)\kappa_{1,c}A_1(j))\cdot\Omega(j)e_1'\}^2 \\ \{(C_{n-1,1}(j)-\gamma e_1-\Gamma_x(j)+(\theta-1)\kappa_{1,c}A_1(j))\cdot\Omega(j)e_2'\}^2 \end{array}\right]'\Big)
\end{aligned}$$

with the initial conditions $C_{0,1}(i) = [0, 0, 0]$ and $C_{0,2}(i) = [0, 0]$ for i=1,. . . ,4. Because of the regime-switching feature, the coefficients are not easy to interpret. However, it is relatively easy to verify that bond prices will respond negatively to positive shocks to long-run growth and the inflation target when $n = 1$.

2.3 State-Space Representation of the LRR Model

To facilitate estimation, it is convenient to cast the LRR model of Section 3.2 into state-space form. The state-space representation consists of a measurement equation that relates the observables to underlying state variables and a transition equation that describes the law of motion of the state variables. I use the superscript o to distinguish observed variables from model-implied ones. The regime-contingent measurement equation can be written as

$$y^o_{t+1} = A_{t+1}\Big(D(S_{t+1}) + F(S_{t+1})f_{t+1} + F^v(S_{t+1})f^v_{t+1} + \Sigma^\varepsilon\varepsilon_{t+1}\Big), \quad \varepsilon_{t+1} \sim iidN(0, I) \quad (2.12)$$

The vector of observables, y^o_{t+1}, contains consumption growth, dividend growth, the log price to dividend ratio, inflation, U.S. Treasury bills with maturities of one and three months, U.S. Treasury bonds with maturities of between one and five years, as well as bonds with maturity of ten years, and measures of one quarter ahead forecasts for real growth from the historical forecasts taken from the Survey of Professional Forecasters (SPF). The vector f_{t+1} stacks state variables that characterize the level of fundamentals. The vector f^v_{t+1} is a function of the log volatilities of long-run growth and the inflation target, h_t and h_{t+1}, in (3.2). Finally, ε_{t+1} is a vector of measurement errors, and A_{t+1} is a selection matrix that accounts for deterministic changes in the data availability.

The solution of the LRR model sketched in Section 2.2.5 provides the link between the state variables and the observables y^o_{t+1}. The state variables themselves follow regime-contingent vector autoregressive processes of the form

$$f_{t+1} = \Phi(S_{t+1})f_t + v_{t+1}(S_{t+1})(h_t), \quad h_{t+1} = \Psi h_t + \Sigma_h w_{t+1}, \quad w_{t+1} \sim iidN(0, I), \quad (2.13)$$

where $v_{t+1}(S_{t+1})$ is an innovation process with a variance that is a function of the log volatility process h_t, and w_{t+1} is the innovation of the stochastic volatility process. Roughly speaking, the vector f_{t+1} consists of the long-run components $x_{c,t}$, $x_{\pi,t}$, and $x_{m,t}$ in Section 3.2. In order to express the observables y^o_{t+1} as a linear function of f_{t+1} and to account for potentially missing observations it is necessary to augment f_{t+1} by lags of $x_{c,t}, x_{\pi,t}, x_{m,t}$ as well as the innovations for the fundamentals. A precise definition of f_{t+1} is included to the Appendix.

The novelty in the estimation is that the state-space representation is set up in a way to incorporate the measurement error modeling of consumption growth outlined in Schorfheide, Song, and Yaron (2013). The authors show that post-1959 monthly consumption series are subject to sizeable measurement errors and argue that accounting for measurement errors is crucial in identifying the predictable component in consumption growth. In addition, the state-space representation exploits the SPF measures that are released in a different (quarterly) frequency. As argued in Bansal and Shaliastovich (2013), survey-based expected measures provide the most accurate forecasts of future growth, which is why bringing this information into the estimation will sharpen the inference on expected terms. For purpose of illustration, I represent the monthly time subscript t as $t = 3(j-1)+m$, where $m = 1,2,3$. Here j indexes the quarter and m the month within the quarter. The formulae below summarize the implementation of measurement error modeling of consumption and exploitation of the SPF measures:

1. A Measurement Equation for Consumption

$$\begin{aligned} g^o_{c,3(j-1)+1} &= g_{c,3(j-1)+1} + \sigma_\epsilon\big(\epsilon_{3(j-1)+1} - \epsilon_{3(j-2)+3}\big) - \frac{1}{3}\sum_{m=1}^{3}\sigma_\epsilon\big(\epsilon_{3(j-1)+m} - \epsilon_{3(j-2)+m}\big) \\ &\quad + \sigma^q_\epsilon\big(\epsilon^q_{(j)} - \epsilon^q_{(j-1)}\big) \\ g^o_{c,3(j-1)+m} &= g_{c,3(j-1)+m} + \sigma_\epsilon\big(\epsilon_{3(j-1)+m} - \epsilon_{3(j-1)+m-1}\big), \quad m = 2,3, \end{aligned}$$

where the monthly and quarterly measurement errors follow $\epsilon_{3(j-1)+m}, \epsilon^q_{(j)} \sim N(0,1)$.

2. Exploiting the SPF Measures

$$x^{q,o}_{c,(j)} = \sum_{\tau=1}^{5} \left(\frac{3 - |\tau - 3|}{3} \right) x_{c,3j-\tau+1} + \sigma^q_{x,\epsilon} \epsilon^q_{x,(j)},$$

where $x^{q,o}_{c,(j)}$ denotes the j^{th} quarter median SPF forecasts for real growth measured at $j-1^{th}$ quarter, and the measurement error follows $\epsilon^q_{x,(j)} \sim N(0,1)$.

2.3.1 Bayesian Inference

The system to be estimated consists of equations (2.12) and (2.13) whose coefficient matrices are functions of the parameter vector

$$\begin{aligned} \Theta_0 &= \left(\delta, \psi, \gamma\right) \\ \Theta_1 &= \left(\{\varphi_k, \bar{\sigma}_k, \mu_k, \nu_k, \sigma_{w_k}\}_{k=c}^{\pi}, \mu_d, \varphi_d, \phi_x, \phi_\eta, \sigma_\epsilon, \rho_m, \sigma_m, \left\{\rho_c^{(i)}, \rho_\pi^{(i)}, \chi_{c,\pi}^{(i)}\right\}_{i=1}^{2}, \left\{\tau_c^{(j)}, \tau_\pi^{(j)}\right\}_{j=1}^{2} \right) \\ \Theta_2 &= \left(\mathbb{P}_{X_1}, \mathbb{P}_{X_2}, \mathbb{P}_{M_1}, \mathbb{P}_{M_2}\right). \end{aligned} \tag{2.14}$$

I will use a Bayesian approach to make inferences about $\Theta = \{\Theta_0, \Theta_1, \Theta_2\}$ and the latent state vector S and study the implications of the model. Bayesian inference requires the specification of prior distributions $p(\Theta)$ and $p(S|\Theta_2)$ and the evaluation of the likelihood function $p(Y^o|\Theta, S)$.

The posterior can be expressed as

$$p(\Theta, S|Y^o) = \frac{p(Y^o|\Theta, S)p(S|\Theta_2)p(\Theta)}{p(Y^o)}, \tag{2.15}$$

which can be factorized as

$$p(\Theta, S|Y^o) = p(\Theta|Y^o)p(S|\Theta, Y^o). \tag{2.16}$$

The practical difficulty is to generate draws from $p(\Theta|Y^o)$ since it requires numerical evaluation of the prior density and the likelihood function $p(Y^o|\Theta)$. Due to the presence of the volatility states and the regime-switching processes, the computation of the likelihood function relies on a sequential Monte Carlo procedure also known as particle filter. To obtain

a computationally efficient filter, I extend the algorithm developed in Schorfheide, Song, and Yaron (2013), in which they exploit the partially linear structure of the state-space model conditional on the volatility states and derive a very efficient particle filter. The key feature of my state-space model is that it is still nonliner conditional on the volatility states. However, conditional on the volatility states, I can apply Kim's Filter in Kim and Nelson (1999) (i.e., an extension of the Kalman filter with a collapsing procedure that is proposed for handling Markov-switching models) to evaluate the likelihood. In essence, I use a swarm of particles to represent the distribution of volatilities and employ Kim's Filter for each particle (i.e., volatility). After resampling step (i.e., eliminating particles with low weights), the filter produces a sequence of likelihood approximations. I embed the likelihood approximation in a fairly standard random-walk Metropolis algorithm and draw the parameter vector $\{\Theta^{(m)}\}_{m=1}^{n_{sim}}$. Conditional on the parameter vector, $\{\Theta^{(m)}\}_{m=1}^{n_{sim}}$, I use Kim's smoothing algorithm in Kim and Nelson (1999) to generate draws from the history of latent states, $\{S^{(m)}\}_{m=1}^{n_{sim}}$. A full description of the particle filter is provided in the Appendix.

2.4 Empirical Results

The data set used in the empirical analysis is described in Section 3.4.1.

2.4.1 Data

Monthly consumption data represent per-capita series of real consumption expenditure on non-durables and services from the National Income and Product Accounts (NIPA) tables available from the Bureau of Economic Analysis. Aggregate stock market data consist of monthly observations of returns, dividends, and prices of the CRSP value-weighted portfolio of all stocks traded on the NYSE, AMEX, and NASDAQ. Price and dividend series are constructed on the per-share basis as in Campbell and Shiller (1988b) and Hodrick (1992). Market data are converted to real using the consumer price index (CPI) from the Bureau of

Labor Statistics. Growth rates of consumption and dividends are constructed by taking the first difference of the corresponding log series. Inflation represents the log difference of the CPI. Monthly observations of U.S. Treasury bills and bonds with maturities at one month, three months, one to five years, and ten years are from CRSP. The time series span of the monthly data is from 1959:M1 to 2011:M12.[12] The quarterly SPF survey forecasts are from the Federal Reserve Bank of Philadelphia. I use the median survey forecasts values for GDP growth that span the period from 1968:Q4 to 2011:Q4. The descriptive data statistics are provided in Table 2.2.

2.4.2 Prior and Posterior Summaries

I begin by estimating the state-space model described in Section 3.3.

Prior Distribution. This section provides a brief discussion of the prior distribution. Percentiles for marginal prior distributions are reported in Table 2.3. The prior distribution for the preference parameters which affect the asset pricing implications of the model are the same as the ones used in Schorfheide, Song, and Yaron (2013). Thus, I focus on the parameters of the fundamental processes specified in (2.1) and (3.2).

The prior 90% credible intervals for average annualized consumption and dividend growth and inflation are fairly wide and agnostic and range from approximately -7% to +7%. The priors for ϕ_x and ϕ_η, parameters that determine the comovement of consumption and dividend growth, are centered at zero and have large variances. $\bar{\sigma}_c$ and $\bar{\sigma}_\pi$ are the average standard deviation of the *iid* component of consumption growth and inflation whose 90% prior intervals range from 1.2% to 7.2% at an annualized rate. The parameters φ_d, φ_c, and φ_π capture the magnitude of innovations to dividend growth and the long-run growth and inflation target component relative to the magnitude of consumption growth

[12]Monthly consumption growth is available from 1959:M2.

innovations. The prior for φ_d covers the interval 0.2 to 12, whereas the priors for φ_c, and φ_π cover the interval 0 to 0.11. Finally, the prior interval for the persistence of the volatility processes ranges from -0.1 to 0.97 and the prior for the standard deviation of the volatility process implies that the volatility may fluctuate either relatively little, within the range of 0.67 to 1.5 times the average volatility, or substantially, within the range of 0.1 to 7 times the average volatility.

The prior distribution for the persistence of the long-run growth, inflation target, and monetary policy shock $x_{c,t}$, $x_{\pi,t}$, $x_{m,t}$ is a normal distribution centered at 0.9 with a standard deviation of 0.5, truncated to the interval $(-1, 1)$. The corresponding 90% credible interval ranges from -0.1 to 0.97, encompassing values that imply *iid* dynamics as well as very persistent local levels. The prior distribution for the parameter that captures contemporaneous correlation between the long-run growth and inflation target shocks is a normal distribution centered at zero with a relatively large standard deviation of 0.5. Sign restrictions are imposed to identify two different correlation regimes: one is truncated below zero, and the other is truncated above zero. The prior intervals for the standard deviation of the monetary policy shock cover the range from 0 to 0.001.

The priors for the monetary policy rule coefficients are normal distributions with range of between ± 4.28, but those for inflation components are truncated above zero, reflecting the view that the central bank *raises* rather than *lowers* the interest rate in response to positive inflation fluctuations. Finally, I employ beta priors for the Markov-chain transition probabilities that cover 0.38 to 1.00.

Posterior Distribution. Percentiles for the posterior distribution are also reported in Table 2.3. The estimated parameters for preferences and dividend growth (first panel) are, by and large, similar to those reported in Schorfheide, Song, and Yaron (2013). Those for the consumption and inflation process (second panel) are consistent with the sample

mean and standard deviation reported in Table 2.2. One interesting feature is that the unconditional standard deviation of the long-run growth is substantially smaller than that of the inflation target, 0.07% versus 0.29% at annualized rates. The estimation results also provide strong evidence for stochastic variation in the long-run growth and inflation target. According to the posteriors reported in Table 2.3, all $\sigma_{c,t}$ and $\sigma_{\pi,t}$ exhibit significant time variation. The posterior medians of ν_c and ν_π are .9952 and .9928, respectively, and the unconditional volatility standard deviations σ_{w_c} and σ_{w_π} are around 0.31 and 0.45.

The most important results for the subsequent analysis are provided in the third and fourth panels of Table 2.3. First, there is strong evidence for parameter instability in the VAR dynamics of the long-run components. Most prominently, the posterior median estimate of $\chi_{c,\pi}$, which captures contemporaneous correlation between the long-run growth and inflation target shocks, is -0.40 in the first regime and 0.15 in the second regime. Another notable difference between the two regimes is the drop in the persistence of the long-run growth and inflation target components. Unlike in their appearance, the process half-life is very different between two regimes: the process half-life for the long-run growth (inflation target) component in the first regime is about 12 (12) years; while that in the second regime is about 1 (3) years. The values of persistence and the standard deviation of the monetary policy shock are 0.9916 and 0.0002, and are assumed to be identical across regimes. In general, the magnitude of the differentials between the two VAR coefficient regimes are small, but the sign change in the correlation structure is notable. Since the group of estimates distinguish themselves as ones that generate negative correlation between long-run growth and inflation target shocks and ones that do not, I label the first regime as the "countercyclical" inflation regime and the second regime as the "procyclical" inflation regime.

Second, two very different posterior estimates of the monetary policy rule in the fourth

panel of Table 2.3 support the view of Clarida, Gali, and Gertler (2000) that there has been a substantial change in the way monetary policy is conducted. One regime is associated with larger monetary policy rule coefficients, which implies that the central bank will respond more aggressively to consumption gap, short-run, and long-run inflation fluctuations. The other regime is characterized by a less responsive monetary policy rule, in which I find much lower loadings on consumption gap and short-run inflation fluctuations. In particular, the magnitude of the loading on short-run inflation fluctuation τ_π is one-third of that in the former regime and is below one. Following the convention in the monetary policy literature, the regimes are distinguished by which has an "active" central bank, and which has a "passive" central bank.

Finally, the bottom panel of Table 2.3 reports posterior estimates of the Markov-chain transition probabilities. The countercyclical inflation regime is most persistent: The probability that it will continue is 99.2%. The procyclical inflation regime, on the contrary, is the less persistent one: Its duration is one-fourteenth of the countercyclical inflation regime. This result can be interpreted as the "risks" of falling back to the countercyclical inflation regime are substantial. The transition probability of the active monetary policy regime is around 0.99, which implies that agents expect its average duration to be about 9 years. For the passive monetary policy regime, the same result is about 3-4 years. Given posterior transition probabilities, it is interesting to look at the smoothed probabilities for transitions between regimes.

Smoothed Posterior Regime Probabilities. Figure 2.1 depicts the smoothed posterior probabilities of the procyclical inflation and active monetary policy regimes. Figure 2.1(a) is consistent with the evidence provided in Table 2.1 that procyclical inflation regimes were prevalent after late-1990s. It also suggests that the switch is not a permanent event, but

rather, an occasional one.[13] Figure 2.1(b) provides the historical paths of monetary policy stance: The active monetary policy appeared in the mid-1960s but was largely dormant during the 1970s; it became active after the appointment of Paul Volcker as Chairman of the Federal Reserve in 1979 and remained active for 20 years (except for short periods in the early 1990s); after that, in response to the economic crisis triggered by the 9/11 attacks in 2001, the central bank lowered interest rates and took a passive stance for 3-4 years; around the mid-2000s, it switched back to a more active stance until the Great Recession started; and finally, post-2008 periods are characterized by the passive regime.[14]

Smoothed Mean and Volatility States. The top panel of Figure 2.2 depicts smoothed estimates of long-run growth $x_{c,t}$ and inflation target $x_{\pi,t}$, which are overlaid with monthly consumption growth and inflation, respectively.[15] $x_{c,t}$ tends to fall in recessions (indicated by the shaded bars in Figure 2.2) but periods of falling $x_{c,t}$ also occur during expansions; the pattern is broadly similar to the one reported in Schorfheide, Song, and Yaron (2013). $x_{\pi,t}$ reaches its peak during the Great Inflation periods and substantially decreases afterwards. It is interesting to note that during the 1970s and 1980s, recessions were accompanied by increases in the inflation target. The pattern clearly reverses starting in the late 1990s. The smoothed volatility processes are plotted below. Recall that my model has two independent volatility processes, $h_{c,t}$ and $h_{\pi,t}$, which are associated with the innovations to the long-run growth and inflation target, respectively. The most notable feature of $h_{c,t}$ is that it captures a drop in growth volatility that occurred in the 1980s, also known as the Great Moderation. The stochastic volatility process for the inflation target displays different properties: It

[13]This evidence is also supported by David and Veronesi (2013).

[14]The smoothed paths for the monetary policy are broadly consistent with those found in Clarida, Gali, and Gertler (2000), Ang, Boivin, Dong, and Loo-Kung (2011), Bikbov and Chernov (2013), and Coibon and Gorodnichenko (2011).

[15]Figure 2.10 provides the path of the estimated monetary policy shock.

jumps around 1970 and remains high for 25 years, and features wide fluctuations in the beginning of the 2000s, that is not apparent in $h_{c,t}$. Overall, the smoothed $h_{\pi,t}$ seems to exhibit more medium and high-frequency movements than $h_{c,t}$. Also, due to the inclusion of a greater amount of bond yields data, $h_{\pi,t}$ is more precisely estimated than $h_{c,t}$, indicated by tighter credible intervals.

2.4.3 Implications for Macroeconomic Aggregates and Asset Prices

It is instructive to examine the extent to which sample moments implied by the estimated state-space model mimic the sample moments computed from the actual data set. To do this, I conduct a posterior predictive check (see Geweke (2005) for a textbook treatment). I use previously generated draws $\Theta^{(s)}, S^{(s)}$, $s = 1, \ldots, n_{sim}$, from the posterior distribution of the model parameters $p(\Theta, S|Y^o)$ and simulate for each $\Theta^{(s)}, S^{(s)}$ the model for 636 periods, which corresponds to the number of monthly observations in the estimation sample.[16] This leads to n_{sim} simulated trajectories, which I denote by $Y^{(s,o)}$. For each of these trajectories, I compute various sample moments, such as means, standard deviations, and cross correlations. Suppose I denote such statistics generically by $\mathcal{S}(Y^{(s,o)})$. The simulations provide a characterization of the posterior predictive distribution $p(\mathcal{S}(Y^{(s,o)})|Y^o)$.

Matching Moments of the Macroeconomic Aggregates and Stock Price. To save space, the model-implied distributions for the first and second moments of the macroeconomic aggregates and stock price are provided in Table 2.12 and Table 2.13 in the Appendix. In sum, the first and second moments for consumption and dividend growth, log price to dividend ratio, and inflation implied by the model replicate the actual counterparts well. Since monetary policy does not affect the cash flows, the sample moments for consumption and dividend growth and log price to dividend ratio do not differ across monetary

[16] To generate the simulated data, I also draw measurement errors.

policy regimes (i.e., column-wise comparisons). Yet the sample moments across inflation regimes (i.e., row-wise comparisons) are quite different: Those in the countercyclical inflation regime are much more volatile. This finding is consistent with the near unit-root estimates of long-run growth and inflation target persistence in the countercyclical inflation regime (see Table 2.3). The sample correlation between consumption and inflation is provided in Table 2.4. While the model-implied numbers are somewhat larger than their data estimates, the model performs well in terms of matching the sign-switching patterns. One notable feature is that monetary policy does seem to matter for the correlation of expected values: Passive monetary policy lowers the correlation of expected values particularly more during the procyclical inflation regime. Overall, I find that $\chi_{c,\pi}$ is the key model ingredient to capturing the sign-switching patterns, and that monetary policy influences the correlation of the expected consumption growth and inflation but on its own cannot change the sign.

Equilibrium Nominal Bond Yield Loadings. It is also instructive to understand the equilibrium bond yield loadings first before looking at the model-implied yield curve. Figure 2.3 shows the regime-contingent bond yield loadings on long-run growth, inflation target, and long-run growth and inflation target volatilities based on the median posterior coefficient estimates.[17] To ease exposition, I use abbreviations for each regime: "CA" stands for the countercyclical inflation and the active monetary policy regimes, while "PP" stands for the procyclical inflation and the passive monetary policy regimes; "CP" and "PA" indicate the remaining combinations of regimes. The CP loading on inflation target for a bond with a maturity of 1 month is normalized to 100% to bring all of the loadings into proportion with one another.[18] It is evident from Figure 2.3 that inflation target is

[17] I do not present the graph for monetary policy since its influence on bond yields is very small compared to these variables.

[18] An easier way to interpret this is to fix one regime and compare loadings across the model state variables.

the most important factor in the term structure analysis. Note that loadings on inflation target volatility increase over maturities and become the second most important factor for longer maturity yields. In terms of patterns of the loadings, I find that they are broadly in line with those found in Bansal and Shaliastovich (2013). The loadings on long-run growth and inflation targets are positive; the loading on long-run growth volatility has a negative decreasing slope; and the loading on inflation target volatility is mostly positive and rises with maturities. However, the loadings across regimes have very different implications. Let us focus on monetary policy regimes. For example, while a positive shock to the inflation target induces an essentially parallel shift in the entire yield curve (loadings are nearly flat across maturities) in the active monetary policy regime, it has disproportionately larger effects on yields with short maturities (loadings decrease substantially over maturities) in the passive case. It seems that in the active monetary policy regime, inflation target behaves like a level factor, but in the passive cases it becomes a slope factor.[19] Moreover, the magnitude of the loadings in the passive monetary policy stance almost doubles. With regard to inflation regimes, the loadings on all model state variables will be uniformly shifted out in the countercyclical inflation regime, implying that the risks associated with the countercyclical inflation regime are much larger than those in the procyclical case.

Matching Moments of the Yield Spread. The estimated model is quite successful at fitting Treasury yields over the entire sample—the yield prediction error in different maturity are generally quite small over the entire sample. To save space, the evidence is provided in Figure 2.14 in the Appendix. Now, in order to evaluate if the model can reproduce key patterns in the data, I focus on posterior predictive assessment in the main text. Distributions generated from the LRR model using the posterior estimates are graphically provided

By focusing on one state variable, you can move across regimes to compare their magnitudes.

[19]Readers are referred to Figure 1 in Rudebusch and Wu (2008).

in Figure 2.4. The top and bottom ends of the boxes correspond to the 5th and 95th percentiles, respectively, of the posterior distribution, and the horizontal lines signify the medians. The first row of Figure 2.4 is simulated conditional on the countercyclical inflation regime while the second row in Figure 2.4 is generated from the procyclical inflation one. For each row, the figure on the left conditions on the active monetary policy regime while the one on the right does the same on the passive monetary policy regime. The figure also depicts the same moments computed from U.S. data (black squares). "Actual" sample moments that fall far into the tails of the posterior predictive distribution provide evidence for model deficiencies. Roughly speaking, the model performs well along this dimension since the model-implied median values are fairly close to their data estimates. Yet important distinctions arise across regimes. Going from left to right (CA to CP or PA to PP), I find that yield spread distributions are more dispersed. The 90% credible intervals in the latter, right-hand figures (CP or PP) are approximately twice as large as those in the left-hand column (CA or PA). This is consistent with the impulse response functions shown in Figure 2.11, in that the passive monetary policy leads to more unstable economic dynamics. From top to bottom (CA to PA or CP to PP), I find that the 10y-3m yield spreads in the countercyclical inflation regime are roughly 150 basis points (annualized) higher than those in the procyclical inflation regime. This implies that agents will demand higher yields as compensation for the risks associated with the countercyclical inflation regimes. An interesting feature of the model is that due to the presence of the countercyclical inflation regimes, agents will still demand inflation premiums, which is shown by the upward slope found in PP of Figure 2.4. This is a prominent feature of the model that generates an upward-sloping yield curve even when the economy is in the procyclical inflation regime. The second moment for the yield spread implied by the model is provided in Figure 2.15 in the Appendix. The model performs well along this dimension and the model-implied

patterns are very similar to the first moment case.

Bond Risk Premia Implications. Under the Expectations Hypothesis (EH), the expected holding returns from long-term and short-term bonds should be the same (strong form) or should only differ by a constant (weak form). However, even the weak form has been consistently rejected by empirical researchers. For example, Campbell and Shiller (1991), Dai and Singleton (2002), Cochrane and Piazzesi (2005), and Bansal and Shaliastovich (2013) all argue that the EH neglects the risks inherent in bonds, and provide strong empirical evidence for predictable changes in future excess returns.

The presence of stochastic volatilities and regime-switching loadings in my model gives rise to time-variations in risk premia which has testable implications for the EH.[20] First, I focus on the term spread regression of Campbell and Shiller (1991) to examine the validity of the EH. The excess log return on buying an n month bond at t and selling it as an $n-12$ month bond at $t+12$ is denoted by

$$rx_{t+12,n} = (n)y_{t,n} - (n-12)y_{t+12,n-12} - 12y_{t+12}.$$

Under the weak form of the EH, the expected excess bond returns are constant, which implies that the theoretical slope coefficient β_n value (below) predicted by the EH is equal to unity for all n

$$y_{t+12,n-12} - y_{t,n} \quad = \quad \alpha_n + \beta_n\left(\left(y_{t,n} - y_{t,12}\right)\frac{12}{n-12}\right) + \epsilon_{t+12}. \tag{2.17}$$

Bansal and Shaliastovich (2013)[21] show that the population value for β_n can be expressed by

$$\beta_n = 1 - \frac{cov(\mathbb{E}_t rx_{t+12,n}, y_{t,n} - y_{t,12})}{var(y_{t,n} - y_{t,12})}. \tag{2.18}$$

[20] My model extends Bansal and Shaliastovich (2013) by allowing regime-switching bond yield loadings which provide additional channels for time variations in risk premia.

[21] The earlier version of their paper considered this explanation.

This means that downward deviation from unity, equivalent to $cov(\mathbb{E}_t rx_{t+12,n}, y_{t,n} - y_{t,12}) > 0$, implies that the term spread contains information about the expected excess bond returns. Put differently, the predictability of excess bond returns (by the term spread) reflects time variations in the expected risk premium.

Figure 2.5 compares model-implied distributions for the slope coefficient, β_n, to the corresponding data estimates. The first thing to note is that the model generates very comparable results. Roughly speaking, the model produces β_ns that are significantly lower than unity and whose absolute magnitudes rise over maturities, as in the data. Second, it is important to understand that the violations of the EH or deviations from unity are less apparent in the passive monetary policy regimes. In particular, the model-implied distributions for β_ns in the PP regime are close to or even greater than zero. The striking feature is that the data estimates for β_n in the PP regime are all greater than zero and even close to unity for maturities of two and three years. It can be deduced from (2.18) that either the term spread contains much less information about the expected excess bond returns, or the variance of the term spread is much larger in the passive monetary policy regime.

In order to understand this feature, I decompose the bond yields into the component implied by the EH, the expected sum of future short rates, and the term premium,

$$y_{t,n} = \underbrace{\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_t(y_{t+i,1})}_{\text{short-rate expectations}} + \text{term premium}_{t,n}. \tag{2.19}$$

Let us focus on the monetary policy regimes and assume that we are in the countercyclical inflation regime. Here are two possible channels through which the passive monetary policy stance can affect the bond yields. In order to generate results that are consistent with Figure 2.4, we would expect to see an increase either in the expected sum of future short rates or in the term premium.

Figure 2.6 compares the model-implied distributions for the term premium to the corresponding data estimates (black squares). Data estimates are within-regime averages from Figure 2.16 where the time-series of the estimated term premia for bonds with maturities of 1–10 years are depicted. It is very interesting to observe that the term premia in the passive monetary policy regime are actually smaller than those in the active regimes (both in the data and model-implied estimates). This implies that the effect of monetary policy is mostly on the expectations component (without affecting the term premium component), which further implies an increase in the variance of the current period's term spread. From (2.18), an increase of the term spread variance will bring the slope coefficient, β_n, closer to 1. The underlying economic intuition is that the future yields will incorporate the expected increase in the future inflation rates as the passive monetary policy stance is more prone to large inflation, which is predicted by the EH. While the estimated model is successful in generating these patterns, it falls short of data estimates found in the CA regime. The model is not able to capture the substantive increase in term premia as in the data.

Similar logic can be applied when the inflation regime is procyclical. The directional influence of the passive monetary policy stance on the expectations component is ambiguous because, on the one hand, the procyclicality will lower the expected inflation, but on the other hand, the risks of falling back to the countercyclical inflation regime will increase the expected inflation. However, the inherent instability associated with the passive monetary policy stance will increase the relative weight on the expectations component, which brings the bond market closer to what the EH predicts.

In contrast to monetary policy, the countercyclical inflation regime affects both terms. It is clear from the row-to-row comparison of Figure 2.6 that the risks associated with the countercyclical inflation regime increase the term premiums, which are on average 50 basis

points higher for 10-year bonds.[22]

A final exercise consists of running regressions that predict excess bond returns. Following Cochrane and Piazzesi (2005), I focus on regressing the excess bond return of an n year bond over the 1 year bond on a linear combination of forward rates that includes a constant term, a one year bond yield, and four forwards rates with maturities of 2 to 5 years. The model-implied 90% distributions for R^2 values (in percents) from the regression are provided in Figure 2.7. Consistent with previous findings, the expected excess returns are less preditable (indicated by about 5% lower R^2 values) in the passive monetary policy stance.[23] This is due to relative decrease in the role played by the risks channel (term premium) in the passive monetary policy regime. Also, I find that the procyclical inflation regimes (PA and PP) deliver, on average, 5–10% lower R^2 values (see the bottom panel in Table 2.1).

Determinants of Asset Price Fluctuations. Table 2.5 provides the contribution of various risk factors, namely the variation in long-run growth, inflation target, monetary policy shock, and the conditional volatility variations of long-run growth and inflation target to asset price volatility. Given the posterior estimates of the state-space model I can compute smoothed estimates of the latent asset price volatilities. Moreover, I can also generate counterfactual volatilities by sequentially shutting down each risk factor. The ratio of the counterfactual and the actual volatilities measures the contribution of the non-omitted risk factors. If I subtract this ratio from one, I obtain the relative contribution of the omitted risk factor, which is shown in Table 2.5. I find that the key risk drivers of stock price variations are long-run growth, long-run growth volatility, and inflation target volatility. Since

[22]Note that the differences are modest because the term premia are generated from the unconditional distributions. Once I condition on different levels of volatilities (the relative magnitude of the conditional heteroscedasticity present is larger in the countercyclical inflation regime), the results will change.

[23]Again, the differences are modest since they are generated from the unconditional distributions.

the shock to the inflation target moves long-run growth (captured by $\chi_{c,\pi}$), it becomes one of the major drivers of stock price variations. Bond yield variations are mostly driven by variations in the inflation target and in its volatility. Going from the short-end to the long-end of the yield curve, the importance of the inflation target volatility increases. My findings demonstrate that the long-term rates are much more sensitive to inflation target volatility fluctuations than the short-term rates. My model also shows that the variations in the short-term rates are not driven by fluctuations in volatilities. Hence, the assumption that the short-rate contains no risk premium seems very plausible (see the Fisher-type asset-pricing equation in Section 2.2.3).

Understanding Stock-Bond Returns Comovement. An important feature of my estimation is that the likelihood also focuses on conditional correlation between stock market returns and bond returns. Figure 2.8 displays the time-series of the estimated stock-bond correlation which is overlaid with monthly realized stock-bond correlation (dashed-line). During the Great Inflation periods (1970s–1980s), returns on both assets were low, which resulted in positive comovements. The striking feature here is that in the beginning and towards the end of the estimation sample, the return performances decoupled, and stock and bond returns started to move in opposite directions. Through the estimation, I have identified that the economy faced changes in the covariance between the inflation target and long-run growth shocks (i.e., transition from the countercyclical inflation regime to the procyclical inflation regime). Hence, from an agent's perspective, positive shocks to the inflation target component are perceived as positive signals to the long-run growth. Thus, stock returns, unlike bond returns, can respond positively to long-run inflation shocks.[24] The regime-switching covariance coefficient in the model, $\chi_{c,\pi}$, is able to capture this data feature. Figure 2.9 displays the unconditional stock-bond correlation implied by the model.

[24]David and Veronesi (2013) support this evidence.

This experiment is useful because it disentangles the role of monetary policy in stock-bond return correlation. I find that the active monetary policy stance tends to generate stronger positive stock-bond comovement, although the effect is small. My results are consistent with the findings in Campbell, Pflueger, and Viceira (2013) in which they argue that a more aggressive response of the central bank to inflation fluctuations will increase stock-bond correlation. However, I find that changes in monetary policy stance alone cannot generate a sign-switch in stock-bond return correlation.[25]

2.5 Conclusion

Building on Bansal and Yaron (2004), I developed an equilibrium term structure model incorporating monetary policy to address the issue of whether the structural changes in the U.S. Treasury yield curve are caused by changes in external shocks or in monetary policy. The model framework is general enough to encompass both Markov-switching coefficients and stochastic volatility processes. To estimate the model, I conditioned on the volatilities states to achieve an efficient implementation of a particle Markov Chain Monte Carlo algorithm and made inferences about the model parameters, volatility states, and Markov states. Through the estimation, I characterized bond market exposures to macroeconomic and monetary policy risks, and identified the changes in the conditional covariance dynamics of long-run growth and the inflation target as the main driver of structural changes in bond markets. I found that the changes in monetary policy affect the volatility of bond yields, while the changes in the correlation between growth and inflation affect both the level as well as the volatility of bond yields. Overall, the model is quite successful in explaining several bond market phenomena.

---

[25] Campbell, Pflueger, and Viceira (2013) find similar results. However, they claim that changes in the persistence of monetary policy can generate sign-switches. Since I do not incorporate the "smoothing" motive in the monetary policy action, my results show a limited role for monetary policy.

2.6 Tables and Figures

Table 2.1: Descriptive Statistics

	Pre-1998	Post-1998	Full Sample
	Annualized Average Bond Yields		
Mean (y_{3m})	6.07	2.64	5.16
Mean (y_{1y})	6.51	2.88	5.55
Mean (y_{3y})	6.87	3.35	5.94
Mean (y_{5y})	7.05	3.78	6.19
Mean (y_{10y})	7.35	4.38	6.57
	Correlation between Growth and Inflation		
Corr($\Delta c, \pi$)	-0.19	0.02	-0.11
Corr($\Delta c, \pi$)^Q	-0.36	0.18	-0.16
Corr(Δgdp, π)^Q	-0.26	0.33	-0.13
Corr($\mathbb{E}\Delta$gdp, $\mathbb{E}\pi$)^Q	-0.43	0.19	-0.31
	Correlation between Stock and Bond Returns		
Corr(r_m, r_{2y})	0.16	-0.13	0.09
Corr(r_m, r_{3y})	0.21	-0.14	0.13
Corr(r_m, r_{4y})	0.22	-0.14	0.14
Corr(r_m, r_{5y})	0.24	-0.14	0.15
	Term Spread Regression, Slope Coefficient		
$r_{2y,t+1y}$ onto spread$_{2y,t}$	-0.95	0.89	-0.62
$r_{3y,t+1y}$ onto spread$_{3y,t}$	-1.37	0.43	-1.00
$r_{4y,t+1y}$ onto spread$_{4y,t}$	-1.77	0.02	-1.40
$r_{5y,t+1y}$ onto spread$_{5y,t}$	-1.69	-0.28	-1.41
	Excess Bond Return Predictability, R^2		
$rx_{2y,t+1y}$ onto forward$_t$	34.34	13.60	20.68
$rx_{3y,t+1y}$ onto forward$_t$	35.29	13.92	21.54
$rx_{4y,t+1y}$ onto forward$_t$	37.72	15.79	24.38
$rx_{5y,t+1y}$ onto forward$_t$	34.49	19.15	22.32

Notes: The top three panels report descriptive statistics for aggregate consumption growth (Δc), gross domestic product (GDP) growth (Δgdp), expected GDP growth ($\mathbb{E}\Delta$gdp), consumer price index (CPI) inflation (π), expected inflation ($\mathbb{E}\pi$), log returns of the aggregate stock market (r_m), the log bond yields (y_n), log bond returns (r_n), and log bond excess returns (rx_n) where $n \in \{3m, 1y, 2y, 3y, 4y, 5y, 10y\}$. It shows mean (Mean) and pairwise correlation (Corr) between growth and inflation and market and bond returns. Measures of expected GDP growth ($\mathbb{E}\Delta$gdp) and expected inflation ($\mathbb{E}\pi$) are based on the Survey of Professional Forecasters historical forecasts, which are available from 1968 to 2011. The remaining variables are available from 1959 to 2011. The numbers in the table are derived from monthly frequency data except for those with the superscript "Q"; those numbers are derived from quarterly frequency data. The fourth panel provides slope coefficient from the term spread regression of Campbell and Shiller (1991). The "spread$_{n,t}$" is the difference between an n year yield and a 1 year yield. I focus on a one year return horizon. r_n (rx_n) denotes return (excess return) on an n year bond. The last panel provides R^2 values (in percent) from the excess bond return predictability regression found in Cochrane and Piazzesi (2005). "forward$_t$" includes a constant term, a one year bond yield, and four forwards rates.

Table 2.2: Descriptive Statistics - Data Moments

(a) Quarterly Frequency: 1968:Q4–2011:Q4

	Δc	Δgdp	$\mathbb{E}\Delta$gdp	π
Mean	0.43	0.68	0.58	1.08
StdDev	0.44	0.86	0.58	0.80
AC1	0.54	0.33	0.71	0.74

(b) Monthly Frequency: 1959:M1–2011:M12

	Δc	Δd	π	r_m	pd	y_{3m}	y_{1y}	y_{2y}	y_{3y}	y_{4y}	y_{5y}	y_{10y}
Mean	0.16	0.11	0.32	0.43	3.57	0.43	0.46	0.48	0.50	0.51	0.52	0.55
StdDev	0.34	1.26	0.32	4.55	0.39	0.25	0.25	0.24	0.24	0.23	0.22	0.22
AC1	-0.16	-0.01	0.63	0.10	0.99	0.98	0.98	0.99	0.99	0.99	0.99	0.99

Notes: I report descriptive statistics for aggregate consumption growth (Δc), gross domestic product (GDP) growth (Δgdp), expected GDP growth ($\mathbb{E}\Delta$gdp), consumer price index (CPI) inflation (π), dividend growth (Δd), log returns of the aggregate stock market (r_m), log price to dividend ratio (pd), and U.S. Treasury yields (y_n) with maturity $n \in \{3m, 1y, 2y, 3y, 4y, 5y, 10y\}$. The table shows mean, standard deviation, and sample first order autocorrelation. Means and standard deviations are expressed in percentage terms.

Table 2.3: Posterior Estimates

		Prior		Posterior					Prior		Posterior		
	Distr.	5%	95%	5%	50%	95%		Distr.	5%	95%	5%	50%	95%
		Preferences							Dividend Process				
δ	B	[.9951	.9999]	.9985	.9989	.9991	μ_d	N	[-.007	.006]	-	.0010	-
ψ	G	[0.31	3.45]	1.80	1.81	1.82	ϕ_x	N	[-13.1	13.4]	2.39	2.51	2.67
γ	G	[2.74	15.45]	10.82	10.99	11.17	ϕ_η	N	[-1.68	1.63]	1.09	1.10	1.13
							φ_d	G	[0.22	11.90]	4.74	5.01	5.19
		Consumption Process							Inflation Process				
μ_c	N	[-.006	.006]	-	.0016	-	μ_π	N	[-.007	.006]	.0027	.0029	.0030
$\bar{\sigma}_c$	IG	[.001	.006]	.0020	.0021	.0021	$\bar{\sigma}_\pi$	N	[.001	.006]	.0015	.0015	.0016
φ_c	G	[0.00	0.11]	.026	.031	.033	φ_π	G	[0.00	0.11]	0.11	0.12	0.12
ν_c	N^T	[-0.08	0.97]	.9906	.9952	.9959	ν_π	N^T	[-0.08	0.97]	.9915	.9928	.9937
σ_{w_c}	IG	[0.22	1.03]	0.30	0.31	0.34	σ_{w_π}	IG	[0.22	1.03]	0.43	0.45	0.46
		Regime-Switching VAR Coefficients											
		Countercyclical Inflation Regime							Procyclical Inflation Regime				
ρ_c	N^T	[-0.08	0.97]	.9957	.9957	.9958	ρ_c	N^T	[-0.08	0.97]	.951	.953	.957
ρ_π	N^T	[-0.08	0.97]	.9957	.9959	.9961	ρ_π	N^T	[-0.08	0.97]	.980	.980	.981
$\chi_{c,\pi}$	N	[-0.80	0.80]	-.40	-.40	-.41	$\chi_{c,\pi}$	N	[-0.80	0.80]	.150	.155	.162
ρ_m	N^T	[-0.08	0.97]	.9906	.9916	.9929	ρ_m	N^T	[-0.08	0.97]	.9906	.9916	.9929
σ_m	IG	[.000	.001]	.0001	.0002	.0003	σ_m	IG	[.000	.001]	.0001	.0002	.0003
		Regime-Switching Monetary Policy Coefficients											
		Active Monetary Policy Regime							Passive Monetary Policy Regime				
τ_c	N	[-4.28	4.28]	.9540	.9543	.9545	τ_c	N	[-4.28	4.28]	.548	.550	.551
τ_π	N^T	[0.00	4.28]	3.09	3.10	3.11	τ_π	N^T	[0.00	4.28]	.960	.960	.961
		Markov-Chain Transition Probabilities											
		Inflation Regime							Monetary Policy Regime				
$\mathbb{P}_{X_1}$	B	[0.38	1.00]	.989	.992	.995	$\mathbb{P}_{M_1}$	B	[0.38	1.00]	.987	.990	.991
$\mathbb{P}_{X_2}$	B	[0.38	1.00]	.938	.941	.945	$\mathbb{P}_{M_2}$	B	[0.38	1.00]	.974	.975	.979

Notes: The estimation results are based on monthly data from 1959:M1 to 2011:M12 with the exception that the consumption series only starts in 1959:M2. For consumption I adopt the measurement error model of Schorfheide, Song, and Yaron (2013) with the modification that the statistical agency uses the proxy series to distribute quarterly (instead of annual) consumption growth over the three months of the quarter (instead of the twelve months of a year). I fix $\mu_c = 0.0016$ and $\mu_d = 0.0010$ in the estimation. B, N, N^T, G, and IG denote beta, normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, and inverse gamma distributions, respectively.

Table 2.4: Model-Generated Correlations between Consumption and Inflation

	Data	Model					
	$\text{corr}(\Delta c_t, \pi_t)$	$\text{corr}(\Delta c_t, \pi_t)$			$\text{corr}(\mathbb{E}\Delta c_{t+1}, \mathbb{E}\pi_{t+1})$		
Regime	Estimate	Median	5%	95%	Median	5%	95%
CA	-0.24	-0.58	[-0.80,	-0.22]	-0.93	[-0.99,	-0.64]
CP	-0.09	-0.48	[-0.78,	0.02]	-0.74	[-0.95,	-0.15]
PA	0.01	0.17	[-0.13,	0.42]	0.59	[0.27,	0.80]
PP	0.03	0.19	[-0.14,	0.47]	0.27	[0.44,	0.84]

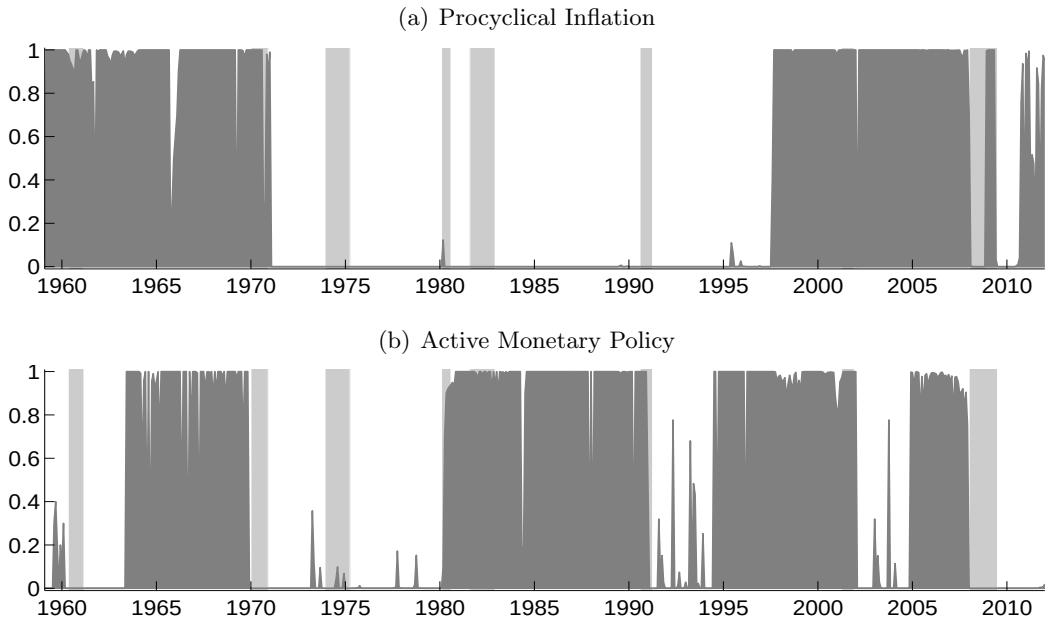
Notes: "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes. Data estimates are based on monthly consumption growth and inflation series.

Table 2.5: Variance Decomposition

Variable Name	Long-Run Growth & Inflation Target			Monetary Policy Shock			Long-Run Growth Vol. & Inflation Target Vol.		
	Median	5%	95%	Median	5%	95%	Median	5%	95%
log Price-Dividend Ratio	51.3	[43.5,	62.7]	-	[-	-]	49.7	[37.1,	57.2]
3-Month Bond Yield	94.5	[91.1,	97.4]	4.2	[2.1,	5.5]	0.2	[0.0,	0.3]
10-Year Bond Yield	80.7	[71.0,	94.3]	5.3	[3.3,	6.2]	14.2	[6.3,	23.7]

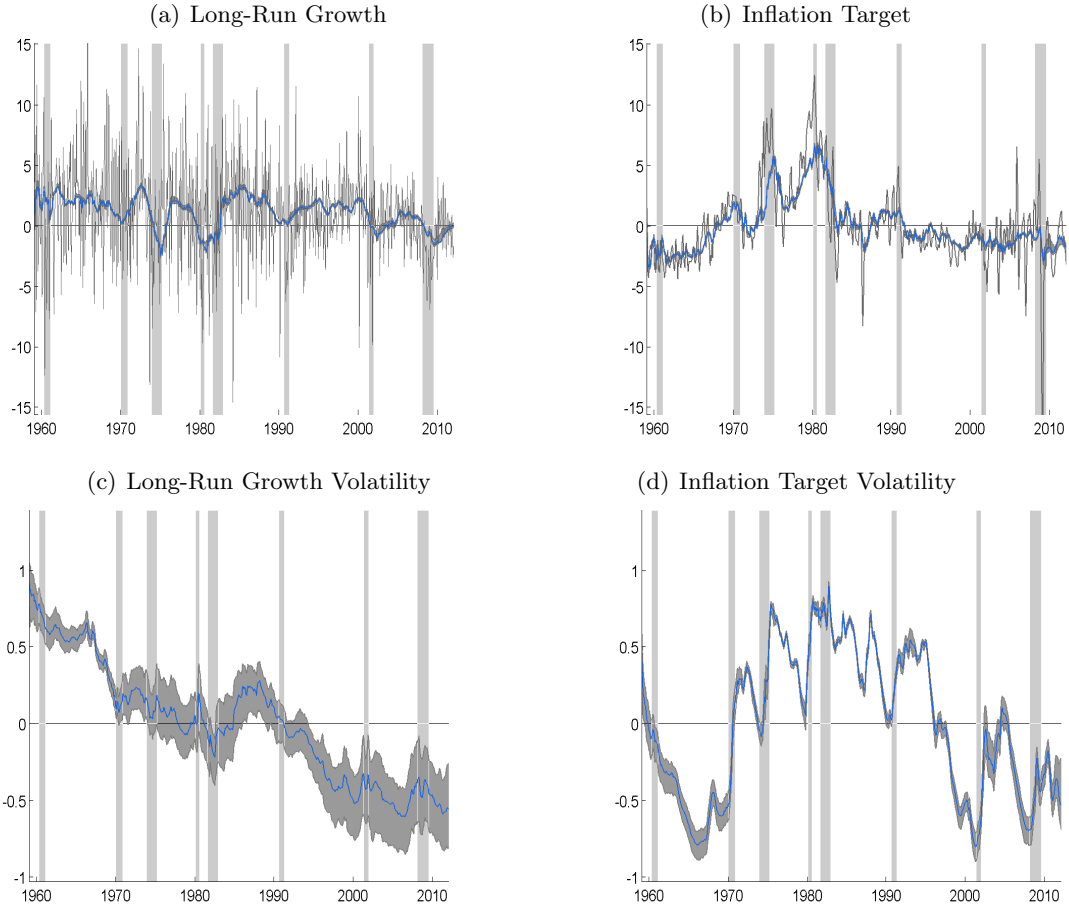
Notes: Fraction of volatility fluctuations (in percents) of the log price dividend ratio, the 3-month nominal bond yield, and the 10-year nominal bond yield that is due to the long-run growth ($x_{c,t}$), inflation target ($x_{\pi,t}$), monetary policy shock ($x_{m,t}$), long-run growth volatility ($\sigma^2_{c,t}$), and inflation target volatility ($\sigma^2_{\pi,t}$), respectively. Note that due to measurement errors, the numbers do not sum to 100%.

Figure 2.1: Smoothed Probabilities for Transitions between Regimes



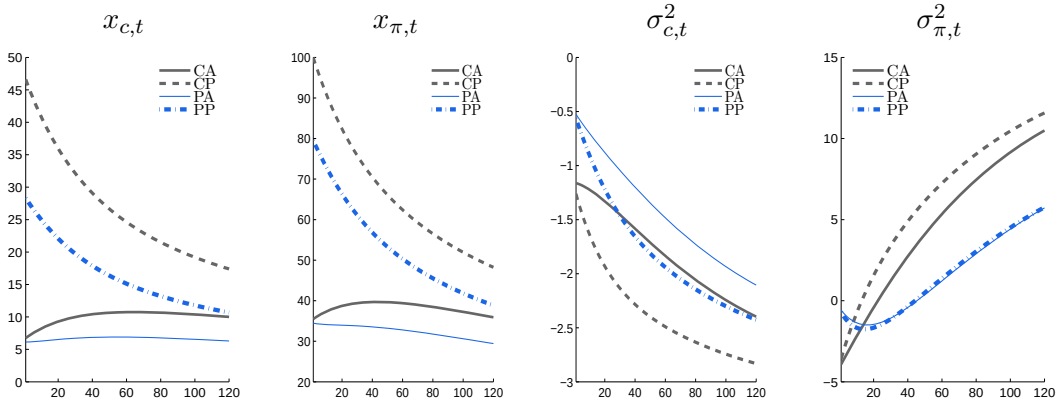
Notes: Dark gray shaded areas represent posterior medians of smoothed regime probabilities. Light gray shaded bars indicate NBER recession dates. Figure 2.1(a) displays the smoothed probabilities of the procyclical inflation regime while Figure 2.1(b) shows the smoothed probabilities of the active monetary policy regime.

Figure 2.2: Smoothed Mean and Volatility States



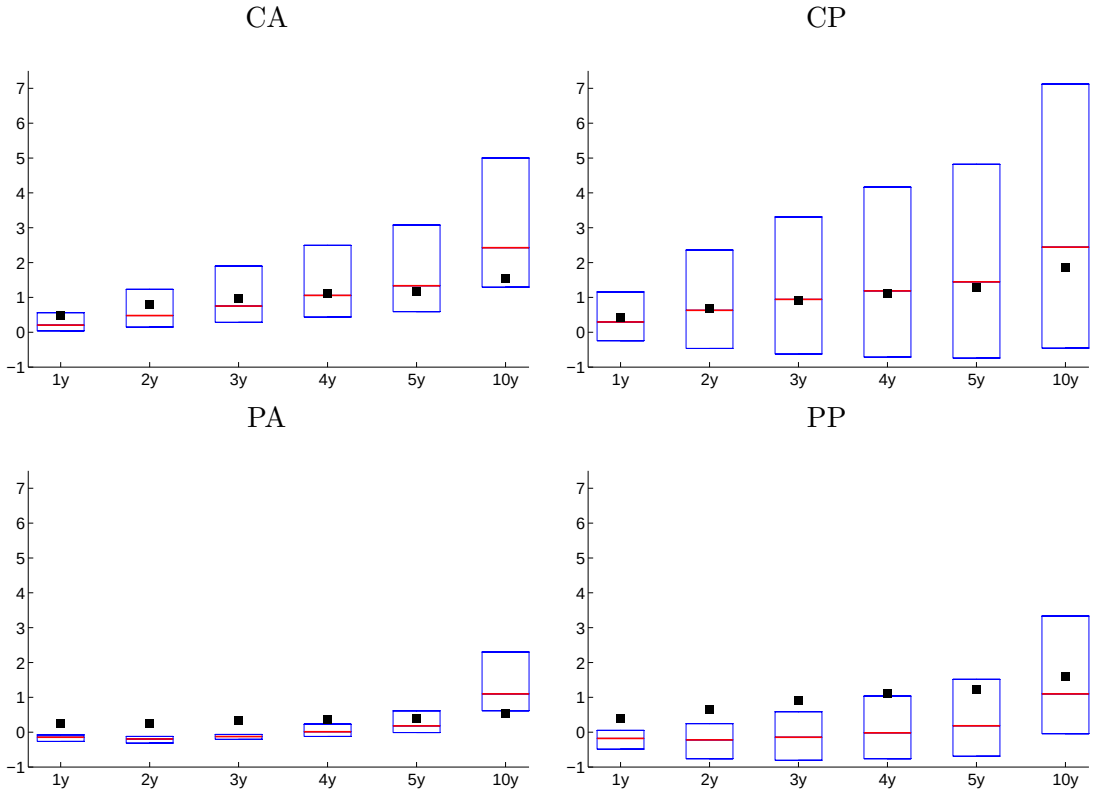
Notes: Blue lines represent posterior medians of smoothed states and dark gray shaded area corresponds to 90% credible intervals. Light gray shaded bars indicate NBER recession dates. In the top panel, I overlay the smoothed states with monthly consumption growth and inflation (gray solid lines).

Figure 2.3: Equilibrium Nominal Bond Yield Loadings



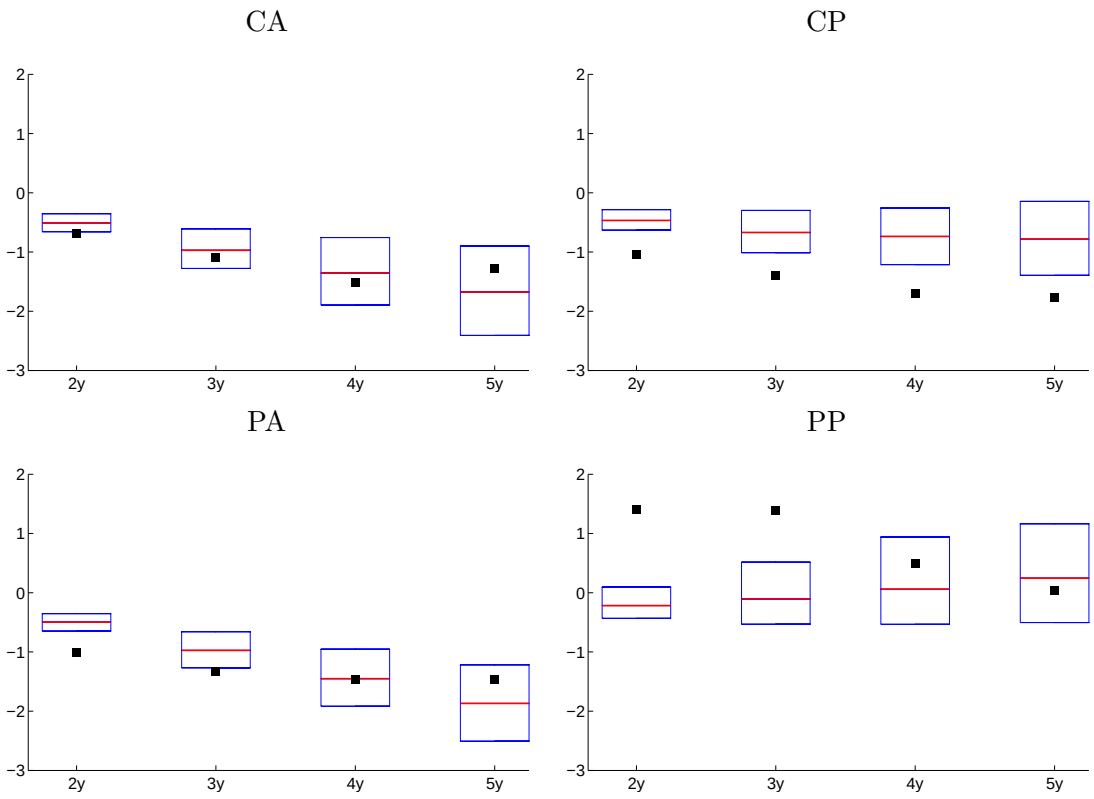
Notes: Model-implied nominal bond yield loadings on the long-run growth ($x_{c,t}$), inflation target ($x_{\pi,t}$), long-run growth volatility ($\sigma^2_{c,t}$), and inflation target volatility ($\sigma^2_{\pi,t}$) are provided. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes. Maturity on the x-axis is in months. Numbers are displayed in percent.

Figure 2.4: Model-Generated Yield Spread



Notes: "Spread" is the difference between 3m yield and yields with maturity at 1y–10y. Black squares indicate values from actual data. Figure also depicts medians (red lines) and 90% credible intervals (top and bottom lines of boxes) of the distribution of yield spreads obtained with model-generated data. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes. Numbers are displayed in percent (annualized).

Figure 2.5: Term Spread Regression

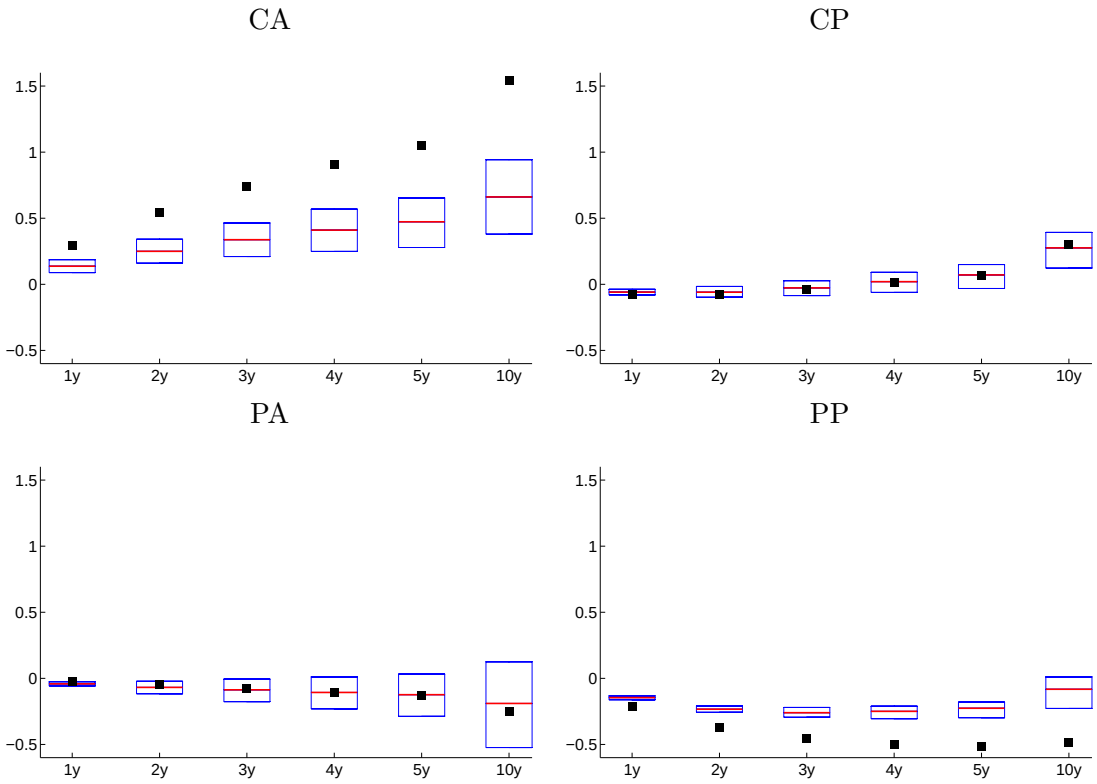


Notes: The model-implied 90% distributions for the slope coefficient, β_n, from the regression below are provided.

$$y_{t+12,n-12} - y_{t,n} \quad = \quad \alpha_n + \beta_n \left(\left(y_{t,n} - y_{t,12} \right) \frac{12}{n-12} \right) + \epsilon_{t+12}, \quad n \in \{24, 36, 48, 60\} .$$

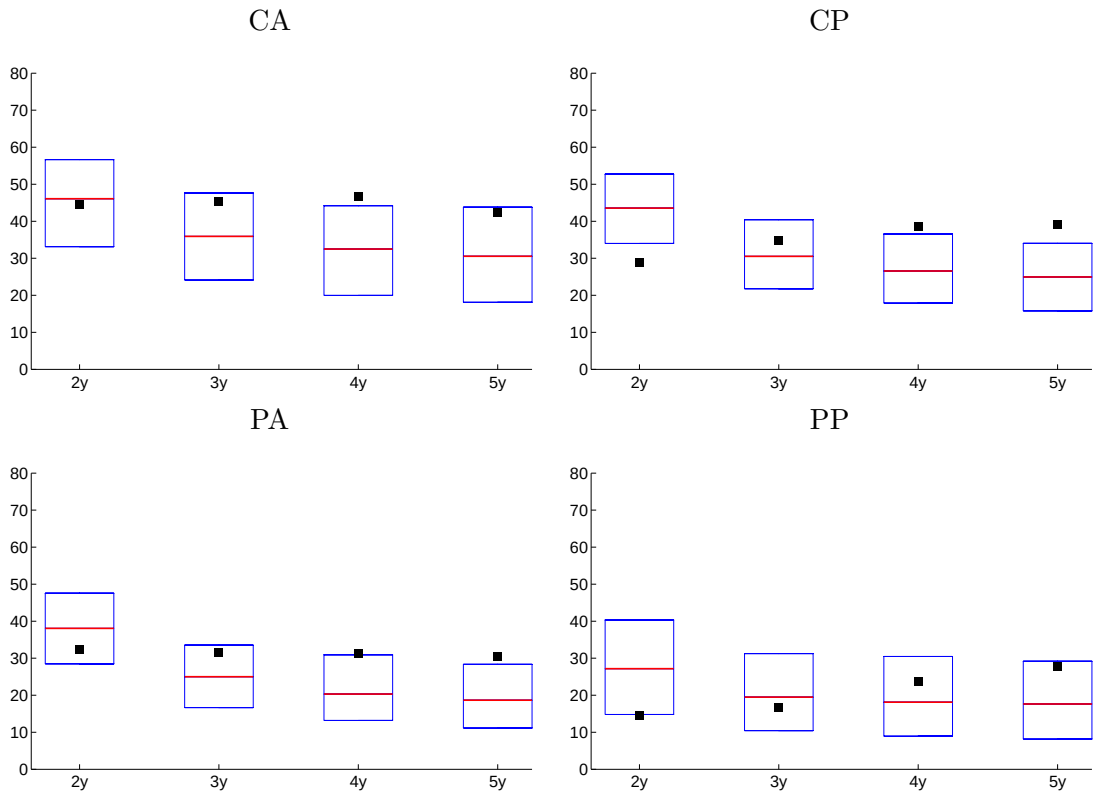
Medians are depicted by red lines. Black squares indicate estimates from actual data. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes.

Figure 2.6: Term Premia



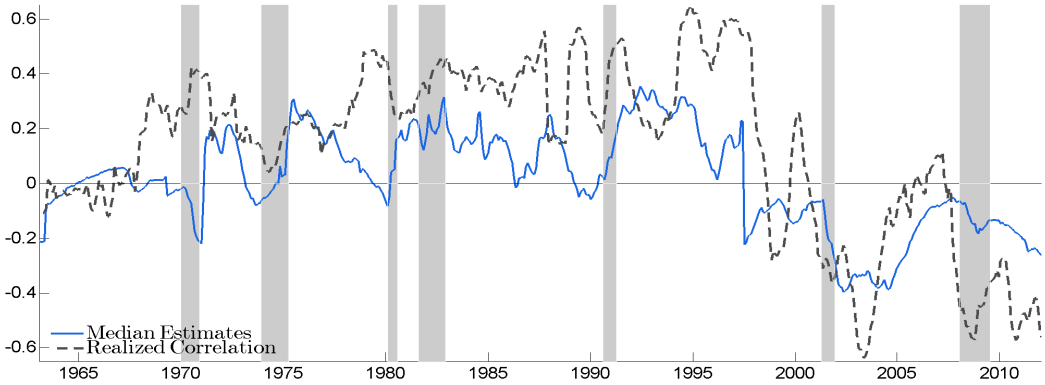
Notes: The model-implied 90% distributions for term premium$_{t,n} = y_{t,n} - \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_t(y_{t+i,1})$ are provided, $n \in \{12, 24, 36, 48, 60, 120\}$. Medians are depicted by red lines. Black squares indicate estimates from actual data. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes.

Figure 2.7: Excess Bond Return Predictability Regression by Cochrane and Piazzesi (2005)



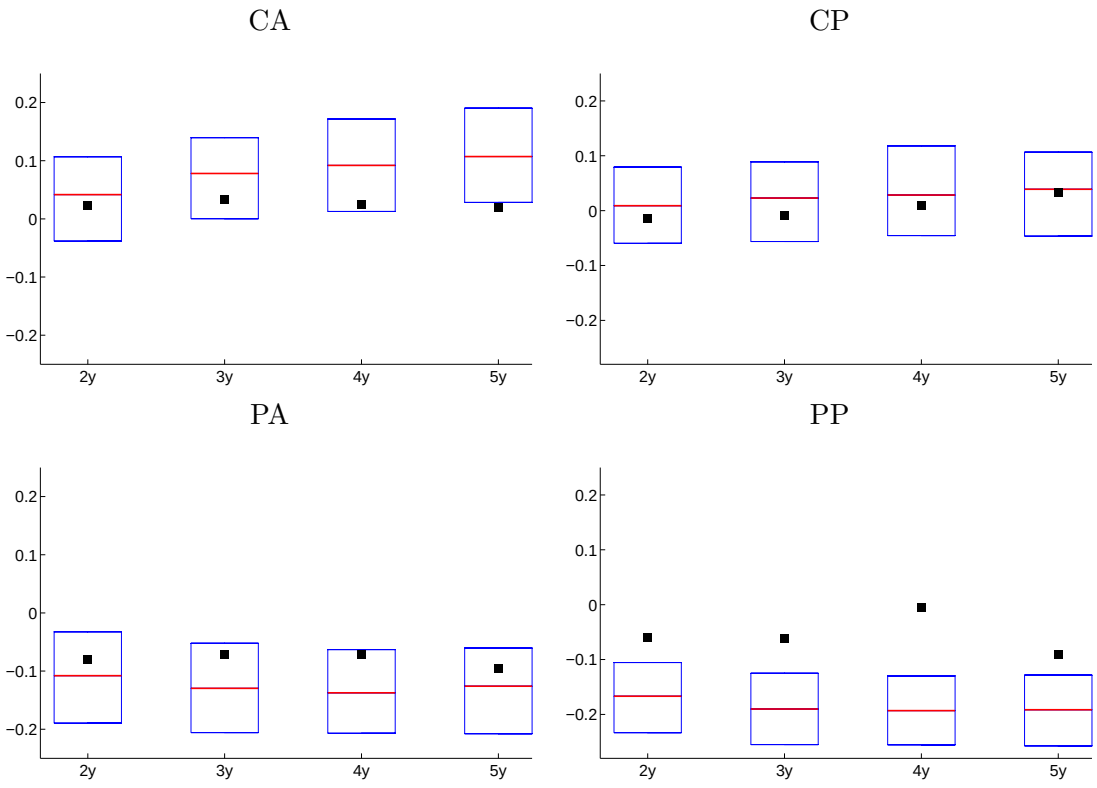
Notes: The model-implied 90% distributions for R^2 values (in percents) from the excess bond return predictability regression by Cochrane and Piazzesi (2005) are provided. Medians are depicted by red lines. Black squares indicate estimates from actual data. I focus on regressing the excess bond return of an n year bond over the 1 year bond on a linear combination of forward rates that includes a constant term, a one year bond yield, and four forwards rates with maturities of 2 to 5 years. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes.

Figure 2.8: Estimated Stock-Bond Return Correlation



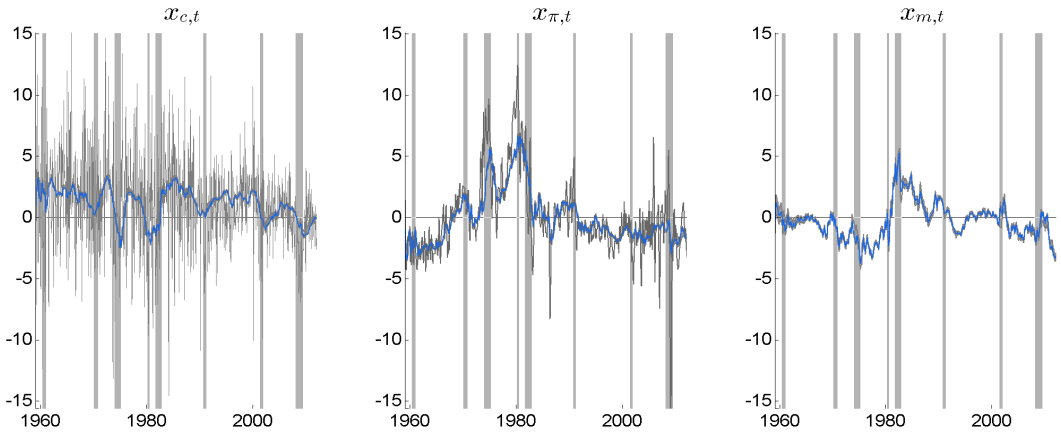
Notes: The correlation between stock market returns and 1 year holding period bond returns for maturity at 10 years is provided. Black dashed line depicts the monthly realized stock-bond correlation obtained from daily data. Blue solid line represents posterior median of correlations. Light gray shaded bars indicate NBER recession dates. The unconditional correlation between two measures are about 0.68.

Figure 2.9: Stock-Bond Return Correlation



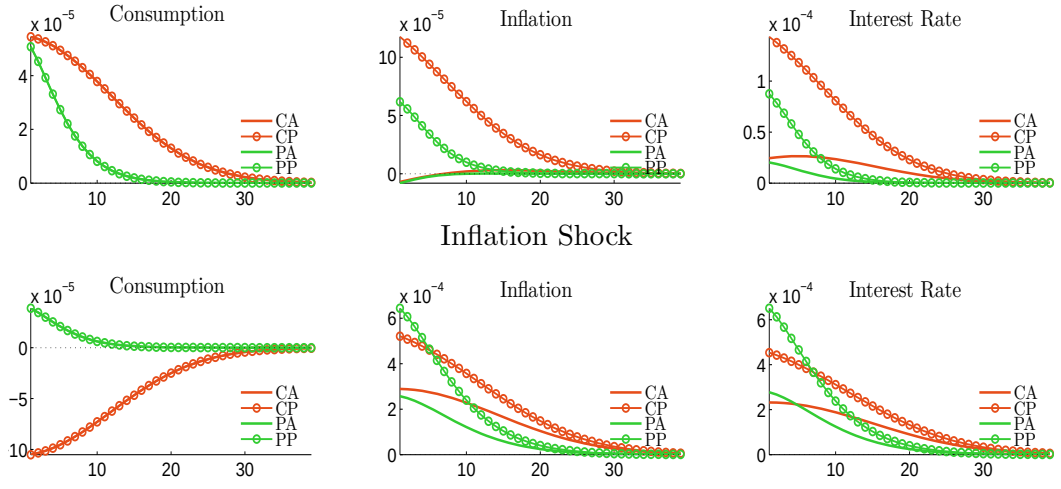
Notes: The estimated correlation between stock market returns and 1 year holding period bond returns for maturities of 2-5 years are provided. Black squares indicate regime-dependent sample correlations of actual data. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes.

Figure 2.10: Smoothed Mean States



Notes: Black lines represent posterior medians of smoothed states and the dark gray shaded area corresponds to 90% credible intervals. Light gray shaded bars indicate NBER recession dates. I overlay the smoothed long-run growth with monthly consumption growth and the smoothed long-run inflation with realized inflation (blue solid lines).

Figure 2.11: Impulse Response Function
Growth Shock



Inflation Shock

Figure 2.12: Model-Generated Unconditional Mean

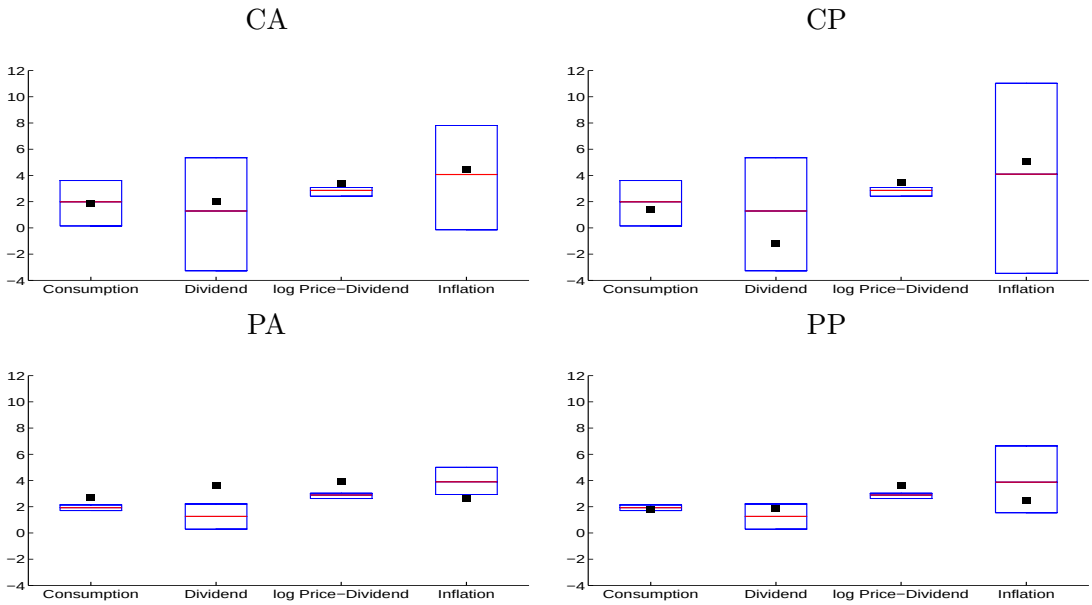
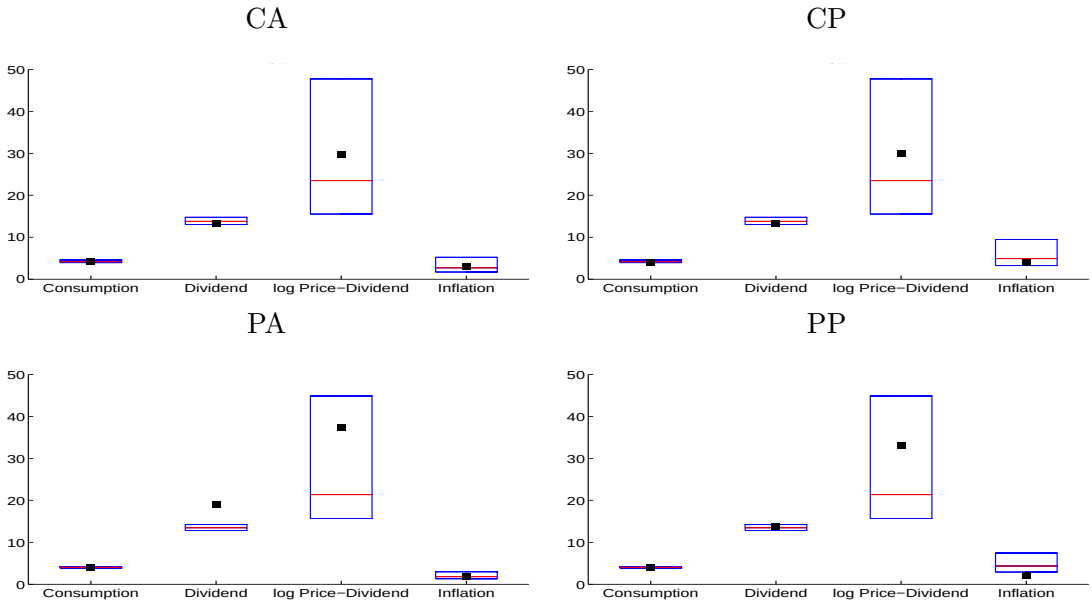
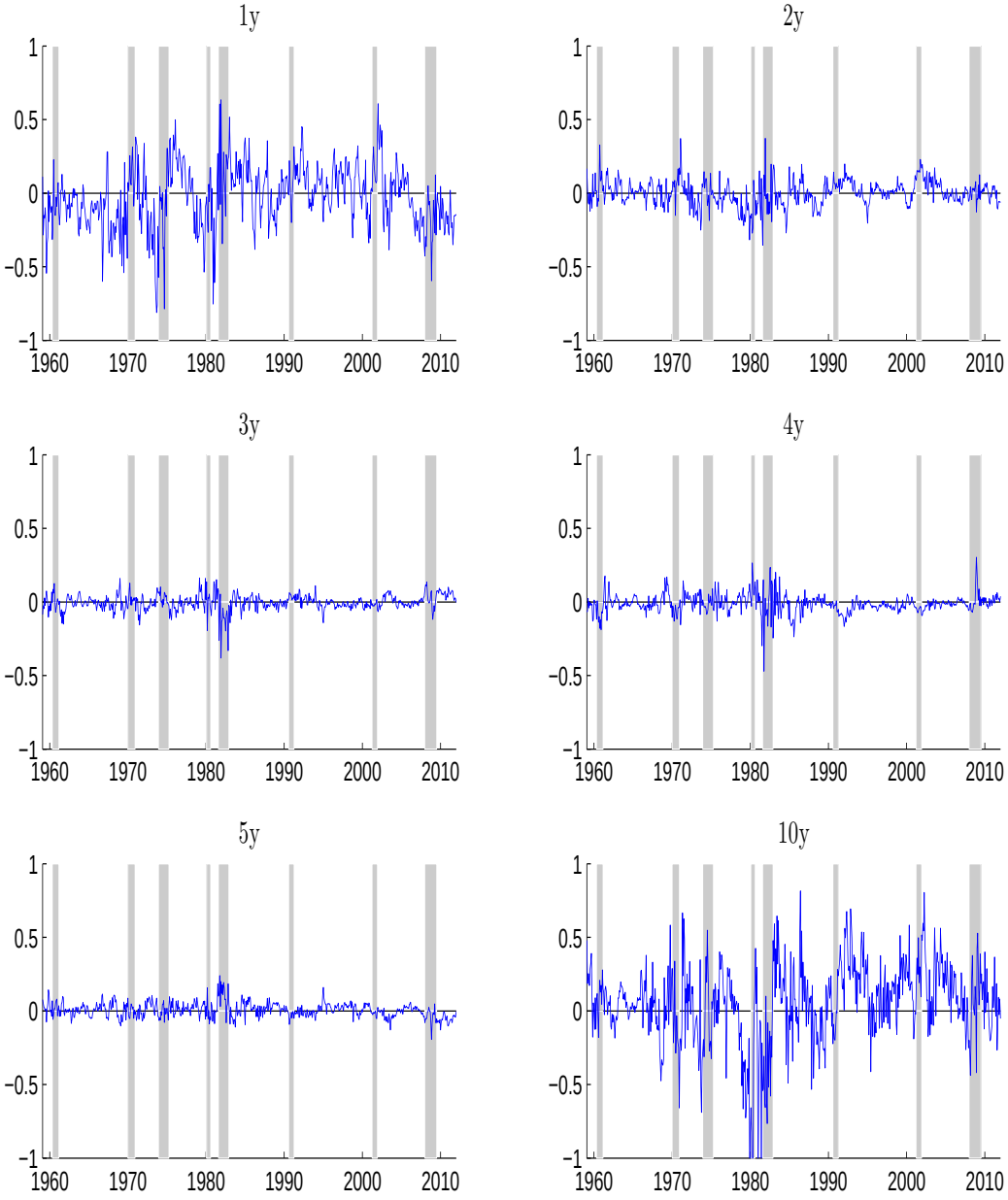


Figure 2.13: Model-Generated Unconditional Standard Deviation



Notes: Black squares indicate values from actual data. The figure also depicts medians (red lines) and 90% credible intervals (top and bottom lines of boxes) of the distribution of yield spreads obtained with model-generated data. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes. Numbers are displayed in percents (annualized).

Figure 2.14: Yield Prediction Errors

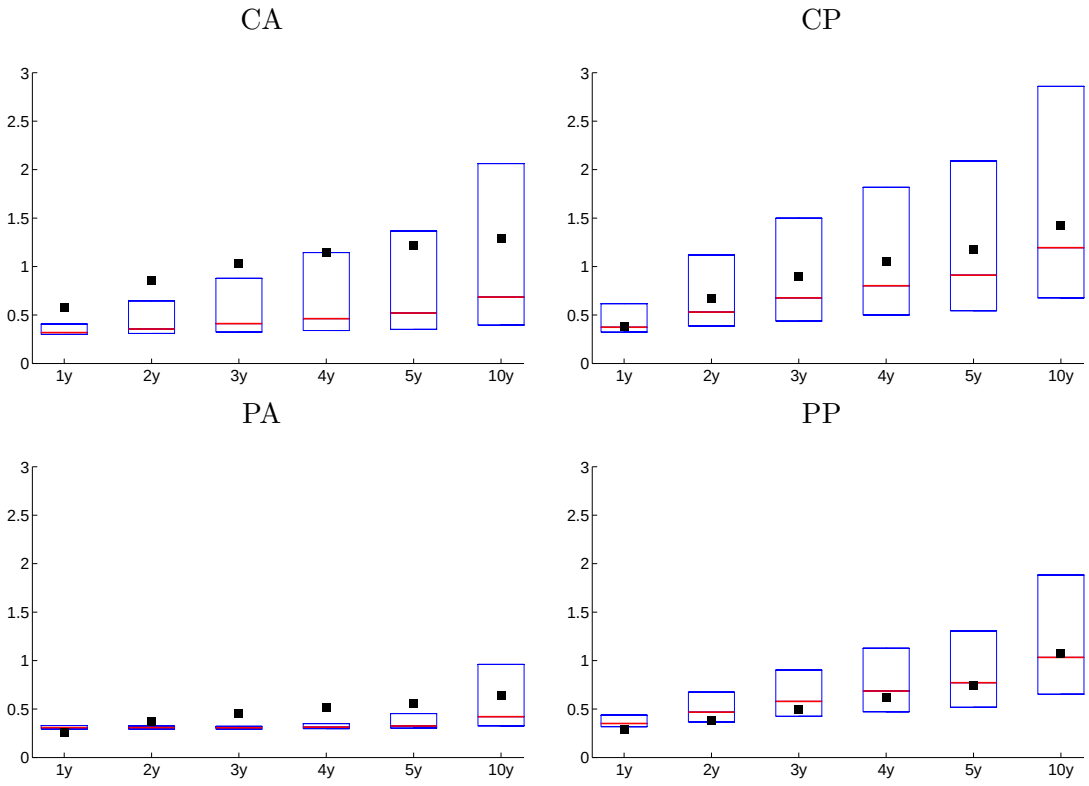


Note: Numbers are displayed in percents (annualized). In-sample RMSE numbers

1y	2y	3y	4y	5y	10y
21	8	5	7	5	28

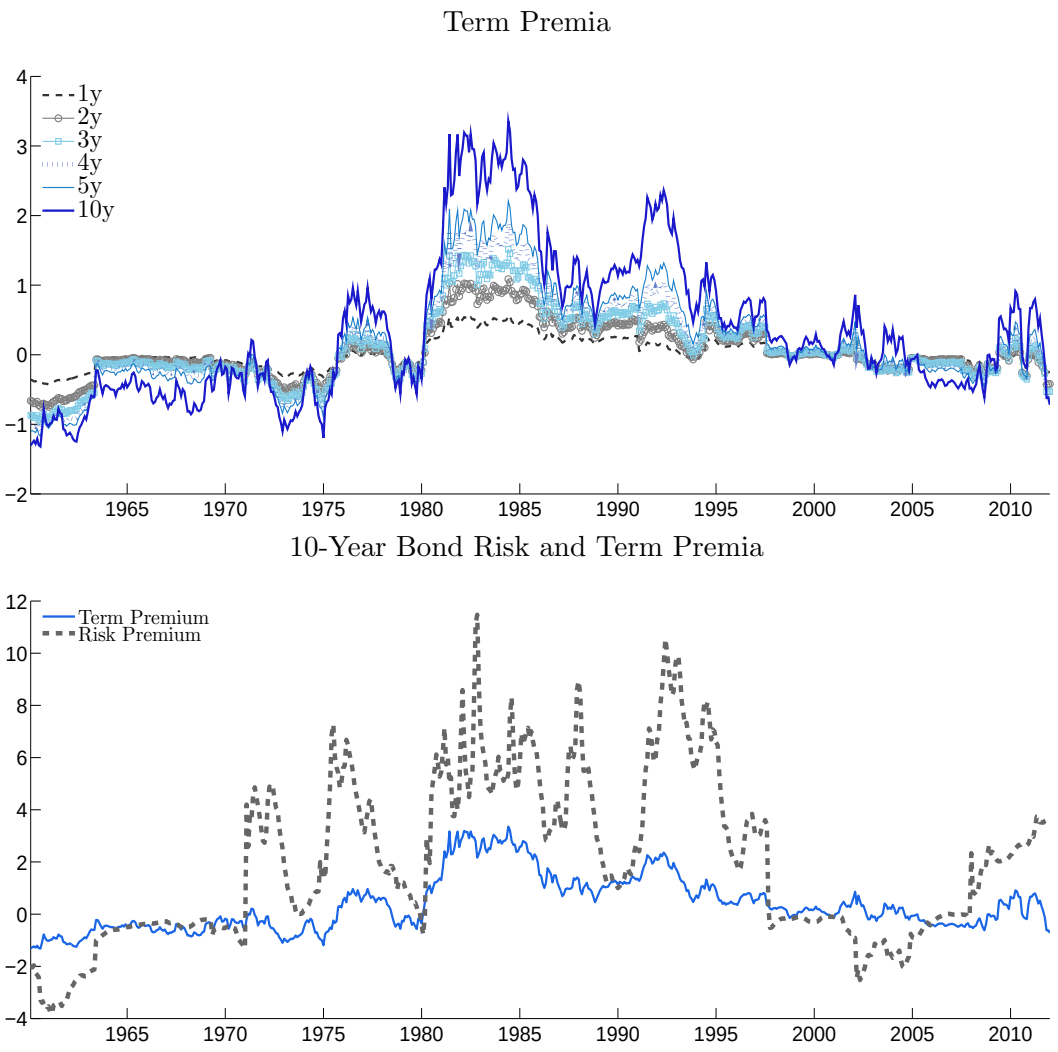
are also provided in basis points (annualized).

Figure 2.15: Model-Generated Yield Spread: Unconditional Standard Deviation



Notes: The "spread" is the difference between the 3m yield and yields with maturities of 1y–10y. Black squares indicate values from actual data. The figure also depicts medians (red lines) and 90% credible intervals (top and bottom lines of boxes) of the distribution of yield spreads obtained with model-generated data. "CA" stands for the countercyclical inflation and the active monetary policy regimes while "PP" stands for the procyclical inflation and the passive monetary policy regimes. "CP" and "PA" indicate the remaining combinations of regimes. Numbers are displayed in percents (annualized).

Figure 2.16: Risk and Term Premia



Note: The median estimates of term premium$_{t,n} = y_{t,n} - \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_t(y_{t+i,1})$ and risk premium$_{t,n} = -cov_t(m_{t+1} - \pi_{t+1}, rx_{t+1,n})$ are provided.

2.7 Appendix

2.7.1 Solving the LRR Model

This section provides approximate analytical solutions for the equilibrium asset prices.

Exogenous Dynamics

The joint dynamics of consumption, dividend growth, and inflation are

$$\begin{bmatrix} g_{c,t+1} \\ g_{d,t+1} \\ \pi_{t+1} \end{bmatrix} = \begin{bmatrix} \mu_c \\ \mu_d \\ \mu_\pi \end{bmatrix} + \begin{bmatrix} e_1 \\ \phi_x e_1 \\ \Gamma_x(S^X_{t+1}, S^M_{t+1}) \end{bmatrix} X_{t+1} + \begin{bmatrix} 1 & 0 & 0 \\ \phi_\eta & 1 & 0 \\ \Gamma_\eta(S^X_{t+1}, S^M_{t+1}) & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{\sigma}_c \eta_{c,t+1} \\ \bar{\sigma}_d \eta_{d,t+1} \\ \bar{\sigma}_\pi \eta_{\pi,t+1} \end{bmatrix} \tag{2.20}$$

The conditional mean and volatility processes evolve according to

$$\begin{aligned} \underbrace{\begin{bmatrix} x_{c,t+1} \\ x_{\pi,t+1} \\ x_{m,t+1} \end{bmatrix}}_{X_{t+1}} &= \underbrace{\begin{bmatrix} \rho_c(S^X_{t+1}) & \rho_{c,\pi}(S^X_{t+1}) & \rho_{c,m}(S^X_{t+1}) \\ \rho_{\pi,c}(S^X_{t+1}) & \rho_\pi(S^X_{t+1}) & \rho_{\pi,m}(S^X_{t+1}) \\ 0 & 0 & \rho_m(S^X_{t+1}) \end{bmatrix}}_{\Upsilon(S^X_{t+1})} \underbrace{\begin{bmatrix} x_{c,t} \\ x_{\pi,t} \\ x_{m,t} \end{bmatrix}}_{X_t} \\ &+ \underbrace{\begin{bmatrix} 1 & \chi_{c,\pi}(S^X_{t+1}) & 0 \\ \chi_{\pi,c}(S^X_{t+1}) & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\Omega(S^X_{t+1})} \underbrace{\begin{bmatrix} \sigma_{c,t} e_{c,t+1} \\ \sigma_{\pi,t} e_{\pi,t+1} \\ \sigma_m e_{m,t+1} \end{bmatrix}}_{E_{t+1}} \end{aligned} \tag{2.21}$$

$$\underbrace{\begin{bmatrix} \sigma^2_{c,t+1} \\ \sigma^2_{\pi,t+1} \end{bmatrix}}_{\Sigma_{t+1}} = \underbrace{\begin{bmatrix} (1-\nu_c)(\varphi_c \bar{\sigma})^2 \\ (1-\nu_\pi)(\varphi_\pi \bar{\sigma})^2 \end{bmatrix}}_{\Phi_\mu} + \underbrace{\begin{bmatrix} \nu_c & 0 \\ 0 & \nu_\pi \end{bmatrix}}_{\Phi_\nu} \underbrace{\begin{bmatrix} \sigma^2_{c,t} \\ \sigma^2_{\pi,t} \end{bmatrix}}_{\Sigma_t} + \underbrace{\begin{bmatrix} \sigma_{w_c} w_{c,t+1} \\ \sigma_{w_\pi} w_{\pi,t+1} \end{bmatrix}}_{W_{t+1}}, \quad W_{t+1} \sim N(0, \Phi_w),$$

where $\eta_{j,t+1}, e_{k,t+1}, w_{l,t+1} \sim N(0,1)$ for $j \in \{c, d, \pi\}$, $k \in \{c, \pi, m\}$, and $l \in \{c, \pi\}$.

Note that the VAR dynamics are generalized to allow for intertemporal feedback effects (captured by off-diagonal coefficients) and that the inflation target can become correlated with long-run growth innovation. Furthermore, the channels through which monetary policy shock affects long-run growth or inflation target, are not restricted to zero as in the main text. (Of course, one could set them equal to zero.)

Derivation of Approximate Analytical Solutions

The Euler equation for the economy is

$$1 = E_t \left[\exp\left(m_{t+1} + r_{k,t+1} \right) \right], \quad k \in \{c, m\}, \tag{2.22}$$

where $m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{c,t+1}$ is the log stochastic discount factor, $r_{c,t+1}$ is the log return on the consumption claim, and $r_{m,t+1}$ is the log market return. All returns are given by the approximation of Campbell and Shiller (1988a):

$$\begin{aligned} r_{c,t+1} &= \kappa_{0,c} + \kappa_{1,c} z_{c,t+1} - z_{c,t} + g_{c,t+1} \\ r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}. \end{aligned} \tag{2.23}$$

Let I_t denote the current information set $\left\{S_{1:t}^X, X_t, \Sigma_t\right\}$ and define $I_{t+1}{=}I_t \cup \{S_{t+1}^X\}$ that includes information regarding S_{t+1}^X in addition to I_t. Suppose $S_t^X = i$ for $i{=}1,\ 2$. Derivation of (2.22) follows Bansal and Zhou (2002), who make repeated use of the law of iterated expectations and log-linearization, and Schorfheide, Song, and Yaron (2013) who utilize log-linear approximation for returns and for volatilities.

$$\begin{aligned} 1 &= \mathbb{E}\Big(\mathbb{E}\left[\exp\left(m_{t+1} + r_{m,t+1} \right) \mid I_{t+1} \right] \mid I_t \Big) \\ &= \sum_{j=1}^{4} \mathbb{P}_{ij} \mathbb{E}\Big(\exp\left(m_{t+1} + r_{m,t+1} \right) \mid S_{t+1} = j, X_t, \Sigma_t \Big) \\ 0 &= \sum_{j=1}^{4} \mathbb{P}_{ij} \underbrace{\left(\mathbb{E}\left[m_{t+1} + r_{m,t+1} \mid S_{t+1} = j \right] + \frac{1}{2} \mathbb{V}\left[m_{t+1} + r_{m,t+1} \mid S_{t+1} = j \right]. \right)}_{B} \end{aligned} \tag{2.24}$$

The first line uses the law of iterated expectations, second line uses the definition of Markov-chain; and the third line applies log-linearization, $\exp(B){-}1 \approx B$, log-normality assumption, and log-linearization for returns and for volatilities.

Real Consumption Claim

Conjecture that the price to consumption ratio follows

$$z_t(S_t^X) = A_0(S_t^X) + A_1(S_t^X) X_t + A_2(S_t^X) \Sigma_t, \tag{2.25}$$

where $A_1(S_t^X) = \begin{bmatrix} A_{1,c}(S_t^X) & A_{1,\pi}(S_t^X) & A_{1,m}(S_t^X) \end{bmatrix}$ and $A_2(S_t^X) = \begin{bmatrix} A_{2,c}(S_t^X) & A_{2,\pi}(S_t^X) \end{bmatrix}$.

From (2.20), (2.21), (2.23), and (2.25),

$$\begin{aligned} r_{c,t+1} &= \kappa_{0,c} + \kappa_{1,c}A_0(S^X_{t+1}) - A_0(S^X_t) + \mu_c + \kappa_{1,c}A_2(S^X_{t+1})\Phi_\mu \qquad (2.26)\\ &+ \left\{(e_1 + \kappa_{1,c}A_1(S^X_{t+1}))\Upsilon(S^X_{t+1}) - A_1(S^X_t)\right\}X_t + \left\{\kappa_{1,c}A_2(S^X_{t+1})\Phi_\nu - A_2(S^X_t)\right\}\Sigma_t \\ &+ \bar{\sigma}_c\eta_{t+1} + (e_1 + \kappa_{1,c}A_1(S^X_{t+1}))\Omega(S^X_{t+1})E_{t+1} + \kappa_{1,c}A_2(S^X_{t+1})W_{t+1} \end{aligned}$$

and from (2.20), (2.21), (2.23), (2.24), and (2.25)

$$\begin{aligned} m_{t+1} &= \theta\log\delta + (\theta-1)\left\{\kappa_{0,c} + \kappa_{1,c}A_0(S^X_{t+1}) - A_0(S^X_t) + \kappa_{1,c}A_2(S^X_{t+1})\Phi_\mu\right\} - \gamma\mu \qquad (2.27)\\ &- \frac{1}{\psi}e_1\Upsilon(S^X_{t+1})X_t + (\theta-1)\left\{((1-\frac{1}{\psi})e_1 + \kappa_{1,c}A_1(S^X_{t+1}))\Upsilon(S^X_{t+1}) - A_1(S^X_t)\right\}X_t \\ &+ (\theta-1)\left\{\kappa_{1,c}A_2(S^X_{t+1})\Phi_\nu - A_2(S^X_t)\right\}\Sigma_t - \gamma\bar{\sigma}_c\eta_{c,t+1} \\ &+ \left\{-\gamma e_1 + (\theta-1)\kappa_{1,c}A_1(S^X_{t+1})\right\}\Omega(S^X_{t+1})E_{t+1} + (\theta-1)\kappa_{1,c}A_2(S^X_{t+1})W_{t+1}. \end{aligned}$$

The solutions for As that describe the dynamics of the price-consumption ratio are determined from (2.24), and they are,

$$\begin{aligned} \begin{bmatrix} A_1(1) & A_1(2) \end{bmatrix} &= (1-\frac{1}{\psi})e_1 \begin{bmatrix} p_{X_1}\Upsilon(1) + (1-p_{X_1})\Upsilon(2) & (1-p_{X_2})\Upsilon(1) + p_{X_2}\Upsilon(2) \end{bmatrix} \qquad (2.28)\\ &\quad \times \begin{bmatrix} \mathbb{I}_2 - p_{X_1}\kappa_{1,c}\Upsilon(1) & -(1-p_{X_2})\kappa_{1,c}\Upsilon(1) \\ -(1-p_{X_1})\kappa_{1,c}\Upsilon(2) & \mathbb{I}_2 - p_{X_2}\kappa_{1,c}\Upsilon(2) \end{bmatrix}^{-1} \\ \begin{bmatrix} A_{2,c}(1) \\ A_{2,c}(2) \end{bmatrix} &= \frac{\theta}{2}\left[\ \mathbb{I}_2 - \kappa_{1,c}\nu_c\mathbb{P}_X\ \right]^{-1} \times \mathbb{P}_X \times \begin{bmatrix} \xi_c(1) \\ \xi_c(2) \end{bmatrix} \\ \begin{bmatrix} A_{2,\pi}(1) \\ A_{2,\pi}(2) \end{bmatrix} &= \frac{\theta}{2}\left[\ \mathbb{I}_2 - \kappa_{1,c}\nu_\pi\mathbb{P}_X\ \right]^{-1} \times \mathbb{P}_X \times \begin{bmatrix} \xi_\pi(1) \\ \xi_\pi(2) \end{bmatrix} \\ \begin{bmatrix} A_0(1) \\ A_0(2) \end{bmatrix} &= \left[\ \mathbb{I}_2 - \kappa_{1,c}\mathbb{P}_X\ \right]^{-1} \times \mathbb{P}_X \times \begin{bmatrix} \bar{A}_0 + \kappa_{1,c}A_2(1)\Phi_\mu + \frac{\theta}{2}\kappa^2_{1,c}A_2(1)\Phi_w A_2(1)' + \frac{\theta}{2}\xi_m(1)\sigma^2_m(1) \\ \bar{A}_0 + \kappa_{1,c}A_2(2)\Phi_\mu + \frac{\theta}{2}\kappa^2_{1,c}A_2(2)\Phi_w A_2(2)' + \frac{\theta}{2}\xi_m(2)\sigma^2_m(2) \end{bmatrix} \end{aligned}$$

where $\bar{A}_0 = \log\delta + \kappa_{0,c} + \mu_c(1-\frac{1}{\psi}) + \frac{\theta}{2}\bar{\sigma}^2_c(1-\frac{1}{\psi})^2$ and

$$\begin{aligned} \xi_c(i) &= \left\{\left((1-\frac{1}{\psi})e_1 + \kappa_{1,c}A_1(i)\right)\cdot\Omega(i)e_1'\right\}^2, \quad \xi_\pi(i) = \left\{\left((1-\frac{1}{\psi})e_1 + \kappa_{1,c}A_1(i)\right)\cdot\Omega(i)e_2'\right\}^2 \\ \xi_m(i) &= \left\{\left((1-\frac{1}{\psi})e_1 + \kappa_{1,c}A_1(i)\right)\cdot\Omega(i)e_3'\right\}^2, \quad i\in\{1,2\}. \end{aligned}$$

Real Market Returns

Similarly, using the conjectured solution to the price-dividend ratio

$$z_{m,t}(S^X_t) = A_{0,m}(S^X_t) + A_{1,m}(S^X_t)X_t + A_{2,m}(S^X_t)\Sigma_t, \qquad (2.29)$$

the market return equation can be expressed as

$$\begin{aligned} r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m}A_{0,m}(S^X_{t+1}) - A_{0,m}(S^X_t) + \mu_d + \kappa_{1,m}A_{2,m}(S^X_{t+1})\Phi_\mu \\ &+ \left\{(\phi_x e_1 + \kappa_{1,m}A_{1,m}(S^X_{t+1}))\Upsilon(S^X_{t+1}) - A_{1,m}(S^X_t)\right\} X_t + \left\{\kappa_{1,m}A_{2,m}(S^X_{t+1})\Phi_\nu - A_{2,m}(S^X_t)\right\}\Sigma_t \\ &+ \phi_\eta\bar{\sigma}_c\eta_{c,t+1} + \bar{\sigma}_d\eta_{d,t+1} + (\phi_x e_1 + \kappa_{1,m}A_{1,m}(S^X_{t+1}))\Omega(S^X_{t+1})E_{t+1} + \kappa_{1,m}A_{2,m}(S^X_{t+1})W_{t+1}. \end{aligned} \tag{2.30}$$

From (2.20), (2.21), (2.23), and (2.29), the solutions for A_m-s that describe the dynamics of the price-dividend ratio are

$$\begin{aligned} \begin{bmatrix} A_{1,m}(1) & A_{1,m}(2) \end{bmatrix} &= (\phi_x - \frac{1}{\psi})e_1 \begin{bmatrix} p_{X_1}\Upsilon(1) + (1-p_{X_1})\Upsilon(2) & (1-p_{X_2})\Upsilon(1) + p_{X_2}\Upsilon(2) \end{bmatrix} \\ &\quad \times \begin{bmatrix} \mathbb{I}_2 - p_{X_1}\kappa_{1,m}\Upsilon(1) & -(1-p_{X_2})\kappa_{1,m}\Upsilon(1) \\ -(1-p_{X_1})\kappa_{1,m}\Upsilon(2) & \mathbb{I}_2 - p_{X_2}\kappa_{1,m}\Upsilon(2) \end{bmatrix}^{-1} \\ \begin{bmatrix} A_{2,c,m}(1) \\ A_{2,c,m}(2) \end{bmatrix} &= \begin{bmatrix} \mathbb{I}_2 - \kappa_{1,m}\nu_c\mathbb{P}_X \end{bmatrix}^{-1} \left(\mathbb{P}_X \begin{bmatrix} (\theta-1)\kappa_{1,c}\nu_c A_{2,c}(1) + \frac{1}{2}f_c(1) \\ (\theta-1)\kappa_{1,c}\nu_c A_{2,c}(2) + \frac{1}{2}f_c(2) \end{bmatrix} - (\theta-1)\begin{bmatrix} A_{2,c}(1) \\ A_{2,c}(2) \end{bmatrix} \right) \\ f_c(i) &= \left((\phi_x - \gamma)e_1 \cdot \Omega(i)e_1' + \begin{bmatrix} A_1(i)\cdot\Omega(i)e_1' & A_{1,m}(i)\cdot\Omega(i)e_1' \end{bmatrix} \begin{bmatrix} (\theta-1)\kappa_{1,c} \\ \kappa_{1,m} \end{bmatrix} \right)^2, \\ \begin{bmatrix} A_{2,\pi,m}(1) \\ A_{2,\pi,m}(2) \end{bmatrix} &= \begin{bmatrix} \mathbb{I}_2 - \kappa_{1,m}\nu_\pi\mathbb{P}_X \end{bmatrix}^{-1} \left(\mathbb{P}_X \begin{bmatrix} (\theta-1)\kappa_{1,c}\nu_\pi A_{2,\pi}(1) + \frac{1}{2}f_\pi(1) \\ (\theta-1)\kappa_{1,c}\nu_\pi A_{2,\pi}(2) + \frac{1}{2}f_\pi(2) \end{bmatrix} - (\theta-1)\begin{bmatrix} A_{2,\pi}(1) \\ A_{2,\pi}(2) \end{bmatrix} \right) \\ f_\pi(i) &= \left((\phi_x - \gamma)e_1 \cdot \Omega(i)e_2' + \begin{bmatrix} A_1(i)\cdot\Omega(i)e_2' & A_{1,m}(i)\cdot\Omega(i)e_2' \end{bmatrix} \begin{bmatrix} (\theta-1)\kappa_{1,c} \\ \kappa_{1,m} \end{bmatrix} \right)^2, \\ \begin{bmatrix} A_{0,m}(1) \\ A_{0,m}(2) \end{bmatrix} &= \begin{bmatrix} \mathbb{I}_2 - \kappa_{1,m}\mathbb{P}_X \end{bmatrix}^{-1} \left(\mathbb{P}_X \begin{bmatrix} \bar{A}_{0,m} + f_0(1) \\ \bar{A}_{0,m} + f_0(2) \end{bmatrix} - (\theta-1)\begin{bmatrix} A_0(1) \\ A_0(2) \end{bmatrix} \right) \\ \bar{A}_{0,m} &= \theta\log\delta + (\theta-1)\kappa_{0,c} - \gamma\mu_c + \kappa_{0,m} + \mu_d + \frac{1}{2}\bar{\sigma}_d^2 + \frac{1}{2}\bar{\sigma}_c^2(\phi_\eta - \gamma)^2 \\ f_0(i) &= (\theta-1)\kappa_{1,c}\Big(A_0(i) + A_2(i)\Phi_\mu\Big) + \frac{\sigma^2_{w_c}}{2}\left(\begin{bmatrix} A_{2,c}(i) & A_{2,c,m}(i) \end{bmatrix} \begin{bmatrix} (\theta-1)\kappa_{1,c} \\ \kappa_{1,m} \end{bmatrix} \right)^2 \\ &+ \frac{\sigma^2_{w_\pi}}{2}\left(\begin{bmatrix} A_{2,\pi}(i) & A_{2,\pi,m}(i) \end{bmatrix} \begin{bmatrix} (\theta-1)\kappa_{1,c} \\ \kappa_{1,m} \end{bmatrix} \right)^2 + \kappa_{1,m}A_{2,m}(i)\Phi_\mu \\ &+ \frac{1}{2}\left((\phi_x - \gamma)e_1 \cdot \Omega(i)e_3' + \begin{bmatrix} A_1(i)\cdot\Omega(i)e_3' & A_{1,m}(i)\cdot\Omega(i)e_3' \end{bmatrix} \begin{bmatrix} (\theta-1)\kappa_{1,c} \\ \kappa_{1,m} \end{bmatrix} \right)^2 \sigma^2_m(i), \end{aligned} \tag{2.31}$$

for $i \in \{1,2\}$.

Linearization Parameters

Let $\bar{p}_j = \frac{1-p_l}{2-p_l-p_j}$. For any asset, the linearization parameters are determined endogenously by the following system of equations

$$\begin{aligned} \bar{z}_i &= \sum_{j=1}^{2} \bar{p}_j \left(A_{0,i}(j) + A_{2,c,i}(j)(\varphi_c\bar{\sigma})^2 + A_{2,\pi,i}(j)(\varphi_\pi\bar{\sigma})^2 \right) \\ \kappa_{1,i} &= \frac{\exp(\bar{z}_i)}{1+\exp(\bar{z}_i)} \\ \kappa_{0,i} &= \log(1+\exp(\bar{z}_i)) - \kappa_{1,i}\bar{z}_i. \end{aligned}$$

The solution is determined numerically by iteration until reaching a fixed point of $\bar{z}_i$ for $i \in \{1,2\}$.

Nominal Bond Prices

Endogenous Inflation Determination under a Regime-Switching Taylor Rule

I consider a version of the model where inflation is endogenous. The natural framework in which to this is a model where monetary policy is implemented by a central bank that follows a Taylor rule

$$\begin{aligned} i_t &= \mu_i^{MP}(S_t^M) + \tau_c(S_t^M)(g_{c,t} - \mu_c) + \tau_\pi(S_t^M)(\pi_t - x_{\pi,t}) + x_{\pi,t} + x_{m,t}, \qquad (2.32) \\ &= \mu_i^{MP}(S_t^M) + \left[\begin{array}{cccc} \tau_c(S_t^M) & 1 - \tau_\pi(S_t^M) & 1 & \tau_c(S_t^M) \end{array}\right] X_t^B + \tau_\pi(S_t^M)\pi_t, \end{aligned}$$

where $g_{c,t}$ is consumption growth, $x_{\pi,t}$ is the long-run inflation, and $x_{m,t}$ is the monetary policy shock. Assume for simplicity that π_t is "demeaned" inflation and $X_t^B = [x_{c,t}, x_{\pi,t}, x_{m,t}, \eta_{c,t}]'$.

The asset pricing equation for the short-rate is

$$\begin{aligned} i_t &= -E_t\left[m_{t+1}\right] + E_t\left[\pi_{t+1}\right] - \frac{1}{2}Var_t\left[m_{t+1}\right] - \frac{1}{2}Var_t\left[\pi_{t+1}\right] + Cov_t\left[m_{t+1}, \pi_{t+1}\right] \qquad (2.33) \\ &= \tilde{\mu}_i^{AP}(S_t^X) + \alpha_{X^B}(S_t^X)X_t^B + \alpha_\Sigma(S_t^X)\Sigma_t \\ &\approx \tilde{\mu}_i^{AP}(S_t^X) + \alpha_{X^B}(S_t^X)X_t^B + \alpha_\Sigma(S_t^X)\bar{\Sigma} \\ &= \mu_i^{AP}(S_t^X) + \left[\frac{1}{\psi}\mathbb{E}_t[e_1\Upsilon(S_{t+1}^X)], 0\right]X_t^B + E_t\left[\pi_{t+1}\right]. \end{aligned}$$

The first to second line uses the log normality assumption, the second to third line uses the fact that stochastic volatility contribute very little to the short-rate, and the third to fourth line rearranges parameter values such that the short-rate is expressed in terms of X_t^B and $E_t\left[\pi_{t+1}\right]$.

S_t^X and S_t^M are discrete-valued random variables that follow a two-state Markov chain,

$$\mathbb{P}_X = \left[\begin{array}{cc} p_{X_1} & 1 - p_{X_1} \\ 1 - p_{X_2} & p_{X_2} \end{array}\right], \quad \mathbb{P}_M = \left[\begin{array}{cc} p_{M_1} & 1 - p_{M_1} \\ 1 - p_{M_2} & p_{M_2} \end{array}\right],$$

where X_1 (X_2) stands for negative (positive) correlation regime and M_1 (M_2) stands for

active (passive) monetary policy regime. For notational convenience, define

$$S_t = \begin{cases} 1 & \text{if} \quad S_t^X = X_1 \quad \text{and} \quad S_t^M = M_1 \\ 2 & \text{if} \quad S_t^X = X_1 \quad \text{and} \quad S_t^M = M_2 \\ 3 & \text{if} \quad S_t^X = X_2 \quad \text{and} \quad S_t^M = M_1 \\ 4 & \text{if} \quad S_t^X = X_2 \quad \text{and} \quad S_t^M = M_2 \end{cases}$$

and $\mathbb{P} = \mathbb{P}_X \otimes \mathbb{P}_M$.

Joint restriction of (2.32) and (2.33) gives

$$\begin{aligned} \tau_\pi(S_t^M)\pi_t &= E_t\left[\pi_{t+1}\right] + \underbrace{\left(\left[\frac{1}{\psi}\mathbb{E}_t[e_1\Upsilon(S_{t+1}^X)], 0\right] - \left[\tau_c(S_t^M), 1-\tau_\pi(S_t^M), 1, \tau_c(S_t^M)\right]\right)}_{\Lambda(S_t^X, S_t^M)} X_t^B \quad (2.34) \\ &= E_t\left[\pi_{t+1}\right] + \Lambda(S_t^X, S_t^M) X_t^B, \end{aligned}$$

assuming $\mu_i^{MP}(S_t^M) = \mu_i^{AP}(S_t^X)$. Since (2.34) is satisfied for each current state, I can express them as

$$\text{Diag}\left(\begin{bmatrix} \tau_\pi(S_t=1) \\ \tau_\pi(S_t=2) \\ \tau_\pi(S_t=3) \\ \tau_\pi(S_t=4) \end{bmatrix}\right) \times \begin{bmatrix} \pi_t(S_t=1) \\ \pi_t(S_t=2) \\ \pi_t(S_t=3) \\ \pi_t(S_t=4) \end{bmatrix} = \begin{bmatrix} E\left[\pi_{t+1}|S_t=1\right] \\ E\left[\pi_{t+1}|S_t=2\right] \\ E\left[\pi_{t+1}|S_t=3\right] \\ E\left[\pi_{t+1}|S_t=4\right] \end{bmatrix} + \begin{bmatrix} \Lambda(S_t=1) \\ \Lambda(S_t=2) \\ \Lambda(S_t=3) \\ \Lambda(S_t=4) \end{bmatrix} X_t. \quad (2.35)$$

In a slight abuse of notation, I use (i) to denote the current state instead of $(S_t = i)$ for i=1,2,3,4. From (2.27), observe that

$$\begin{bmatrix} \Lambda(1) \\ \Lambda(2) \\ \Lambda(3) \\ \Lambda(4) \end{bmatrix} = \mathbb{P} \times \begin{bmatrix} \frac{1}{\psi}e_1\Upsilon(1) & 0 \\ \frac{1}{\psi}e_1\Upsilon(2) & 0 \\ \frac{1}{\psi}e_1\Upsilon(3) & 0 \\ \frac{1}{\psi}e_1\Upsilon(4) & 0 \end{bmatrix} - \begin{bmatrix} \tau_c(1) & 1-\tau_\pi(1) & 1 & \tau_c(1) \\ \tau_c(2) & 1-\tau_\pi(2) & 1 & \tau_c(2) \\ \tau_c(3) & 1-\tau_\pi(3) & 1 & \tau_c(3) \\ \tau_c(4) & 1-\tau_\pi(4) & 1 & \tau_c(4) \end{bmatrix}. \quad (2.36)$$

I posit regime-dependent linear solutions of the form as in Davig and Leeper (2007).

$$\begin{bmatrix} \pi_t(1) \\ \pi_t(2) \\ \pi_t(3) \\ \pi_t(4) \end{bmatrix} = \begin{bmatrix} \Gamma(1) \\ \Gamma(2) \\ \Gamma(3) \\ \Gamma(4) \end{bmatrix} X_t^B \quad (2.37)$$

where $\Xi(i) = \begin{bmatrix} \Gamma_{x,c}(i) & \Gamma_{x,\pi}(i) & \Gamma_{x,m}(i) & \Gamma_\eta(i) \end{bmatrix}$ for i=1,2,3,4.

Necessary and Sufficient Conditions for the Existence of a Unique Bounded Solution. According to Proposition 2 of Davig and Leeper (2007), there exists a unique bounded solution if the following conditions are satisfied:

1. $\tau_\pi(i) > 0$, for i=1,2,3,4,

2. All the eigenvalues of $\left(\begin{bmatrix} \tau_\pi(1) & 0 & 0 & 0 \\ 0 & \tau_\pi(2) & 0 & 0 \\ 0 & 0 & \tau_\pi(3) & 0 \\ 0 & 0 & 0 & \tau_\pi(4) \end{bmatrix}^{-1} \times \mathbb{P} \right)$ lie inside the unit circle.

Solution. Substituting (2.37) to (2.35) yields

$$\begin{bmatrix} \tau_\pi(1) & 0 & 0 & 0 \\ 0 & \tau_\pi(2) & 0 & 0 \\ 0 & 0 & \tau_\pi(3) & 0 \\ 0 & 0 & 0 & \tau_\pi(4) \end{bmatrix} \begin{bmatrix} \Gamma(1) \\ \Gamma(2) \\ \Gamma(3) \\ \Gamma(4) \end{bmatrix} X_t^B = \mathbb{P} \times \begin{bmatrix} \Gamma(1)\Upsilon(1) \\ \Gamma(2)\Upsilon(2) \\ \Gamma(3)\Upsilon(3) \\ \Gamma(4)\Upsilon(4) \end{bmatrix} X_t^B + \begin{bmatrix} \Lambda(1) \\ \Lambda(2) \\ \Lambda(3) \\ \Lambda(4) \end{bmatrix} X_t^B \quad (2.38)$$

Analytical expressions for $\Gamma(i)$s are quite difficult to interpret, but are easily obtained from solving (2.38).

Nominal Bond Prices

Define $m_{t+1}^{\$} = m_{t+1} - \pi_{t+1}$. Let $P_{n,t}$ be the price at date t of a nominal bond with n periods to maturity. Conjecture that $p_{n,t}$ depends on the regime S_t and the current state variables,

$$p_{n,t} = C_{n,0}(S_t) + C_{n,1}(S_t)X_t + C_{n,2}(S_t)\Sigma_t \quad (2.39)$$

where $C_{n,1}(S_t) = \begin{bmatrix} C_{n,1,c}(S_t) & C_{n,1,\pi}(S_t) & C_{n,1,m}(S_t) \end{bmatrix}$ and $C_{n,2}(S_t) = \begin{bmatrix} C_{n,2,c}(S_t) & C_{n,2,\pi}(S_t) \end{bmatrix}$.

Exploit the law of iterated expectations

$$P_{n,t} = E_t\Big(E[\exp(m_{t+1}^{\$} + p_{n-1,t+1})|I_{t+1}] \Big)$$

and log-linearization to solve for $p_{n,t}$

$$p_{n,t} \approx \sum_{j=1}^{4} \mathbb{P}_{ij} \log\Big(E[\exp(m_{t+1}^{\$} + p_{n-1,t+1}^{\$})|S_t = i, S_{t+1} = j] \Big).$$

The solution to (2.39) is

$$
\begin{aligned}
C_{n,1}(i) &= \sum_{j=1}^{4} \mathbb{P}_{ij} \Big(C_{n-1,1}(j) - \frac{1}{\psi} e_1 - \Gamma_x(j) \Big) \Upsilon(j) \\
C_{n,2}(i) &= \sum_{j=1}^{4} \mathbb{P}_{ij} \Bigg(C_{n-1,2}(j) \Phi_\nu + (\theta - 1) \left\{ \kappa_{1,c} A_2(j) \Phi_\nu - A_2(i) \right\} \\
&+ \frac{1}{2} \begin{bmatrix} \left\{ (C_{n-1,1}(j) - \gamma e_1 - \Gamma_x(j) + (\theta-1)\kappa_{1,c} A_1(j)) \cdot \Omega(j) e_1' \right\}^2 \\ \left\{ (C_{n-1,1}(j) - \gamma e_1 - \Gamma_x(j) + (\theta-1)\kappa_{1,c} A_1(j)) \cdot \Omega(j) e_2' \right\}^2 \end{bmatrix}' \Bigg) \\
C_{n,0}(i) &= \sum_{j=1}^{4} \mathbb{P}_{ij} \Bigg(\theta \log \delta + (\theta - 1) \left\{ \kappa_{0,c} + \kappa_{1,c} A_0(j) + \kappa_{1,c} A_2(j) \Phi_\mu \right\} - (\theta - 1) A_0(i) - \gamma \mu - \mu_\pi \\
&+ C_{n-1,0}(j) + C_{n-1,2}(j) \Phi_\mu + \frac{1}{2} \bar{\sigma}_c^2 (\Gamma_\eta(j) + \gamma)^2 + \frac{1}{2} \bar{\sigma}_\pi^2 \\
&+ \frac{1}{2} \left\{ (C_{n-1,2,c}(j) + (\theta - 1) \kappa_{1,c} A_{2,c}(j)) \sigma_{w_c} \right\}^2 + \frac{1}{2} \left\{ (C_{n-1,2,\pi}(j) + (\theta - 1) \kappa_{1,c} A_{2,\pi}(j)) \sigma_{w_\pi} \right\}^2 \\
&+ \frac{1}{2} \left\{ (C_{n-1,1}(j) - \gamma e_1 - \Gamma_x(j) + (\theta - 1) \kappa_{1,c} A_1(j)) \cdot \Omega(j) e_3' \right\}^2 \sigma_m(j)^2 \Bigg),
\end{aligned}
$$

with initial conditions $C_{0,0}(i) = 0$, $C_{0,1}(i) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$, and $C_{0,2}(i) = \begin{bmatrix} 0 & 0 \end{bmatrix}$ for $i \in \{1, 2, 3, 4\}$.

Chapter 3

Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach

3.1 Introduction

Financial economists seek to understand the sources underlying risk and return in the economy. In equilibrium models this endeavor hinges on the preference specification and the joint dynamics of cash flows, which in an endowment economy correspond to consumption and dividends. There are many equilibrium models that appeal to low-frequency components in these cash flows as well as important time variation in the fundamentals (e.g., models of long-run risks (LRR) as in Bansal and Yaron (2004), and models of rare disasters as in Barro (2009)). Identifying both of these components is challenging. To measure the small persistent component in, say, consumption and dividend growth one would want the longest span of data. On the other hand, to estimate the time variation in second moments of cash flows one would ideally like to use high-frequency data. The empirical analysis is constrained by the availability of consumption data. For the U.S., the longest span of available data for consumption growth is at the annual frequency starting in 1929. The highest-frequency consumption data is available at the monthly frequency from 1959. To exploit all the available information in mixed-frequency data, this paper develops a Bayesian state-space model

that prominently features stochastic volatility and time-aggregates consumption whenever it is observed only at a low frequency.

Our state-space model is designed to capture the joint dynamics of consumption, dividend growth, and asset returns. Building on the work of Bansal and Yaron (2004), the core of our model consists of an endowment economy that is, in part, driven by a common predictable component for consumption and dividend growth. The economy delivers a stochastic discount factor that is used to price equities and a risk-free asset. Our model distinguishes itself from the existing LRR literature in several important dimensions. First, our state-space representation contains measurement equations that time-aggregate consumption to the observed frequency, yet allow us to maintain the likelihood representation (see Bansal, Kiku, and Yaron (2012b) for a generalized methods-of-moments (GMM) approach using time aggregation). Our measurement-error specification accounts for different types of measurement errors at monthly and annual frequencies while respecting the constraint that monthly growth rates have to be consistent with annual growth rates.

Second, we generalize the volatility dynamics of Bansal and Yaron (2004)'s model specification by allowing for three separate volatility processes — one capturing long-run consumption innovations, one capturing short-run consumption innovations, and a separate process for dividend dynamics. We do so since our estimation procedure, which focuses on the joint distribution of consumption, dividends, and asset prices, requires separate stochastic volatility processes to fit the data. Third, we specify an additional process for variation in the time rate of preference (see Albuquerque, Eichenbaum, and Rebelo (2012)), which generates risk-free rate variation that is independent of cash flows and leads to an improved fit for the risk-free rate.

The estimation of the state-space model generates several important empirical findings. First, we find strong evidence for a small predictable component in consumption growth.

This evidence consists of two parts. We begin by estimating the state-space model on cash flow growth data only. Our carefully specified measurement-error model for cash flow data allows us to measure this component which otherwise is difficult to detect. We then proceed by adding asset return data to the estimation and, in line with the existing LRR literature, find even stronger evidence for this predictable component. The Bayesian approach allows us to characterize the uncertainty about the persistence of the conditional mean growth process. We find that in spite of using a prior with a mean of 0.9 and a standard deviation of 0.5 our estimation yields a posterior distribution that is tightly centered around 0.99. Second, our estimated measurement errors for consumption growth are consistent with the common view (see Wilcox (1992)) that consumption growth is measured more precisely at an annual rather than monthly frequency.

Third, all three stochastic volatility processes display significant time variation yet behave distinctly over time. The volatility processes partly capture heteroskedasticity of innovations, and in part they break some of the tight links that the model imposes on the conditional mean dynamics of asset prices and cash flows. This feature significantly improves the model implications for consumption and return predictability. As emphasized by the LRR literature, the volatility processes have to be very persistent in order to have significant quantitative effects on asset prices. An important feature of our estimation is that the likelihood focuses on conditional correlations between the risk-free rate and consumption — a dimension often not directly targeted in the literature. We show that because consumption growth and its volatility determine the risk-free rate dynamics, one requires another independent volatility process to account for the weak correlation between consumption growth and the risk-free rate. In the generalized specification of the model in which there are independent time rate of preference shocks, this correlation is further muted and the model fit for the dynamics of the risk-free rate is improved.

Fourth, it is worth noting that the median posterior estimate for risk aversion is 10-11 while it is around 1.5 for the intertemporal elasticity of substitution (IES). These estimates are consistent with the parameters highlighted in the LRR literature (see Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2012a), and Bansal, Kiku, and Yaron (2012b)). Fifth, at the estimated preference parameters and those characterizing the consumption and dividend dynamics, the model is able to successfully generate many key asset-pricing moments. In particular, as in the data, the posterior median for the equity premium generated by the model is 6.0%.

Our paper also contains a number of technical innovations. First, in the specification of our state-space model we follow the stochastic volatility literature and assume that volatilities evolve according to exponential Gaussian processes that guarantee nonnegativity. While the cash flows in our state-space model evolve exogenously, the law of motion of the financial variables is determined endogenously from the economic structure. In order to solve the model, we approximate the exponential Gaussian volatility processes by linear Gaussian processes such that the standard analytical solution techniques that have been widely used in the LRR literature can be applied. However, the approximation of the exponential volatility process is used only to derive the coefficients in the law of motion of the asset prices.

Second, we use a Markov chain Monte Carlo (MCMC) algorithm to generate parameter draws from the posterior distribution. This algorithm requires us to evaluate the likelihood function of our state-space model with a nonlinear filter. Due to the high-dimensional state space that arises from the mixed-frequency setting, this nonlinear filtering is a seemingly daunting task. We show how to exploit the partially linear structure of the state-space model to derive a very efficient sequential Monte Carlo (particle) filter.

Our paper is related to several strands of the literature. In terms of the LRR literature,

our paper is closely related to that of Bansal, Kiku, and Yaron (2012b) who utilize time aggregation and GMM to estimate the LRR model (see also Bansal, Gallant, and Tauchen (2007) for an approach using the efficient method of moments (EMM)). As noted above, our likelihood-based approach provides evidence which is broadly consistent with the results highlighted in that paper and other calibrated LRR models (see Bansal, Kiku, and Yaron (2012a)). Our likelihood function implicitly utilizes a broader set of moments than earlier GMM or EMM estimation approaches. These moments include the entire sequence of autocovariances as well as higher-order moments of the time series used in the estimation and let us measure the time path of the predictable component of cash flows as well as the time path of the innovation volatilities. Rather than asking the model to fit a few selected moments, we are raising the bar and force the model to track cash flow and asset return time series.

To implement Bayesian inference, we embed a particle-filter-based likelihood approximation into a Metropolis-Hastings algorithm as in Fernández-Villaverde and Rubio-Ramírez (2007) and Andrieu, Doucet, and Holenstein (2010). Since our state-space system is linear conditional on the volatility states, we can use Kalman-filter updating to integrate out a subset of the state variables. The genesis of this idea appears in the auxiliary particle filter of Pitt and Shephard (1999) and Chen and Liu (2000) and is recently discussed in Shephard (2013). Particle filter methods are also utilized in Johannes, Lochstoer, and Mou (2013), who estimate an asset pricing model in which agents have to learn about the parameters of the cash flow process from consumption growth data. While Johannes, Lochstoer, and Mou (2013) examine the role of parameter uncertainty for asset prices, which is ignored in our analysis, they use a more restrictive version of the cash flow process and do not utilize mixed-frequency observations.

Our state-space setup makes it relatively straightforward to utilize data that are avail-

able at different frequencies. The use of state-space systems to account for missing monthly observations dates back to at least Harvey (1989a) and has more recently been used in the context of dynamic factor models (see, e.g., Mariano and Murasawa (2003a) and Aruoba, Diebold, and Scotti (2009a)) and VARs (see, e.g., Schorfheide and Song (2012)). Finally, there is a growing and voluminous literature in macro and finance that highlights the importance of volatility for understanding the macroeconomy and financial markets (see, e.g., Bloom (2009), Fernández-Villaverde and Rubio-Ramírez (2011), Bansal, Kiku, and Yaron (2012a), and Bansal, Kiku, Shaliastovich, and Yaron (2013)). Our volatility specification that accommodates three processes further contributes to identifying the different uncertainty shocks in the economy.

The remainder of the paper is organized as follows. Section 3.2 introduces the model environment and describes the model solution. Section 3.3 presents the empirical state-space model and describes the estimation procedure. Section 3.4 discusses the empirical findings and Section 5.5 provides concluding remarks.

3.2 The Long-Run Risks (LRR) Model

Our baseline LRR model is described in Section 3.2.1. The solution of the model is outlined in Section 3.2.2. Section 3.2.3 presents a generalized version of the LRR with an exogenous shock to the time rate of preference.

3.2.1 Model Statement

We consider an endowment economy with a representative agent that has Epstein and Zin (1989) recursive preferences and maximizes her lifetime utility,

$$V_t = \max_{C_t} \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \big(\mathbb{E}_t [V_{t+1}^{1-\gamma}] \big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$

subject to budget constraint

$$W_{t+1} = (W_t - C_t)R_{c,t+1}$$

, where W_t is the wealth of the agent, $R_{c,t+1}$ is the return on all invested wealth, γ is risk aversion, $\theta = \frac{1-\gamma}{1-1/\psi}$, and ψ is intertemporal elasticity of substitution.

Following Bansal and Yaron (2004),we decompose consumption growth, $g_{c,t+1}$, into a persistent component, x_t, and a transitory component, $\sigma_{c,t}\eta_{c,t+1}$. The dynamics for the persistent conditional mean follow an AR(1) with its own stochastic volatility process. Dividend streams have levered exposures to both the persistent and transitory component in consumption which is captured by the parameters ϕ and π, respectively. We allow $\sigma_{d,t}\eta_{d,t+1}$ to capture idiosyncratic movements in dividend streams. Overall, the dynamics for the cash flows are

$$\begin{aligned} g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \\ x_{t+1} &= \rho x_t + \sigma_{x,t}\eta_{x,t+1} \\ g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1}, \end{aligned} \tag{3.1}$$

where the conditional volatilities evolve according to

$$\sigma_{i,t} = \varphi_i\bar{\sigma}\exp(h_{i,t}), \quad h_{i,t+1} = \rho_{h_i}h_{i,t} + \sigma_{h_i}\sqrt{1-\rho_{h_i}^2}w_{i,t+1}, \quad i = \{c, x, d\} \tag{3.2}$$

and the shocks are assumed to be

$$\eta_{i,t+1}, w_{i,t+1} \sim N(0,1), \quad i = \{c, x, d\}.$$

Relative to Bansal and Yaron (2004), the volatility dynamics contain three separate volatility processes. More importantly, the logarithm of the volatility process is assumed to be normal, which ensures that the standard deviation of the shocks remains positive at every point in time.

3.2.2 Solution

The Euler equation for any asset $r_{i,t+1}$ takes the form

$$E_t\left[\exp\left(m_{t+1}+r_{i,t+1}\right)\right]=1, \tag{3.3}$$

where $m_{t+1}=\theta\log\delta-\frac{\theta}{\psi}g_{c,t+1}+(\theta-1)r_{c,t+1}$ is the log of the real stochastic discount factor (SDF), and $r_{c,t+1}$ is the log return on the consumption claim. We reserve $r_{m,t+1}$ for the log market return — the return on a claim to the dividend cash flows. Given the cash flow dynamics in (3.1) and the Euler equation (3.3), we derive asset prices using the approximate analytical solution described in Bansal, Kiku, and Yaron (2012a) which utilizes the Campbell and Shiller (1988a) log-linear approximation for returns.

However, since the volatility processes in (3.2) do not follow normal distributions, an analytical expression to (3.3) is infeasible. To accommodate an analytical solution, we utilize a linear approximation to (3.2) and express volatility in (3.4) as a process that follows Gaussian dynamics:

$$\begin{aligned}
\sigma_{i,t}^2-(\varphi_i\bar{\sigma})^2 &= 2(\varphi_i\bar{\sigma})^2h_{i,t}+O(|h_{i,t}^2|), \quad h_{i,t+1}=\rho_{h_i}h_{i,t}+\sigma_{h_i}\sqrt{1-\rho_{h_i}^2}w_{i,t+1}\\
\sigma_{i,t+1}^2 &\approx (\varphi_i\bar{\sigma})^2(1-\rho_{h_i})+\rho_{h_i}\sigma_{i,t}^2+(2(\varphi_i\bar{\sigma})^2\sigma_{h_i}\sqrt{1-\rho_{h_i}^2})w_{i,t+1}\\
&= (\varphi_i\bar{\sigma})^2(1-\nu_i)+\nu_i\sigma_{i,t}^2+\sigma_{w_i}w_{i,t+1}, \quad i=\{c,x,d\}.
\end{aligned} \tag{3.4}$$

The analytical solution afforded via this pseudo-volatility process is important since it facilitates estimation (see details below).

The solution to the log price-consumption ratio follows, $z_t=A_0+A_1x_t+A_{2,c}\sigma_{c,t}^2+A_{2,x}\sigma_{x,t}^2$. As discussed in Bansal and Yaron (2004), $A_1=\frac{1-\frac{1}{\psi}}{1-\kappa_1\rho}$, the elasticity of prices with respect to growth prospects, will be positive whenever the IES, ψ, is greater than 1. Further, the elasticity of z_t with respect to the two volatility processes $\sigma_{c,t}^2$ and $\sigma_{x,t}^2$ is $\frac{\theta}{2}\frac{(1-\frac{1}{\psi})^2}{1-\kappa_1\nu_c}$ and $\frac{\theta}{2}\frac{(\kappa_1A_1)^2}{1-\kappa_1\nu_x}$ respectively; both will be negative — namely, prices will decline with uncertainty — whenever θ is negative.

State prices in the economy are reflected in the innovations to the stochastic discount factor (SDF),

$$m_{t+1} - E_t\left[m_{t+1}\right] = \underbrace{\lambda_c \sigma_{c,t} \eta_{c,t+1}}_{\text{short-run consumption risk}} + \underbrace{\lambda_x \sigma_{x,t} \eta_{x,t+1}}_{\text{long-run growth risk}} + \underbrace{\lambda_{w_x} \sigma_{w_x} w_{x,t+1} + \lambda_{w_c} \sigma_{w_c} w_{c,t+1}}_{\text{volatility risks}},$$

where the derivation and λs are given in Appendix 3.6.1. It is instructive to note that $\lambda_c = -\gamma$, $\lambda_x = \frac{-(\gamma - \frac{1}{\psi})\kappa_1}{1-\kappa_1 \rho}$ (and λ_{w_c} and λ_{w_x}) is negative (positive) whenever preferences exhibit early resolution of uncertainty $\gamma > 1/\psi$. Furthermore the λs (except λ_c) will be zero when preferences are time separable, namely, when $\theta = 1$.

Risk premia are determined by the negative covariation between the innovations to returns and the innovations to the SDF. It follows that the risk premium for the market return, $r_{m,t+1}$, is

$$\begin{aligned} E_t(r_{m,t+1} - r_{f,t}) + \frac{1}{2} var_t(r_{m,t+1}) &= -cov_t(m_{t+1}, r_{m,t+1}) \\ &= \underbrace{\beta_{m,c} \lambda_c \sigma^2_{c,t}}_{\text{short-run risk}} + \underbrace{\beta_{m,x} \lambda_x \sigma^2_{x,t}}_{\text{long-run growth risk}} \\ &\quad + \underbrace{\beta_{m,w_x} \lambda_{w_x} \sigma^2_{w_x} + \beta_{m,w_c} \lambda_{w_c} \lambda_c \sigma^2_{w_c}}_{\text{volatility risks}}, \end{aligned} \tag{3.5}$$

where the βs are given in Appendix 3.6.1 and reflect the exposures of the market return to the underlying consumption risks. Equation (3.5) highlights that the conditional equity premium can be attributed to (i) short-run consumption growth, (ii) long-run growth, (iii) short-run and long-run volatility risks.

A key variable for identifying the model parameters is the risk-free rate. Under the assumed dynamics in (3.1), the risk-free rate is affine in the state variables and follows

$$r_{f,t} = B_0 + B_1 x_t + B_{2,c} \sigma^2_{c,t} + B_{2,x} \sigma^2_{x,t},$$

where the Bs are derived in Appendix 3.6.1. It is worth noting that $B_1 = \frac{1}{\psi} > 0$ and the risk-free rate rises with good economic prospects, while under $\psi > 1$, $\gamma > 1$ and whenever preferences exhibit early resolution of uncertainty, $B_{2,c} = -\frac{1}{2}(\frac{\gamma - 1}{\psi} + \gamma)$ and $B_{2,x} = -\frac{(1-\frac{1}{\psi})(\gamma - \frac{1}{\psi})\kappa_1^2}{2(1-\kappa_1 \rho)^2}$ are negative so the risk-free rate declines with a rise in economic uncertainty.

3.2.3 Generalized Model

In this section we augment the baseline model, as highlighted in Albuquerque, Eichenbaum, and Rebelo (2012), by allowing for a preference shock to the time rate of preference. Specifically, now the utility function contains a time rate of preference shock, λ_t, so the lifetime utility is

$$V_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta\big(\mathbb{E}_t[V_{t+1}^{1-\gamma}]\big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}.$$

The resulting SDF equals the SDF described in equation (3.5) plus the term $\theta x_{\lambda,t}$, where $x_{\lambda,t} = \lambda_{t+1}/\lambda_t$ is the growth rate of the preference shock, which is assumed to follow an AR(1) process with persistence parameter ρ_λ (see Appendix 3.6.1 for derivation of this augmented SDF). Since $x_{\lambda,t}$ is known at time t, the risk-free rate will incorporate its values and consequently allow this generalized model to fit the risk-free rate dynamics better than the benchmark model.

3.3 State-Space Representation of the LRR Model

In order to conduct our empirical analysis, we cast the LRR model of Section 3.2 into state-space form. The state-space representation consists of a measurement equation that expresses the observables as a function of underlying state variables and a transition equation that describes the law of motion of the state variables. The measurement equation takes the form

$$y_{t+1} = A_{t+1}\big(D + Z s_{t+1} + Z^v s_{t+1}^v + \Sigma^u u_{t+1}\big), \quad u_{t+1} \sim iidN(0, I). \tag{3.6}$$

In our application, y_{t+1} consists of consumption growth, dividend growth, market returns, and the risk-free rate. The vector s_{t+1} stacks state variables that characterize the level of cash flows. The vector s_{t+1}^v is a function of the log volatilities of cash flows, h_t and h_{t+1}, in (3.2). Finally, u_{t+1} is a vector of measurement errors and A_{t+1} is a selection matrix that

accounts for deterministic changes in the data availability. The solution of the LRR model sketched in Section 3.2.2 provides the link between the state variables and the observables y_{t+1}.

The state variables themselves follow vector autoregressive processes of the form

$$s_{t+1} = \Phi s_t + v_{t+1}(h_t), \quad h_{t+1} = \Psi h_t + \Sigma_h w_{t+1}, \quad w_{t+1} \sim iidN(0, I), \tag{3.7}$$

where $v_{t+1}(\cdot)$ is an innovation process with a variance that is a function of the log volatility process h_t and w_{t+1} is the innovation of the stochastic volatility process. Roughly speaking, the vector s_{t+1} consists of the persistent cash flow component x_t (see (3.1)) as well as $x_{\lambda,t}$ in the generalized model of Section 3.2.3. However, in order to express the observables y_{t+1} as a linear function of s_{t+1} and to account for potentially missing observations, it is necessary to augment s_{t+1} by lags of x_t and $x_{\lambda,t}$ as well as the innovations for the cash flow process. Since the details are cumbersome and at this stage non essential, a precise definition of s_{t+1} is relegated to the appendix.

A novel feature of our empirical analysis is the mixed-frequency approach. While dividend growth, equity return, and risk-free rate data are available at a monthly frequency from 1929 onwards, consumption data prior to 1959 are not available at a monthly frequency. Moreover, post-1959 monthly consumption growth data are subject to sizeable measurement errors, which is why many authors prefer to estimate consumption-based asset pricing models based on time-aggregated data. Our state-space approach avoids the loss of information due to time aggregation, yet we can allow for imprecisely measured consumption data at a monthly frequency. We discuss the measurement equations for consumption in Section 3.3.1 and the other observables in Section 3.3.2. Section 3.3.3 describes the implementation of Bayesian inference.

3.3.1 A Measurement Equation for Consumption

In our empirical analysis we use annual consumption growth rates prior to 1959 and monthly consumption growth rates subsequently.[1] The measurement equation for consumption in our state-space representation has to be general enough to capture two features: (i) the switch from annual to monthly observations in 1959, and (ii) measurement errors that are potentially larger at a monthly frequency than an annual frequency. To describe the measurement equation for consumption growth data, we introduce some additional notation. We use the superscript o to distinguish observed consumption and consumption growth, C_t^o and $g_{c,t}^o$, from model-implied consumption and consumption growth, C_t and $g_{c,t}$. Moreover, we represent the monthly time subscript t as $t = 12(j-1)+m$, where $m = 1, \ldots, 12$. Here j indexes the year and m the month within the year.

We define annual consumption as the sum of monthly consumption over the span of one year, i.e.:

$$C_{(j)}^a = \sum_{m=1}^{12} C_{12(j-1)+m}.$$

Log-linearizing this relationship around a monthly value C_* and defining lowercase c as percentage deviations from the log-linearization point, i.e., $c = \log C/C_*$, we obtain

$$c_{(j)}^a = \frac{1}{12} \sum_{m=1}^{12} c_{12(j-1)+m}.$$

Thus, monthly consumption growth rates can be defined as

$$g_{c,t} = c_t - c_{t-1}$$

and annual growth rates are given by

$$g_{c,(j)}^a = c_{(j)}^a - c_{(j-1)}^a = \sum_{\tau=1}^{23} \left(\frac{12 - |\tau - 12|}{12} \right) g_{c,12j-\tau+1}. \tag{3.8}$$

[1] In principle we could utilize the quarterly consumption growth data from 1947 to 1959, but we do not in this version of the paper.

We assume a multiplicative *iid* measurement-error model for the level of annual consumption, which implies that, after taking log differences,

$$g^{a,o}_{c,(j)} = g^{a}_{c,(j)} + \sigma^{a}_{\epsilon}\big(\epsilon^{a}_{(j)} - \epsilon^{a}_{(j-1)}\big). \tag{3.9}$$

Moreover, consistent with the practice of the Bureau of Economic Analysis, we assume that the levels of monthly consumption are constructed by distributing annual consumption over the 12 months of a year. This distribution is based on an observed monthly proxy series z_t that is assumed to provide a noisy measure of monthly consumption. The monthly levels of consumption are determined such that the growth rates of monthly consumption are proportional to the growth rates of the proxy series and monthly consumption adds up to annual consumption. A measurement-error model that is consistent with this assumption is the following:

$$\begin{aligned} g^{o}_{c,12(j-1)+1} &= g_{c,12(j-1)+1} + \sigma_{\epsilon}\big(\epsilon_{12(j-1)+1} - \epsilon_{12(j-2)+12}\big) \\ &\quad -\frac{1}{12}\sum_{m=1}^{12}\sigma_{\epsilon}\big(\epsilon_{12(j-1)+m} - \epsilon_{12(j-2)+m}\big) + \sigma^{a}_{\epsilon}\big(\epsilon^{a}_{(j)} - \epsilon^{a}_{(j-1)}\big) \\ g^{o}_{c,12(j-1)+m} &= g_{c,12(j-1)+m} + \sigma_{\epsilon}\big(\epsilon_{12(j-1)+m} - \epsilon_{12(j-1)+m-1}\big), \quad m = 2,\ldots,12 \end{aligned} \tag{3.10}$$

The term $\epsilon_{12(j-1)+m}$ can be interpreted as the error made by measuring the level of monthly consumption through the monthly proxy variable, that is, in log deviations $c_{12(j-1)+m} = z_{12(j-1)+m} + \epsilon_{12(j-1)+m}$. The summation of monthly measurement errors in the second line of (3.10) ensures that monthly consumption sums up to annual consumption. It can be verified that converting the monthly consumption growth rates into annual consumption growth rates according to (3.8) averages out the measurement errors and yields (3.9).

We operate under the assumption that the agents in the model observe consumption growth, dividend growth, and asset returns in every period. As econometricians who are estimating the model, we have to rely on the statistical agency to release the consumption growth data. While the statistical agency may have access to the monthly proxy series z_t

in real time, it can only release the monthly consumption series that is consistent with the annual observations on consumption at the end of each year. Thus, for months $m = 1, \ldots, 11$ the vector $y_{12(j-1)+m}$ in (3.6) does not contain any observations on consumption growth. At the end of each year, in month $m = 12$, the vector $y_{12(j-1)+m}$ contains the 12 monthly growth rates of year j and (3.10) provides the portion of the measurement equation for the consumption data. The vector s_t has to contain sufficiently many lags of the model states as well as some lagged measurement errors such that it is possible to write (3.10) as a linear function of s_t. For the earlier part of the sample in which monthly consumption growth observations are not available, (3.10) is replaced by (3.8) and (3.9). The matrix M_t in (3.6) adapts the system to the availability of consumption data and the changing dimension of the vector y_t. Further details are provided in the appendix.

3.3.2 Measurement Equations for Dividend Growth and Asset Returns

It is reasonable to believe that consumption measurement errors are large, but those for financial variables (e.g., dividend streams, market returns, risk-free rates) are negligible. However, to be chary, we introduce idiosyncratic components for dividend growth, market returns, and the risk-free rate as well:

$$\begin{aligned} g^o_{d,t+1} &= g_{d,t+1} + \sigma^d_\epsilon \epsilon_{d,t+1} \\ r^o_{m,t+1} &= r_{m,t+1} + \sigma^{r_m}_\epsilon \epsilon_{r_m,t+1} \\ r^o_{f,t+1} &= r_{f,t+1} + \sigma^{r_f}_\epsilon \epsilon_{r_f,t+1}. \end{aligned} \tag{3.11}$$

In the subsequent empirical analysis we consider a version of the model in which only the risk-free rate is measured with error, i.e., $\sigma^d_\epsilon = 0$, $\sigma^{r_m}_\epsilon = 0$, $\sigma^{r_f}_\epsilon > 0$. We believe that aggregate dividend growth and stock market data are cleanly measured and we do not want to deviate too much from the original Bansal and Yaron (2004) framework.

3.3.3 Bayesian Inference

Equations (3.6) and (3.7) define a nonlinear state-space system in which the size of the vector of observables y_t changes in a deterministic manner. The system matrices of the system are functions for the parameter vector

$$\Theta = \left(\delta, \psi, \gamma, \rho, \phi, \varphi_x, \varphi_d, \bar{\sigma}, \mu, \mu_d, \pi, \sigma_\epsilon, \sigma_\epsilon^a, \rho_\lambda, \sigma_\lambda, \sigma_\epsilon^{r_f}, \rho_{h_c}, \sigma_{h_c}, \rho_{h_x}, \sigma_{h_x}, \rho_{h_d}, \sigma_{h_d} \right). \quad (3.12)$$

We will use a Bayesian approach to make inference about Θ and to study the implications of our model. Bayesian inference requires the specification of a prior distribution $p(\Theta)$ and the evaluation of the likelihood function $p(Y|\Theta)$. The posterior can be expressed as

$$p(\Theta|Y) = \frac{p(Y|\Theta)p(\Theta)}{p(Y)}. \quad (3.13)$$

We will use MCMC methods to generate a sequence of draws $\{\Theta^{(s)}\}_{s=1}^{n_{sim}}$ from the posterior distribution.

To generate the draws from the posterior distribution, we need to be able to numerically evaluate the prior density and the likelihood function $p(Y|\Theta)$. Since our state-space system is nonlinear, it is not possible to evaluate the likelihood function using the Kalman filter. Instead, we use a sequential Monte Carlo procedure also known as particle filter. The particle filter creates an approximation $\hat{p}(Y|\Theta)$ of the likelihood function $p(Y|\Theta)$. It has been shown in Andrieu, Doucet, and Holenstein (2010) that the use of $\hat{p}(Y|\Theta)$ in MCMC algorithms can still deliver draws from the actual posterior $p(\Theta|Y)$ because these approximation errors essentially average out as the Markov chain progresses.

Capturing the annual release schedule for the monthly consumption data described in Section 3.3.1 requires a high-dimensional state vector s_t. This creates a computational challenge for the evaluation of the likelihood function because the accuracy of particle filter approximations tends to decrease as the dimension of the latent state vector increases. In order to obtain a computationally efficient filter, we exploit the fact that our state-space

model is linear and Gaussian conditional on the volatility states (h_t, h_{t-1}). We use a swarm of particles to represent the distribution of $(h_t, h_{t-1})|Y_{1:t}$ and employ the Kalman filter to characterize the conditional distribution of $s_t|(h_t, h_{t-1}, Y_{1:t})$. This idea has been used by Chen and Liu (2000) and more recently by Shephard (2013). A full description of the particle filter is provided in the appendix. We embed the likelihood approximation in a fairly standard random-walk Metropolis algorithm that is widely used in the DSGE model literature; see for instance Del Negro and Schorfheide (2010).

3.4 Empirical Results

The data set used in the empirical analysis is described in Section 3.4.1. The subsequent analysis is divided into two parts. In Section 3.4.2 we use consumption and dividend growth data to estimate the persistent components in conditional mean and volatility dynamics of cash flows. In Section 3.4.3 we include the market return and risk-free rate data in the estimation and analyze the asset pricing implications of our model.

3.4.1 Data

We use the per capita series of real consumption expenditure on nondurables and services from the NIPA tables available from the Bureau of Economic Analysis. Annual observations are available from 1929 to 2011, quarterly from 1947:Q1 to 2011:Q4, and monthly from 1959:M1 to 2011:M12. We also use monthly observations of returns, dividends, and prices of the CRSP value-weighted portfolio of all stocks traded on the NYSE, AMEX, and NASDAQ. Price and dividend series are constructed on the per share basis as in Campbell and Shiller (1988b) and Hodrick (1992). The stock market data are converted to real using the consumer price index (CPI) from the Bureau of Labor Statistics. Finally, the ex ante real risk-free rate is constructed as a fitted value from a projection of the ex post real rate on the current nominal yield and inflation over the previous year. To run the predictive regression, we use

Table 3.1: Descriptive Statistics - Data Moments

Annual Frequency: 1930 to 2011					
	Δc	Δd	r_m	r_f	pd
Mean	1.83	0.98	5.43	0.46	3.36
StdDev	2.19	11.24	19.98	2.78	0.43
AC1	0.48	0.21	0.01	0.72	0.90
AC2	0.18	-0.21	-0.16	0.40	0.81
AC3	-0.07	-0.15	0.01	0.31	0.75
Corr	1.00	0.56	0.12	-0.26	0.07
Monthly Frequency: 1929:M1 to 2011:M12					
	Δc	Δd	r_m	r_f	pd
Mean	-	0.09	0.45	0.04	3.36
StdDev	-	1.68	5.50	0.24	0.44
AC1	-	0.20	0.11	0.98	0.99
Monthly Frequency: 1959:M2 to 2011:M12					
	Δc	Δd	r_m	r_f	pd
Mean	0.16	0.11	0.43	0.10	3.57
StdDev	0.34	1.26	4.55	0.14	0.39
AC1	-0.16	-0.01	0.10	0.96	0.99
Corr	1.00	0.04	0.16	0.13	0.00

Notes: We report descriptive statistics for aggregate consumption growth (Δc), dividend growth (Δd), log returns of the aggregate stock market (r_m), the log risk-free rate (r_f), and log price-dividend ratio (pd). It shows mean, standard deviation, sample autocorrelations up to order three, and correlation with aggregate consumption growth. Means and standard deviations are expressed in percentage terms.

monthly observations on the three-month nominal yield from the CRSP Fama Risk Free Rate tapes and CPI series. A more detailed explanation of the data sources is provided in Appendix 3.6.2. Growth rates of consumption and dividends are constructed by taking the first difference of the corresponding log series. The time-series span of the stock market data and the risk-free rate is from 1929:M1 to 2011:M12.

Table 3.1 presents descriptive statistics for aggregate consumption growth, dividend growth, aggregate stock market returns, the risk-free rate, and the log price-dividend ratio. The statistics are computed for a sample of annual observations from 1930 to 2011, a sample of monthly observations from 1929:M1 to 2011:M12, and a sample of monthly observations

from 1959:M2 to 2011:M12. Consumption data is only available for the shorter of the two monthly samples. For our subsequent analysis, a few features of the data turn out to be important. First, the sample first autocorrelation function of monthly and annual consumption have different signs. Second, consumption and dividend growth are highly correlated at the low (annual) frequency but not at the high (monthly) frequency. Third, the sample standard deviations for the long monthly sample starting in 1929:M1 are larger than the sample standard deviations for the post-1958 sample.

3.4.2 Estimation with Cash Flow Data Only

We begin by estimating the state-space model described in Section 3.3 based only on consumption and dividend growth data, dropping market returns and the risk-free rate from the measurement equation. We employ the mixed-frequency approach by utilizing annual consumption growth data from 1929 to 1959 and monthly data from 1960:M1 to 2011:M12.

Prior Distribution. We begin with a brief discussion of the prior distribution for the parameters of the cash flow process specified in (3.1) and (3.2). In general, our prior attempts to restrict parameter values to economically plausible magnitudes. The judgment of what is *economically plausible* is, of course, informed by some empirical observations, in the same way the choice of the model specification is informed by empirical observations. Percentiles of marginal prior distributions are reported in Table 3.2.

The prior 90% credible intervals for average annualized consumption and dividend growth range from approximately $\pm$ 7%. In view of the sample statistics reported in Table 3.1, this range is fairly wide and agnostic. The prior distribution for the persistence of the predictable cash flow growth component x_t is a normal distribution centered at 0.9 with a standard deviation of 0.5, truncated to the interval $(-1, 1)$. The corresponding 90% credible interval ranges from -0.1 to 0.97, encompassing values that imply *iid* cash flow growth dynamics as well as very persistent local levels. The priors for ϕ and π, parameters

that determine the comovement of cash flows, are centered at zero and have large variances. $\bar{\sigma}$ is, roughly speaking, the average standard deviation of the *iid* component of consumption growth. At an annualized rate, our 90% credible interval ranges from 1.2% to 7.2%. For comparison, the sample standard deviation of annual consumption growth and annualized monthly consumption growth are approximately 2% and 4%, respectively (see Table 3.1).

The parameters φ_d and φ_x capture the magnitude of innovations to dividend growth and the persistent cash flow component relative to the magnitude of consumption growth innovations. The prior for the former covers the interval 0.2 to 12, whereas the prior for the latter captures the interval 0 to 0.11. Thus, *a priori* we expect dividends to be more volatile than consumption and the persistent component of cash flow growth to be much smoother than the *iid* component. Our prior interval for the persistence of the volatility processes ranges from -0.1 to 0.97 and the prior for the standard deviation of the volatility process implies that the volatility may fluctuate either relatively little, over the range of 0.67 to 1.5 times the average volatility, or substantially, over the range of 0.1 to 7 times the average volatility.

Posterior Distribution. Percentiles of the posterior distribution are also reported in Table 3.2. The most important result for the subsequent analysis of the asset pricing implications of the LRR model is the large estimate of ρ, the autocorrelation coefficient of the persistent cash flow component x_t. The posterior median of ρ is 0.97. Thus, according to our estimate, cash flow growth dynamics are very different from *iid* dynamics; the half-life of the persistent component is about three years; and the magnitude of the parameter estimate is quite close to the values used in the LRR literature (see Bansal, Kiku, and Yaron (2012a)).

At first glance, the large estimate of ρ may appear inconsistent with the negative sample autocorrelation of consumption growth and the near-zero autocorrelation of dividend

Table 3.2: Posterior Estimates: Cashflows Only

	Prior			Posterior				Prior			Posterior		
	Distr.	5%	95%	5%	50%	95%		Distr.	5%	95%	5%	50%	95%
	Consumption Process							Dividend Process					
μ_c	N	[-.006	.006]	.0014	.0016	.0019	μ_d	N	[-.007	.006]	.0000	.0006	.0013
ρ	N^T	[-0.08	0.97]	0.95	0.97	0.98	ϕ	N	[-13.1	13.4]	2.04	2.11	2.20
φ_x	G	[0.00	0.11]	0.17	0.20	0.22	π	N	[-1.68	1.63]	- 0.18	0.03	0.14
$\bar{\sigma}$	IG	[.001	.006]	.0021	.0024	.0026	φ_d	G	[0.22	11.9]	4.92	5.30	5.78
ρ_{h_c}	N^T	[-0.08	0.97]	.993	.995	.997	ρ_{h_d}	N^T	[-0.08	0.97]	0.83	0.89	0.94
σ_{h_c}	IG	[0.22	1.03]	0.31	0.39	0.49	σ_{h_d}	IG	[0.22	1.03]	0.47	0.53	0.61
ρ_{h_x}	N^T	[-0.08	0.97]	.979	.992	.998							
σ_{h_x}	IG	[0.22	1.03]	0.23	0.43	1.07							

Notes: We utilize the mixed-frequency approach in the estimation: For consumption we use annual data from 1929 to 1959 and monthly data from 1960:M1 to 2011:M12; we use monthly dividend growth data from 1929:M1 to 2011:M12. For consumption we adopt the measurement error model of Section 3.3.1. We fix φ_c in (3.2) at $\varphi_c = 1$. N, N^T, G, and IG denote normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, and inverse gamma distributions, respectively.

growth at the monthly frequency reported in the third panel of Table 3.1. However, these sample moments confound the persistence of the "true" cash flow processes and the dynamics of the measurement errors. Our state-space framework is able to disentangle the various components of observed cash flow growth, thereby detecting a highly persistent predictable component x_t that is hidden under a layer of measurement errors. Based on our measurement-error model, we can compute the fraction of the variance of observed consumption growth that is due to measurement errors. In a constant-volatility version of our state-space model, 46% of the observed consumption growth variation at the monthly frequency is due to measurement errors. For annualized consumption growth data, this fraction drops below 1%.

The estimation results also provide strong evidence for stochastic volatility. According to the posteriors reported in Table 3.2, all $\sigma_{c,t}$ and $\sigma_{d,t}$ exhibit significant time variation. The posterior medians of ρ_{h_c} and ρ_{h_d} are .995 and 0.89, respectively, and the unconditional volatility standard deviations σ_{h_c} and σ_{h_d} are 0.39 and 0.53, respectively. Also, the

volatility of the growth prospect component, $\sigma_{x,t}$, shows clear evidence for time variation: the posterior medians of ρ_{h_x} and σ_{h_x} are 0.992 and 0.43, respectively. It is evident that the estimation supports three independent volatility processes for consumption growth and dividend growth.

Robustness. The evidence for a persistent component in consumption and dividend growth is robust to the choice of estimation sample. We shift the beginning of our estimation sample from 1929:M1 to 1959:M1 and use only monthly data. Given that this shorter sample is dominated by the Great Moderation and does not contain the fluctuations associated with the Great Depression, this sample should be conservative in terms of providing evidence for predictable component and aggregate stochastic volatility. Interestingly, we obtain similar estimates of ρ and find that changes in the estimates of the other parameters are generally small. In all, this sample also provides strong evidence for a predictable component as well as stochastic volatility in consumption and dividends.

3.4.3 Estimation with Cash Flow and Asset Return Data

We now include data on market returns and the risk-free rate in the estimation of our state-space model. Recall from Section 3.2 that we distinguish between a benchmark model and a generalized model that allows for a shock to the time rate of preference. We will report estimates for both specifications and discuss the role played by the preference shock in fitting our observations.

Prior Distribution. The prior distribution for the parameters associated with the exogenous cash flow process are the same as the ones used in Section 3.4.2. Thus, we focus on the preference parameters that affect the asset pricing implications of the model. Percentiles for the prior are reported in the left-side columns of Table 3.3. The prior for the discount rate δ reflects beliefs about the magnitude of the risk-free rate. For the asset pricing implications of our model, it is important whether the IES is below or above 1. Thus, we choose a prior

that covers the range from 0.3 to 3.5. The 90% prior credible interval for the risk-aversion parameter γ ranges from 3 to 15, encompassing the values that are regarded as reasonable in the asset pricing literature. We also use the same prior for the persistence and the innovation standard deviation of the preference shock as we did for the cash flow parameters ρ and $\bar{\sigma}$. Finally, we assume that consumption growth is measured without error at the annual frequency. We estimate measurement errors only for monthly consumption growth rates and the risk-free rates, using the same prior distributions as for $\bar{\sigma}$.

Posterior Distribution. The remaining columns of Table 3.3 summarize the percentiles of the posterior distribution for the parameters of the benchmark model and the generalized model. While the estimated cash flow parameters are, by and large, similar to those reported in Table 3.2 when asset prices are not utilized, a few noteworthy differences emerge. The estimate of ρ, the persistence of the predictable cash flow component, increases from 0.97 to 0.99 to capture part of the equity premium. The time variation in the volatility of the long-run risk innovation, σ_{h_x}, also increases, reflecting the information in asset prices about growth uncertainty. The estimate of φ_x drops from 0.20 to 0.03, which reduces the model-implied predictability of asset returns and consumption growth and brings it more in line with the data. Finally, the estimate of $\bar{\sigma}$ increases by a factor of 2 to explain the highly volatile asset prices data.

Overall, the information from the market returns and risk-free rate reduces the posterior uncertainty about the cash flow parameters and strengthens the evidence in favor of a time-varying conditional mean of cash flow growth rates as well as time variation in the volatility components. Table 3.3 also provides the estimated preference parameters. The IES is estimated above 1 with a relatively tight credible band. Risk aversion is estimated at 11 for the benchmark model and 10 for the generalized model.

Matching First and Second Moments. Much of the asset pricing literature, e.g.,

Table 3.3: Posterior Estimates

		Prior		Benchmark Model Posterior			Generalized Model Posterior		
	Distr.	5%	95%	5%	50%	95%	5%	50%	95%
				Preferences					
δ	B	[.9951	.9999]	.9992	.9996	.9998	.9990	.9992	.9996
ψ	G	[0.31	3.45]	1.62	1.70	1.75	1.33	1.36	1.44
γ	G	[2.74	15.45]	10.14	10.84	11.37	9.88	9.97	10.32
ρ_λ	N^T	[-0.08	0.97]	-	-	-	.935	.936	.938
σ_λ	IG	[.001	.006]	-	-	-	.0003	.0004	.0005
				Consumption					
μ_c	N	[-.006	.006]	-	.0016	-	-	.0016	-
ρ	N^T	[-0.08	0.97]	.989	.993	.994	.990	.992	.994
φ_x	G	[0.00	0.11]	0.03	0.04	0.04	0.02	0.02	0.03
$\bar{\sigma}$	IG	[.001	.006]	.004	.005	.006	.003	.004	.005
ρ_{h_c}	N^T	[-0.08	0.97]	.944	.956	.967	0.943	.946	.951
σ_{h_c}	IG	[0.22	1.03]	0.55	0.60	0.67	0.83	0.84	0.84
ρ_{h_x}	N^T	[-0.08	0.97]	.981	.990	.993	.990	.992	.994
σ_{h_x}	IG	[0.22	1.03]	0.50	0.53	0.54	0.56	0.57	0.57
				Dividend					
μ_d	N	[-.007	.006]	-	.0010	-	-	.0010	-
ϕ	N	[-13.07	13.40]	3.01	3.20	3.45	3.09	3.11	3.13
π	N	[-1.68	1.63]	1.08	1.17	1.25	1.13	1.19	1.31
φ_d	G	[0.22	11.90]	5.39	5.46	5.68	6.27	6.30	6.48
ρ_{h_d}	N^T	[-0.08	0.97]	.936	.940	.947	.939	.949	.952
σ_{h_d}	IG	[0.22	1.03]	0.44	0.45	0.46	0.56	0.57	0.57

Notes: The estimation results are based on annual consumption growth data from 1930 to 1960 and monthly consumption growth data from 1960:M1 to 2011:M12. For the other three series we use monthly data from 1929:M1 to 2011:M12. We fix $\mu_c = 0.0016$, $\mu_d = 0.0010$, and $\varphi_c = 1$ in the estimation. B, N, N^T, G, and IG are beta, normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, and inverse gamma distributions, respectively.

Bansal, Gallant, and Tauchen (2007), Bansal, Kiku, and Yaron (2012a), and Beeler and Campbell (2012), uses unconditional moments to estimate model parameters and judge model fit. While these moments implicitly enter the likelihood function of our state-space model, it is instructive to examine the extent to which sample moments implied by the estimated state-space model mimic the sample moments computed from our actual data set. To do so, we conduct a posterior predictive check (see, for instance, Geweke (2005)

for a textbook treatment). We use previously generated draws $\Theta^{(s)}$, $s = 1, \ldots, n_{sim}$, from the posterior distribution of the model parameters $p(\Theta|Y)$ and simulate for each $\Theta^{(s)}$ the benchmark and the generalized LLR models for 996 periods, which corresponds to the number of monthly observations in our estimation sample.[2] This leads to n_{sim} simulated trajectories, which we denote by $Y^{(s)}$. For each of these trajectories, we compute various sample moments, such as means, standard deviations, cross correlations, and autocorrelations. Suppose we denote such statistics generically by $\mathcal{S}(Y^{(s)})$. The simulations provide a characterization of the posterior predictive distribution $p(\mathcal{S}(Y^{(s)})|Y)$. Percentiles of this distribution for various sample moments are reported in Table 3.4. The table also lists the same moments computed from U.S. data. "Actual" sample moments that fall far into the tails of the posterior predictive distribution provide evidence for model deficiencies. The moments reported in Table 3.4 are computed for year-on-year cash flow growth rates. Market returns, the risk-free rate, and the price-dividend ratio are 12-month averages.

We first focus on the results from the benchmark model. Except for the first-order autocorrelation (AC1) of the risk-free rate r_f, all of the "actual" sample moments are within the 5th and the 95th percentile of the corresponding posterior predictive distribution. The variance of the posterior predictive distribution reflects the uncertainty about model parameters as well as the variability of the sample moments. The 90% credible intervals for the consumption growth and risk-free rate moments are much smaller than the intervals for the dividend growth and market-return moments, indicating that much of the uncertainty in the posterior predictive moments is due to the variability of the sample moments themselves. The high volatility of dividend growth and market returns translates into a large variability of their sample moments.

More specifically, the benchmark model replicates well the first two moments of con-

[2]To generate the simulated data, we also draw measurement errors.

Table 3.4: Moments of Cash Flow Growth and Asset Prices

	Data	Benchmark Model			Generalized Model		
		5%	50%	95%	5%	50%	95%
Mean (Δc)	1.83	0.91	1.89	2.79	0.88	1.93	2.86
StdDev (Δc)	2.19	1.65	2.19	2.99	1.52	2.22	3.49
AC1 (Δc)	0.48	0.09	0.32	0.56	0.08	0.33	0.57
Mean (Δd)	1.00	-2.55	1.02	4.61	-2.27	1.30	4.68
StdDev (Δd)	11.15	11.01	13.29	16.60	10.35	12.97	16.99
AC1 (Δd)	0.20	-0.19	0.03	0.23	-0.20	0.03	0.27
Corr ($\Delta c, \Delta d$)	0.55	0.12	0.32	0.51	0.13	0.34	0.56
Mean (r_m)	5.71	1.88	5.10	8.46	2.40	5.61	9.64
StdDev (r_m)	19.95	14.70	20.30	38.04	13.38	19.99	46.21
AC1 (r_m)	-0.01	-0.28	-0.06	0.17	-0.28	-0.05	0.17
Corr ($\Delta c, r_m$)	0.12	-0.03	0.18	0.39	-0.06	0.17	0.40
Mean (r_f)	0.44	-0.44	0.46	1.21	-0.39	0.67	1.49
StdDev (r_f)	2.88	2.47	2.87	3.45	1.26	1.96	4.29
AC1 (r_f)	0.64	-0.13	0.07	0.30	0.13	0.43	0.66
Mean (pd)	3.36	2.90	3.24	3.41	2.72	3.15	3.36
StdDev (pd)	0.45	0.15	0.27	0.64	0.13	0.27	0.86
AC1 (pd)	0.86	0.50	0.74	0.87	0.47	0.74	0.89

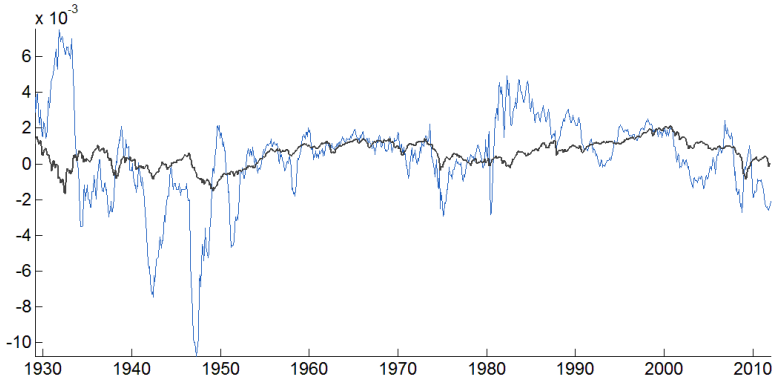
Notes: We present descriptive statistics for aggregate consumption growth (Δc), dividends growth (Δd), log returns of the aggregate stock market (r_m), the log risk-free rate (r_f), and the log price-dividend ratio (pd). We report means (Mean), standard deviations (StdDev), first-order sample autocorrelations (AC1), and correlations (Corr). Cash flow growth rates are year-on-year (in percent); market returns, the risk-free rate, and the price-dividend ratio refer to 12-month averages (in percent).

sumption and dividend growth and their correlation. The benchmark model also generates a sizable equity risk premium with a median value of 6%. The model's return variability is about 20% with the market return being not highly autocorrelated. As in the data, the model generates both a highly variable and persistent price-dividend ratio. It is particularly noteworthy that the median and 95th percentile of the price-dividend volatility distribution are significantly larger than in other LRR calibrated models with Gaussian shocks. This feature owes in part to the fact that the models contain three volatility components with underlying log-volatility dynamics, thus accommodating some non-Gaussian features.

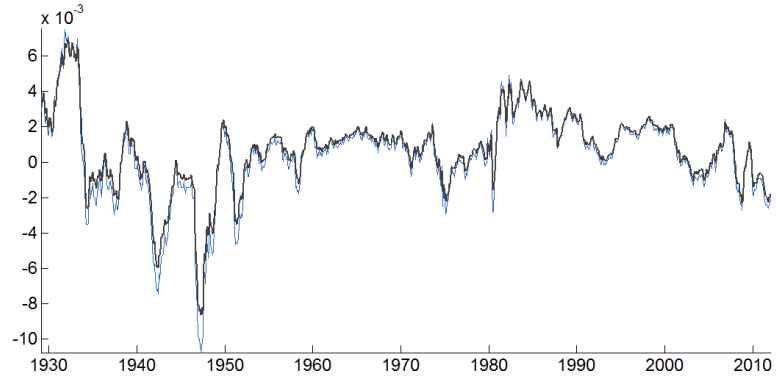
The sample moments implied by the generalized model are very similar to those of the benchmark model, except for the moments associated with the risk-free rate. Most notably,

Figure 3.1: Model-Implied Risk-Free Rate

(a) Benchmark Model



(b) Generalized Model



Notes: Blue lines depict the actual risk-free rate, and black lines depict the smoothed, model-implied risk-free rate without measurement errors.

the benchmark model generates a slightly negative autocorrelation of the risk-free rate, whereas the generalized model with the preference shock is able to reproduce the strongly positive serial correlation in the data.

Risk-Free Rate Dynamics. Our estimated state-space model can be used to decompose the observed risk-free rate into the "true" risk-free rate and a component that is due to measurement errors. Figure 3.1 overlays the actual risk rate and the smoothed "true" or model-implied risk-free rate. It is clear from the top panel of the figure that the model has difficulties generating the high volatility of the risk-free rate in the pre-1960 sample,

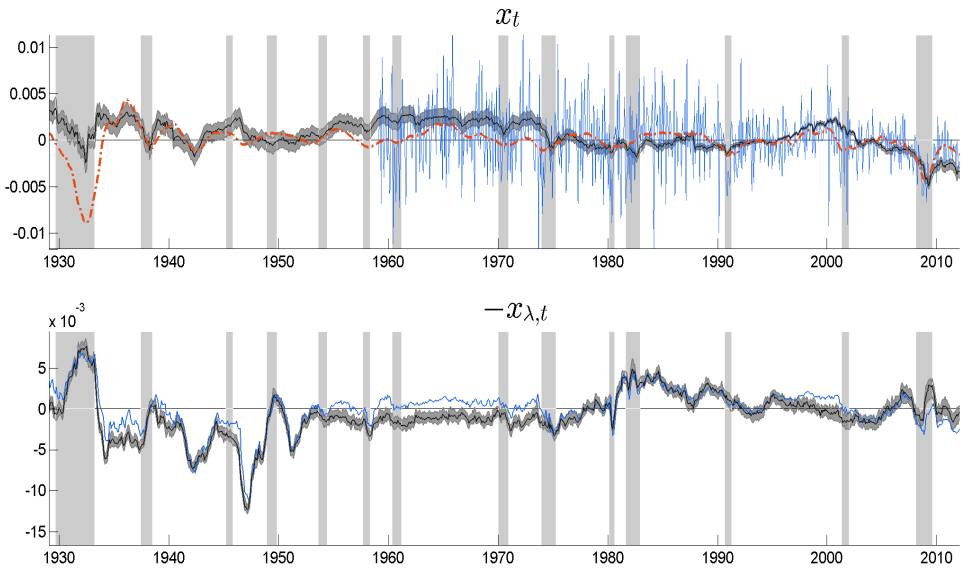
the 1980s, and the period since 2002. The benchmark model attributes these fluctuations to measurement errors. Recall that the risk-free rate series is constructed by subtracting random-walk inflation forecasts from a nominal interest rate series, which makes the presence of measurement errors plausible. In particular, our nominal interest rate series includes several periods with negative nominal yields in the period from 1938 to 1941. The pre-1960 sample also contains periods with artificially large inflation rates, which are partly due to price adjustments following price controls after World War II. Overall, the estimated benchmark model implies that about 70-80% of the fluctuations in the risk-free rate are due to measurement errors.

The generalized model with the preference shock λ_t is able to track the risk-free rate much better than the benchmark model. By construction, λ_t generates additional fluctuations in the model-implied expected stochastic discount factor and hence the model-implied risk-free rate. The likelihood-based estimation procedure reverses this logic. Persistent movements in the observed risk-free rate suggest that λ_t fluctuated substantially between 1929 and 2011. The fraction of the fluctuations attributed to measurement errors is now much smaller. In fact, the bottom panel of Figure 3.1 illustrates that the difference between the observed series and the smoothed, model-implied series is now very small. This is consistent with the predictive checks reported in Table 3.4. Since the generalized model is more successful at tracking the observed risk-free rate, we focus on the model specification with preference shock in the remainder of this section unless otherwise noted.[3]

Smoothed Mean and Volatility States. Figure 3.2 depicts smoothed estimates of the predictable component x_t and the preference shock process $x_{\lambda,t}$. Since the estimate of x_t is, to a large extent, determined by the time path of consumption, the 90% credible

[3] An alternative way to interpret the preference shocks is that the model requires correlated measurement errors to capture the time series dynamics of the real risk-free rate.

Figure 3.2: Smoothed Mean States

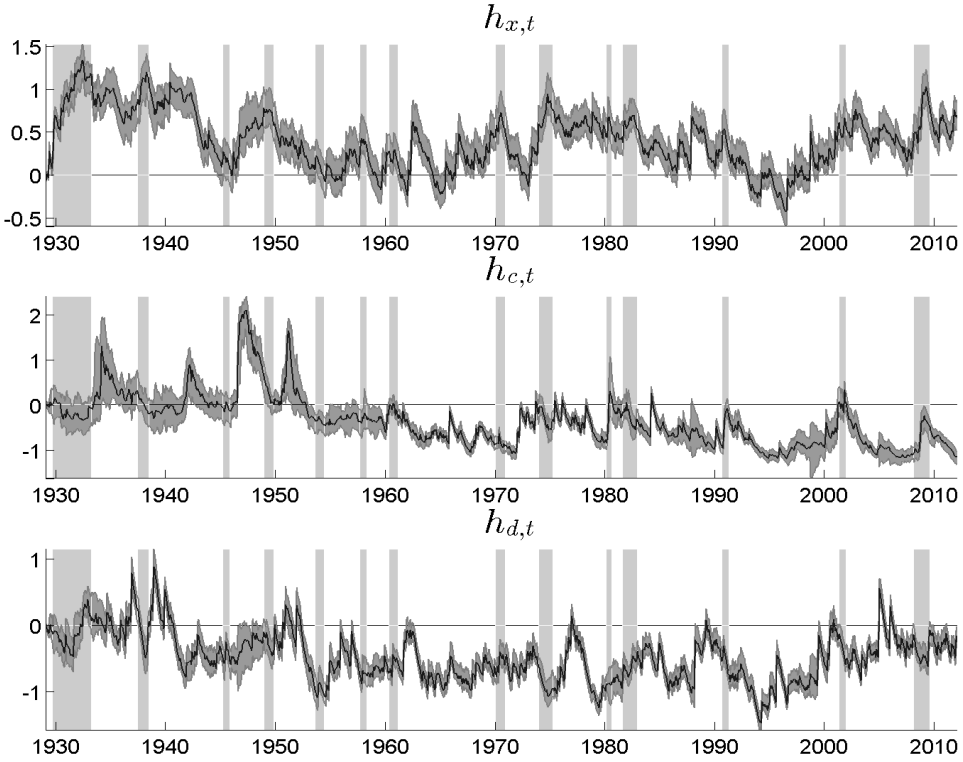


Notes: Black lines represent posterior medians of smoothed states and gray-shaded areas correspond to 90% credible intervals. In the top panel we overlay the smoothed state x_t obtained from the estimation without asset prices (red dashed line) and monthly consumption growth data (blue solid line). In the bottom panel we overlay a standardized version of the risk-free rate (blue solid line). Shaded bars indicate NBER recession dates.

bands are much wider prior to 1960, when only annual consumption growth data were used in the estimation. Post 1959, x_t tends to fall in recessions (indicated by the shaded bars in Figure 3.2), but periods of falling x_t also occur during expansions. We overlay the smoothed estimate of x_t obtained from the estimation without asset price data (see Section 3.4.2). It is very important to note that the two estimates are similar, which highlights that x_t is, in fact, detectable based on cash flow data only. We also depict the monthly consumption growth data post 1959, which confirms that x_t indeed captures low-frequency movements in consumption growth. A visual comparison of the smoothed $x_{\lambda,t}$ process with the standardized risk-free rate in the bottom panel of Figure 3.2 confirms that the preference shock in the generalized model mainly helps track the observed risk-free rate.

The smoothed volatility processes are plotted in Figure 3.3. Recall that our model has

Figure 3.3: Smoothed Volatility States



Notes: Black lines represent posterior medians of smoothed states and gray-shaded areas correspond to 90% credible intervals. Shaded bars indicate NBER recession dates.

three independent volatility processes, $h_{c,t}$, $h_{d,t}$, and $h_{x,t}$, associated with the innovations to consumption growth, dividend growth, and the predictable component, respectively. The most notable feature of $h_{c,t}$ is that it captures a drop in consumption growth volatility that occurred between 1950 and 1965. In magnitude, this drop in volatility is much larger than a subsequent decrease around 1984, the year typically associated with the Great Moderation. The stochastic volatility process for dividend growth shows a drop around 1955, but it also features an increase in volatility starting in 2000, which is not apparent in $h_{c,t}$. Overall, the smoothed $h_{d,t}$ seems to exhibit more medium- and high-frequency movements than $h_{c,t}$. Finally, the volatility of the persistent component, $h_{x,t}$, exhibits substantial fluctuations over our sample period, and it tends to peak during NBER recessions.

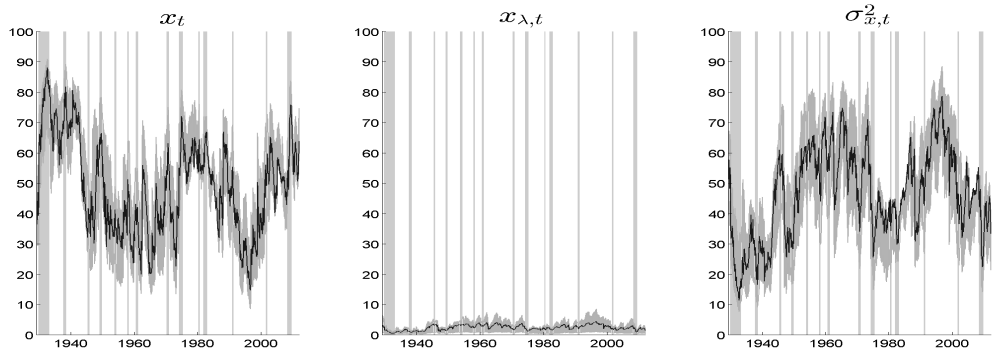
Determinants of Asset Price Fluctuations. After a visual inspection of the latent mean and volatility processes in Figures 3.2 and 3.3, we now examine their implications for asset prices. In equilibrium, the market returns, the risk-free rate, and the price-dividend ratios are functions of the mean and volatility states. Figure 3.4 depicts the contribution of various risk factors: namely, the variation in growth prospects, x_t, the preference shock, $x_{\lambda,t}$, and the conditional variability of growth prospects, $\sigma_{x,t}$, to asset price volatility. Given the posterior estimates of our state-space model, we can compute smoothed estimates of the latent asset price volatilities at every point in time. Moreover, we can also generate counterfactual volatilities by shutting down x_t, $x_{\lambda,t}$, or $\sigma_{x,t}$. The ratio of the counterfactual and the actual volatilities measures the contribution of the non-omitted risk factors. If we subtract this ratio from 1, we obtain the relative contribution of the omitted risk factor, which is shown in Figure 3.4.

While the preference shocks are important for the risk-free rate, they contribute very little to the variance of the price-dividend ratio and the market return. The figure shows that most of the variability of the price-dividend ratio is, in equal parts, due to the variation in x_t and $\sigma_{x,t}$. As Appendix 3.6.1 shows, the risk premium on the market return is barely affected by the preference shocks and consequently its variation is almost entirely attributable to the time variation in the stochastic volatility $\sigma_{x,t}^2$ and the growth prospect x_t. The remaining risk factors $\sigma_{c,t}^2$ and $\sigma_{d,t}^2$ have negligible effects (less than 1% on average) on asset price volatilities.

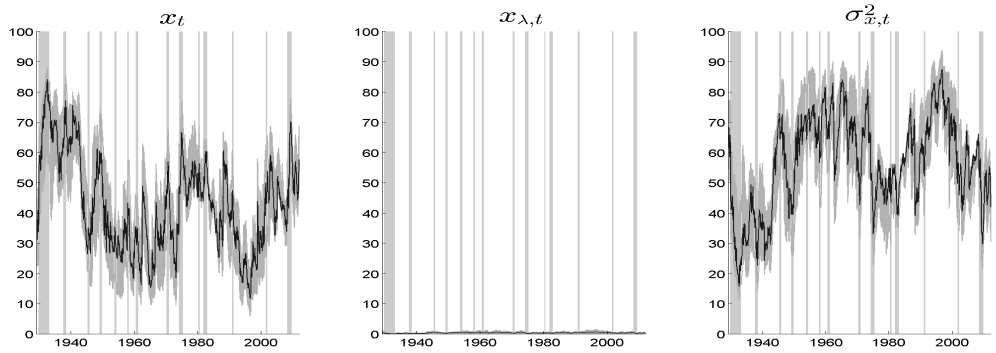
We assumed that in our endowment economy the preference shock is uncorrelated with cash flows. In a production economy this assumption will typically not be satisfied. Stochastic fluctuations in the discount factor generate fluctuations in consumption and investment, which in turn affect cash flows. To assess whether our assumption of uncorrelated shocks is contradicted by the data, we computed the correlation between the smoothed preference

Figure 3.4: Variance Decomposition for Market Returns and Risk-Free Rate

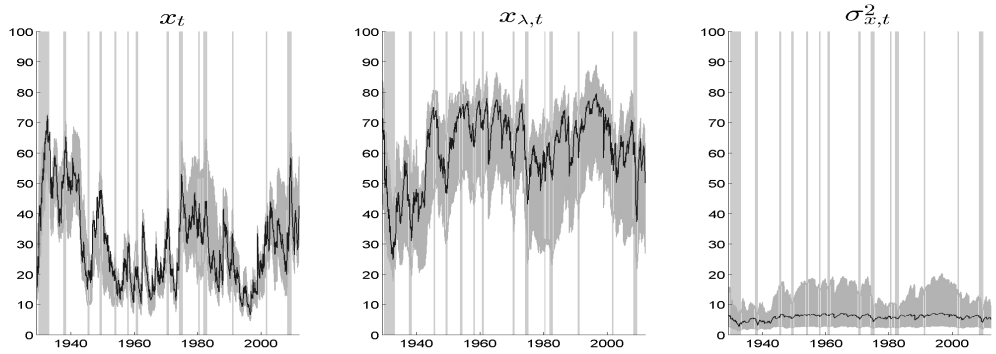
(a) Market Returns: r_m



(b) Log Price Dividend Ratio: pd



(c) Risk-Free Rate: r_f



Notes: Fraction of volatility fluctuations (in percent) in the market returns, the price-dividend ratio, and the risk-free rate that is due to x_t, $x_{\lambda,t}$, and $\sigma^2_{x,t}$, respectively. We do not present the graphs for $\sigma^2_{c,t}, \sigma^2_{d,t}$ since their time-varying shares are less than 1% on average. See the main text for computational details.

shock innovations $\eta_{\lambda,t}$ and the cash flow innovations $\eta_{c,t}$ and $\eta_{x,t}$. We can do so for every parameter draw $\Theta^{(s)}$ from the posterior distribution. The 90% posterior predictive intervals range from -0.09 to 0.03 for the correlation between $\eta_{\lambda,t}$ and $\eta_{c,t}$ and from 0 to 0.2 for the

correlation between $\eta_{\lambda,t}$ and $\eta_{x,t}$. Based on these results we conclude that there is no strong evidence that contradicts the assumption of uncorrelated preference shocks.

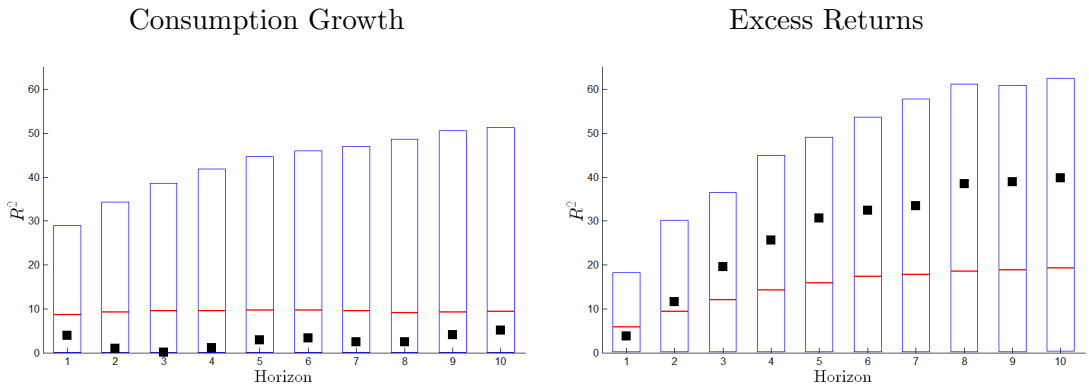
Predictability. One aspect of the data that is often discussed in the context of asset pricing models — and in particular, in the context of models featuring long-run risks — is the low predictability of future consumption growth by the current price-dividend ratio. Another key issue in the asset pricing literature is return predictability by the price-dividend ratio (e.g., Hodrick (1992)). We address these two issues in Figure 3.5 where we regress cumulative consumption growth and multi-period excess returns on the current price-dividend ratio using OLS:

$$\begin{aligned} \sum_{h=1}^{H} \Delta c_{t+h} &= \alpha + \beta pd_t + \text{resid}_{t+H} \\ \sum_{h=1}^{H} (r_{m,t+h} - r_{f,t+h-1}) &= \alpha + \beta pd_t + \text{resid}_{t+H}. \end{aligned}$$

The results are presented as posterior predictive checks, similar to those in Table 3.4, but now depicted graphically. The statistics $\mathcal{S}(Y)$ considered are the R^2 values obtained from the two regressions. The top and bottom ends of the boxes correspond to the 5th and 95th percentiles, respectively, of the posterior distribution, and the horizontal bars signify the medians. Finally, the small squares correspond to statistics computed from "actual" U.S. data.

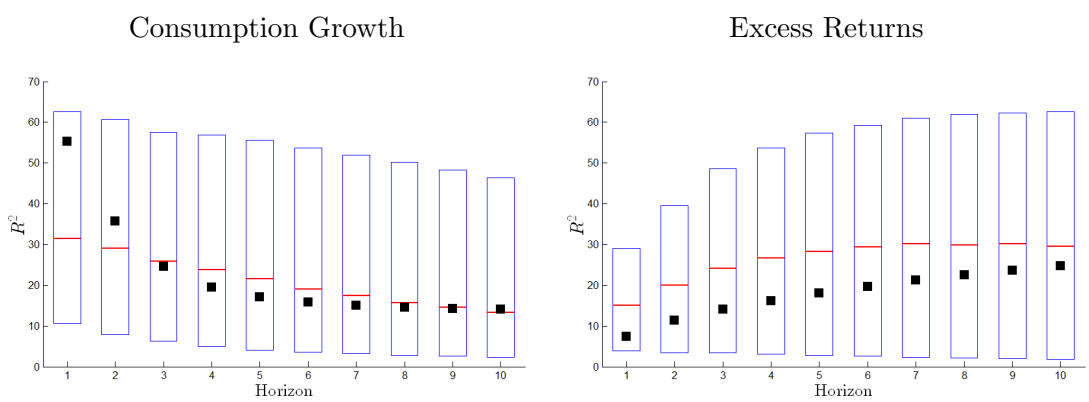
The left panel of Figure 3.5 documents the predictability of consumption growth. While the model's median R^2 value is somewhat larger (red lines) than the corresponding data estimate, the model's finite sample R^2 distribution for consumption growth encompasses the low data estimate. In terms of return predictability, depicted in the right panel of Figure 3.5, the model's median R^2 for the five-year horizon R^2 is large at 15%, with a 95 percentile value of 47% that clearly contains the data estimate. These model-implied R^2s are larger than what is typically found in models with long-run risks (e.g., Bansal, Kiku, and

Figure 3.5: Univariate Predictability Checks



Notes: Black boxes indicate regression R^2 values from actual data. Figure also depicts medians (red lines) and 90% credible intervals (top and bottom lines of boxes) of distribution of R^2 values obtained with model-generated data.

Figure 3.6: VAR Predictability Checks



Notes: Black boxes indicate VAR R^2 values from actual data. The figure also depicts medians (red lines) and 90% credible intervals (top and bottom lines of boxes) of distribution of R^2 values obtained with model-generated data.

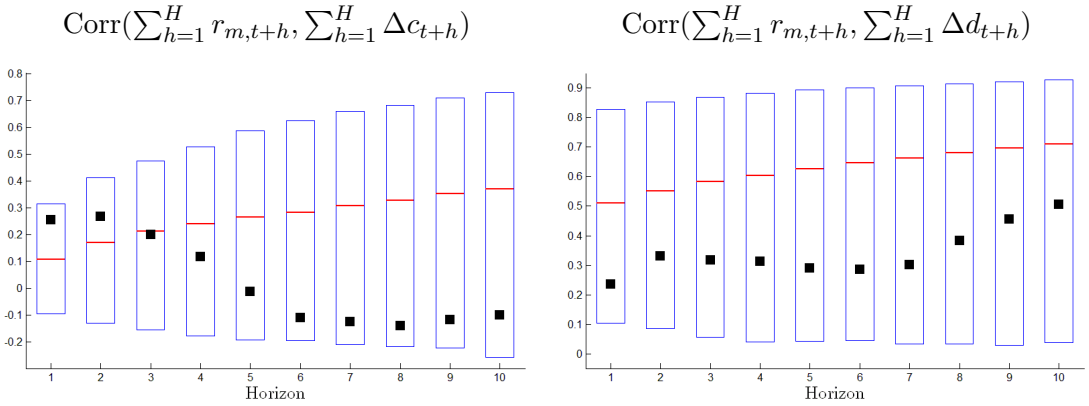
Yaron (2012a)) — a feature attributable to the presence of the three exponential volatility processes that allow this model specification for an improved fit.

It is well known that, in the data, the price-dividend ratio is very persistent, a feature that can render the aforementioned regressions spurious (see Hodrick (1992) and Stambaugh (1999)). In the model, and possibly in the data, the price-dividend ratio reflects

multiple state variables. Consequently, a VAR-based predictive regression may offer a more robust characterization. As in Bansal, Kiku, and Yaron (2012a), Figure 3.6 displays the predictability of consumption growth and the market excess returns based on a first-order VAR that includes consumption growth, the price-dividend ratio, the real risk-free rate, and the market excess return. The first thing to note is that, with multiple predictive variables, consumption growth is *highly* predictable. The VAR provides quite a different view on consumption predictability relative to the case of using the price-dividend ratio as a univariate regressor. In particular, now consumption growth predictability at the one-year horizon is very large with an R^2 of about 55% (see also Bansal, Kiku, Shaliastovich, and Yaron (2013)). While the predictability diminishes over time, it is still nontrivial with an R^2 of 14% at the 10-year horizon. It is important to note that the model-based VAR yields very comparable results (and in fact yields a median R^2 for the one-year horizon that is somewhat *lower* than its data estimate). On the other hand, since long-horizon return predictability is highly influenced by the price-dividend ratio, the VAR-based implications for excess return predictability do not change much relative to the univariate estimates. Nonetheless, the model performs well along this dimension and its generated VAR-based R^2s are closer to their VAR data estimates, relative to the R^2s based on univariate price-dividend regressor.

One additional feature in which the generalized model performance is improved relative to the benchmark model is the correlation between long-horizon return and long-horizon dividend and consumption growth. Figure 3.7 presents these correlations in the data and the model. In the model, the 10-year consumption growth and 10-year return have a correlation of 0.3, but with enormous standard deviations that encompass -0.3 to 0.7, which contain the data estimate. The analogous correlations for dividend growth are 0 to 0.9 with the data at 0.5 close to the model median estimate. In the benchmark model, without the x_λ process, this last correlation will be quite a bit larger for all percentiles and would be a

Figure 3.7: Correlation between Market Return and Growth Rates of Fundamentals



Notes: Black boxes indicate sample correlations of actual data. The figure also depicts medians (red lines) and 90% credible intervals (top and bottom lines of boxes) of distribution of sample correlations obtained with model-generated data.

challenging dimension for the model.

It is well understood that whether the IES parameter is above or below 1 plays a significant role in the asset pricing implication of the model. One common approach for estimating the IES has been to regress the growth rate of consumption on the risk-free rate (e.g., Hall (1988)). Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a) show that this regression is misspecified in the presence of stochastic volatility and leads to downward-biased estimates of the IES. Given that our estimation formally ascribes measurement errors to both consumption and the risk-free rate, we revisit the implication of this regression for inference on the IES. For completeness, we also run the reverse equation of regressing the risk-free rate on consumption growth. We use the two-year lagged consumption growth, log price-dividend ratio, market return, and risk-free rate as instrumental variables. As shown in Table 3.5, in both regression approaches the data based estimates are in fact negative, but lie well within the very wide 90% model-based credible band, even though in all model simulations the IES was set to its median estimate of 1.36. In totality, this evidence shows that, with the estimated levels of measurement errors, it is very difficult to

Table 3.5: ψ (IES) from Instrumental Variables Estimation

Specification	Data	Generalized Model		
		5%	50%	95%
Δc onto r_f	-0.30	-0.45	0.56	1.50
r_f onto Δc	-0.90	-5.53	1.26	6.80

Notes: The first row provides finite sample values of the ψ from the regression $\Delta c_{t+1} = \text{constant} + \psi r_{f,t+1} + \text{resid}_{t+1}$, while the second row provides the ψ values from the regression $r_{f,t+1} = \text{constant} + \frac{1}{\psi}\Delta c_{t+1} + \text{resid}_{t+1}$. The instruments are lagged (two years) consumption growth, log price-dividend ratio, market return, and risk-free rate. The "true" ψ value in the model is 1.36 from Table 3.3. Regressions are implemented at an annual frequency.

precisely estimate the IES via this regression approach.

3.5 Conclusion

We developed a nonlinear state-space model to capture the joint dynamics of consumption, dividend growth, and asset returns. Building on Bansal and Yaron (2004), our model consists of an economy containing a common predictable component for consumption and dividend growth and multiple stochastic volatility processes. To maximize the economic span of the data for recovering the predictable components and maximizing the frequency of data for efficiently identifying the volatility processes, we use mixed-frequency data. Our econometric framework is general enough to encompass other asset pricing models that can be written as state-space models that are linear conditional on the volatility states. A careful modeling of measurement errors in consumption growth reveals that the predictable cash flow component can be identified from consumption and dividend growth data only. The additional inclusion of asset prices sharpens the inference. The inclusion of two additional volatility processes improves the model fit considerably. The preference shock included in the generalized version of our model mostly captures the dynamics of the risk-free rate, but has little effect on market returns and price-dividend ratios. Overall, the estimated model is able to capture key asset-pricing facts of the data.

3.6 Appendix

3.6.1 Solving the Long-Run Risks Model

This section provides solutions for the consumption and dividend claims for the endowment process:

$$\begin{aligned}
g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \\
g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} \\
x_{t+1} &= \rho x_t + \sigma_{x,t}\eta_{x,t+1} \\
x_{\lambda,t+1} &= \rho_\lambda x_{\lambda,t} + \sigma_\lambda \eta_{\lambda,t+1} \\
\sigma^2_{c,t+1} &= (1-\nu_c)(\varphi_c\bar{\sigma})^2 + \nu_c\sigma^2_{c,t} + \sigma_{w_c} w_{c,t+1} \\
\sigma^2_{x,t+1} &= (1-\nu_x)(\varphi_x\bar{\sigma})^2 + \nu_x\sigma^2_{x,t} + \sigma_{w_x} w_{x,t+1} \\
\sigma^2_{d,t+1} &= (1-\nu_d)(\varphi_d\bar{\sigma})^2 + \nu_d\sigma^2_{d,t} + \sigma_{w_d} w_{d,t+1} \\
&\quad \eta_{i,t+1}, \eta_{\lambda,t+1}, w_{i,t+1} \sim N(0,1), \quad i \in \{c,x,d\}.
\end{aligned} \tag{3.14}$$

The Euler equation for the economy is

$$E_t\left[\exp\left(m_{t+1} + r_{i,t+1}\right)\right] = 1, \quad i \in \{c, m\}, \tag{3.15}$$

where

$$m_{t+1} = \theta \log\delta + \theta x_{\lambda,t+1} - \frac{\theta}{\psi} g_{c,t+1} + (\theta - 1) r_{c,t+1} \tag{3.16}$$

is the log of the real stochastic discount factor (SDF), $r_{c,t+1}$ is the log return on the consumption claim, and $r_{m,t+1}$ is the log market return. (3.16) is derived in Section 3.6.1 below. Returns are given by the approximation of Campbell and Shiller (1988a):

$$\begin{aligned}
r_{c,t+1} &= \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{c,t+1} \\
r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}.
\end{aligned} \tag{3.17}$$

The risk premium on any asset is

$$E_t(r_{i,t+1} - r_{f,t}) + \frac{1}{2}Var_t(r_{i,t+1}) = -Cov_t(m_{t+1}, r_{i,t+1}). \tag{3.18}$$

In Section 3.6.1 we solve for the law of motion for the return on the consumption claim, $r_{c,t+1}$. In Section 3.6.1 we solve for the law of motion for the market return, $r_{m,t+1}$. The risk-free rate is derived in Section 3.6.1. All three solutions depend on linearization parameters that are derived in Section 3.6.1. Finally, as mentioned above, the SDF is derived in Section 3.6.1.

Consumption Claim

In order to derive the dynamics of asset prices, we rely on approximate analytical solutions. Specifically, we conjecture that the price-consumption ratio follows

$$z_t = A_0 + A_1 x_t + A_{1,\lambda} x_{\lambda,t} + A_{2,c}\sigma_{c,t}^2 + A_{2,x}\sigma_{x,t}^2 \tag{3.19}$$

and solve for A's using (3.14), (3.15), (3.17), and (3.19).

From (3.14), (3.17), and (3.19)

$$\begin{aligned} r_{c,t+1} &= \left\{\kappa_0 + A_0(\kappa_1 - 1) + \mu_c + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\ &+ \frac{1}{\psi}x_t + A_{1,\lambda}(\kappa_1\rho_\lambda - 1)x_{\lambda,t} + A_{2,x}(\kappa_1\nu_x - 1)\sigma_{x,t}^2 + A_{2,c}(\kappa_1\nu_c - 1)\sigma_{c,t}^2 \\ &+ \sigma_{c,t}\eta_{c,t+1} + \kappa_1 A_1\sigma_{x,t}\eta_{x,t+1} + \kappa_1 A_{1,\lambda}\sigma_\lambda\eta_{\lambda,t+1} + \kappa_1 A_{2,x}\sigma_{w_x}w_{x,t+1} + \kappa_1 A_{2,c}\sigma_{w_c}w_{c,t+1} \end{aligned} \tag{3.20}$$

and from (3.14), (3.15), (3.17), and (3.19)

$$\begin{aligned} m_{t+1} &= (\theta-1)\left\{\kappa_0 + A_0(\kappa_1 - 1) + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\ &- \gamma\mu + \theta\log\delta - \frac{1}{\psi}x_t + \rho_\lambda x_{\lambda,t} + (\theta-1)A_{2,x}(\kappa_1\nu_x - 1)\sigma_{x,t}^2 + (\theta-1)A_{2,c}(\kappa_1\nu_c - 1)\sigma_{c,t}^2 \\ &- \gamma\sigma_{c,t}\eta_{c,t+1} + (\theta-1)\kappa_1 A_1\sigma_{x,t}\eta_{x,t+1} + \left\{(\theta-1)\kappa_1 A_{1,\lambda} + \theta\right\}\sigma_\lambda\eta_{\lambda,t+1} \\ &+ (\theta-1)\kappa_1 A_{2,x}\sigma_{w_x}w_{x,t+1} + (\theta-1)\kappa_1 A_{2,c}\sigma_{w_c}w_{c,t+1}. \end{aligned} \tag{3.21}$$

The solutions for A's that describe the dynamics of the price-consumption ratio are determined from

$$E_t\left[m_{t+1}+r_{c,t+1}\right]+\frac{1}{2}Var_t\left[m_{t+1}+r_{c,t+1}\right]=0$$

and they are

$$A_1=\frac{1-\frac{1}{\psi}}{1-\kappa_1\rho},\quad A_{1,\lambda}=\frac{\rho_\lambda}{1-\kappa_1\rho_\lambda},\quad A_{2,x}=\frac{\frac{\theta}{2}(\kappa_1 A_1)^2}{1-\kappa_1\nu_x},\quad A_{2,c}=\frac{\frac{\theta}{2}(1-\frac{1}{\psi})^2}{1-\kappa_1\nu_c} \tag{3.22}$$

and $A_0=\frac{A_0^1+A_0^2}{1-\kappa_1}$, where

$$\begin{aligned}A_0^1 &= \log\delta+\kappa_0+\mu(1-\frac{1}{\psi})+\kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2+\kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\\ A_0^2 &= \frac{\theta}{2}\left\{(\kappa_1 A_{1,\lambda}+1)^2\sigma_\lambda^2+(\kappa_1 A_{2,x}\sigma_{w_x})^2+(\kappa_1 A_{2,c}\sigma_{w_c})^2\right\}.\end{aligned}$$

For convenience, (3.21) can be rewritten as

$$m_{t+1}-E_t[m_{t+1}] = \lambda_c\sigma_{c,t}\eta_{c,t+1}+\lambda_x\sigma_{x,t}\eta_{x,t+1}+\lambda_\lambda\sigma_\lambda\eta_{\lambda,t+1}+\lambda_{w_x}\sigma_{w_x}w_{x,t+1}+\lambda_{w_c}\sigma_{w_c}w_{c,t+1}.$$

Note that λs represent the market price of risk for each source of risk. To be specific,

$$\begin{aligned}\lambda_c &= -\gamma,\quad \lambda_x=-(\gamma-\frac{1}{\psi})\frac{\kappa_1}{1-\kappa_1\rho},\quad \lambda_\lambda=\frac{\theta-\kappa_1\rho_\lambda}{1-\kappa_1\rho_\lambda},\\ \lambda_{w_x} &= -\frac{\theta(\gamma-\frac{1}{\psi})(1-\frac{1}{\psi})\kappa_1}{2(1-\kappa_1\nu_x)}(\frac{\kappa_1}{1-\kappa_1\rho})^2,\quad \lambda_{w_c}=-\frac{\theta(\gamma-\frac{1}{\psi})(1-\frac{1}{\psi})\kappa_1}{2(1-\kappa_1\nu_c)}.\end{aligned} \tag{3.23}$$

Similarly, rewrite (3.20) as

$$r_{c,t+1}-E_t[r_{c,t+1}] = -\beta_{c,c}\sigma_{c,t}\eta_{c,t+1}-\beta_{c,x}\sigma_{x,t}\eta_{x,t+1}-\beta_{c,\lambda}\sigma_\lambda\eta_{\lambda,t+1}-\beta_{c,w_x}\sigma_{w_x}w_{x,t+1}-\beta_{c,w_c}\sigma_{w_c}w_{c,t+1}$$

where

$$\beta_{c,c}=-1,\quad \beta_{c,x}=-\kappa_1 A_1,\quad \beta_{c,\lambda}=-\kappa_1 A_{1,\lambda},\quad \beta_{c,w_x}=-\kappa_1 A_{2,x},\quad \beta_{c,w_c}=-\kappa_1 A_{2,c}. \tag{3.24}$$

The risk premium for the consumption claim is

$$\begin{aligned}E_t(r_{c,t+1}-r_{f,t})+\frac{1}{2}Var_t(r_{c,t+1}) &= -Cov_t(m_{t+1},r_{c,t+1}) \qquad (3.25)\\ &= \beta_{c,x}\lambda_x\sigma_{x,t}^2+\beta_{c,c}\lambda_c\sigma_{c,t}^2+\beta_{c,\lambda}\lambda_\lambda\sigma_\lambda^2+\beta_{c,w_x}\lambda_{w_x}\sigma_{w_x}^2+\beta_{c,w_c}\lambda_{w_c}\sigma_{w_c}^2.\end{aligned}$$

Market Return

Similarly, using the conjectured solution to the price-dividend ratio

$$z_{m,t} = A_{0,m} + A_{1,m}x_t + A_{1,\lambda,m}x_{\lambda,t} + A_{2,x,m}\sigma^2_{x,t} + A_{2,c,m}\sigma^2_{c,t} + A_{2,d,m}\sigma^2_{d,t} \tag{3.26}$$

the market return can be expressed as

$$\begin{aligned} r_{m,t+1} &= \kappa_{0,m} + A_{0,m}(\kappa_{1,m}-1) + \mu_d + \kappa_{1,m}A_{2,x,m}(1-\nu_x)(\varphi_x\bar{\sigma})^2 \\ &+ \kappa_{1,m}A_{2,c,m}(1-\nu_c)(\varphi_c\bar{\sigma})^2 + \kappa_{1,m}A_{2,d,m}(1-\nu_d)(\varphi_d\bar{\sigma})^2 + \{\phi + A_{1,m}(\kappa_{1,m}\rho-1)\}\, x_t \\ &+ (\kappa_{1,m}\rho_\lambda - 1)A_{1,\lambda,m}x_{\lambda,t} + A_{2,x,m}(\kappa_{1,m}\nu_x - 1)\sigma^2_{x,t} + A_{2,c,m}(\kappa_{1,m}\nu_c - 1)\sigma^2_{c,t} \\ &+ A_{2,d,m}(\kappa_{1,m}\nu_d - 1)\sigma^2_{d,t} + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} + \kappa_{1,m}A_{1,m}\sigma_{x,t}\eta_{x,t+1} + \kappa_{1,m}A_{1,\lambda,m}\sigma_\lambda\eta_{\lambda,t+1} \\ &+ \kappa_{1,m}A_{2,x,m}\sigma_{w_x}w_{x,t+1} + \kappa_{1,m}A_{2,c,m}\sigma_{w_c}w_{c,t+1} + \kappa_{1,m}A_{2,d,m}\sigma_{w_d}w_{d,t+1}. \end{aligned} \tag{3.27}$$

Given the solution for A's, A_m's can be derived as follows:

$$\begin{aligned} A_{0,m} &= \frac{A^{1st}_{0,m} + A^{2nd}_{0,m}}{1-\kappa_{1,m}} \\ A_{1,m} &= \frac{\phi - \frac{1}{\psi}}{1-\kappa_{1,m}\rho} \\ A_{1,\lambda,m} &= \frac{\rho_\lambda}{1-\kappa_{1,m}\rho_\lambda} \\ A_{2,x,m} &= \frac{\frac{1}{2}\left\{(\theta-1)\kappa_1 A_1 + \kappa_{1,m}A_{1,m}\right\}^2 + (\theta-1)(\kappa_1\nu_x - 1)A_{2,x}}{1-\kappa_{1,m}\nu_x} \\ A_{2,c,m} &= \frac{\frac{1}{2}(\pi-\gamma)^2 + (\theta-1)(\kappa_1\nu_c - 1)A_2}{1-\kappa_{1,m}\nu_c} \\ A_{2,d,m} &= \frac{\frac{1}{2}}{1-\kappa_{1,m}\nu_d}, \end{aligned} \tag{3.28}$$

where

$$\begin{aligned}
A_{0,m}^{1st} &= \theta \log \delta + (\theta - 1)\left\{\kappa_0 + A_0(\kappa_1 - 1) + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x \bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c \bar{\sigma})^2\right\} \\
&- \gamma\mu + \kappa_{0,m} + \mu_d + \kappa_{1,m} A_{2,x,m}(1-\nu_x)(\varphi_x \bar{\sigma})^2 + \kappa_{1,m} A_{2,c,m}(1-\nu_c)(\varphi_c \bar{\sigma})^2 \\
&+ \kappa_{1,m} A_{2,d,m}(1-\nu_d)(\varphi_d \bar{\sigma})^2 \\
A_{0,m}^{2nd} &= \frac{1}{2}\left(\kappa_{1,m} A_{2,x,m}\sigma_{w_x} + (\theta-1)\kappa_1 A_{2,x}\sigma_{w_x}\right)^2 + \frac{1}{2}\left(\kappa_{1,m} A_{2,c,m}\sigma_{w_c} + (\theta-1)\kappa_1 A_{2,c}\sigma_{w_c}\right)^2 \\
&+ \frac{1}{2}\left(\kappa_{1,m} A_{2,d,m}\sigma_{w_d}\right)^2 + \frac{1}{2}\left(\kappa_{1,m} A_{1,\lambda,m}\sigma_\lambda + (\theta-1)\kappa_1 A_{1,\lambda}\sigma_\lambda + \theta\sigma_\lambda\right)^2.
\end{aligned}$$

Rewrite market-return equation (3.27) as

$$r_{m,t+1} - E_t[r_{m,t+1}] = -\beta_{m,c}\sigma_{c,t}\eta_{c,t+1} - \beta_{m,x}\sigma_{x,t}\eta_{x,t+1} - \beta_{m,\lambda}\sigma_\lambda\eta_{\lambda,t+1} - \beta_{m,w_x}\sigma_{w_x}w_{x,t+1} - \beta_{m,w_c}\sigma_{w_c}w_{c,t+1},$$

where

$$\begin{aligned}
\beta_{m,c} &= -\pi, \quad \beta_{m,x} = -\kappa_{1,m}A_{1,m}, \quad \beta_{m,\lambda} = -\kappa_{1,m}A_{1,\lambda,m}, \qquad (3.29)\\
\beta_{m,w_x} &= -\kappa_{1,m}A_{2,x,m}, \quad \beta_{m,w_c} = -\kappa_{1,m}A_{2,c,m}.
\end{aligned}$$

The risk premium for the dividend claim is

$$\begin{aligned}
E_t(r_{m,t+1} - r_{f,t}) + \frac{1}{2}Var_t(r_{m,t+1}) &= -Cov_t(m_{t+1}, r_{m,t+1}) \qquad (3.30)\\
&= \beta_{m,x}\lambda_x\sigma_{x,t}^2 + \beta_{m,c}\lambda_c\sigma_{c,t}^2 \\
&\quad + \beta_{m,\lambda}\lambda_\lambda\sigma_\lambda^2 + \beta_{m,w_x}\lambda_{w_x}\sigma_{w_x}^2 + \beta_{m,w_c}\lambda_{w_c}\sigma_{w_c}^2.
\end{aligned}$$

Risk-Free Rate

The model-driven equation for the risk-free rate is

$$\begin{aligned}
r_{f,t} &= -E_t\left[m_{t+1}\right] - \frac{1}{2}var_t\left[m_{t+1}\right] \qquad (3.31)\\
&= -\theta\log\delta - E_t\left[x_{\lambda,t+1}\right] + \frac{\theta}{\psi}E_t\left[g_{c,t+1}\right] + (1-\theta)E_t\left[r_{c,t+1}\right] - \frac{1}{2}var_t\left[m_{t+1}\right].
\end{aligned}$$

Subtract $(1-\theta)r_{f,t}$ from both sides and divide by θ,

$$r_{f,t} = -\log\delta - \frac{1}{\theta}E_t\left[x_{\lambda,t+1}\right] + \frac{1}{\psi}E_t\left[g_{c,t+1}\right] + \frac{(1-\theta)}{\theta}E_t\left[r_{c,t+1} - r_{f,t}\right] - \frac{1}{2\theta}var_t\left[m_{t+1}\right] \qquad (3.32)$$

From (3.14) and (3.21)

$$r_{f,t} = B_0 + B_1 x_t + B_{1,\lambda} x_{\lambda,t} + B_{2,x}\sigma^2_{x,t} + B_{2,c}\sigma^2_{c,t},$$

where

$$B_1 = \frac{1}{\psi}, \quad B_{1,\lambda} = -\rho_\lambda, \quad B_{2,x} = -\frac{(1-\frac{1}{\psi})(\gamma-\frac{1}{\psi})\kappa_1^2}{2(1-\kappa_1\rho)^2}, \quad B_{2,c} = -\frac{1}{2}(\frac{\gamma-1}{\psi}+\gamma) \quad (3.33)$$

and

$$\begin{aligned} B_0 &= -\theta\log\delta - (\theta-1)\left\{\kappa_0 + (\kappa_1-1)A_0 + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\ &+ \gamma\mu - \frac{1}{2}\left\{(\theta-1)\kappa_1 A_{2,x}\sigma_{w_x}\right\}^2 - \frac{1}{2}\left\{(\theta-1)\kappa_1 A_{2,c}\sigma_{w_c}\right\}^2 - \frac{1}{2}\left\{((\theta-1)\kappa_1 A_{1,\lambda}+\theta)^2\sigma^2_\lambda\right\}. \end{aligned}$$

Linearization Parameters

For any asset, the linearization parameters are determined endogenously by the following system of equations:

$$\begin{aligned} \bar{z}_i &= A_{0,i}(\bar{z}_i) + \sum_{j\in\{c,x,d\}} A_{2,i,j}(\bar{z}_i)\times(\varphi_j\bar{\sigma})^2 \\ \kappa_{1,i} &= \frac{\exp(\bar{z}_i)}{1+\exp(\bar{z}_i)} \\ \kappa_{0,i} &= \log(1+\exp(\bar{z}_i)) - \kappa_{1,i}\bar{z}_i. \end{aligned}$$

The solution is determined numerically by iteration until reaching a fixed point of $\bar{z}_i$.

Deriving the Intertemporal Marginal Rate of Substitution (MRS)

We consider a representative-agent endowment economy modified to allow for time-preference shocks. The representative agent has Epstein and Zin (1989) recursive preferences and maximizes her lifetime utility

$$V_t = \max_{C_t}\left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta\big(\mathbb{E}_t[V_{t+1}^{1-\gamma}]\big)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

subject to budget constraint

$$W_{t+1} = (W_t - C_t)R_{c,t+1},$$

where W_t is the wealth of the agent, $R_{c,t+1}$ is the return on all invested wealth, γ is risk aversion, $\theta = \frac{1-\gamma}{1-1/\psi}$, and ψ is intertemporal elasticity of substitution. The ratio $\frac{\lambda_{t+1}}{\lambda_t}$ determines how agents trade off current versus future utility and is referred to as the time-preference shock (see Albuquerque, Eichenbaum, and Rebelo (2012)).

First conjecture a solution for $V_t = \phi_t W_t$. The value function is homogenous of degree 1 in wealth; it can now be written as

$$\phi_t W_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta\big(\mathbb{E}_t[(\phi_{t+1}W_{t+1})^{1-\gamma}]\big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \tag{3.34}$$

subject to

$$W_{t+1} = (W_t - C_t)R_{c,t+1}.$$

Epstein and Zin (1989) show that the above dynamic program has a maximum.

Using the dynamics of the wealth equation, we substitute W_{t+1} into (3.34) to derive

$$\phi_t W_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta(W_t - C_t)^{\frac{1-\gamma}{\theta}} \big(\mathbb{E}_t[(\phi_{t+1}R_{c,t+1})^{1-\gamma}]\big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}. \tag{3.35}$$

At the optimum, $C_t = b_t W_t$, where b_t is the consumption-wealth ratio. Using (3.35) and shifting the exponent on the braces to the left-hand side, and dividing by W_t, yields

$$\phi_t^{\frac{1-\gamma}{\theta}} = (1-\delta)\lambda_t \left(\frac{C_t}{W_t}\right)^{\frac{1-\gamma}{\theta}} + \delta\left(1 - \frac{C_t}{W_t}\right)^{\frac{1-\gamma}{\theta}} \big(\mathbb{E}_t[(\phi_{t+1}R_{c,t+1})^{1-\gamma}]\big)^{\frac{1}{\theta}} \tag{3.36}$$

or simply

$$\phi_t^{\frac{1-\gamma}{\theta}} = (1-\delta)\lambda_t b_t^{\frac{1-\gamma}{\theta}} + \delta(1-b_t)^{\frac{1-\gamma}{\theta}} \big(\mathbb{E}_t[(\phi_{t+1}R_{c,t+1})^{1-\gamma}]\big)^{\frac{1}{\theta}}. \tag{3.37}$$

The first-order condition with respect to the consumption choice yields

$$(1-\delta)\lambda_t b_t^{\frac{1-\gamma}{\theta}-1} = \delta(1-b_t)^{\frac{1-\gamma}{\theta}-1} \big(\mathbb{E}_t[(\phi_{t+1}R_{c,t+1})^{1-\gamma}]\big)^{\frac{1}{\theta}}. \tag{3.38}$$

Plugging (3.38) into (3.37) yields

$$\phi_t = (1-\delta)^{\frac{\theta}{1-\gamma}} \lambda_t^{\frac{\theta}{1-\gamma}} \left(\frac{C_t}{W_t}\right)^{\frac{1-\gamma-\theta}{1-\gamma}} = (1-\delta)^{\frac{\psi}{\psi-1}} \lambda_t^{\frac{\psi}{\psi-1}} \left(\frac{C_t}{W_t}\right)^{\frac{1}{1-\psi}}. \tag{3.39}$$

The lifetime value function is $\phi_t W_t$, with the solution to ϕ_t stated above. This expression for ϕ_t is important: It states that the maximized lifetime utility is determined by the consumption-wealth ratio.

(3.38) can be rewritten as

$$(1-\delta)^{\theta} \lambda_t^{\theta} \left(\frac{b_t}{1-b_t}\right)^{-\frac{\theta}{\psi}} = \delta^{\theta} \mathbb{E}_t[(\phi_{t+1} R_{c,t+1})^{1-\gamma}]. \tag{3.40}$$

Consider the term $\phi_{t+1} R_{c,t+1}$:

$$\phi_{t+1} R_{c,t+1} = (1-\delta)^{\frac{\psi}{\psi-1}} \lambda_{t+1}^{\frac{\psi}{\psi-1}} \left(\frac{C_{t+1}}{W_{t+1}}\right)^{\frac{1}{1-\psi}} R_{c,t+1}. \tag{3.41}$$

After substituting the wealth constraint, $\frac{C_{t+1}}{W_{t+1}} = \frac{C_{t+1}/C_t}{W_t/C_t - 1} \cdot \frac{1}{R_{c,t+1}} = \frac{G_{t+1}}{R_{c,t+1}} \cdot \frac{b_t}{1-b_t}$, into the above expression, it follows that

$$\phi_{t+1} R_{c,t+1} = (1-\delta)^{\frac{\psi}{\psi-1}} \lambda_{t+1}^{\frac{\psi}{\psi-1}} \left(\frac{b_t}{1-b_t}\right)^{\frac{1}{1-\psi}} \left(\frac{G_{t+1}}{R_{c,t+1}}\right)^{\frac{1}{1-\psi}} R_{c,t+1}. \tag{3.42}$$

After some intermediate tedious manipulations,

$$\delta^{\theta} (\phi_{t+1} R_{c,t+1})^{1-\gamma} = \delta^{\theta} (1-\delta)^{\theta} \lambda_{t+1}^{\theta} \left(\frac{b_t}{1-b_t}\right)^{-\frac{\theta}{\psi}} G_{t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta}. \tag{3.43}$$

Taking expectations and substituting the last expression into (3.40) yields

$$\delta^{\theta} \mathbb{E}_t[\left(\frac{\lambda_{t+1}}{\lambda_t}\right)^{\theta} G_{t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta-1} R_{c,t+1}] = 1. \tag{3.44}$$

From here we see that the MRS in terms of observables is

$$M_{t+1} = \delta^{\theta} \left(\frac{\lambda_{t+1}}{\lambda_t}\right)^{\theta} G_{t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta-1}. \tag{3.45}$$

The log of MRS is

$$m_{t+1} = \theta \log \delta + \theta x_{\lambda,t+1} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{c,t+1}, \tag{3.46}$$

where $x_{\lambda,t+1} = \log(\frac{\lambda_{t+1}}{\lambda_t})$.

3.6.2 Data Source

Nominal PCE

We download seasonally adjusted data for nominal PCE from NIPA Tables 2.3.5 and 2.8.5. We then compute within-quarter averages of monthly observations and within-year averages of quarterly observations.

Real PCE

We use Table 2.3.3., Real Personal Consumption Expenditures by Major Type of Product, Quantity Indexes (A:1929-2011)(Q:1947:Q1-2011:Q4) to extend Table 2.3.6., Real Personal Consumption Expenditures by Major Type of Product, Chained Dollars (A:1995-2011) (Q:1995:Q1-2011:Q4). Monthly data are constructed analogously using Table 2.8.3. and Table 2.8.6.

Real Per Capita PCE: ND+S

The LRR model defines consumption as per capita consumer expenditures on nondurables and services. We download mid-month population data from NIPA Table 7.1.(A:1929-2011)(Q:1947:Q1-2011:Q4) and from Federal Reserve Bank of St. Louis' FRED database (M:1959:M1-2011:M12). We convert consumption to per capita terms.

Dividend and Market Returns Data

Data are from the Center for Research in Security Prices (CRSP). The three monthly series from CRSP are the value-weighted with-, RN_t, and without-dividend nominal returns, RX_t, of CRSP stock market indexes (NYSE/AMEX/NASDAQ/ARCA), and the CPI inflation rates, π_t. The sample period is from 1928:M1 to 2011:M12. The monthly real dividend series are constructed as in Hodrick (1992):

1. A normalized nominal value-weighted price series is produced by initializing $P_0 = 1$

and recursively setting $P_t = (1 + RX_t)P_{t-1}$.

2. A normalized nominal divided series, d_t, is obtained by recognizing that $d_t = (RN_t - RX_t)P_{t-1}$.

3. The annualized dividend is $D_t = \sum_{j=0}^{11} d_{t-j}$, which sums the previous 11 months of dividends with the current dividend. The first observation is 1928:M12.

Both dividend growth, $\log(\frac{D_{t+1}}{D_t})$, and market returns, RN_{t+1}, are converted from nominal to real terms using the CPI inflation rates, which are denoted by $g_{d,t+1}$ and $r_{m,t+1}$ respectively. They are available from 1929:M1 to 2011:M12.

Ex Ante Risk-Free Rate

The ex ante risk-free rate is constructed as in the online appendix of Beeler and Campbell (2012). Nominal yields to calculate risk-free rates are the CRSP Fama Risk Free Rates. Even though our model runs in monthly frequencies, we use the three-month yield because of the larger volume and higher reliability. We subtract annualized three-month inflation, $\pi_{t,t+3}$, from the nominal yield, $i_{f,t}$, to form a measure of the ex post (annualized) real three-month interest rate. The ex ante real risk-free rate, $r_{f,t}$, is constructed as a fitted value from a projection of the ex post real rate on the current nominal yield, $i_{f,t}$, and inflation over the previous year, $\pi_{t-12,t}$:

$$\begin{aligned} i_{f,t} - \pi_{t,t+3} &= \beta_0 + \beta_1 i_{f,t} + \beta_2 \pi_{t-12,t} + \varepsilon_{t+3} \\ r_{f,t} &= \hat{\beta}_0 + \hat{\beta}_1 i_{f,t} + \hat{\beta}_2 \pi_{t-12,t}. \end{aligned}$$

The ex ante real risk-free rates are available from 1929:M1 to 2011:M12.

3.6.3 The State-Space Representation of the LRR Model

Measurement Equations

In order to capture the correlation structure between the measurement errors at monthly frequency, we assumed in the main text that 12 months of consumption growth data are released at the end of each year. We will now present the resulting measurement equation. To simplify the exposition, we assume that the monthly consumption data are released at the end of the quarter (rather than at the end of the year). In the main text, the measurement equation is written as

$$y_{t+1} = A_{t+1}\Big(D + Zs_{t+1} + Z^v s^v_{t+1} + \Sigma^u u_{t+1}\Big), \quad u_{t+1} \sim N(0, I). \tag{3.47}$$

The selection matrix A_{t+1} accounts for the deterministic changes in the vector of observables, y_{t+1}. Recall that monthly observations are available only starting in 1959:M1. For the sake of exposition, suppose prior to 1959:M1 consumption growth was available at a quarterly frequency. Then:

1. Prior to 1959:M1:

   (a) If $t+1$ is the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g^q_{c,t+1} \\ g_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

   (b) If $t+1$ is not the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

2. From 1959:M1 to present:

(a) If $t+1$ is the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g_{c,t+1} \\ g_{c,t} \\ g_{c,t-1} \\ g_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) If $t+1$ is not the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The relationship between observations and states (ignoring the measurement errors) is given by the approximate analytical solution of the LRR model described in Section 3.6.1:

$$\begin{aligned}
g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \qquad (3.48)\\
g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} \\
r_{m,t+1} &= \{\kappa_{0,m} + (\kappa_{1,m}-1)A_{0,m} + \mu_d\} + +(\kappa_{1,m}A_{1,m})x_{t+1} + (\phi - A_{1,m})x_t \\
&+ (\kappa_{1,m}A_{1,\lambda,m})x_{\lambda,t+1} - A_{1,\lambda,m}x_{\lambda,t} + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} \\
&+ (\kappa_{1,m}A_{2,x,m})\sigma^2_{x,t+1} - A_{2,x,m}\sigma^2_{x,t} + (\kappa_{1,m}A_{2,c,m})\sigma^2_{c,t+1} - A_{2,c,m}\sigma^2_{c,t} \\
&+ (\kappa_{1,m}A_{2,d,m})\sigma^2_{d,t+1} - A_{2,d,m}\sigma^2_{d,t} \\
r_{f,t} &= B_0 + B_1 x_t + B_{1,\lambda}x_{\lambda,t} + B_{2,x}\sigma^2_{x,t} + B_{2,c}\sigma^2_{c,t} \\
&\eta_{i,t+1}, \eta_{\lambda,t+1}, w_{i,t+1} \sim N(0,1), \quad i \in \{c,x,d\}.
\end{aligned}$$

In order to reproduce (3.48) and the measurement-error structure described in Sections 3.3.1 and 3.3.2, we define the vectors of states s_{t+1} and s^v_{t+1} as

$$s_{t+1} = \begin{bmatrix} x_{t+1} \\ x_t \\ x_{t-1} \\ x_{t-2} \\ x_{t-3} \\ x_{t-4} \\ \sigma_{c,t}\eta_{c,t+1} \\ \sigma_{c,t-1}\eta_{c,t} \\ \sigma_{c,t-2}\eta_{c,t-1} \\ \sigma_{c,t-3}\eta_{c,t-2} \\ \sigma_{c,t-4}\eta_{c,t-3} \\ \sigma_\epsilon\epsilon_{t+1} \\ \sigma_\epsilon\epsilon_t \\ \sigma_\epsilon\epsilon_{t-1} \\ \sigma_\epsilon\epsilon_{t-2} \\ \sigma_\epsilon\epsilon_{t-3} \\ \sigma_\epsilon\epsilon_{t-4} \\ \sigma_\epsilon^q\epsilon_{t+1}^q \\ \sigma_\epsilon^q\epsilon_t^q \\ \sigma_\epsilon^q\epsilon_{t-1}^q \\ \sigma_\epsilon^q\epsilon_{t-2}^q \\ \sigma_{d,t}\eta_{d,t+1} \\ x_{\lambda,t+1} \\ x_{\lambda,t} \end{bmatrix}, \quad s_{t+1}^v = \begin{bmatrix} \sigma_{x,t+1}^2 \\ \sigma_{x,t}^2 \\ \sigma_{c,t+1}^2 \\ \sigma_{c,t}^2 \\ \sigma_{d,t+1}^2 \\ \sigma_{d,t}^2 \end{bmatrix}. \tag{3.49}$$

It can be verified that the coefficient matrices D, Z, Z^v, and Σ^e are given by

$$Z = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi & 0 & 0 & 0 & 0 & \pi & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \mu_{r,1} & \mu_{r,2} & 0 & 0 & 0 & 0 & \mu_{r,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_{r,4} & \mu_{r,5} & \mu_{r,6} \\ 0 & B_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_{1,\lambda} \end{bmatrix}$$

$$Z^v = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_{r,7} & \mu_{r,8} & \mu_{r,9} & \mu_{r,10} & \mu_{r,11} & \mu_{r,12} \\ 0 & B_{2,x} & 0 & B_{2,c} & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \\ \mu \\ \mu_d \\ \mu_{r,0} \\ B_0 \end{bmatrix}, \quad \Sigma^u = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_\epsilon^{rf} \end{bmatrix}.$$

The coefficients $\mu_{r,0}$ to $\mu_{r,12}$ are obtained from the solution of the LRR model:

$$\begin{bmatrix} \mu_{r,0} \\ \mu_{r,1} \\ \mu_{r,2} \\ \mu_{r,3} \\ \mu_{r,4} \\ \mu_{r,5} \\ \mu_{r,6} \end{bmatrix} = \begin{bmatrix} \kappa_{0,m} + A_{0,m}(\kappa_{1,m} - 1) + \mu_d \\ \kappa_{1,m}A_{1,m} \\ \phi - A_{1,m} \\ \pi \\ 1 \\ \kappa_{1,m}A_{1,\lambda,m} \\ -A_{1,\lambda,m} \end{bmatrix}, \quad \begin{bmatrix} \mu_{r,7} \\ \mu_{r,8} \\ \mu_{r,9} \\ \mu_{r,10} \\ \mu_{r,11} \\ \mu_{r,12} \end{bmatrix} = \begin{bmatrix} \kappa_{1,m}A_{2,x,m} \\ -A_{2,x,m} \\ \kappa_{1,m}A_{2,m} \\ -A_{2,m} \\ \kappa_{1,m}A_{2,d,m} \\ -A_{2,d,m} \end{bmatrix}.$$

State Transition Equations

Using the definition of s_{t+1} in (3.49), we write the state-transition equation as

$$s_{t+1} = \Phi s_t + v_{t+1}(h_t). \tag{3.50}$$

Conditional on the volatilities h_t, this equation reproduces the law of motion of the two persistent conditional mean processes

$$\begin{aligned} x_{t+1} &= \rho x_t + \sigma_{x,t}\eta_{x,t+1} \\ x_{\lambda,t+1} &= \rho_\lambda x_{\lambda,t} + \sigma_\lambda \eta_{\lambda,t+1} \end{aligned} \tag{3.51}$$

and it contains some trivial relationships among the measurement-error states. The matrices Φ and $v_{t+1}(h_t)$ are defined as

$$\Phi = \begin{bmatrix}
\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_\lambda & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}$$

and

$$v_{t+1}(h_t) = \begin{bmatrix} \sigma_{x,t}\eta_{x,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_{c,t}\eta_{c,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_\epsilon \epsilon_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_\epsilon^q \epsilon_{t+1}^q \\ 0 \\ 0 \\ 0 \\ \sigma_{d,t}\eta_{d,t+1} \\ \sigma_\lambda \eta_{\lambda,t+1} \\ 0 \end{bmatrix}.$$

The law of motion of the three persistent conditional log volatility processes is given by

$$h_{t+1} = \Psi h_t + \Sigma_h w_{t+1}, \tag{3.52}$$

where

$$h_{t+1} = \begin{bmatrix} h_{x,t+1} \\ h_{c,t+1} \\ h_{d,t+1} \end{bmatrix}, \quad \Psi = \begin{bmatrix} \rho_{h_x} & 0 & 0 \\ 0 & \rho_{h_c} & 0 \\ 0 & 0 & \rho_{h_d} \end{bmatrix}$$

$$\Sigma_h = \begin{bmatrix} \sigma_{h_x}\sqrt{1-\rho_{h_x}^2} & 0 & 0 \\ 0 & \sigma_{h_c}\sqrt{1-\rho_{h_c}^2} & 0 \\ 0 & 0 & \sigma_{h_d}\sqrt{1-\rho_{h_d}^2} \end{bmatrix}, \quad w_{t+1} = \begin{bmatrix} w_{x,t+1} \\ w_{c,t+1} \\ w_{d,t+1} \end{bmatrix}.$$

We express

$$\sigma_{x,t} = \varphi_x \bar{\sigma} \exp(h_{x,t}), \quad \sigma_{c,t} = \varphi_c \bar{\sigma} \exp(h_{c,t}), \quad \sigma_{d,t} = \varphi_d \bar{\sigma} \exp(h_{d,t}),$$

which delivers the dependence on h_t in the above definition of $v_{t+1}(\cdot)$.

3.6.4 Posterior Inference

As discussed in the main text, we use a particle-filter approximation of the likelihood function and embed this approximation into a fairly standard random walk Metropolis algorithm.

Particle Filter

Our state-space representation, given by equations (3.47), (3.50), and (3.52), is linear conditional on the volatility states h_t. Thus, following Chen and Liu (2000), we update s_{t+1} conditional on h_t using Kalman filter iterations, which improves the efficiency of the filter substantially. In the subsequent exposition we omit the dependence of all densities on the parameter vector Θ. The particle filter approximates the sequence of distributions $\{p(z_t|Y_{1:t})\}_{t=1}^T$ by a set of pairs $\left\{z_t^{(i)}, \pi_t^{(i)}\right\}_{i=1}^N$, where $z_t^{(i)}$ is the i'th particle vector, $\pi_t^{(i)}$ is its weight, and N is the number of particles. As a by-product, the filter produces a sequence of likelihood approximations $\hat{p}(y_t|Y_{1:t-1})$, $t = 1, \ldots, T$.

- **Initialization:** We generate the particle values $z_0^{(i)}$ by drawing the volatilities (h_0, h_{-1}) from the unconditional distribution associated with (3.52). Conditional on the volatility state $(h_0^{(i)}, h_{-1}^{(i)})$, $s_0^{(i)}$ is generated from the unconditional distribution associated with (3.50). We set $\pi_0^{(i)} = 1/N$ for each i.

- **Propagation of particles:** We simulate (3.52) forward to generate $(h_t^{(i)}, h_{t-1}^{(i)})$ conditional on $(h_{t-1}^{(i)}, h_{t-2}^{(i)})$. Taking $s_{t-1}^{(i)}$ and $(h_t^{(i)}, h_{t-1}^{(i)})$ as given, for each particle we run one iteration of the Kalman filter based on the linear state-space system comprised of (3.47)and (3.50) to determine $p(s_t|y_t, s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)})$. This distribution is normal with mean $s_{t|t}^{(i)}$ and $P_{t|t}^{(i)}$. We sample $s_t^{(i)}$ from $N\big(s_{t|t}^{(i)}, P_{t|t}^{(i)}\big)$. We use $q(z_t|z_{t-1}^{(i)}, y_t)$ to represent the distribution from which we draw $z_t^{(i)}$.

- **Correction of particle weights:** Define the unnormalized particle weights for period t as

$$\begin{aligned}
\tilde{\pi}_t^{(i)} &= \pi_{t-1}^{(i)} \times \frac{p(y_t|z_t^{(i)})p(z_t^{(i)}|z_{t-1}^{(i)})}{q(z_t^{(i)}|z_{t-1}^{(i)}, y_t)} \\
&= \pi_{t-1}^{(i)} \times \frac{p(y_t|z_t^{(i)})p(z_t^{(i)}|z_{t-1}^{(i)})}{p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)}, y_t)q(h_t^{(i)}|h_{t-1}^{(i)})} \\
&= \pi_{t-1}^{(i)} \times \frac{p(y_t|z_t^{(i)})p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)})p(h_t^{(i)}|h_{t-1}^{(i)})}{p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)}, y_t)q(h_t^{(i)}|h_{t-1}^{(i)})} \\
&= \pi_{t-1}^{(i)} \times \frac{p(y_t|z_t^{(i)})p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)})}{p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)}, y_t)}.
\end{aligned} \tag{3.53}$$

  The term $\pi_{t-1}^{(i)}$ is the initial particle weight and the ratio $p(y_t|z_t^{(i)})p(z_t^{(i)}|z_{t-1}^{(i)})/q(z_t^{(i)}|z_{t-1}^{(i)}, y_t)$ is the importance weight of the particle. The second equality is obtained by factorizing $q(z_t^{(i)}|z_{t-1}^{(i)}, y_t)$ into the density of $h_t^{(i)}$ associated with the forward simulation of the volatility states, and the conditional density of $s_t|(y_t, s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)})$ is obtained from the Kalman filter updating step. The third equality is obtained by factorizing the joint density of $(s_t^{(i)}, h_t^{(i)})$, $p(z_t^{(i)}|z_{t-1}^{(i)})$, into a marginal density for $h_t^{(i)}$ and a conditional density for $s_t^{(i)}|h_t^{(i)}$. The last equality follows from the fact that we chose $q(h_t^{(i)}|h_{t-1}^{(i)}) = p(h_t^{(i)}|h_{t-1}^{(i)})$. We further simplify the expression in the last line of (3.53) in the next subsection.

  The log likelihood function approximation is given by

$$\log \hat{p}(y_t|Y_{1:t-1}) = \log \hat{p}(y_{t-1}|Y_{1:t-2}) + \log\left(\sum_{i=1}^{N} \tilde{\pi}_t^{(i)}\right).$$

- **Resampling:** Define the normalized weights

$$\pi_t^{(i)} = \frac{\tilde{\pi}_t^{(i)}}{\sum_{j=1}^{N} \tilde{\pi}_t^{(j)}}$$

and generate N draws from the distribution $\{s_t^{(i)}, \pi_t^{(i)}\}_{i=1}^N$ using multinomial resampling. In slight abuse of notation, we denote the resampled particles and their weights also by $s_t^{(i)}$ and $\pi_t^{(i)}$, where $\pi_t^{(i)} = 1/N$.

Further Details on the Correction and Updating Step

We now derive the density $p(s_t|y_t, s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)})$ as well as a simplified expression for the density ratio in the last line of (3.53). Recall that, conditional on the volatilities (h_t, h_{t-1}), the state-space representation of our model takes the form

$$y_t = A_t\big(D + Zs_t + Z^v s_t^v + \Sigma^u u_t\big), \quad u_t \sim N(0, I) \tag{3.54}$$

$$s_t = \Phi s_{t-1} + v_t(h_{t-1}). \tag{3.55}$$

We now proceed with a Kalman filter forecasting and updating step. Conditional on $(s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)})$, the state-transition equation can be used to forecast s_t:

$$s_t\big|(s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)}) \sim N\big(s_{t|t-1}^{(i)}, P_{t|t-1}^{(i)}\big),$$

where

$$s_{t|t-1}^{(i)} = \Phi s_{t-1}^{(i)}, \quad P_{t|t-1}^{(i)} = \mathbb{E}[v_t(h_{t-1}^{(i)}) v_t'(h_{t-1}^{(i)})].$$

Using the measurement equation we can forecast y_t, conditional on $(s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)})$, as follows:

$$y_t\big|(s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)}) \sim N\big(\hat{y}_{t|t-1}^{(i)}, F_{t|t-1}^{(i)}\big), \tag{3.56}$$

where

$$\hat{y}_{t|t-1}^{(i)} = A_t\big(D + Zs_{t|t-1}^{(i)} + Z^v s_t^v(h_t^{(i)}, h_{t-1}^{(i)})\big), \quad F_{t|t-1}^{(i)} = (A_t Z) P_{t|t-1}^{(i)} (A_t Z)' + (A_t \Sigma^u)(A_t \Sigma^u)'.$$

Finally, we can apply the Kalman filter updating step to obtain

$$s_t\big|(y_t, s_{t-1}^{(i)}, h_t^{(i)}, h_{t-1}^{(i)}) \sim N\big(s_{t|t}^{(i)}, P_{t|t}^{(i)}\big), \tag{3.57}$$

where

$$\begin{aligned} s_{t|t}^{(i)} &= s_{t|t-1}^{(i)} + (A_t Z P_{t|t-1}^{(i)})'(F_{t|t-1}^{(i)})^{-1}(y_t - \hat{y}_{t|t-1}^{(i)}) \\ P_{t|t}^{(i)} &= P_{t|t-1}^{(i)} - (A_t Z P_{t|t-1}^{(i)})'(F_{t|t-1}^{(i)})^{-1}(A_t Z P_{t|t-1}^{(i)}). \end{aligned}$$

Define $\mathcal{F}^{(i)} = \{h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)}\}$ and consider the density ratio used to update the particle weights:

$$\begin{aligned} \frac{p(y_t|z_t^{(i)})p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)})}{p(s_t^{(i)}|h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)}, y_t)} &= \frac{p(y_t|s_t^{(i)}, \mathcal{F}^{(i)})p(s_t^{(i)}|\mathcal{F}^{(i)})}{p(s_t^{(i)}|y_t, \mathcal{F}^{(i)})} \\ &= \frac{p(s_t^{(i)}|y_t, \mathcal{F}^{(i)})p(y_t|\mathcal{F}^{(i)})}{p(s_t^{(i)}|y_t, \mathcal{F}^{(i)})} \\ &= p(y_t|\mathcal{F}^{(i)}). \end{aligned} \tag{3.58}$$

The first equality in (3.58) follows from

$$p(y_t|z_t^{(i)}) = p(y_t|s_t^{(i)}, h_t^{(i)}, h_{t-1}^{(i)}) = p(y_t|s_t^{(i)}, h_t^{(i)}, h_{t-1}^{(i)}, s_{t-1}^{(i)})$$

and the second equality in (3.58) is an application of Bayes' Theorem. The expression for $p(y_t|\mathcal{F}^{(i)}$ was previously derived in (3.56).

3.6.5 The Measurement-Error Model for Consumption

Monthly Interpolation and Adjustment of Consumption

For expositional purposes, we assume that the accurately measured low-frequency observations are available at quarterly frequency (instead of annual frequency as in the main text). Correspondingly, we define the time subscript $t = 3(j-1)+m$, where month $m = 1, 2, 3$ and quarter $j = 1, \ldots$. We use uppercase C to denote the level of consumption and lowercase c to denote percentage deviations from some log-linearization point. Growth rates are approximated as log differences and we use a superscript o to distinguish observed from "true" values.

The measurement-error model presented in the main text can be justified by assuming that the statistical agency uses a high-frequency proxy series to determine monthly consumption growth rates. We use $Z_{3(j-1)+m}$ to denote the monthly value of the proxy series and $Z^q_{(j)}$ the quarterly aggregate. Suppose the proxy variable provides a noisy measure of monthly consumption. More specifically, we consider a multiplicative error model of the form

$$Z_{3(j-1)+m} = C_{3(j-1)+m} \exp(\epsilon_{3(j-1)+m}). \tag{3.59}$$

The interpolation is executed in two steps. In the first step we construct a series $\tilde{C}^o_{3(j-1)+m}$, and in the second step we rescale the series to ensure that the reported monthly consumption data add up to the reported quarterly consumption data within the period. In Step 1, we start from the level of consumption in quarter $j-1$, $C^q_{(j-1)}$, and define

$$\begin{aligned}
\tilde{C}^o_{3(j-1)+1} &= C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} \right) \\
\tilde{C}^o_{3(j-1)+2} &= C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} \right) \left(\frac{Z_{3(j-1)+2}}{Z_{3(j-1)+1}} \right) = C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+2}}{Z^q_{(j-1)}} \right) \\
\tilde{C}^o_{3(j-1)+3} &= C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} \right) \left(\frac{Z_{3(j-1)+2}}{Z_{3(j-1)+1}} \right) \left(\frac{Z_{3(j-1)+3}}{Z_{3(j-1)+2}} \right) = C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+3}}{Z^q_{(j-1)}} \right).
\end{aligned} \tag{3.60}$$

Thus, the growth rates of the proxy series are used to generate monthly consumption data for quarter q. Summing over the quarter yields

$$\tilde{C}^{q,o}_{(j)} = \sum_{m=1}^{3} \tilde{C}^{o}_{3(j-1)+m} = C^{q,o}_{(j-1)} \left[\frac{Z_{3(j-1)+1}}{Z^{q}_{(j-1)}} + \frac{Z_{3(j-1)+2}}{Z^{q}_{(j-1)}} + \frac{Z_{3(j-1)+3}}{Z^{q}_{(j-1)}} \right] = C^{q,o}_{(j-1)} \frac{Z^{q}_{(j)}}{Z^{q}_{(j-1)}} \quad (3.61)$$

In Step 2, we adjust the monthly estimates $\tilde{C}^{o}_{3(j-1)+m}$ by the factor $C^{q,o}_{(j)}/\tilde{C}^{q,o}_{(j)}$, which leads to

$$\begin{aligned} C^{o}_{3(j-1)+1} &= \tilde{C}^{o}_{3(j-1)+1} \left(\frac{C^{q,o}_{(j)}}{\tilde{C}^{q,o}_{(j)}} \right) = C^{q,o}_{(j)} \frac{Z_{3(j-1)+1}}{Z^{q}_{(j)}} \\ C^{o}_{3(j-1)+2} &= \tilde{C}^{o}_{3(j-1)+2} \left(\frac{C^{q,o}_{(j)}}{\tilde{C}^{q,o}_{(j)}} \right) = C^{q,o}_{(j)} \frac{Z_{3(j-1)+2}}{Z^{q}_{(j)}} \\ C^{o}_{3(j-1)+3} &= \tilde{C}^{o}_{3(j-1)+3} \left(\frac{C^{q,o}_{(j)}}{\tilde{C}^{q,o}_{(j)}} \right) = C^{q,o}_{(j)} \frac{Z_{3(j-1)+3}}{Z^{q}_{(j)}} \end{aligned} \quad (3.62)$$

and guarantees that

$$C^{q,o}_{(j)} = \sum_{m=1}^{3} C^{o}_{3(j-1)+m}.$$

We now define the growth rates $g^{o}_{c,t} = \log C^{o}_{t} - \log C^{o}_{t-1}$ and $g_{c,t} = \log C_t - \log C_{t-1}$. By taking logarithmic transformations of (3.59) and (3.62) and combining the resulting equations, we can deduce that the growth rates for the second and third month of quarter q are given by

$$\begin{aligned} g^{o}_{c,3(j-1)+2} &= g_{c,3(j-1)+2} + \epsilon_{3(j-1)+2} - \epsilon_{3(j-1)+1} \\ g^{o}_{c,3(j-1)+3} &= g_{c,3(j-1)+3} + \epsilon_{3(j-1)+3} - \epsilon_{3(j-1)+2}. \end{aligned} \quad (3.63)$$

The derivation of the growth rate between the third month of quarter $j-1$ and the first month of quarter j is a bit more cumbersome. Using (3.62), we can write the growth rate as

$$\begin{aligned} g^{o}_{c,3(j-1)+1} &= \log C^{q,o}_{(j)} + \log Z_{3(j-1)+1} - \log Z^{q}_{(j)} \\ &\quad - \log C^{q,o}_{(j-1)} - \log Z_{3(j-2)+3} + \log Z^{q}_{(j-1)}. \end{aligned} \quad (3.64)$$

To simplify (3.64) further, we are using a log-linear approximation. Suppose we log-linearize an equation of the form

$$X^q_{(j)} = X_{3(j-1)+1} + X_{3(j-1)+2} + X_{3(j-1)+3}$$

around X^q_* and $X_* = X^q_*/3$, using lowercase variables to denote percentage deviations from the log-linearization point. Then,

$$x^q_{(j)} \approx \frac{1}{3}(x_{3(j-1)+1} + x_{3(j-1)+2} + x_{3(j-1)+3}).$$

Using (3.59) and the definition of quarterly variables as sums of monthly variables, we can apply the log-linearization as follows:

$$\log C^{q,o}_{(j)} - \log Z^q_{(j)} = \log(C^q_*/Z^q_*) + \epsilon^q_{(j)} - \frac{1}{3}\big(\epsilon_{3(j-1)+1} + \epsilon_{3(j-1)+2} + \epsilon_{3(j-1)+3}\big). \qquad (3.65)$$

Substituting (3.65) into (3.64) yields

$$\begin{aligned} g^o_{c,3(j-1)+1} \;=\; & g_{c,3(j-1)+1} + \epsilon_{3(j-1)+1} - \epsilon_{3(j-2)+3} + \epsilon^q_{(j)} - \epsilon^q_{(j-1)} \qquad (3.66) \\ & -\frac{1}{3}\big(\epsilon_{3(j-1)+1} + \epsilon_{3(j-1)+2} + \epsilon_{3(j-1)+3}\big) + \frac{1}{3}\big(\epsilon_{3(j-2)+1} + \epsilon_{3(j-2)+2} + \epsilon_{3(j-2)+3}\big). \end{aligned}$$

An "annual" version of this equation appears in the main text.

Chapter 4

Improving GDP Measurement: A Measurement-Error Perspective

4.1 Introduction

Aggregate real output is surely the most fundamental and important concept in macroeconomic theory. Surprisingly, however, significant uncertainty still surrounds its measurement. In the U.S., in particular, two often-divergent GDP estimates exist, a widely-used expenditure-side version, GDP_E, and a much less widely-used income-side version, GDP_I.[1] Nalewaik (2010) and Fixler and Nalewaik (2009) make clear that, at the very least, GDP_I deserves serious attention and may even have properties in certain respects superior to those of GDP_E.[2] That is, if forced to choose between GDP_E and GDP_I, a surprisingly strong case exists for GDP_I. But of course one is *not* forced to choose between GDP_E and GDP_I, and a GDP estimate based on *both* GDP_E and GDP_I may be superior to either one alone. In this paper we propose and implement a framework for obtaining such a blended estimate.

Our work is related to, and complements, Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012). There we took a forecast-error perspective, whereas here we take a

[1] Indeed we will focus on the U.S. because it is a key egregious example of unreconciled GDP_E and GDP_I estimates.

[2] For additional informative background on GDP_E, GDP_I, the statistical discrepancy, and the national accounts more generally, see of Economic Analysis (2006), McCulla and Smith (2007), Landefeld, Seskin, and Fraumeni (2008), and Rassier (2012).

measurement-error perspective.[3] In particular, we work with a dynamic factor model in the tradition of Geweke (1977) and Sargent and Sims (1977), as used and extended by Watson and Engle (1983), Edwards and Howrey (1991), Harding and Scutella (1996), Jacobs and van Norden (2011), Kishor and Koenig (2011), and Fleischman and Roberts (2011), among others.[4] That is, we view "true GDP" as a latent variable on which we have several indicators, the two most obvious being GDP_E and GDP_I, and we then extract true GDP using optimal filtering techniques.

The measurement-error approach is time honored, intrinsically compelling, and very different from the forecast-combination perspective of Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012), for several reasons.[5] First, it enables extraction of latent true GDP using a model with parameters *estimated* with exact likelihood or Bayesian methods, whereas the forecast-combination approach forces one to use calibrated parameters. Second, it delivers not only point extractions of latent true GDP but also interval extractions, enabling us to assess the associated uncertainty. Third, the state-space framework in which the measurement-error models are embedded facilitates exploration of the relationship between GDP measurement errors and the economic environment, such as stage of the business cycle, which is of special interest. Fourth, the state-space framework facilitates real-time analysis and forecasting, despite the fact that preliminary GDP_I data are not available as quickly as those for GDP_E.

We proceed as follows. In section 4.2 we consider several measurement-error models and assess their identification status, which turns out to be challenging and interesting in

[3]Hence the pair of papers roughly parallels the well-known literature on "forecast error" and "measurement error" properties of of data revisions; see Mankiw, Runkle, and Shapiro (1984), Mankiw and Shapiro (1986), Faust, Rogers, and Wright (2005), and Aruoba (2008).

[4]See also Smith, Weale, and Satchell (1998), who take a different but related approach, and the independent work of Greenaway-McGrevy (2011), who takes a closely-related approach but unfortunately estimates a model that we will show to be unidentified.

[5]On the time-honored aspect, see, for example, Gartaganis and Goldberger (1955).

the most realistic and hence compelling case. In section 4.3 we discuss the data, estimation framework and estimation results. In section 4.4 we explore the properties of our new GDP series. We conclude in section 4.5.

4.2 Five Measurement-Error Models of GDP

We use dynamic-factor measurement-error models, which embed the idea that both GDP_E and GDP_I are noisy measures of latent true GDP. We work throughout with growth rates of GDP_E, GDP_I and GDP (hence, for example, GDP_E denotes a growth rate).[6]

We assume throughout that true GDP growth transitions with simple $AR(1)$ dynamics, and we entertain several measurement structures, to which we now turn.

4.2.1 (Identified) 2-Equation Model: Σ Diagonal

We begin with the simplest 2-equation model; the measurement errors are orthogonal to each other and to transition shocks at all leads and lags.[7] The model has a natural state-space structure, and we write

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{4.1}$$

$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where GDP_{Et} and GDP_{It} are expenditure- and income-side estimates, respectively, GDP_t is latent true GDP, and all shocks are Gaussian and uncorrelated at all leads and lags.

[6] We will elaborate on the reasons for this choice later in section 4.3.

[7] Here and throughout, when we say "N-equation" state-space model, we mean that the measurement equation is an N-variable system.

That is, $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It})' \sim iid\, N(\underline{0}, \Sigma)$, where

$$\Sigma = \begin{bmatrix} \sigma_{GG}^2 & 0 & 0 \\ 0 & \sigma_{EE}^2 & 0 \\ 0 & 0 & \sigma_{II}^2 \end{bmatrix}. \tag{4.2}$$

The Kalman smoother will deliver optimal extractions of GDP_t conditional upon observed expenditure- and income-side measurements. Moreover, the model can be easily extended, and some of its restrictive assumptions relaxed, with no fundamental change. We now proceed to do so.

4.2.2 (Identified) 2-Equation Model: Σ Block-Diagonal

The first extension is to allow for correlated measurement errors. This is surely important, as there is roughly a 25 percent overlap in the counts embedded in GDP_E and GDP_I, and moreover, the same deflator is used for conversion from nominal to real magnitudes.[8] We write

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{4.3}$$

$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where now ϵ_{Et} and ϵ_{It} may be correlated contemporaneously but are uncorrelated at all other leads and lags, and all other definitions and assumptions are as before; in particular, ϵ_{Gt} and $(\epsilon_{Et}, \epsilon_{It})'$ are uncorrelated at all leads and lags. That is, $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It})' \sim iid\, N(\underline{0}, \Sigma)$, where

$$\Sigma = \begin{bmatrix} \sigma_{GG}^2 & 0 & 0 \\ 0 & \sigma_{EE}^2 & \sigma_{EI}^2 \\ 0 & \sigma_{IE}^2 & \sigma_{II}^2 \end{bmatrix}. \tag{4.4}$$

Nothing is changed, and the Kalman filter retains its optimality properties.

[8] See Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012) for more. Many of the areas of overlap are particularly poorly measured, such as imputed financial services, housing services, and government output.

4.2.3 (Unidentified) 2-Equation Model, Σ Unrestricted

The second key extension is motivated by Fixler and Nalewaik (2009) and Nalewaik (2010), who document cyclicality in the statistical discrepancy ($GDP_E - GDP_I$), which implies failure of the assumption that $(\epsilon_{Et}, \epsilon_{It})'$ and ϵ_{Gt} are uncorrelated at all leads and lags. Of particular concern is contemporaneous correlation between ϵ_{Gt} and $(\epsilon_{Et}, \epsilon_{It})'$. Hence we allow the measurement errors $(\epsilon_{Et}, \epsilon_{It})'$ to be correlated with GDP_t, or more precisely, correlated with GDP_t innovations, ϵ_{Gt}. We write

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{4.5}$$

$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It})' \sim iid\, N(\underline{0}, \Sigma)$, with

$$\Sigma = \begin{bmatrix} \sigma^2_{GG} & \sigma^2_{GE} & \sigma^2_{GI} \\ \sigma^2_{EG} & \sigma^2_{EE} & \sigma^2_{EI} \\ \sigma^2_{IG} & \sigma^2_{IE} & \sigma^2_{II} \end{bmatrix}. \tag{4.6}$$

In this environment the standard Kalman filter is rendered sub-optimal for extracting GDP, due to correlation between ϵ_{Gt} and $(\epsilon_{Et}, \epsilon_{It})$, but appropriately-modified optimal filters are available.

Of course in what follows we will be concerned with *estimating* our measurement-equation models, so we will be concerned with identification. The diagonal-Σ model (4.1)-(4.2) and the block-diagonal-Σ model (4.3)-(4.4) are identified. Identification of less-restricted dynamic factor models, however, is a very delicate matter. In particular, it is not obvious that the unrestricted-Σ model (4.5)-(4.6) is identified. Indeed it is not, as we prove in Appendix 4.6.1. Hence we now proceed to determine minimal restrictions that achieve identification.

4.2.4 (Identified) 2-Equation Model: Σ Restricted

The identification problem with the general model (4.5)-(4.6) stems from the fact that we can make true GDP more volatile (increase σ^2_{GG}) and make the measurement errors more volatile (increase σ^2_{EE} and σ^2_{II}), but reduce the covariance between the fundamental shocks and the measurement errors (reduce σ^2_{EG} and σ^2_{IG}), without changing the distribution of observables.

Restricting the Original Parameterization

But we can achieve identification by slightly restricting parameterization (4.5)-(4.6). In particular, as we show in Appendix 4.6.1, the unrestricted system (4.5)-(4.6) is unidentified because the Σ matrix has six free parameters with only five moment conditions to determine them. Hence we can achieve identification by restricting any single element of Σ. Imposing any such restriction would seem challenging, however, as we have no strong prior views directly on any single element of Σ. Fortunately, however, a simple re-parameterization exists about which we have a more natural prior view, to which we now turn.

A Useful Re-Parameterization

Let

$$\zeta = \frac{\frac{1}{1-\rho^2}\sigma^2_{GG}}{\frac{1}{1-\rho^2}\sigma^2_{GG} + 2\sigma^2_{GE} + \sigma^2_{EE}}, \tag{4.7}$$

the variance of latent true GDP relative to the variance of expenditure-side measured GDP_E. Then, rather than fixing an element of Σ to achieve identification, we can fix ζ, about which we have a more natural prior view. In particular, at first pass we might take $\sigma^2_{GE} \approx 0$, in which case $0 < \zeta < 1$. Or, put differently, $\zeta > 1$ would require a *very* negative σ^2_{GE}, which seems unlikely. All told, we view a ζ value less than, but close to, 1.0 as most natural. We take $\zeta = 0.80$ as our benchmark in the empirical work that follows, although we explore a wide range of ζ values both below and above 1.0.

4.2.5 (Identified) 3-Equation Model: Σ Unrestricted

Thus far we showed how to achieve identification by fixing a parameter, ζ, and we noted that our prior is centered around $\zeta = 0.80$. It is of also of interest to know whether we can get some complementary data-based guidance on choice of ζ. The answer turns out to be yes, by adding a third measurement equation with a certain structure.

Suppose, in particular, that we have an additional observable variable U_t that loads on true GDP_t with measurement error orthogonal to those of GDP_I and GDP_E. In particular, consider the 3-equation model

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \\ U_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \lambda \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \tag{4.8}$$

$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It}, \epsilon_{Ut})' \sim iid\, N(\underline{0}, \Omega)$, with

$$\Omega = \begin{bmatrix} \sigma^2_{GG} & \sigma^2_{GE} & \sigma^2_{GI} & \sigma^2_{GU} \\ \sigma^2_{EG} & \sigma^2_{EE} & \sigma^2_{EI} & 0 \\ \sigma^2_{IG} & \sigma^2_{IE} & \sigma^2_{II} & 0 \\ \sigma^2_{UG} & 0 & 0 & \sigma^2_{UU} \end{bmatrix}. \tag{4.9}$$

Note that the upper-left 3x3 block of Ω is just Σ, which is now unrestricted. Nevertheless, as we prove in Appendix 4.6.2, the 3-equation model (4.8)-(4.9) is identified. Of course some of the remaining elements of the overall 4x4 covariance matrix Ω *are* restricted, which is how we achieve identification in the 3-equation model, but the economically interesting sub-matrix, which the 3-equation model leaves completely unrestricted, is Σ.

Depending on the application, of course, it is not obvious that an identifying variable U_t with measurement errors orthogonal to those of GDP_E and GDP_I (i.e., with stochastic properties that satisfy (4.9)), is available. Hence it is not obvious that estimation of the 3-equation model (4.8)-(4.9) is feasible in practice, despite the model's appeal in principle. Indeed, much of the data collected from business surveys is used in the BEA's estimates,

invalidating use of that data as U_t since any measurement error in that data appears directly in either GDP_E or GDP_I, producing correlation across the measurement errors. Moreover, variables drawn from business surveys similar to those used to produce GDP_E and GDP_I, even if they are not used directly in the estimation of GDP_E and GDP_I, might still be invalid identifying variables if the survey methodology itself produces similar measurement errors.[9]

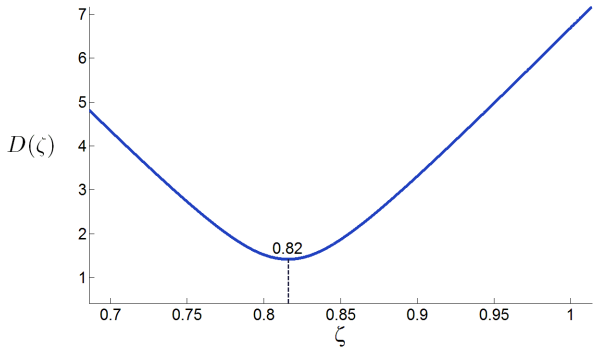
Fortunately, however, some important macroeconomic data is collected not from surveys of businesses, but from samples of households. A sample of data drawn from a universe of households seems likely to have measurement errors that are different than those contaminating a data sample drawn from a universe of businesses, especially when the "universes" of businesses and households are not complete census counts, as is the case here. For example, the universe of business surveys is derived from tax records, so businesses not paying taxes will not appear on that list, but individuals working at that business may appear in the universe of households.

Importantly, very little data collected from household surveys are used to construct GDP_E and GDP_I, so a U_t variable computed from a household survey seems most likely to meet our identification conditions. The change in the unemployment rate is a natural choice (hence our notational choice U_t). U_t arguably loads on true GDP with a measurement error orthogonal to those of GDP_E and GDP_I, because the U_t data is being produced independently (by the BLS rather than BEA) from different types of surveys. In addition, virtually all of the GDP_E and GDP_I data are estimated in nominal dollars and then converted to real dollars using a price deflator, whereas U_t is estimated directly with no deflation.

All told,we view "3-equation identification" as a useful complement to the "ζ-identification"

[9] For example, if the business surveys used to produce GDP_E and GDP_I tend to oversample large firms, variables drawn from a business survey that also oversamples large firms may have measurement errors that are correlated with those in GDP_E and GDP_I, absent appropriate corrections.

Figure 4.1: Divergence Between $\hat{\Sigma}_\zeta$ and $\hat{\Sigma}_3$



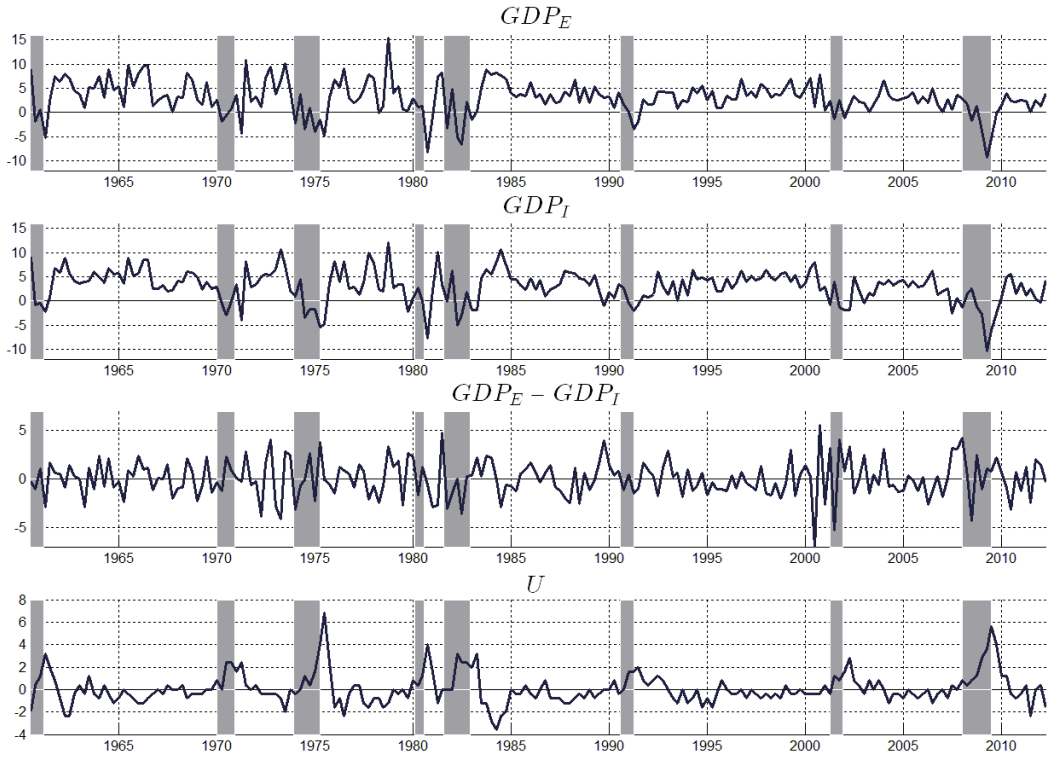
Notes: We show the Frobenius-norm divergence $D(\zeta)$ between $\widehat{\Sigma}_\zeta$ and $\widehat{\Sigma}_3$ as a function of ζ. The optimum is $\zeta = 0.82$. See text for details.

discussed earlier in section 4.2.4. All identifications involve assumptions. ζ-identification involves introspection about likely values of ζ, given its structure and components, and that introspection is of course subject to error. 3-equation identification involves introspection about various measurement-error correlations involving the newly-introduced third variable, which is of course also subject to error. Indeed the two approaches to identification are usefully used in tandem, and compared.

One can even view the 3-equation approach as a device for implicitly selecting ζ. In particular, we can find the ζ implied by the 3-equation model estimate, that is, find the ζ that minimizes the divergence between $\hat{\Sigma}_\zeta$ and $\hat{\Sigma}_3$, in an obvious notation.[10] For example, using the Frobenius matrix-norm to measure divergence, we obtain an optimum of $\zeta^* = 0.82$. We show the full surface in Figure 4.1, and the minimum is sharp and unique. The implied ζ^* of 0.82 is of course quite close to the directly-assessed value of 0.80 at which we arrived earlier, which lends additional credibility to the earlier assessment.

[10] We will discuss subsequently the estimation procedure used to obtain $\hat{\Sigma}_\zeta$ and $\hat{\Sigma}_3$.

Figure 4.2: GDP and Unemployment Data



Notes: GDP_E and GDP_I are in growth rates and U_t is in changes. All are measured in annualized percent.

4.3 Data and Estimation

We intentionally work with a stationary system in growth rates, because we believe that measurement errors are best modeled as *iid* in growth rates rather than in levels, due to BEA's devoting maximal attention to estimating the "best change." [11] In its above-cited "Concepts and Methods ..." document, for example,the BEA emphasizes that:

> Best change provides the most accurate measure of the period-to-period movement in an economic statistic using the best available source data. In an annual revision of the NIPAs, data from the annual surveys of manufacturing and trade

[11] For example, see "Concepts and Methods in the U.S. National Income and Product Accounts," available at `http://www.bea.gov/national/pdf/methodology/chapters1-4.pdf`.

> are generally incorporated into the estimates on a best-change basis. In the current quarterly estimates, most of the components are estimated on a best-change basis from the annual levels established at the most recent annual revision.

The monthly source data used to estimate GDP_E (such as retail sales) and GDP_I (such as nonfarm payroll employment) are generally produced on a best-change basis as well, using a so-called "link-relative estimator." This estimator computes growth rates using firms in the sample in both the current and previous months, in contrast to a best-level estimator, which would generally use all the firms in the sample in the current month regardless of whether or not they were in the sample in the previous month. For example, for retail sales the BEA notes that: [12]

> Advance sales estimates for the most detailed industries are computed using a type of ratio estimator known as the link-relative estimator. For each detailed industry, we compute a ratio of current-to-previous month weighted sales using data from units for which we have obtained usable responses for both the current and previous month.

Indeed the BEA produces estimates on a best-level basis only at 5-year benchmarks. These best-level benchmark revisions should drive only the very-low frequency variation in GDP_E, and thus probably matter very little for the quarterly growth rates estimated on a best-change basis.

4.3.1 Descriptive Statistics

We show time-series plots of the "raw" GDP_E and GDP_I data in Figure 4.2, and we show summary statistics in the top panel of Table 4.1. Not captured in the table but also true is that the raw data are highly correlated; the simple correlations are $corr(GDP_E, GDP_I) =$

[12] See `http://www.census.gov/retail/marts/how_surveys_are_collected.html`.

Table 4.1: Descriptive Statistics for Various GDP Series

	$\bar{x}$	50%	$\hat{\sigma}$	Sk	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	Q_{12}	$\hat{\sigma}_e$	R^2	$\hat{V}_e$
GDP_E	3.03	3.04	3.49	-0.31	.33	.27	.08	.09	47.07	3.28	.06	12.12
GDP_I	3.02	3.39	3.40	-0.55	.47	.27	.22	.08	81.60	2.99	.12	11.43
GDP_M 2-eqn, Σ diag	3.02	3.22	3.00	-0.56	.56	.34	.21	.09	108.25	2.48	.18	8.92
GDP_M 2-eqn, Σ block	3.02	3.35	2.64	-0.64	.70	.45	.28	.13	170.08	1.89	.29	6.90
GDP_M 2-eqn, $\zeta = 0.65$	3.02	3.32	2.61	-0.64	.67	.43	.27	.12	157.56	1.92	.26	6.73
GDP_M 2-eqn, $\zeta = 0.75$	3.02	3.30	2.77	-0.63	.65	.41	.26	.11	148.23	2.08	.25	7.60
GDP_M 2-eqn, $\zeta = 0.80$	3.02	3.29	2.87	-0.62	.64	.39	.25	.11	141.14	2.19	.24	8.16
GDP_M 2-eqn, $\zeta = 0.85$	3.02	3.31	2.89	-0.64	.66	.41	.28	.12	153.27	2.15	.25	8.29
GDP_M 2-eqn, $\zeta = 0.95$	3.02	3.26	3.02	-0.64	.66	.40	.28	.12	149.61	2.27	.25	9.07
GDP_M 2-eqn, $\zeta = 1.05$	3.01	3.22	3.12	-0.65	.67	.40	.28	.12	155.60	2.30	.26	9.69
GDP_M 2-eqn, $\zeta = 1.15$	3.04	3.34	3.07	-0.67	.76	.47	.31	.15	201.15	1.99	.35	9.46
GDP_M 3-eqn	3.02	3.37	3.02	-1.14	.63	.37	.21	.03	141.79	2.33	.23	9.03
GDP_F	3.02	3.29	3.30	-0.51	.46	.29	.19	.07	78.28	2.92	.12	10.80

Notes: The sample period is 1960Q1-2011Q4. In the top panel we show statistics for the raw data. In the middle panel we show statistics for various posterior-median measurement-error-based ("M") estimates of true GDP, where all estimates are smoothed extractions. In the bottom panel we show statistics for the forecast-error-based estimate of true GDP produced by Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012), GDP_F. $\bar{x}$, 50%, $\hat{\sigma}$ and Sk are sample mean, median, standard deviation and skewness, respectively, and $\hat{\rho}_\tau$ is a sample autocorrelation at a displacement of τ quarters. Q_{12} is the Ljung-Box serial correlation test statistic calculated using $\hat{\rho}_1$, ..., $\hat{\rho}_{12}$. $R^2 = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}^2}$, where $\hat{\sigma}_e$ denotes the estimated disturbance standard deviation from a fitted $AR(1)$ model, is a predictive R^2. $\hat{V}_e$ is the unconditional variance implied by a fitted AR_1 model, $\hat{V}_e = \frac{\hat{\sigma}_e^2}{1-\hat{\rho}^2}$.

0.85, $corr(GDP_E, U) = -0.67$, and $corr(GDP_I, U) = -0.73$. Median GDP_I growth is a bit higher than that of GDP_E, and GDP_I growth is noticeably more persistent than that of GDP_E. Related, GDP_I also has smaller $AR(1)$ innovation variance and greater predictability as measured by the predictive R^2.[13]

4.3.2 Bayesian Analysis of Measurement-Error Models

Here we describe Bayesian analysis of our three-equation model, which of course also includes our various two-equation models as special cases. Bayesian estimation involves parameter estimation and latent state smoothing. First, we generate draws from the posterior distribution of the model parameters using a Random-Walk Metropolis-Hastings algorithm.

[13] On this and related predictability measures, see Diebold and Kilian (2001).

Next, we apply a simulation smoother as described in Durbin and Koopman (2001a) to obtain draws of the latent states conditional on the parameters.

State-Space Representation

We proceed by introducing a state-space representation of (4.8) for estimation. Let $y_t = [GDP_{Et}, GDP_{It}, U_t]'$, $C = [0, 0, \kappa]'$, $s_t = [GDP_t, \epsilon_{Et}, \epsilon_{It}, \epsilon_{Ut}]'$, $D = [\mu(1-\rho), 0, 0, 0]'$, $\epsilon_t = [\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It}, \epsilon_{Ut}]'$ and

$$Z = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \lambda & 0 & 0 & 1 \end{bmatrix}, \quad \Phi = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Our state-space model is

$$y_t = C + Z s_t \tag{4.10}$$

$$s_t = D + \Phi s_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \Omega).$$

We collect the parameters in (4.10) in $\Theta = (\mu, \rho, \sigma^2_{GG}, \sigma^2_{GE}, \sigma^2_{GI}, \sigma^2_{EE}, \sigma^2_{EI}, \sigma^2_{II}, \sigma^2_{GU}, \sigma^2_{UU}, \kappa, \lambda)$.

Metropolis-Hastings MCMC Algorithm

Now let us proceed to our implementation of the Metropolis-Hastings MCMC Algorithm. Denote the number of MCMC draws by N. We first maximize the posterior density

$$p(\Theta|Y_{1:T}) \propto p(Y_{1:T}|\Theta)p(\Theta) \tag{4.11}$$

to obtain the mode Θ^0 and construct a covariance matrix for the proposal density, Σ_Θ, from the inverse Hessian of the log posterior density evaluated at Θ^0. We also use Θ^0 to initialize the algorithm. At each iteration j we draw a proposed parameter vector $\Theta^* \sim N(\Theta^{j-1}, c\Sigma_\Theta)$, where c is a scalar tuning parameter that we calibrate to achieve an acceptance rate of 25-30%. We accept the proposed parameter vector, that is, we set $\Theta^j = \Theta^*$, with probability min $\left\{1, \frac{p(Y_{1:T}|\Theta^*)p(\Theta^*)}{p(Y_{1:T}|\Theta^{j-1})p(\Theta^{j-1})}\right\}$, and set $\Theta^j = \Theta^{j-1}$ otherwise. We

adopt the convention that $p(\Theta^*) = 0$ if the covariance matrix Ω implied by Θ^* is not positive definite. The results reported subsequently are based on $N = 50,000$ iterations of the algorithm. We discard the first 25,000 draws and use the remaining draws to compute summary statistics for the posterior distribution.

Filtering and Smoothing

The evaluation of the likelihood function $p(Y_{1:T}|\Theta)$ requires the use of the Kalman filter. The Kalman filter recursions take the following form. Suppose that

$$s_{t-1}|(Y_{1:t-1}, \Theta) \sim N(s_{t-1|t-1}, P_{t-1|t-1}), \tag{4.12}$$

where $s_{t-1|t-1}$ and $P_{t-1|t-1}$ are the mean and variance of the latent state at $t-1$. Then the means and variances of the predictive densities $p(s_t|Y_{1:t-1}, \Theta)$ and $p(y_t|Y_{1:t-1}, \Theta)$ are

$$\begin{aligned} s_{t|t-1} &= D + \Phi s_{t-1|t-1}, \quad P_{t|t-1} = \Phi P_{t-1|t-1}\Phi' + \Omega \\ y_{t|t-1} &= C + Z s_{t|t-1}, \quad F_{t|t-1} = Z P_{t|t-1} Z', \end{aligned}$$

respectively. The contribution of observation y_t to the likelihood function $p(Y_{1:T}|\Theta)$ is given by $p(y_t|Y_{1:t-1}, \Theta)$. Finally, the updating equations are

$$\begin{aligned} s_{t|t} &= s_{t|t-1} + (ZP_{t|t-1})' F_{t|t-1}^{-1} \left(y_t - \hat{y}_{t|t-1}\right) \\ P_{t|t} &= P_{t|t-1} - (ZP_{t|t-1})'(ZP_{t|t-1}Z')^{-1}(ZP_{t|t-1}), \end{aligned}$$

leading to

$$s_t|(Y_{1:t}, \Theta) \sim N(s_{t|t}, P_{t|t}). \tag{4.13}$$

We initialize the Kalman filter by drawing $s_{0|0}$ from a mean-zero Gaussian stationary distribution whose covariance matrix, $P_{0|0}$, is the solution of the underlying Ricatti equation.

Because we are interested in inference for the latent GDP, we use the backward-smoothing algorithm of Carter and Kohn (1994a) to generate draws recursively from $s_t|(S_{t+1:T}, Y_{1:T}, \Theta)$,

$t = T-1, T-2, \ldots, 1$, where the last iteration of the Kalman filter recursion provides the initialization for the backward simulation smoother,

$$s_{t|t+1} = s_{t|t} + P_{t|t}\Phi' P_{t+1|t}^{-1}\left(s_{t+1} - D - \Phi s_{t|t}\right) \tag{4.14}$$

$$P_{t|t+1} = P_{t|t} - P_{t|t}\Phi' P_{t+1|t}^{-1}\Phi P_{t|t}$$

$$\text{draw } s_t|(S_{t+1:T}, Y_{1:T}, \Theta) \sim N(s_{t|t+1}, P_{t|t+1}),$$

$t = T-1, T-2, ..., 1$.

4.3.3 Parameter Estimation Results

Here we present and discuss estimation results for our various models. In Table 4.2 we show details of parameter prior and posterior distributions, as well as statistics describing the overall posterior and likelihood, for various 2-equation models, and in Table 4.3 we provide the same information for the 3-equation model.

The complete estimation information in the tables can be difficult to absorb fully, however, so here we briefly present aspects of the results in a more revealing way. For the 2-equation models, the parameters to be estimated are those in the transition equation and those in the covariance matrix Σ, which includes variances and covariances of both transition and measurement shocks. Hence we simply display the estimated transition equation and the estimated Σ matrices. For the 3-equation model, we also need to estimate a factor loading in the measurement equation, so we display the estimated measurement equation as well. Below each posterior median parameter estimate, we show the posterior interquartile range in brackets.

For the 2-equation model with Σ diagonal, we have

$$GDP_t = \underset{[2.81,3.33]}{3.07}(1 - 0.53) + \underset{[0.48,0.57]}{0.53} GDP_{t-1} + \epsilon_{Gt}, \tag{4.15}$$

Table 4.2: Priors and Posteriors, 2-Equation Models, 1960Q1-2011Q4

			Diagonal			Block Diagonal	
	Prior		Posterior			Posterior	
	(Mean,Std.Dev)	25%	50%	75%	25%	50%	75%
μ	N(3,10)	2.81	3.07	3.33	2.77	3.06	3.34
ρ	N(0.3,1)	0.48	0.53	0.57	0.57	0.62	0.68
σ^2_{GG}	IG(10,15)	6.39	6.90	7.44	4.39	5.17	5.95
σ^2_{GE}	N(0,10)	-	-	-	-	-	-
σ^2_{GI}	N(0,10)	-	-	-	-	-	-
σ^2_{EE}	IG(10,15)	2.12	2.32	2.55	3.34	3.86	4.48
σ^2_{EI}	N(0,10)	-	-	-	0.96	1.43	1.95
σ^2_{II}	IG(10,15)	1.52	1.68	1.85	2.25	2.70	3.22
posterior	-	-984.57	-983.46	-982.60	-986.23	-985.00	-984.01
likelihood	-	-951.68	-950.41	-949.43	-950.70	-949.49	-948.60
			$\zeta = 0.75$			$\zeta = 0.80$	
	Prior		Posterior			Posterior	
	(Mean,Std.Dev)	25%	50%	75%	25%	50%	75%
μ	N(3,10)	2.75	3.03	3.31	2.79	3.08	3.35
ρ	N(0.3,1)	0.53	0.59	0.64	0.51	0.57	0.62
σ^2_{GG}	IG(10,15)	5.78	6.31	6.92	6.54	7.09	7.70
σ^2_{GE}	N(0,10)	-0.76	-0.29	0.15	-1.15	-0.69	-0.29
σ^2_{GI}	N(0,10)	-0.34	0.01	0.34	-0.74	-0.38	-0.04
σ^2_{EE}	IG(10,15)	3.08	3.88	4.75	3.14	3.90	4.77
σ^2_{EI}	N(0,10)	0.73	1.23	1.78	0.80	1.29	1.85
σ^2_{II}	IG(10,15)	1.94	2.30	2.76	1.98	2.36	2.82
posterior	-	-982.50	-980.99	-979.87	-982.48	-981.05	-979.91
likelihood	-	-950.93	-949.55	-948.40	-950.85	-949.44	-948.41
			$\zeta = 0.85$			$\zeta = 0.95$	
	Prior		Posterior			Posterior	
	(Mean,Std.Dev)	25%	50%	75%	25%	50%	75%
μ	N(3,10)	2.72	2.96	3.14	2.84	3.03	3.25
ρ	N(0.3,1)	0.51	0.56	0.60	0.49	0.54	0.60
σ^2_{GG}	IG(10,15)	6.67	7.19	7.76	7.69	8.43	9.28
σ^2_{GE}	N(0,10)	-2.17	-1.98	-1.77	-2.88	-2.73	-2.50
σ^2_{GI}	N(0,10)	-0.97	-0.80	-0.53	-1.99	-1.58	-1.22
σ^2_{EE}	IG(10,15)	5.36	5.79	6.28	5.64	6.10	6.39
σ^2_{EI}	N(0,10)	2.04	2.33	2.63	2.43	2.64	2.93
σ^2_{II}	IG(10,15)	2.36	2.65	3.04	2.45	3.22	3.81
posterior	-	-982.62	-981.40	-980.48	-984.09	-982.80	-981.57
likelihood	-	-949.42	-948.25	-947.49	-950.19	-948.84	-947.81
			$\zeta = 1.05$			$\zeta = 1.15$	
	Prior		Posterior			Posterior	
	(Mean,Std.Dev)	25%	50%	75%	25%	50%	75%
μ	N(3,10)	2.85	3.07	3.33	2.55	2.89	3.21
ρ	N(0.3,1)	0.48	0.53	0.58	0.52	0.56	0.61
σ^2_{GG}	IG(10,15)	8.92	9.57	10.25	9.07	9.88	10.73
σ^2_{GE}	N(0,10)	-4.04	-3.88	-3.70	-5.61	-5.50	-5.22
σ^2_{GI}	N(0,10)	-3.09	-2.65	-2.29	-4.38	-4.21	-4.01
σ^2_{EE}	IG(10,15)	6.74	7.13	7.41	8.51	9.07	9.30
σ^2_{EI}	N(0,10)	3.23	3.46	4.13	5.29	5.52	5.89
σ^2_{II}	IG(10,15)	3.27	3.66	4.43	5.68	6.00	6.31
posterior	-	-984.89	-983.63	-982.49	-988.63	-987.18	-986.32
likelihood	-	-949.31	-948.30	-947.53	-949.82	-948.51	-947.67

Table 4.3: Priors and Posteriors, 3-Equation Model, 1960Q1-2011Q4

Parameter	Prior	Posterior		
	(Mean, Std)	25%	50%	75%
μ	N(3,10)	2.60	2.78	2.95
ρ	N(0.3,1)	0.54	0.58	0.63
σ^2_{GG}	IG(10,15)	6.73	6.96	7.35
σ^2_{GE}	N(0,10)	-1.27	-1.10	-0.84
σ^2_{GI}	N(0,10)	-1.03	-0.82	-0.59
σ^2_{EE}	IG(10,15)	4.17	4.57	4.79
σ^2_{EI}	N(0,10)	1.70	1.95	2.12
σ^2_{II}	IG(10,15)	2.54	3.07	3.27
σ^2_{GU}	N(0,10)	1.27	1.46	1.66
σ^2_{UU}	IG(0.3,10)	0.50	0.59	0.71
κ	N(0,10)	1.53	1.62	1.71
λ	N(-0.5,10)	-0.55	-0.52	-0.50
posterior	-	-1251.1	-1249.6	-1248.3
likelihood	-	-1199.0	-1197.5	-1196.2

$$\Sigma = \begin{bmatrix} \underset{[6.39,7.44]}{6.90} & 0 & 0 \\ 0 & \underset{[2.12,2.55]}{2.32} & 0 \\ 0 & 0 & \underset{[1.52,1.85]}{1.68} \end{bmatrix}. \tag{4.16}$$

For the 2-equation model with Σ block-diagonal, we have

$$GDP_t = \underset{[2.77,3.34]}{3.06}(1 - 0.62) + \underset{[0.57,0.68]}{0.62} GDP_{t-1} + \epsilon_{Gt}, \tag{4.17}$$

$$\Sigma = \begin{bmatrix} \underset{[4.39,5.95]}{5.17} & 0 & 0 \\ 0 & \underset{[3.34,4.48]}{3.86} & \underset{[0.96,1.95]}{1.43} \\ 0 & \underset{[0.96,1.95]}{1.43} & \underset{[2.25,3.22]}{2.70} \end{bmatrix}. \tag{4.18}$$

For the 2-equation model with benchmark $\zeta = 0.80$, we have

$$GDP_t = \underset{[2.79,3.35]}{3.08}(1 - 0.57) + \underset{[0.51,0.62]}{0.57} GDP_{t-1} + \epsilon_{Gt}, \tag{4.19}$$

$$\Sigma = \begin{bmatrix} \underset{[6.54,7.70]}{7.09} & \underset{[-1.15,-0.29]}{-0.69} & \underset{[-0.74,-0.04]}{-0.38} \\ \underset{[-1.15,-0.29]}{-0.69} & \underset{[3.14,4.77]}{3.90} & \underset{[0.80,1.85]}{1.29} \\ \underset{[-0.74,-0.04]}{-0.38} & \underset{[0.80,1.85]}{1.29} & \underset{[1.98,2.82]}{2.36} \end{bmatrix}. \tag{4.20}$$

Finally, for the 3-equation model, we have

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \\ U_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \underset{[1.53,1.71]}{1.62} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \underset{[-0.55,-0.50]}{-0.52} \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \tag{4.21}$$

$$GDP_t = \underset{[2.60,2.95]}{2.78}(1 - 0.58) + \underset{[0.54,0.63]}{0.58} GDP_{t-1} + \epsilon_{Gt}, \tag{4.22}$$

$$\begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \underset{[6.73,7.35]}{6.96} & \underset{[-1.27,-0.84]}{-1.10} & \underset{[-1.03,-0.59]}{-0.82} & \underset{[1.27,1.66]}{1.46} \\ \underset{[-1.27,-0.84]}{-1.10} & \underset{[4.17,4.79]}{4.57} & \underset{[1.70,2.12]}{1.95} & 0 \\ \underset{[-1.03,-0.59]}{-0.82} & \underset{[1.70,2.12]}{1.95} & \underset{[2.54,3.27]}{3.07} & 0 \\ \underset{[1.27,1.66]}{1.46} & 0 & 0 & \underset{[0.50,0.71]}{0.59} \end{bmatrix} \right) \tag{4.23}$$

Many aspects of the results are noteworthy; here we simply mention a few. First, every posterior interval in every model reported above excludes zero. Hence the diagonal and block diagonal models do not appear satisfactory.

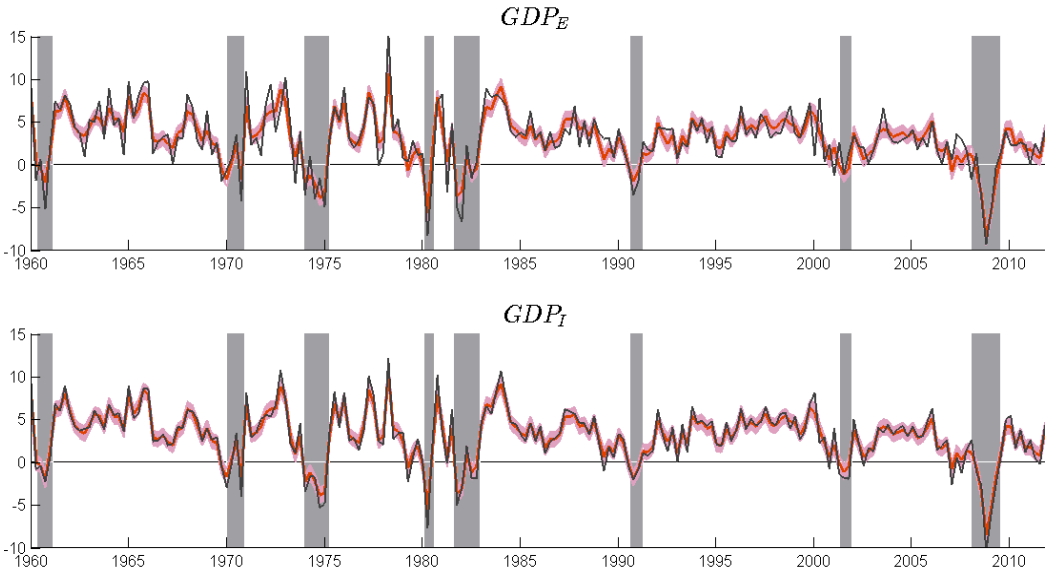
Second, the Σ estimates are qualitatively similar across specifications. Covariances are always negative, as per our conjecture based on the counter-cyclicality in the statistical discrepancy ($GDP_E - GDP_I$) documented by Fixler and Nalewaik (2009) and Nalewaik (2010). Shock variances always satisfy $\hat{\sigma}^2_{GG} > \hat{\sigma}^2_{EE} > \hat{\sigma}^2_{II}$.

Finally, GDP_M is highly serially correlated across all specifications ($\rho \approx .6$), much more so than the current "consensus" based on GDP_E ($\rho \approx .3$). We shall have much more to say about these and other results in the next section.

4.4 New Perspectives on the Properties of *GDP*

Our various extracted GDP_M series differ in fundamental ways from other measures, such as GDP_E and GDP_I. Here we discuss some of the most important differences.

Figure 4.3: GDP Sample Paths, 1960Q1-2011Q4



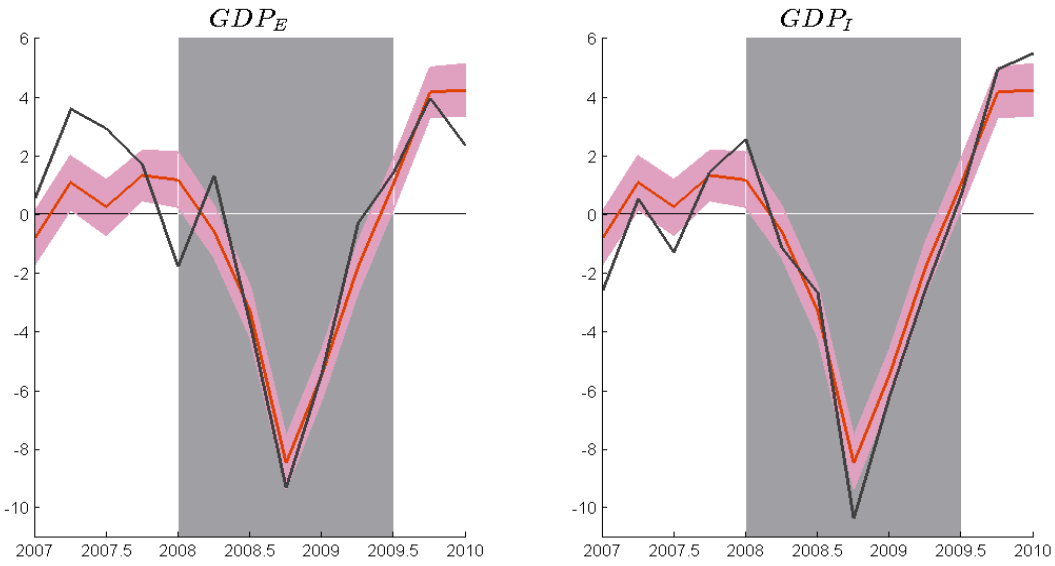
Notes: In each panel we show the sample path of GDP_M in red together with a light-red posterior interquartile range, and we show one of the competitor series in black. For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

4.4.1 *GDP* Sample Paths

Let us begin by highlighting the sample-path differences between our GDP_M and the obvious competitors GDP_E and GDP_I. We make those comparisons in Figure 4.3. In each panel we show the sample path of GDP_M in red together with a light-red posterior interquartile range, and we show one of the competitor series in black.[14] In the top panel we show GDP_M vs. GDP_E. There are often wide divergences, with GDP_E well outside the posterior interquartile range of GDP_M. Indeed GDP_E is substantially more volatile than GDP_M. In the bottom panel of Figure 4.3 we show GDP_M vs. GDP_I. Noticeable divergences again appear often, with GDP_I also outside the posterior interquartile range of GDP_M. The divergences are not as pronounced, however, and the "excess volatility" apparent in GDP_E

[14]For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

Figure 4.4: GDP Sample Paths, 2007Q1-2009Q4



Notes: In each panel we show the sample path of GDP_M in red together with a light-red posterior interquartile range, and we show one of the competitor series in black. For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

is less apparent in GDP_I. That is because, as we will show later, GDP_M loads relatively more heavily on GDP_I.

To emphasize the economic importance of the differences in competing real activity assessments, in Figure 4.4 we focus on the tumultuous period 2007Q1-2009Q4. The figure makes clear not only that *both* GDP_E and GDP_I can diverge substantially from GDP, but also that the timing and nature of their divergences can be very different. In 2007Q3, for example, GDP_E growth was strongly positive and GDP_I growth was negative.

4.4.2 Linear *GDP* Dynamics

In our framework, the data-generating process for true GDP_t is completely characterized by the pair, (σ^2_{GG}, ρ). In Figure 4.5 we show those pairs across MCMC draws for all of our measurement-error models, and for comparison we show (σ^2, ρ) values corresponding

to $AR(1)$ models fit to GDP_E alone and GDP_I alone. In addition, in Table 4.1 we show a variety of statistics quantifying the properties of our various GDP_M measures vs. those of GDP_E, GDP_I and GDP_F, the forecast-error-based estimate of true GDP produced by Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012).

First consider Figure 4.5. Across measurement-error models M, GDP_M is robustly more serially correlated than both GDP_E and GDP_I, and it also has a smaller innovation variance. Hence most of our models achieve closely-matching unconditional variances, but they are composed of very different underlying (σ^2, ρ) values from those corresponding to GDP_E. GDP_M has smaller shock volatility, but much more shock persistence – roughly double that of GDP_E (ρ of roughly 0.60 for GDP_M vs. 0.30 for GDP_E).

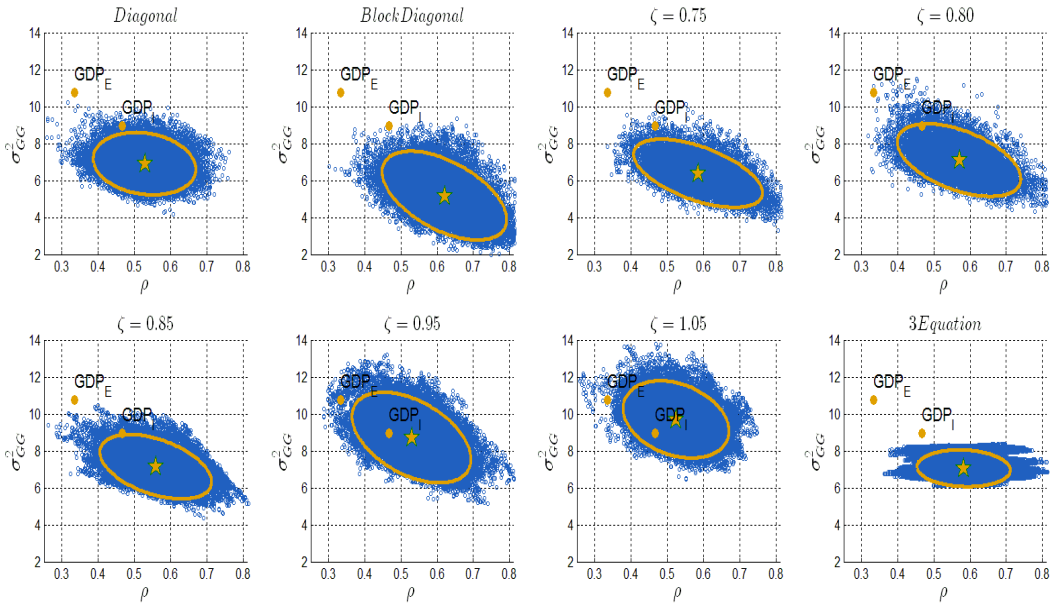
Now consider Table 4.1. The various GDP_M series are all less volatile than each of GDP_E, GDP_I and GDP_F, and a bit more skewed left. Most noticeably, the GDP_M series are much more strongly serially correlated than the GDP_E, GDP_I and GDP_F series, and with smaller innovation variances. This translates into much higher predictive R^2's for GDP_M. Indeed GDP_M is twice as predictable as GDP_I or GDP_F, which in turn are twice as predictable as GDP_E.

4.4.3 Non-Linear *GDP* Dynamics

In Table 4.4 we show Markov-switching $AR(1)$ model results for a variety of GDP series. The model allows for simultaneous switching in both mean and serial-correlation parameters. The model switches between high- and low-growth states, with low-growth states generally including recessions as defined by the National Bureau of Economic Research's Business Cycle Dating Committee (see also Nalewaik (2012)). The most interesting aspect of the results concerns the estimated low- and high-state serial-correlation parameters ($\hat{\rho}_0$ and $\hat{\rho}_1$, respectively).

First, always and everywhere, $\hat{\rho}_0 > \hat{\rho}_1$; that is, a disproportionate share of overall serial

Figure 4.5: $(\hat{\rho}, \hat{\sigma}^2_{GG})$ Pairs Across MCMC Draws



Notes: Solid lines indicate 90% (σ^2_{GG}, ρ) posterior coverage ellipsoids for the various models. Stars indicate posterior median values. The sample period is 1960Q1-2011.Q4. For comparison we show (σ^2, ρ) values corresponding to $AR(1)$ models fit to GDP_E alone and GDP_I alone.

correlation comes from low-growth states. This interesting result parallels recent work indicating that a disproportionate share of stock market return predictability comes from recessions (Rapach, Strauss, and Zhou (2010)), as well as work showing that shocks to business orders for capital goods are more persistent in downturns (Nalewaik and Pinto (2012)).

Second, comparison of GDP_I to GDP_E reveals that they have *identical* $\hat{\rho}_0$ values (0.55), but that $\hat{\rho}_1$ is much bigger for GDP_I than for GDP_E (0.31 vs. 0.14). Hence the stronger overall serial correlation of GDP_I comes entirely from its stronger serial correlation during expansions.

Finally, comparison of GDP_M to GDP_E reveals much bigger $\hat{\rho}_0$ *and* $\hat{\rho}_1$ values for GDP_M than for GDP_E, regardless of the particular measurement-error model M examined. The

Table 4.4: Regime-Switching Model Estimates, 1960Q1-2011Q4

	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\rho}_0$	$\hat{\rho}_1$	$\hat{\sigma}^2_H$	$\hat{\sigma}^2_L$	$\hat{p}_{00}$	$\hat{p}_{11}$
GDP_E	1.31	4.71	0.55	0.14	16.55	4.81	0.81	0.88
GDP_I	1.28	4.87	0.55	0.31	12.07	5.51	0.82	0.87
GDP_M 2-eqn, Σ diag	1.76	5.12	0.73	0.41	9.81	3.37	0.83	0.85
GDP_M 2-eqn, Σ block	1.75	4.72	0.83	0.63	6.22	2.41	0.81	0.86
GDP_M 2-eqn, $\zeta = 0.80$	1.79	4.95	0.78	0.55	7.96	3.04	0.82	0.85
GDP_M 3-eqn	1.88	5.32	0.88	0.39	7.85	2.95	0.80	0.85
GDP_F	1.51	4.93	0.64	0.30	13.20	4.17	0.82	0.87

Notes: In the top panel we show posterior median estimates for two-state regime-switching $AR(1)$ models fit to raw data. In the middle panel we show posterior median estimates for Regime-switching models fit to GDP_M. In the bottom panel we show posterior median estimates for regime-switching models fit to GDP_F, the forecast-error-based estimate of true GDP produced by Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012). We allow for a one-time structural break in volatility in 1984 (the "Great Moderation").

general finding of $\hat{\rho}_0 > \hat{\rho}_1$ is preserved, but *both* $\hat{\rho}_0$ and $\hat{\rho}_1$ are much larger for GDP_M than for GDP_E. In our benchmark 2-equation model with $\zeta = 0.80$, for example, we have $\hat{\rho}_0 = 0.78$ and $\hat{\rho}_1 = 0.55$.

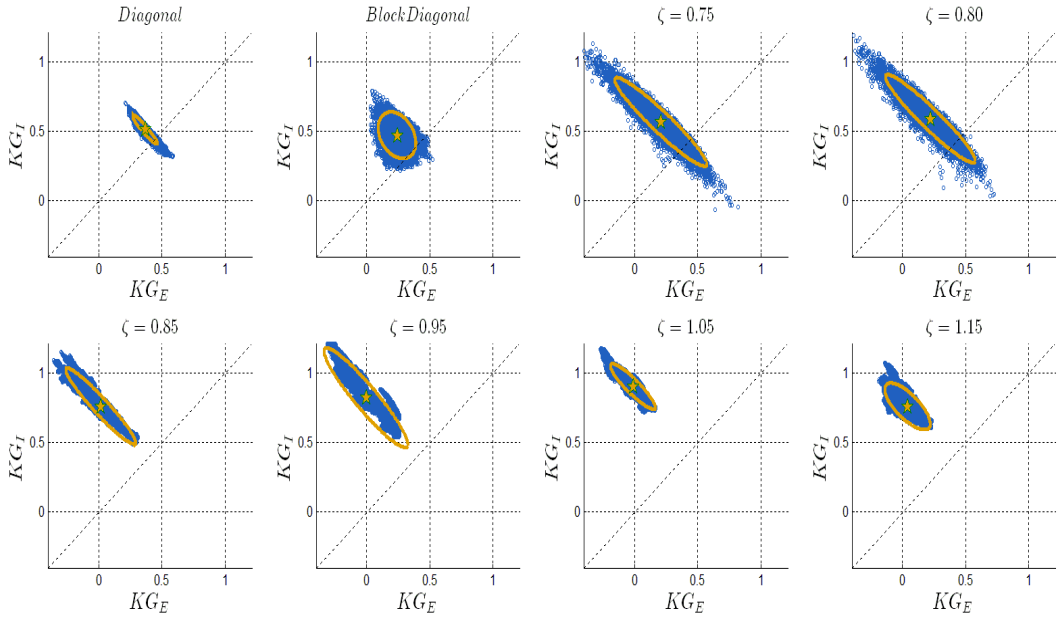
4.4.4 On the Relative Contributions of GDP_E and GDP_I to GDP_M

It is of interest to know how the observed indicators GDP_E and GDP_I contribute to our extracted true GDP. We do this in two ways; in section 4.4.4 we examine Kalman gains, and in section 4.4.4 we find the convex combination of GDP_E and GDP_I closest to our extracted GDP.

Kalman Gains

The Kalman gains associated with GDP_E and GDP_I govern the amount by which news about GDP_E and GDP_I, respectively, causes the optimal extraction of GDP_t (conditional on time-t information) to differ from the earlier optimal prediction of GDP_t (conditional on time-$(t-1)$ information). Put more simply, the Kalman gain of GDP_E (resp. GDP_I) measures its importance in influencing GDP_M, and hence in informing our views about

Figure 4.6: (KG_E, KG_I) Pairs Across MCMC Draws



Notes: Solid lines indicate 90% posterior coverage ellipsoids. Stars indicate posterior median values.

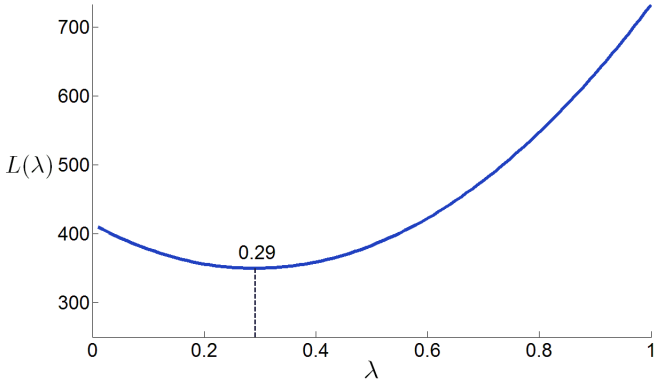
latent true GDP.

We summarize the posterior distributions of Kalman gains in Figure 4.6. Posterior median GDP_I Kalman gains are large in absolute terms, and most notably, very large relative to those for GDP_E. Indeed posterior median GDP_E Kalman gains are zero in several specifications. In any event, it is clear that GDP_I plays a larger role in informing us about GDP than does GDP_E. For our benchmark ζ-model with $\zeta = 0.80$, the posterior median GDP_I and GDP_E Kalman gains are 0.59 and 0.23, respectively.

Closest Convex Combination

The Kalman filter extractions average not only over space, but also over time. Nevertheless, we can ask what contemporaneous convex combination of GDP_E and GDP_I, $\lambda GDP_E + (1-\lambda)GDP_I$, is closest to the extracted GDP_M. That is, we can find $\lambda^* = argmin_\lambda \, L(\lambda)$,

Figure 4.7: Closest Convex Combination



Notes: We show quadratic loss, $L(\lambda) = \sum_{t=1960Q1}^{2011Q4} [(\lambda GDP_{Et} + (1-\lambda)GDP_{It}) - GDP_{Mt}]^2$, as a function of λ, where where GDP_{Mt} is our smoothed extraction of true GDP_t, obtained from the 3-equation model.

where $L(\lambda)$ is a loss function. Under quadratic loss we have

$$\lambda^* = argmin_\lambda \sum_{t=1}^{T} [(\lambda GDP_{Et} + (1-\lambda)GDP_{It}) - GDP_{Mt}]^2 ,$$

where GDP_{Mt} is our smoothed extraction of true GDP_t. Over our sample of 1960Q1-2011Q4, the optimum under quadratic loss is $\lambda^* = 0.29$. The minimum is quite sharp, as we show in Figure 4.7, and it puts more than twice as much weight on GDP_I than on GDP_E. That weighting accords closely with both the Kalman gain results discussed above and the forecast-combination calibration results in Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012). It does not, of course, mean that time series of GDP_M will "match" time series of GDP_F, because the Kalman filter does much more than simple contemporaneous averaging of GDP_E and GDP_I in its extraction of latent true GDP.

4.4.5 A Final Remark on the Serial Correlation in GDP_M

Obviously a key result of our analysis is the strong serial correlation (persistence, forecastability, ...) of our extracted GDP_M, regardless of the particular specification. One might perhaps wonder whether this is a spurious artifact of our signal- extraction approach. We

hasten to add that the answer is no. Indeed optimal extractions of covariance stationary series (in our, case latent true GDP growth) *must* be less variable than the series being extracted, because the optimal extraction is a conditional expectation.[15] Given our models, under quadratic loss any other GDP extractions are sub-optimal and hence inferior.

4.5 Concluding Remarks and Directions for Future Research

We produce several estimates of GDP that blend both GDP_E and GDP_I. All estimates feature GDP_I prominently, and our blended GDP estimate has properties quite different from those of the "traditional" GDP_E (as well as GDP_I). In a sense we build on the literature on "balancing" the national income accounts, which extends back almost as far as national income accounting itself, as for example in Stone, Champernowne, and Meade (1942). We do not, however, advocate that the U.S. publish only GDP_M, as there may at times be value in being able to see the income and expenditure sides separately. But we would advocate the additional calculation of GDP_M and using it as the benchmark GDP estimate.

Many interesting extensions are possible. Perhaps the two most interesting concern forecasting and real-time analysis. First consider forecasting. When forecasting a "traditional" GDP series such as GDP_E, we must take it as given (i.e., we must ignore measurement error). The analogous procedure in our framework would take GDP_M as given, modeling and forecasting it directly, ignoring the fact that it is based on a first-stage extraction subject to error. Fortunately, however, in our framework we need not do that. Instead we can estimate and forecast directly from the dynamic factor model, accounting for all sources of uncertainty, which should translate into superior interval and density forecasts. Related, it

[15]The forecast-error approach of Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012) also has optimality properties, but of a different sort, and there is no reason why in the forecast-error framework the optimal combination should be smoother than latent true GDP growth. Instead it could go either way, depending on the correlation of the forecast errors in GDP_E and GDP_I.

would be interesting to calculate directly the point, interval and density forecast functions corresponding to our measurement-error model.

Second, consider real-time analysis. Although GDP_I data are not as timely as GDP_E data, our filtering framework still uses all available data efficiently, appropriately handling any missing data. A key insight is that when using simple convex combinations as in the forecast-error approach of Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2012), missing GDP_I data for the most-recent quarter(s) forces all weight to be put on GDP_E. Our measurement-error framework is very different, however, because the Kalman filter averages not just over space, but also over time, and earlier quarters for which we *do* have GDP_I data are informative for the most-recent quarters with "missing" GDP_I data.

4.6 Appendix

Here we report various details of theory, establishing identification results for the two- and three-variable models in appendices 4.6.1 and 4.6.2, respectively. The identification analysis is based on Komunjer and Ng (2011).

4.6.1 Identification in the Two-Variable Model

The constants in the state-space model can be identified from the means of GDP_{Et} and GDP_{It}. To simplify the subsequent exposition we now set the constant terms to zero:

$$GDP_t = \rho GDP_{t-1} + \epsilon_{Gt} \tag{4.24}$$

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{4.25}$$

and the joint distribution of the errors is

$$\epsilon_t = \begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \sim iidN\big(0, \Sigma\big), \quad \text{where} \quad \Sigma = \begin{bmatrix} \Sigma_{GG} & \cdot & \cdot \\ \Sigma_{EG} & \Sigma_{EE} & \cdot \\ \Sigma_{IG} & \Sigma_{IE} & \Sigma_{II} \end{bmatrix}.$$

Using the notation in Komunjer and Ng (2011), we write the system as

$$s_{t+1} = A(\theta)s_t + B(\theta)\epsilon_{t+1} \tag{4.26}$$

$$y_{t+1} = C(\theta)s_t + D(\theta)\epsilon_{t+1}, \tag{4.27}$$

where

$$\begin{aligned} s_t &= GDP_t, \quad y_t = \begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} \\ A(\theta) &= \rho, \quad B(\theta) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \\ C(\theta) &= \begin{bmatrix} \rho \\ \rho \end{bmatrix}, \quad D(\theta) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{aligned} \tag{4.28}$$

and $\theta = [\rho, vech(\Sigma)']'$. Note that only $A(\theta)$ and $C(\theta)$ are non-trivial functions of θ.

Assumption 1 *The parameter vector θ satisfies the following conditions: (i) Σ is positive definite; (ii)* $0 \leq \rho < 1$.

Because the rows of D are linearly independent, Assumption 1(i) implies that $D\Sigma D'$ is non-singular. In turn, we deduce that Assumptions 1, 2, and 4-NS of Komunjer and Ng (2011) are satisfied.

We now express the state-space system in terms of its innovation representation

$$\begin{aligned} s_{t+1|t+1} &= A(\theta)s_{t|t} + K(\theta)a_{t+1} \\ y_{t+1} &= C(\theta)\hat{s}_{t|t} + a_{t+1}, \end{aligned} \tag{4.29}$$

where a_{t+1} is the one-step-ahead forecast error of the system whose variance we denote by $\Sigma_a(\theta)$. The innovation representation is obtained from the Kalman filter as follows. Suppose that conditional on time t information $Y_{1:t}$ the distribution of $s_t|Y_{1:t} \sim N(s_{t|t}, P_{t|t})$. Then the joint distribution of $[s_{t+1}, y'_{t+1}]'$ is

$$\left[\begin{array}{c} s_{t+1} \\ y_{t+1} \end{array} \right] \Bigg| Y_{1:T} \sim \left(\left[\begin{array}{c} As_{t|t} \\ Cs_{t|t} \end{array} \right], \left[\begin{array}{cc} AP_{t|t}A' + B\Sigma B' & AP_{t|t}C' + B\Sigma D' \\ CP_{t|t}A' + D\Sigma B' & CP_{t|t}C' + D\Sigma D' \end{array} \right] \right).$$

In turn, the conditional distribution of $s_{t+1}|Y_{1:t+1}$ is

$$s_{t+1}|Y_{1:t+1} \sim N\big(s_{t+1|t+1}, P_{t+1|t+1}\big),$$

where

$$\begin{aligned} s_{t+1|t+1} &= As_{t|t} + (AP_{t|t}C + B\Sigma D')(CP_{t|t}C' + D\Sigma D')^{-1}(y_t - Cs_{t|t}) \\ P_{t+1|t+1} &= AP_{t|t}A' + B\Sigma B' - (AP_{t|t}C' + B\Sigma D')(CP_{t|t}C' + D\Sigma D')^{-1}(CP_{t|t}A' + D\Sigma B'). \end{aligned}$$

Now let P be the matrix that solves the Riccati equation,

$$P = APA' + B\Sigma B' - (APC' + B\Sigma D')(CPC' + D\Sigma D')^{-1}(CPA' + D\Sigma B'), \tag{4.30}$$

and let K be the Kalman gain matrix

$$K = (APC' + B\Sigma D')(CPC' + D\Sigma D')^{-1}. \tag{4.31}$$

Then the one-step-ahead forecast error matrix is given by

$$\Sigma_a = CPC' + D\Sigma D'. \tag{4.32}$$

Equations (4.30) to (4.32) determine the matrices that appear in the innovation-representation of the state-space system (4.29).

In order to be able to apply Proposition 1-NS of Komunjer and Ng (2011) we need to express P, K, and Σ_a in terms of θ. While solving Riccati equations analytically is in general not feasible, our system is scalar, which simplifies the calculation considerably. Replacing A by ρ and P by p such that scalars appear in lower case, and defining

$$\Sigma_{BB} = B\Sigma B', \quad \Sigma_{BD} = B\Sigma D', \quad \text{and} \quad \Sigma_{DD} = D\Sigma D',$$

we can write (4.30) as

$$p = p\rho^2 + \Sigma_{BB} - (p\rho C' + \Sigma_{BD})(pCC' + \Sigma_{DD})^{-1}(p\rho C + \Sigma_{DB}). \tag{4.33}$$

Likewise,

$$K = (p\rho C' + \Sigma_{BD})(pCC' + \Sigma_{DD})^{-1} \quad \text{and} \quad \Sigma_a = pCC' + \Sigma_{DD}. \tag{4.34}$$

Because $\Sigma_{BB} - \Sigma_{BD}\Sigma'_{DD}\Sigma_{DB} > 0$ we can deduce that $p > 0$. Moreover, because $A = \rho \geq 0$ and $C \geq 0$, we deduce that $K \neq 0$ and therefore Assumption 5-NS of Komunjer and Ng (2011) is satisfied. According to Proposition 1-NS in Komunjer and Ng (2011), two vectors θ and θ_1 are observationally equivalent if and only if there exists a scalar $\gamma \neq 0$ such that

$$A(\theta_1) = \gamma A(\theta)\gamma^{-1} \tag{4.35}$$

$$K(\theta_1) = \gamma K(\theta) \tag{4.36}$$

$$C(\theta_1) = C(\theta)\gamma^{-1} \tag{4.37}$$

$$\Sigma_a(\theta_1) = \Sigma_a(\theta). \tag{4.38}$$

Define $\theta = [\rho, vech(\Sigma)']'$ and $\theta_1 = [\rho_1, vech(\Sigma_1)']'$. Using the definition of the scalar $A(\theta)$ in (4.28) we deduce from (4.35) that $\rho_1 = \rho$. Since $C(\theta)$ depends on θ only through ρ we can deduce from (4.37) that $\gamma = 1$. Thus, given θ and ρ, the elements of the vector $vech(\Sigma_1)$ have to satisfy conditions (4.36) and (4.38), which, using (4.34), can be rewritten as

$$\Sigma_a = \Sigma_{a1} = p_1 CC' + \Sigma_{DD1} \tag{4.39}$$

$$K = K_1 = (p_1 \rho C' + \Sigma_{BD1})\Sigma_a^{-1}. \tag{4.40}$$

Moreover, p_1 has to solve the Riccati equation (4.33):

$$p_1 = p_1 \rho^2 + \Sigma_{BB1} - K_0(p_1 \rho C + \Sigma_{BD}). \tag{4.41}$$

Equations (4.39) to (4.41) are satisfied if and only if

$$pCC' + \Sigma_{DD} = p_1 CC' + \Sigma_{DD1} \tag{4.42}$$

$$p\rho C' + \Sigma_{BD} = p_1 \rho C' + \Sigma_{BD1} \tag{4.43}$$

$$p(1-\rho^2) - \Sigma_{BB} = p_1(1-\rho^2) - \Sigma_{BB1}. \tag{4.44}$$

We proceed by deriving expressions for the Σ_{xx} matrices that appear in (4.42) to (4.44):

$$\begin{aligned}
\Sigma_{BB} &= \Sigma_{GG} \\
\Sigma_{BD} &= \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} \end{bmatrix} \\
\Sigma_{DD} &= \begin{bmatrix} \Sigma_{GG} + \Sigma_{EE} + 2\Sigma_{EG} & \cdot \\ \Sigma_{GG} + \Sigma_{GE} + \Sigma_{GI} + \Sigma_{EI} & \Sigma_{GG} + \Sigma_{II} + 2\Sigma_{GI} \end{bmatrix}
\end{aligned}$$

Without loss of generality let

$$\Sigma_{GG1} = \Sigma_{GG} + (1-\rho^2)\delta, \tag{4.45}$$

which implies that

$$\Sigma_{BB1} = \Sigma_{BB} + (1-\rho^2)\delta.$$

We now distinguish the cases $\delta = 0$ and $\delta \neq 0$.

Case 1: $\delta = 0$**.** (4.44) implies $p_1 = p$. It follows from (4.43) that $\Sigma_{BD1} = \Sigma_{BD}$. In turn, $\Sigma_{GE1} = \Sigma_{GE}$ and $\Sigma_{GI1} = \Sigma_{GI}$. Finally, to satisfy (4.42) it has to be the case that $\Sigma_{DD1} = \Sigma_{DD}$, which implies that the remaining elements of Σ and Σ_1 are identical. We conclude that $\theta_1 = \theta$.

Case 2: $\delta \neq 0$**.** (4.44) implies $p_1 = p + \delta$. Now consider (4.43):

$$\begin{aligned} p\rho C' + \Sigma_{BD} &= p\rho^2 \left[\begin{array}{cc} 1 & 1 \end{array}\right] + \left[\begin{array}{cc} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} \end{array}\right] \\ &\stackrel{!}{=} p\rho^2 \left[\begin{array}{cc} 1 & 1 \end{array}\right] + \delta\rho^2 \left[\begin{array}{cc} 1 & 1 \end{array}\right] \\ &\quad + \left[\begin{array}{cc} \Sigma_{GG} + \Sigma_{GE1} & \Sigma_{GG} + \Sigma_{GI1} \end{array}\right] \\ &\quad + \delta(1-\rho^2) \left[\begin{array}{cc} 1 & 1 \end{array}\right] \end{aligned}$$

We deduce that

$$\Sigma_{GE1} = \Sigma_{GE} - \delta, \quad \Sigma_{GI1} = \Sigma_{GI} - \delta. \tag{4.46}$$

Finally, consider (4.42), which can be rewritten as

$$0 = \Sigma_{DD1} - \Sigma_{DD} + \delta CC'.$$

Using the previously derived expressions for Σ_{DD} and Σ_{DD1} we obtain the following three conditions

$$\begin{aligned} 0 &= (1-\rho^2)\delta + (\Sigma_{EE1} - \Sigma_{EE}) - 2\delta + \rho^2\delta = \Sigma_{EE1} - \Sigma_{EE} - \delta \\ 0 &= (1-\rho^2)\delta - 2\delta + (\Sigma_{EI1} - \Sigma_{EI}) + \rho^2\delta = \Sigma_{EI1} - \Sigma_{EI} - \delta \\ 0 &= (1-\rho^2)\delta + (\Sigma_{II1} - \Sigma_{II}) - 2\delta + \rho^2\delta = \Sigma_{II1} - \Sigma_{II} - \delta. \end{aligned}$$

Thus, we deduce that

$$\Sigma_{EE1} = \Sigma_{EE} + \delta, \quad \Sigma_{EI1} = \Sigma_{EI} + \delta, \quad \text{and} \quad \Sigma_{II1} = \Sigma_{II} + \delta. \tag{4.47}$$

Combining (4.45), (4.46), and (4.47) we find that

$$\Sigma_1 = \begin{bmatrix} \Sigma_{GG} + \delta(1-\rho^2) & \Sigma_{GE} - \delta & \Sigma_{GI} - \delta \\ \Sigma_{GE} - \delta & \Sigma_{EE} + \delta & \Sigma_{EI} + \delta \\ \Sigma_{GI} - \delta & \Sigma_{EI} + \delta & \Sigma_{II} + \delta \end{bmatrix}. \tag{4.48}$$

Thus, we have proved the following theorem:

Theorem 4.6.1 *Suppose Assumption 1 is satisfied. Then the two-variable model is*

(i) identified if Σ *is diagonal as in section 4.2.1;*

(ii) identified if Σ *is block-diagonal as in section 4.2.2;*

(iii) not identified if Σ *is unrestricted as in section 4.2.3;*

(iv) identified if Σ *is restricted as in section 4.2.4.*

4.6.2 Identification in the Three-Variable Model

The identification analysis of the three-variable is similar to the analysis of the two-variable model in the previous section. The system is given by

$$GDP_t = \rho GDP_{t-1} + \epsilon_{Gt} \tag{4.49}$$

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \\ U_t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \lambda \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix}, \tag{4.50}$$

and the joint distribution of the errors is

$$\epsilon_t = \begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \sim iidN\big(0, \Sigma\big), , \quad \text{where} \quad \Sigma = \begin{bmatrix} \Sigma_{GG} & \cdot & \cdot & \cdot \\ \Sigma_{EG} & \Sigma_{EE} & \cdot & \cdot \\ \Sigma_{IG} & \Sigma_{IE} & \Sigma_{II} & \cdot \\ \Sigma_{UG} & \Sigma_{UE} & \Sigma_{UI} & \Sigma_{UU} \end{bmatrix}.$$

The matrices $A(\theta)$, $B(\theta)$, $C(\theta)$, and $D(\theta)$ are now given by

$$A(\theta) = \rho, \quad B(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

$$C(\theta) = \begin{bmatrix} \rho \\ \rho \\ \lambda\rho \end{bmatrix}, \quad D(\theta) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \lambda & 0 & 0 & 1 \end{bmatrix},$$

where $\theta = [\rho, \lambda, vech(\Sigma)']'$.

Assumption 2 *The parameter vector θ satisfies the following conditions: (i) Σ is positive definite; (ii) $0 < \rho < 1$; (iii) $\lambda \neq 0$; (iv) $\Sigma_{UE} = \Sigma_{UI} = 0$.*

Condition (4.35) implies that $\rho_1 = \rho$. Moreover, (4.37) implies that $\gamma = 1$ and that $\lambda_1 = \lambda$ provided that $\rho \neq 0$. As for the two-variable model, we have to verify that (4.42) to (4.44) are satisfied. The matrices Σ_{xx} that appear in these equations are given by

$$\begin{aligned}
\Sigma_{BB} &= \Sigma_{GG} \\
\Sigma_{BD} &= \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} & \lambda\Sigma_{GG} + \Sigma_{GU} \end{bmatrix} \\
\Sigma_{DD} &= \begin{bmatrix} \Sigma_{GG} + \Sigma_{EE} + 2\Sigma_{GE} & \cdot & \cdot \\ \Sigma_{GG} + \Sigma_{GE} + \Sigma_{GI} + \Sigma_{EI} & \Sigma_{GG} + \Sigma_{II} + 2\Sigma_{GI} & \cdot \\ \lambda\Sigma_{GG} + \lambda\Sigma_{GE} + \Sigma_{GU} & \lambda\Sigma_{GG} + \lambda\Sigma_{GI} + \Sigma_{GU} & \lambda^2\Sigma_{GG} + 2\lambda\Sigma_{GU} + \Sigma_{UU} \end{bmatrix}.
\end{aligned}$$

Without loss of generality, let

$$\Sigma_{GG,1} = \Sigma_{GG} + (1 - \rho^2)\delta,$$

which implies that

$$\Sigma_{BB,1} = \Sigma_{BB} + (1 - \rho^2)\delta.$$

Case 1: $\delta = 0$**.** (4.44) implies $p_1 = p$. It follows from (4.43) that $\Sigma_{BD,1} = \Sigma_{BD}$. In turn, $\Sigma_{GE,1} = \Sigma_{GE}$, $\Sigma_{GI,1} = \Sigma_{GI}$, and $\Sigma_{GU,1} = \Sigma_{GU}$. Finally, to satisfy (4.40) it has to be the case that $\Sigma_{DD,1} = \Sigma_{DD}$, which implies that the remaining elements of Σ and Σ_1 are identical for the two parameterizations. We conclude that it has to be the case that $\theta_1 = \theta$.

Case 2: $\delta \neq 0$**.** (4.44) implies $p_1 = p + \delta$. Now consider (4.43):

$$\begin{aligned}
p\rho C' + \Sigma_{BD} &= p\rho^2 \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix} + \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} & \lambda\Sigma_{GG} + \Sigma_{GU} \end{bmatrix} \\
&\stackrel{!}{=} p\rho^2 \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix} + \delta\rho^2 \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix} \\
&\quad + \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE,1} & \Sigma_{GG} + \Sigma_{GI,1} & \lambda\Sigma_{GG} + \Sigma_{GU,1} \end{bmatrix} \\
&\quad + (1 - \rho^2)\delta \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix}.
\end{aligned}$$

We deduce that

$$\Sigma_{GE,1} = \Sigma_{GE} - \delta, \quad \Sigma_{GI,1} = \Sigma_{GI} - \delta, \quad \Sigma_{GU,1} = \Sigma_{GU} - \delta.$$

Finally, consider (4.42), which can be rewritten as

$$0 = \Sigma_{DD,1} - \Sigma_{DD} + \delta CC'.$$

Using the previously derived expressions for Σ_{DD} and Σ_{DD1} we obtain the following five conditions

$$\begin{aligned}
0 &= (1-\rho^2)\delta + (\Sigma_{EE1} - \Sigma_{EE}) - 2\delta + \rho^2\delta = \Sigma_{EE1} - \Sigma_{EE} - \delta \\
0 &= (1-\rho^2)\delta - 2\delta + (\Sigma_{EI1} - \Sigma_{EI}) + \rho^2\delta = \Sigma_{EI1} - \Sigma_{EI} - \delta \\
0 &= (1-\rho^2)\delta + (\Sigma_{II1} - \Sigma_{II}) - 2\delta + \rho^2\delta = \Sigma_{II1} - \Sigma_{II} - \delta \\
0 &= \lambda(1-\rho^2)\delta - \lambda\delta - \delta + \lambda\rho^2\delta = \delta \\
0 &= \lambda^2(1-\rho^2)\delta - 2\lambda\delta + (\Sigma_{UU1} - \Sigma_{UU}) + \lambda^2\rho^2\delta = \Sigma_{UU1} - \Sigma_{UU} - \lambda(2-\lambda)\delta.
\end{aligned}$$

Thus, we deduce that

$$\delta = 0, \quad , \Sigma_{EE1} = \Sigma_{EE}, \quad \Sigma_{EI1} = \Sigma_{EI}, \quad \Sigma_{II1} = \Sigma_{II}, \quad \text{and} \quad \Sigma_{UU1} = \Sigma_{UU}.$$

This proves the following theorem:

Theorem 4.6.2 *Suppose Assumption 2 is satisfied. Then the three-variable model is identified.*

Chapter 5

Real-Time Forecasting with a Mixed-Frequency VAR

5.1 Introduction

In macroeconomic applications, vector autoregressions (VARs) are typically estimated either exclusively based on quarterly observations or exclusively based on monthly observations. In a forecasting setting, the advantage of using quarterly observations is that the set of macroeconomic series that could potentially be included in the VAR is larger. Gross domestic product (GDP), as well as many other series that are published as part of the national income and product accounts (NIPA), are only available at quarterly frequency. The advantage of using monthly information, on the other hand, is that the VAR is able to track the economy more closely in real time.

To exploit the respective advantages of both monthly and quarterly VARs, this paper develops a mixed-frequency VAR (MF-VAR) that allows some series to be observed at monthly and others at quarterly frequency. The MF-VAR can be conveniently represented as a state-space model, in which the state-transition equations are given by a VAR at monthly frequency and the measurement equations relate the observed series to the underlying, potentially unobserved, monthly variables that are stacked in the state vector. The MF-VAR is meant to be an attractive alternative to a standard VAR in which all series are

time-aggregated to quarterly frequency (QF-VAR). To cope with the high dimensionality of the parameter space, the MF-VAR is equipped with a Minnesota prior and estimated using Bayesian methods. The Minnesota prior is indexed by a vector of hyperparameters that determine the relative weight of *a priori* and sample information.

This paper makes contributions in two areas. On the methodological front we show how to numerically approximate the marginal data density (MDD) of a linear Gaussian MF-VAR. The MDD can be used for a data-based selection of hyperparameters, which is essential to achieve a good forecasting performance with a densely parameterized VAR. The second set of contributions is empirical. We compile a real-time data set for an eleven-variable VAR that includes observations on real aggregate activity, prices, and financial variables, including GDP, unemployment, inflation, and the federal funds rate. Using this data set, we recursively estimate the MF-VAR and assess its forecasting performance. The comparison to a QF-VAR is the main focus of the empirical analysis.

First, we ask the following very basic question: what is the gain, if any, from utilizing within-quarter monthly information in a VAR framework? To answer this question, we group our end-of-month forecast origins in three bins. Given the release schedule of macroeconomic data in the U.S., at the end of the first month of the quarter, no additional monthly observations of non-financial variables for the current quarter are available. At the end of the second and third month either one or two within-quarter monthly observations are available. We find that during the third month of the quarter the switch from a QF-VAR to a MF-VAR improves the one-step-ahead forecast (nowcast) accuracy on average by 60% for nonfinancial variables observed at the monthly frequency and by 11% for variables observed at quarterly frequency. In the first month of the quarter the improvements are about 6%. The improvement in forecast accuracy is most pronounced for short-horizon forecasts and tempers off in the medium and long run. Thus, if the goal is to generate VAR

nowcasts or forecasts of one- or two-quarters ahead, it is well worth switching to from a QF-VAR to a MF-VAR. If the focus is on a one- to two-year horizon, the QF-VAR is likely to suffice.

Second, we generate real-time forecasts of macroeconomic aggregates for the 2008-09 (Great) recession period. This episode is of great interest to macroeconomists, because the large drop in aggregate real activity poses a challenge for existing structural and non-structural models. We document that the monthly information helped the MF-VAR track the economic downturn more closely in real time than the QF-VAR supporting the view that the MF-VAR is an attractive alternative to a standard QF-VAR. Third, as a by-product of the MF-VAR estimation, we generate an estimate of monthly GDP growth rates, which may be of independent interest to business cycle researchers. Finally, we also provide a comparison of bivariate MF-VARs to mixed data sampling (MIDAS) regressions. We find for GDP forecasts that the percentage differential in forecast accuracy is the same, regardless whether the forecast is made at the end of the first, second, or third month of the quarter. We are interpreting this finding as both approaches being able to exploit the information differentials between the three months of the quarter in relative terms equally well. In absolute terms, the MF-VARs tend to outperform the MIDAS regressions in our particular implementation.

This paper focuses on VARs which are time series models that generate multivariate predictive distributions. VARs have been an important forecasting tool in practice (see Litterman (1986) for an early assessment and Giannone, Lenza, and Primiceri (2012) for a recent assessment) and there is strong evidence that they perform well in high-dimensional environments if estimated with shrinkage estimation techniques (see, e.g., De Mol, Giannone, and Reichlin (2008) and Banbura, Giannone, and Reichlin (2010)). Moreover, in addition to generating unconditional forecasts, they are widely used to produce conditional

forecasts, e.g., conditional on an interest rate path (see Doan, Litterman, and Sims (1984) and Waggoner and Zha (1999)), which do require a multivariate framework. In our comparison between MF-VARs and QF-VARs we mostly study univariate root-mean-squared errors (RMSEs), though we also consider log determinants of (multivariate) forecast error covariance matrices. To the extent that we are considering univariate RMSEs one could conduct comparisons with univariate predictive regressions. However, comparisons of VAR forecasts to forecasts from other classes of time series models are not the focus of this paper and can be found elsewhere in the literature (see, e.g., Chauvet and Potter (2013) for forecasting output and Faust and Wright (2013) for forecasting inflation).

To cope with the high dimensionality of the parameter space, the MF-VAR is equipped with a Minnesota prior and estimated with Bayesian methods. Our version of the Minnesota prior is based on Sims and Zha (1998). This prior is also used, for instance, in Banbura, Giannone, and Reichlin (2010) and Giannone, Lenza, and Primiceri (2012) and the authors document that the forecasting performance of the Bayesian VAR dominates that of an unrestricted VAR by a large margin. Alternative prior specifications for Bayesian VARs are surveyed in Karlsson (2013) and the effect of specification choice on forecast accuracy is studied in Carriero, Clark, and Marcellino (2011). In order to generate accurate forecasts it is important that the prior covariance matrix is properly configured. A set of hyperparameters controls the degree of shrinkage toward the prior mean and we choose the hyperparameter to maximize the log MDD. MDD-based hyperparameter selection has been discussed, for instance, in Phillips (1996), used in Del Negro and Schorfheide (2004) and, most recently, studied in Giannone, Lenza, and Primiceri (2012).

We are building on existing approaches of treating missing observations in state-space models (see, for instance, the books by Harvey (1989b) and Durbin and Koopman (2001b)). We are employing modern Bayesian computational tools, in particular the method of data

augmentation. We construct a Gibbs sampler along the lines of Carter and Kohn (1994b) that alternates between the conditional distribution of the VAR parameters given the unobserved monthly series, and the conditional distribution of the missing monthly observations given the VAR parameters. Draws from the former distribution are generated by direct sampling from a Normal-Inverted Wishart distribution, whereas draws from the latter are obtained by applying a simulation smoother to the state-space representation of the MF-VAR. Our numerical approximation of the log MDD is based on the modified harmonic mean estimator proposed by Geweke (1999).

An alternative Gibbs sampling approach for the coefficients in an MF-VAR is explored in Chiu, Foerster, Kim, and Seoane (2012). Their algorithm also iterates over the conditional posterior distributions of the VAR parameters and the missing monthly observations, but utilizes a different procedure to draw the missing observations. The focus of their paper is on parameter estimation rather than forecasting. The authors link the coefficients of the MF-VAR to the coefficients of a QF-VAR via a transformation. Chiu, Foerster, Kim, and Seoane (2012) then compare the posterior distributions of parameters and impulse response functions obtained from the estimation of the two models to document the value of the monthly observations.

Mixed-frequency observations have also been utilized in the estimation of dynamic factor models (DFMs). Mariano and Murasawa (2003b) apply maximum-likelihood factor analysis to a mixed-frequency series of quarterly real GDP and monthly business cycle indicators to construct an index that is related to monthly real GDP. Aruoba, Diebold, and Scotti (2009b) develop a DFM to construct a broad index of economic activity in real time using a variety of data observed at different frequencies. Giannone, Reichlin, and Small (2008) use a mixed-frequency DFM to evaluate the marginal impact that intra-monthly data releases have on current-quarter forecasts (nowcasts) of real GDP growth.

When using our MF-VAR to forecast quarterly GDP growth, we are essentially predicting a quarterly variable based on a mixture of quarterly and monthly regressors. Ghysels, Sinko, and Valkanov (2007) propose a simple univariate regression model, called a mixed data sampling (MIDAS) regression, to exploit high-frequency information without having to estimate a state-space model. To cope with potentially large numbers of regressors, the coefficients for the high-frequency regressors are tightly restricted through distributed lag polynomials that are indexed by a small number of hyperparameters. Bayesian versions of the MIDAS approach are developed in Rodriguez and Puggioni (2010) and Carriero, Clark, and Marcellino (2012).

Ghysels (2012) generalizes the MIDAS approach to a VAR setting. Unlike our MF-VAR, his MIDAS VAR is an observation-driven model that does not require numerical techniques to integrate out unobserved monthly variables. As in Chiu, Foerster, Kim, and Seoane (2012), the empirical analysis focuses on impulse responses but not on real-time forecasting. In our view, the state-space setup pursued in this paper is more transparent and flexible and the computational advances of the last decade make it feasible to estimate Bayesian state-space models with code written in high-level languages such as MATLAB in a short amount of time.

Bai, Ghysels, and Wright (2013) examine the relationship between MIDAS regressions and state-space models applied to mixed-frequency data. They consider dynamic factor models and characterize conditions under which the MIDAS regression exactly replicates the steady state Kalman filter weights on lagged observables. They conclude that Kalman filter forecasts are typically a little better, but MIDAS regressions can be more accurate if the state-space model is misspecified or over-parameterized. Kuzin, Marcellino, and Schumacher (2011) compare the accuracy of Euro Area GDP growth forecasts from MIDAS regressions and MF-VARs estimated by maximum likelihood. The authors find that the

relative performances of MIDAS and MF-VAR forecasts differ depending on the predictors and forecast horizons. Overall, the authors do not find a clear winner in terms of forecasting performance.

The remainder of this paper is organized as follows. Section 5.2 presents the state-space representation of the MF-VAR and discusses Bayesian inference and forecasting. The real-time data sets used for the forecast comparison of MF-VAR and QF-VAR, as well as the timing of within-quarter monthly information, are discussed in Section 5.3. The empirical results are presented in Section 5.4. Finally, Section 5.5 concludes. The Online Appendix provides detailed information about the Bayesian computations, the construction of the data set, as well as additional empirical results.

5.2 A Mixed-Frequency Vector Autoregression

The MF-VAR considered in this paper is based on a standard constant-parameter VAR in which the length of the time period is one month. Since some macroeconomic time series, e.g., GDP, are measured only at quarterly frequency, we treat the corresponding monthly values as unobserved. To cope with the missing observations, the MF-VAR is represented as a state-space model in Section 5.2.1. In order to ease the exposition, we use a representation with a state vector that includes even those variables that are observable at monthly frequency, e.g., the aggregate price level, the unemployment rate, and the interest rate. A computationally more efficient representation in which variables observed at monthly frequency are dropped from the state vector is presented in the Online Appendix. Bayesian inference and forecasting are discussed in Section 5.2.2.

Throughout this paper, we use $Y_{t_0:t_1}$ to denote the sequence of observations or random variables $\{y_{t_0}, \ldots, y_{t_1}\}$. If no ambiguity arises, we sometimes drop the time subscripts and abbreviate $Y_{1:T}$ by Y. If θ is the parameter vector, then we use $p(\theta)$ to denote the prior

density, $p(Y|\theta)$ is the likelihood function, and $p(\theta|Y)$ the posterior density. We use *iid* to abbreviate independently and identically distributed, and $N(\mu, \Sigma)$ denotes a multivariate normal distribution with mean μ and covariance matrix Σ. Let $\otimes$ be the Kronecker product. If $X|\Sigma \sim MN_{p\times q}(M, \Sigma \otimes P)$ is matricvariate Normal and $\Sigma \sim IW_q(S, \nu)$ has an Inverted Wishart distribution, we say that (X, Σ) has a Normal-Inverted Wishart distribution: $(X, \Sigma) \sim MNIW(M, P, S, \nu)$.

5.2.1 State-Transitions and Measurement

We assume that the economy evolves at monthly frequency according to the following VAR(p) dynamics:

$$x_t = \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \Phi_c + u_t, \quad u_t \sim iidN\big(0, \Sigma\big). \tag{5.1}$$

The $n \times 1$ vector of macroeconomic variables x_t can be composed into $x_t = [x'_{m,t}, x'_{q,t}]'$, where the $n_m \times 1$ vector $x_{m,t}$ collects variables that are observed at monthly frequency, e.g., the consumer price index and the unemployment rate, and the $n_q \times 1$ vector $x_{q,t}$ comprises the unobserved monthly variables that are only published at quarterly frequency, e.g., GDP. Define $z_t = [x'_t, \ldots, x'_{t-p+1}]'$ and $\Phi = [\Phi_1, \ldots, \Phi_p, \Phi_c]'$. Write the VAR in (5.1) in companion form as

$$z_t = F_1(\Phi) z_{t-1} + F_c(\Phi) + v_t, \quad v_t \sim iidN\big(0, \Omega(\Sigma)\big), \tag{5.2}$$

where the first n rows of $F_1(\Phi)$, $F_c(\Phi)$, and v_t are defined to reproduce (5.1) and the remaining rows are defined to deliver the identities $x_{q,t-l} = x_{q,t-l}$ for $l = 1, \ldots, p-1$. The $n \times n$ upper-left submatrix of Ω equals Σ and all other elements are zero. Equation (5.2) is the state-transition equation of the MF-VAR.

We proceed by describing the measurement equation. One can handle the unobserved variables in several ways: by imputing zeros and modifying the measurement equation by

setting the loadings on the state variables to zero (e.g., Mariano and Murasawa (2003b)); by setting the measurement error variance to infinity (e.g., Giannone, Reichlin, and Small (2008)); or by varying the dimension of the vector of observables as a function of time t (e.g., Durbin and Koopman (2001b)). We employ the latter approach. To do so, some additional notation is useful. Let T denote the forecast origin and let $T_b \leq T$ be the last period that corresponds to the last month of the quarter for which all quarterly observations are available. The subscript b stands for *balanced* sample. Up until period T_b the vector of monthly series $x_{m,t}$ is observed every month. We denote the actual observations by $y_{m,t}$ and write

$$y_{m,t} = x_{m,t}, \quad t = 1, \ldots, T_b. \tag{5.3}$$

Assuming that the underlying monthly VAR has at least three lags, that is, $p \geq 3$, we express the three-month average of $x_{q,t}$ as

$$\tilde{y}_{q,t} = \frac{1}{3}(x_{q,t} + x_{q,t-1} + x_{q,t-2}) = \Lambda_{qz} z_t. \tag{5.4}$$

For variables measured in logs, e.g., $\ln GDP$, the formula can be interpreted as a log-linear approximation to an arithmetic average of GDP that preserves the linear structure of the state-space model. For flow variables such as GDP, we adopt the NIPA convention and annualize high-frequency flows. As a consequence, quarterly flows are the average and not the sum of monthly flows. This three-month average, however, is only observed for every third month, which is why we use a tilde superscript. Let $M_{q,t}$ be a selection matrix that equals the identity matrix if t corresponds to the last month of a quarter and is empty otherwise. Adopting the convention that the dimension of the vector $y_{q,t}$ is n_q in periods in which quarterly averages are observed and zero otherwise, we write

$$y_{q,t} = M_{q,t}\tilde{y}_{q,t} = M_{q,t}\Lambda_{qz} z_t, \quad t = 1, \ldots, T_b. \tag{5.5}$$

For periods $t = T_b + 1, \ldots, T$ no additional observations of the quarterly time series are available. Thus, for these periods the dimension of $y_{q,t}$ is zero and the selection matrix $M_{q,t}$ in (5.5) is empty. However, the forecaster might observe additional monthly variables. Let $y_{m,t}$ denote the subset of monthly variables for which period t observations are reported by the statistical agency after period T, and let $M_{m,t}$ be a deterministic sequence of selection matrices such that (5.3) can be extended to

$$y_{m,t} = M_{m,t} x_{m,t}, \quad t = T_b + 1, \ldots, T. \tag{5.6}$$

Notice that the dimension of the vector $y_{m,t}$ is potentially time varying and less than n_m. The measurement equations (5.3) to (5.6) can be written more compactly as

$$y_t = M_t \Lambda_z z_t, \quad t = 1, \ldots, T. \tag{5.7}$$

Here, M_t is a sequence of selection matrices that selects the time t variables that have been observed by period T and are part of the forecaster's information set. In sum, the state-space representation of the MF-VAR is given by (5.2) and (5.7).

5.2.2 Bayesian Inference

The starting point of Bayesian inference for the MF-VAR is a joint distribution of observables $Y_{1:T}$, latent states $Z_{0:T}$, and parameters (Φ, Σ), conditional on a pre-sample $Y_{-p+1:0}$ to initialize lags. Using a Gibbs sampler, we generate draws from the posterior distributions of $(\Phi, \Sigma)|(Z_{0:T}, Y_{-p+1:T})$ and $Z_{0:T}|(\Phi, \Sigma, Y_{-p+1:T})$. Based on these draws, we are able to simulate future trajectories of y_t to characterize the predictive distribution associated with the MF-VAR and to calculate point and density forecasts.

Prior Distribution. An important challenge in practical work with VARs is to cope with the dimensionality of the coefficient matrix Φ. Informative prior distributions can often mitigate the curse of dimensionality. A widely used prior in the VAR literature is the so-called Minnesota prior. This prior dates back to Litterman (1980) and Doan, Litterman,

and Sims (1984). We use the version of the Minnesota prior described in Del Negro and Schorfheide (2011)'s handbook chapter, which in turn is based on Sims and Zha (1998). The main idea of the Minnesota prior is to center the distribution of Φ at a value that implies a random-walk behavior for each of the components of x_t in (5.1). Our version of the Minnesota prior for (Φ, Σ) is proper and belongs to the family of MNIW distributions. We implement the Minnesota prior by mixing artificial (or *dummy*) observations into the estimation sample. The artificial observations are computationally convenient and allow us to generate plausible a priori correlations between VAR parameters. The variance of the prior distribution is controlled by a low-dimensional vector of hyperparameters λ. Details of the prior are relegated to the Online Appendix, and the choice of hyperparameters is discussed below.

Posterior Inference. The joint distribution of data, latent variables, and parameters conditional on some observations to initialize lags can be factorized as follows:

$$\begin{aligned} & p(Y_{1:T}, Z_{0:T}, \Phi, \Sigma | Y_{-p+1:0}, \lambda) \\ = \;& p(Y_{1:T}|Z_{0:T}) p(Z_{1:T}|z_0, \Phi, \Sigma) p(z_0|Y_{-p+1:0}) p(\Phi, \Sigma|\lambda). \end{aligned} \tag{5.8}$$

The distribution of $Y_{1:T}|Z_{1:T}$ is given by a point mass at the value of $Y_{1:T}$ that satisfies (5.7). The density $p(Z_{1:T}|z_0, \Phi, \Sigma)$ is obtained from the linear Gaussian regression (5.2). The conditional density $p(z_0|Y_{-p+1:0})$ is chosen to be Gaussian and specified in the Online Appendix. Finally, $p(\Phi, \Sigma|\lambda)$ represents the prior density of the VAR parameters. The factorization (5.8) implies that the conditional posterior densities of the VAR parameters and the latent states of the MF-VAR take the form

$$\begin{aligned} p(\Phi, \Sigma|Z_{0:T}, Y_{-p+1:T}) &\propto p(Z_{1:T}|z_0, \Phi, \Sigma) p(\Phi, \Sigma|\lambda) \\ p(Z_{0:T}|\Phi, \Sigma, Y_{-p+1:T}) &\propto p(Y_{1:T}|Z_{1:T}) p(Z_{1:T}|z_0, \Phi, \Sigma) p(z_0|Y_{-p+1}). \end{aligned} \tag{5.9}$$

We follow Carter and Kohn (1994b) and use a Gibbs sampler that iterates over the two

conditional posterior distributions in (5.9). Conditional on $Z_{0:T}$, the companion-form state transition (5.2) is a multivariate linear Gaussian regression. Since our prior for (Φ, Σ) belongs to the MNIW family, so does the posterior and draws from this posterior can be obtained by direct Monte Carlo sampling. Likewise, since the MF-VAR is set up as a linear Gaussian state-space model, a standard simulation smoother can be used to draw the sequence $Z_{0:T}$ conditional on the VAR parameters. The distribution $p(z_0|Y_{-p+1})$ provides the initialization for the Kalman-filtering step of the simulation smoother. A detailed discussion of these computations can be found in textbook treatments of the Bayesian analysis of state-space models, e.g., the handbook chapters by Del Negro and Schorfheide (2011) and Giordani, Pitt, and Kohn (2011).

Computational Considerations. For expositional purposes, it has been convenient to define the vector of state variables as $z_t = [x_t', \ldots, x_{t-p+1}]'$, which includes the variables observed at monthly frequency. From a computational perspective, this definition is inefficient because it enlarges the state space of the model unnecessarily. We show in the Online Appendix how to rewrite the state-space representation of the MF-VAR in terms of a lower-dimensional state vector $s_t = [x_{q,t}', \ldots, x_{q,t-p}]'$ that only includes the variables (and their lags) observed at quarterly frequency. Our simulation smoother uses the small state vector s_t for $t = 1, \ldots, T_b$ and then switches to the larger state vector z_t for $t = T_b+1, \ldots, T$ to accommodate missing monthly observations toward the end of the sample.

Forecasting. For each draw $(\Phi, \Sigma, Z_{0:T})$ from the posterior distribution we simulate a trajectory $Z_{T+1:T+H}$ based on the state-transition equation (5.2). Since we evaluate forecasts of quarterly averages in our empirical analysis, we time-aggregate the simulated trajectories accordingly. Based on the simulated trajectories (approximate) point forecasts can be obtained by computing means or medians. Interval forecasts and probability integral transformations (see Section 5.6.3) can be computed from the empirical distribution of the

simulated trajectories.

5.2.3 Marginal Likelihood Function and Hyperparameter Selection

The empirical performance of the MF-VAR is sensitive to the choice of hyperparameters. The prior is parameterized such that $\lambda = 0$ corresponds to a flat (and therefore improper) prior for (Φ, Σ). As $\lambda \longrightarrow \infty$, the MF-VAR is estimated subject to the random-walk restriction implied by the Minnesota prior. The best forecasting performance of the MF-VAR is likely to be achieved for values of λ that are in between the two extremes. In a Bayesian framework the hyperparameter, λ can be interpreted as a model index (since a Bayesian model is the product of likelihood function and prior distribution). We consider a grid $\lambda \in \Lambda$ and assign equal prior probability to each value on the grid. Thus, the posterior probability of λ is proportional to the MDD

$$p(Y_{1:T}|Y_{-p+1:0}, \lambda) = \int p(Y_{1:T}, Z_{0:T}, \Phi, \Sigma|Y_{-p+1:0}, \lambda) d(Z_{0:T}, \Phi, \Sigma). \tag{5.10}$$

The log MDD can be interpreted as the sum of one-step-ahead predictive scores:

$$\ln p(Y_{1:T}|Y_{-p+1:0}, \lambda) = \sum_{t=1}^{T} \ln \int p(y_t|Y_{-p+1:t-1}, \Phi, \Sigma) p(\Phi, \Sigma|Y_{-p+1:t-1}, \lambda) d(\Phi, \Sigma). \tag{5.11}$$

The terms on the right-hand side of (5.11) provide a decomposition of the one-step-ahead predictive densities $p(y_t|Y_{1-p:t-1}, \lambda)$. This decomposition highlights the fact that inference about the parameter is based on time $t-1$ information, when making a one-step-ahead prediction for y_t.

Hyperparameter Selection. To generate the MF-VAR forecasts, for each forecast origin we condition on the value $\hat{\lambda}_T$ that maximizes the log MDD. This procedure can be viewed as an approximation to a model averaging procedure that integrates out λ based on the posterior $p(\lambda|Y_{-p+1:T})$. The MDD-based selection of VAR hyperparameters has a fairly long history and tends to work well for forecasting purposes (see Giannone, Lenza, and Primiceri (2012) for a recent study).

Marginal Data Density Approximation. From (5.10) we see that the computation of the MDD involves integrating out the latent states. In the remainder of this section we describe how we compute the integral. To simplify the exposition we consider the special case of $n = 2$, $p = 1$, and $T = 3$. We assume that one of the variables is observed at monthly frequency and the other as a quarterly average. Thus, we can write $z_t = [x_{m,t}, x_{q,t}]'$. The observations $Y_{1:3}$ are related to the states $Z_{1:3}$ as follows:

$$y_1 = x_{m,1}, \quad y_2 = x_{m,2}, \quad y_3 = \begin{bmatrix} x_{m,3} \\ \frac{1}{3}(x_{q,1} + x_{q,2} + x_{q,3}) \end{bmatrix}. \tag{5.12}$$

Using a change of variable of the form

$$Z_{1:3} = J \begin{bmatrix} Y_{1:3} \\ W_{1:3} \end{bmatrix} \tag{5.13}$$

where $Z_{1:3} = [z_1', z_2', z_3]'$, $Y_{1:3} = [y_1, y_2, y_3']'$, $W_{1:3} = [x_{q,1}, x_{q,2}]'$ (note that despite the $1:3$ subscript, $W_{1:3}$ is a 2×1 vector in this example), and J is a 6×6 non-singular matrix of constants. Thus, we can replace $p(Z_{1:3}|\lambda)$ by $p(Y_{1:3}, W_{1:3}|\lambda) = p(Y_{1:3}|W_{1:3})p(W_{1:3}|\lambda)$. Using Bayes Theorem, we can write (abstracting from the initialization of the VAR)

$$\frac{1}{p(Y_{1:3}|\lambda)} = \frac{p(W_{1:3}|Y_{1:3}, \lambda)}{p(Y_{1:3}, W_{1:3}|\lambda)}. \tag{5.14}$$

Suppose that $f(W_{1:3})$ has the property that $\int f(W_{1:3})dW_{1:3} = 1$, and let $\{W_{1:3}^{(i)}\}_{i=1}^N$ denote a sequence of draws from the posterior distribution of $W_{1:3}|(Y_{1:3}, \lambda)$. Then the MDD can be approximated using Geweke (1999)'s harmonic mean estimator, which is widely used in the DSGE model literature to approximate MDDs in high-dimensional settings:

$$\hat{p}(Y_{1:3}|\lambda) = \left[\frac{1}{N} \sum_{i=1}^{N} \frac{f(W_{1:3}^{(i)})}{p(Y_{1:3}, W_{1:3}^{(i)}|\lambda)} \right]^{-1}. \tag{5.15}$$

The draws from the distribution of $W_{1:3}|(Y_{1:3}, \lambda)$ can be obtained by transforming the draws from $Z_{1:3}|(Y_{1:3}, \lambda)$, which are generated as a by-product of the posterior sampler described in Section 5.2.2. Using the properties of the MNIW distribution, it is straightforward to

compute

$$p(Z_{1:3}|\lambda) = \int p(Z_{1:3}|\Phi,\Sigma)p(\Phi,\Sigma|\lambda)d(\Phi,\Sigma) \tag{5.16}$$

analytically. A straightforward change of variables based on (5.13) leads from $p(Z_{1:3}|\lambda)$ to $p(Y_{1:3}, W_{1:3}^{(i)}|\lambda)$. Note that the Jacobian of this transformation is simply a constant term.

Generalization. Taking the initialization of the VAR into account, the identity provided in (5.14) can be generalized as follows:

$$\frac{1}{p(Y_{1:T}|Y_{-p+1:0},\lambda)} = \frac{p(W_{1:T},w_0|Y_{1:T},Y_{-p+1:0},\lambda)}{p(W_{1:T},Y_{1:T},w_0|Y_{-p+1:0},\lambda)}, \tag{5.17}$$

with the understanding that $W_{1:T}$ stacks the unobserved values of $x_{q,t}$ for the first and second month of each quarter of the estimation sample and w_0 contains the corresponding values for the initialization period $t = -p+1,\ldots,0$. The approximation of the MDD becomes:

$$\hat{p}(Y_{1:T}|Y_{-p+1:0},\lambda) = c\left[\frac{1}{N}\sum_{i=1}^{N}\frac{f_0(w_0^{(i)})f(W_{1:T}^{(i)})}{p(Z_{1:T}^{(i)}|z_0^{(i)},\lambda)p(z_0^{(i)}|Y_{-p+1:0},\lambda)}\right]^{-1}, \tag{5.18}$$

The constant c in (5.18) captures the Jacobian term associated with the change-of-variables from $(w_0, W_{1:T}, Y_{1:T})$ to $(z_0, Z_{1:T})$. For the function $f(\cdot)$ we follow Geweke (1999) and use a trimmed multivariate normal distribution with mean $\hat{\mu}_{W_{1:T}} = \frac{1}{N}\sum_{i=1}^{N} W_{1:T}^{(i)}$ and variance $\hat{\Sigma}_{W_{1:T}} = \frac{1}{N}\sum_{i=1}^{N} W_{1:T}^{(i)}W_{1:T}^{(i)'} - \hat{\mu}_{W_{1:T}}\hat{\mu}'_{W_{1:T}}$. This normal distribution approximates $p(W_{1:T}|Y_{-p+1:T})$ and stabilizes the ratio in (5.18). We set $f_0(w_0^{(i)}) = p(z_0^{(i)}|Y_{-p+1:0},\lambda)$ such that the two terms cancel. To evaluate the denominator, we use the analytical expression for $p(Z_{1:T}^{(i)}|z_0^{(i)},Y_{-p+1:0},\lambda)$, which is obtained from the the normalization constants for the MNIW distribution and is provided, for instance, in Section 2 of Del Negro and Schorfheide (2011).

5.3 Real-Time Data Sets and Information Structure

We subsequently conduct a pseudo-out-of-sample forecast experiment with real-time data to study the extent to which the incorporation of monthly observations via an MF-VAR model improves upon forecasts generated with a VAR that is based on time-aggregated quarterly data (QF-VAR). We consider VARs for eleven macroeconomic variables, which are summarized in Section 5.3.1. The construction of the real-time data sets and the classification of forecast origins based on within-quarter monthly information are described in Section 5.3.2. Section 5.3.3 explains our choice of actual values that are used to compute forecast errors.

5.3.1 Macroeconomic Variables

We consider VARs for eleven macroeconomic variables, of which three are observed at quarterly frequency and eight are observed at monthly frequency. The quarterly series are GDP, Fixed Investment (INVFIX), and Government Expenditures (GOV). The monthly series are the Unemployment Rate (UNR), Hours Worked (HRS), Consumer Price Index (CPI), Industrial Production Index (IP), Personal Consumption Expenditure (PCE), Federal Funds Rate (FF), Treasury Bond Yield (TB), and S&P 500 Index (SP500). Precise data definitions are provided in the Online Appendix. Series that are observed at a higher than monthly frequency are time-aggregated to monthly frequency. The variables enter the VARs in log levels with the exception of UNR, FF, and TB, which are divided by 100 in order to make them commensurable in scale to the other log-transformed variables.

5.3.2 Real-Time Data for End-of-Month Forecasts

We consider an increasing sequence of estimation samples $Y_{-p+1:T}$, $T = T_{min}, \ldots, T_{max}$, and generate forecasts for periods $T+1, \ldots, T+H$. The maximum forecast horizon H is chosen to be 24 months. The period $t = 1$ corresponds to 1968:M1, T_{min} is 1997:M7,

and T_{max} is 2010:M1, which yields 151 estimation samples. We eliminated four of the 151 samples because the real-time data for PCE and INVFIX were incomplete. The estimation samples are constructed from real-time data sets, assuming that the forecasts are generated on the last day of each month. Due to data revisions by statistical agencies, observations of $Y_{1:T-1}$ published in period T are potentially different from the observations that had been published in period $T-1$. For this reason, real-time data are often indexed by a superscript, say $\tau \geq T$, which indicates the vintage or data release date. Using this notation, a forecaster at time T potentially has access to a triangular array of data $Y^1_{-p+1:1}, Y^2_{-p+1:2}, \ldots, Y^T_{-p+1:T}$. Rather than using the entire triangular array and trying to exploit the information content in data revisions, we estimate the MF-VAR and QF-VAR for each forecast origin T based on the information set $Y^T_{-p+1:T} = \{y^T_{-p+1}, \ldots, y^T_T\}$. As in Section 5.2, we are using the convention that the vector y^T_t contains only the subset of the eleven variables listed above for which observations are available at the end of month T.

In order to assess the usefulness of within-quarter information from monthly variables, we sort the forecast origins $T_{min}, \ldots, T_{max}$ into three groups that reflect different within-quarter information sets. Forecast error statistics will be computed for each group separately. The grouping of forecast origins is best explained in a concrete example. Consider the January 31, 1998 forecast origin. By the end of January, the Bureau of Economic Analysis (BEA) has just published an *advance* estimate of 1997:Q4 GDP. In addition, the forecaster has access to nonfinancial monthly indicators from December 1997 and earlier. A similar situation arises at the end of April, July, and October. We refer to this group of forecast origins as "+0 months," because the current-quarter forecasts do not use any additional nonfinancial monthly variables.

At the end of February 1998, the forecaster has access to an *preliminary* estimate of 1997:Q4 GDP and to observations for unemployment, industrial production, and so forth,

Table 5.1: Illustration of Information Sets

January (+0 Months)

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M12	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q1	M1	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

February (+1 Month)

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M12	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q1	M1	X	X	X	X	X	X	X	X	∅	∅	∅
Q1	M2	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

March (+2 Month)

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M12	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q1	M1	X	X	X	X	X	X	X	X	∅	∅	∅
Q1	M2	X	X	X	X	X	X	X	X	∅	∅	∅
Q1	M3	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

Notes: ∅ indicates that the observation is missing. X denotes monthly observation and QAv denotes quarterly average. "+0 Months" group: January, April, July, October; "+1 Month" group: February, May, August, November; "+2 Month" group: March, June, September, December.

for January 1998. Thus, we group February, May, August, and November forecasts and refer to them as "+1 month." Following the same logic, the last subgroup of forecast origins has two additional monthly indicators ("+2 months") and the *final* release of GDP for 1997:Q4 in the information set. Unlike the non-financial variables, which are released with a lag, financial variables are essentially available instantaneously. In particular, at the end of each month, the forecaster has access to average interest rates (FF and TB) and stock prices (SP500). The typical information sets for the three subgroups of forecast origins are summarized in Table 5.1.

Unfortunately, due to variation in release dates, not all 151 estimation samples mimic the information structure in Table 5.1. For 47 samples the last PCE figure is released with a two-period (approximately five weeks) instead of one-period (approximately four weeks) lag. This exception occurs for 28 samples of the "+0 months" group. For these samples

a late release of PCE implies the quarterly consumption for the last completed quarter is not available. In turn, the QF-VAR could only be estimated based on information up to $T-4$ instead of $T-1$ and would be at a severe disadvantage compared to the MF-VAR. Since PCE is released only a few days after the period T forecasts are made, we pre-date its release. Thus, for the 28 samples of the "+0 months" group that are subject to the irregular timing, we use PCE_{T-1} in the estimation of both the QF-VAR and MF-VAR. No adjustments are made for the "+1 month" and "+2 months" groups. Further details about these exceptions are provided in the Online Appendix.

5.3.3 Actuals for Forecast Evaluation

The real-time-forecasting literature is divided as to whether forecast errors should be computed based on the first release following the forecast date, say y_{T+h}^{T+h}, or based on the most recent vintage, say $y_{t+h}^{T_*}$. The former might do a better job of capturing the forecaster's loss, whereas the latter is presumably closer to the underlying "true" value of the time series. We decided to follow the second approach and evaluate the forecasts based on actual values from the $T_* =$ 2012:M1 data vintage. While the MF-VAR in principle generates predictions at the monthly frequency, we focus on the forecasts of quarterly averages, which can be easily compared to forecasts from the QF-VAR.

5.4 Empirical Results

The empirical analysis proceeds in four parts. The hyperparameter selection is discussed in Section 5.4.1. Section 5.4.2 compares root mean squared error (RMSE) statistics from the MF-VAR to a QF-VAR and a set of MIDAS regressions. Section 5.4.3 contrasts MF-VAR density forecasts during the 2008-9 (Great) recession with QF-VAR forecasts. Finally, in Section 5.4.4 we present a monthly GDP series that arises as a by-product of the MF-VAR estimation. Based on some preliminary exploration of the MDDs, we set the number of lags

in the (monthly) state transition of the MF-VAR to $p_{(m)} = 6$ and the number of lags in the QF-VAR to $p_{(q)} = 2$.

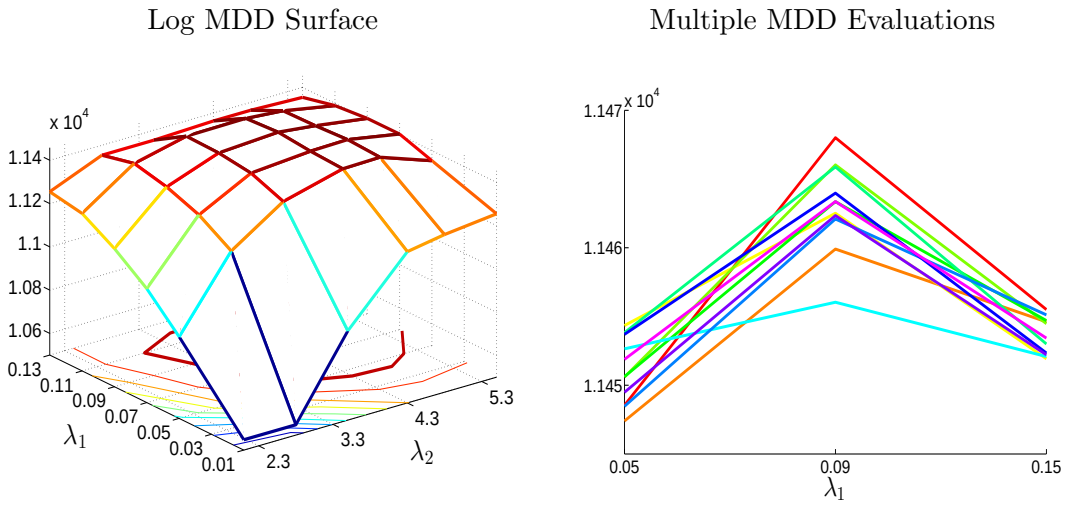
Unless otherwise noted, for each estimation sample we generate 20,000 draws from the posterior distribution of the VAR parameters using the MCMC algorithm described in Section 5.2.2. We discard the first 10,000 draws and use the remaining 10,000 to calculate Monte Carlo approximations of posterior moments. The Online Appendix provides some information on the accuracy of the MCMC. The Monte Carlo standard deviation of the posterior mean forecasts (output, inflation, interest rates, and unemployment), computed across independent runs of the MCMC, is generally less than 0.5 basis points. For comparison, the RMSE associated with these forecasts ranges from 10 to 200 basis points.

5.4.1 Hyperparameter Selection

We will subsequently compare MF-VAR and QF-VAR forecasts. Both VARs are equipped with a Minnesota prior that is represented in terms of dummy observations and indexed by a vector of hyperparameters λ. We use the same set of dummy observations for both types of VAR. However, the hyperparameters are chosen for each type of VAR separately. The careful choice of this hyperparameter vector is crucial for obtaining accurate forecasts. As explained in Section 5.2.3, we determine the hyperparameters by maximizing the log MDD. For the QF-VAR the MDD can be computed analytically (see, e.g., Del Negro and Schorfheide (2011)) and the maximization is straightforward. Thus, we will focus on the hyperparameter selection for the MF-VAR.

The hyperparameter vector consists of five element, controlling: the overall tightness of the prior (λ_1); the rate at which the prior variance on higher-order lag coefficients decays (λ_2); the dispersion of the prior on the innovation covariance matrix (λ_3); the extent to which the sum-of-coefficient on the lags of a variable $x_{i,t}$ is tilted toward unity (λ_4); and the extent to which co-persistence restrictions are imposed on the VAR coefficients (λ_5).

Figure 5.1: Log Marginal Data Density for 11-Variable MF-VAR



Notes: The two plots depict $\ln \hat{p}(Y_{1:T}|Y_{-p+1:0,\lambda})$. In the left panel, we condition on $\lambda_3 = 1$, $\lambda_4 = 2.7$, and $\lambda_5 = 4.3$. In the right panel we condition on $\lambda_2 = 4.3$, $\lambda_3 = 1$, $\lambda_4 = 2.7$, and $\lambda_5 = 4.3$. Each "hair" corresponds to a separate run of the MCMC algorithm.

In general, the larger λ_i the smaller the prior variance and the more informative the prior. From a preliminary analysis based on the QF-VAR, we conclude that λ_3 is not particularly important for the forecasting performance and fix it as $\hat{\lambda}_3 = 1$. Based on a preliminary search over a grid $\Lambda^{(1)}$ we determine suitable values for λ_4 and λ_5 for the first recursive sample, which ranges from 1968:M1 to 1997:M7. These values are $\hat{\lambda}_4 = 2.7$ and $\hat{\lambda}_5 = 4.3$. Conditioning on $\hat{\lambda}_3$ to $\hat{\lambda}_5$, we use a second grid, $\hat{\Lambda}^{(2)}$ to refine the choice of λ_1 and λ_2.

The log MDD surface is depicted in the left panel of Figure 5.1 as function of λ_1 and λ_2, holding the remaining three hyperparameters fixed at $\lambda_3 = 1$, $\lambda_4 = \hat{\lambda}_4$, and $\lambda_5 = \hat{\lambda}_5$. The surface has a convex shape and is maximized at $\hat{\lambda}_1 = 0.09$ and $\hat{\lambda}_2 = 4.3$. At its peak the value of the log MDD is approximately 11,460. While the surface is fairly flat near the peak, e.g. for $\lambda_1 \in [0.05, 0.15]$ and $\lambda_2 \in [4, 4.5]$, the MDD values drop substantially for values of λ outside of these intervals. To assess the accuracy of the MDD evaluation, which involves the numerical evaluation of a high-dimensional integral, we display a hairplot of a slice of the

MDD surface in the right panel of Figure 5.1, fixing λ_2 at 4.3. Each hairline corresponds to a separate run of the MCMC algorithm. We focus on the interval $\lambda_1 \in [0.05, 0.15]$. While there is some noticeable Monte Carlo variation with respect to the absolute magnitude of the log MDD, this variation does not affect inference with respect to the optimal value of λ on the grid. For each simulation, the log MDD peaks at 0.09. The accuracy of the approximation can be improved by increasing the number of MCMC draws.

The re-optimization of the hyperparameters for the MF-VAR is computationally costly. Because we expect the optimal hyperparameter choices to evolve smoothly over time, we are reoptimizing with respect to λ approximately every three years, namely for the 40th, the 75th, the 110th, and the 151th recursive sample. During this reoptimization we keep $\hat{\lambda}_3$, $\hat{\lambda}_4$, and $\hat{\lambda}_5$ fixed. The reoptimization essentially left the choice of hyperparameters unchanged. We obtained a similar result for the QF-VAR and decided to keep the MF-VAR and the QF-VAR hyperparameters constant for all recursive sample.

The hyperparameter estimates for the MF-VAR and the QF-VAR are summarized in Table 5.2. While the overall tightness of the prior, controlled by λ_1, is larger for the QF-VAR than the MF-VAR, the MF-VAR strongly shrinks the coefficients on higher-order lags to zero. The QF-VAR only uses two lags which are associated with 22 regression coefficients for each endogenous variable. The MF-VAR, on the other hand, uses six lags which are associated with 66 regression coefficients. Roughly 30% of these coefficients are associated with regressors that are only observed a quarterly frequency. The hyperparameters for the QF-VAR are broadly in line with the results in Giannone, Lenza, and Primiceri (2012).

5.4.2 MF-VAR Point Forecasts

MF-VAR versus QF-VAR. We begin by comparing RMSEs for MF-VAR and QF-VAR forecasts of quarterly averages to assess the usefulness of monthly information. The RMSEs are computed separately for the "+0 months," "+1 month," and "+2 months" forecast

Table 5.2: Hyperparameters

	λ_1	λ_2	λ_3	λ_4	λ_5
MF-VAR(11)	0.09	4.30	1.0	2.70	4.30
QF-VAR(11)	3.08	0.01	1.0	1.12	1.62

origins defined in the previous section. Results for GDP growth (GDP), unemployment (UNR), inflation (INF), and the federal funds rate (FF) are reported in Figure 5.2. The figure depicts relative RMSEs defined as

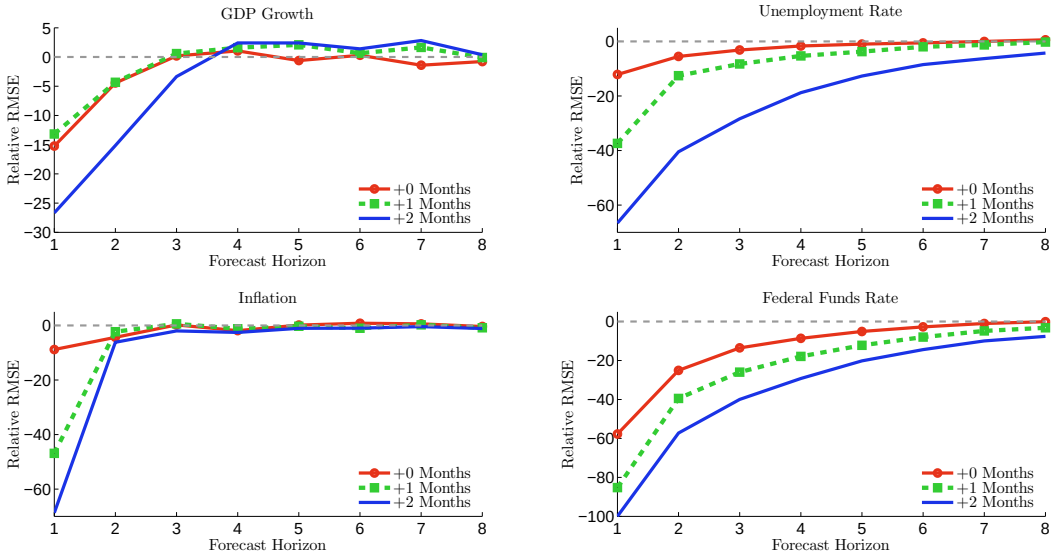
$$\text{Relative RMSE}(i|h) = 100 \times \frac{RMSE(i|h) - RMSE_{Benchmark}(i|h)}{RMSE_{Benchmark}(i|h)}, \tag{5.19}$$

where i denotes the variable and we adopt the convention (in slight abuse of notation) that the forecast horizon h is measured in quarters. The QF-VAR serves as a benchmark model and $h = 1$ corresponds to the quarter in which the forecast is generated. The $h = 1$ forecast is often called a nowcast. Absolute RMSEs for the 11-variable MF-VAR are tabulated in the Online Appendix.

For all four series, the use of monthly information via the MF-VAR leads to a substantial RMSE reduction in the short run. Consider the GDP growth forecasts. The "+2" nowcasts have a 27% lower RMSE than the QF-VAR nowcasts. For the "+1 month" group and the "+0 months" group, the reductions are both 15%. While the "+2 months" group forecasts clearly dominate at the nowcast horizon $h = 1$, the relative ranking among the three sets of MF-VAR forecasts becomes ambiguous for $h \geq 2$. As the forecast horizon increases to $h = 4$, the QF-VAR catches up with the MF-VAR. For horizons $h \geq 4$, the RMSE differentials between QF-VAR and MF-VAR GDP growth forecasts are negligible.

For the monthly unemployment, inflation, and federal funds rate series, the short-run RMSE reductions attained by the MF-VAR for the monthly series are even stronger than for GDP growth, which is observed at quarterly frequency. This is, of course, not surprising.

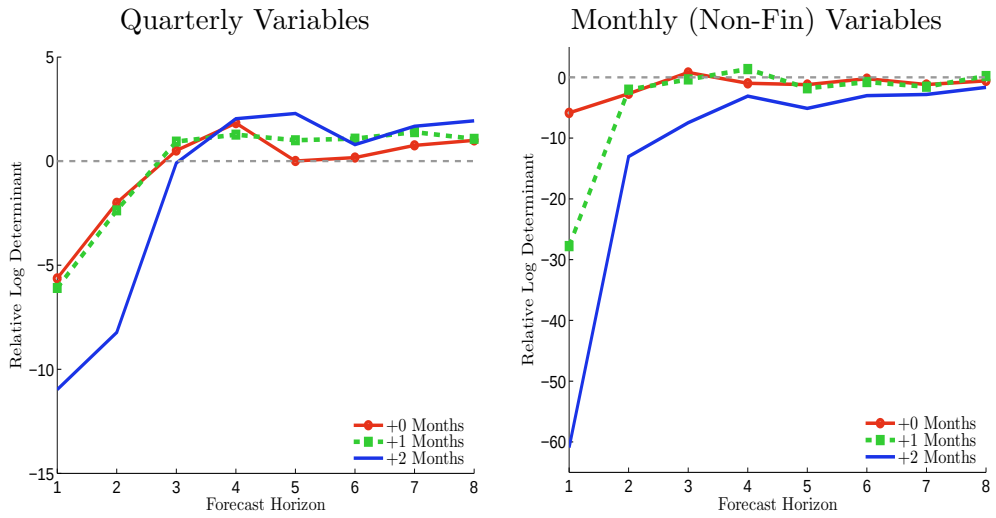
Figure 5.2: Relative RMSEs of 11-Variable MF-VAR versus QF-VAR



At the nowcast horizon, the MF-VAR is able to improve over the precision of the QF-VAR for the "+2 months" forecasts by 65% for unemployment, 70% for inflation, and 100% for the federal funds rate. Recall that "+2 months" corresponds to the last month of the quarter, which means that at the end of the last month, the average quarterly interest rate is known. Thus, by construction the RMSE reduction for the federal funds rate is 100%. The RMSE reductions for the "+1 month" group range from 40% (unemployment) to 80% (federal funds rate). For the "+0 months" group the improvement of the nowcast from using the MF-VAR is about 10% for inflation and the unemployment rate and 60% for the federal funds rate. While the gains from using monthly information tend to persist for unemployment and interest rates as the forecast horizon h increases, for inflation, monthly observations generate no improvements of forecast performance beyond the nowcast horizon.

To summarize the multivariate forecast performance of the VARs and aggregate the univariate RMSE differentials across quarterly and monthly nonfinancial variables we consider the log-determinant of the forecast error covariance matrix, proposed by Doan, Litterman,

Figure 5.3: Log Determinant of MF-VAR versus QF-VAR



Notes: The relative log determinant is defined as Relative Log Determinant $= (100 \cdot 0.5/n_{var})[f(\hat{\epsilon}_{t,MF}) - f(\hat{\epsilon}_{t,QF})]$, where $f(\cdot)$ is given in (5.20) and $n_{var} = 3$ for quarterly variables and $n_{var} = 5$ for monthly nonfinancial variables. The forecast horizon h is measured in quarters and $h = 1$ corresponds to the quarter in which the forecast is generated.

and Sims (1984):

$$f(\hat{\epsilon}_t) = \ln(|\ \frac{1}{T_{max} - T_{min}} \sum_{t=T_{min}}^{T_{max}} \hat{\epsilon}_t \hat{\epsilon}_t'\ |), \tag{5.20}$$

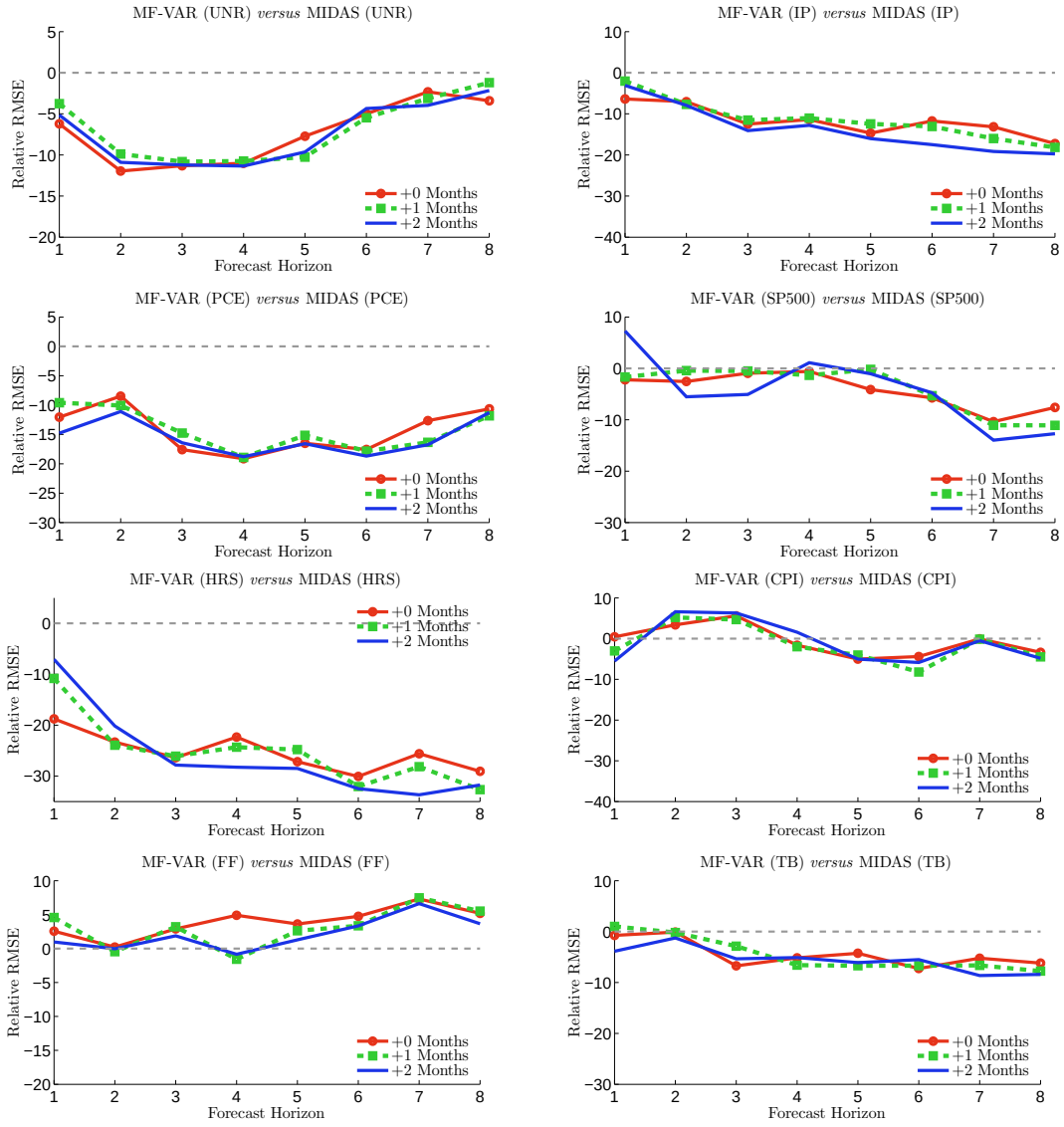
where $\hat{\epsilon}_t$ is a vector of forecast errors. Log-determinant differentials of MF-VAR versus QF-VAR forecasts are depicted in Figure 5.3. We scale the log-determinant differentials by $100 \cdot 0.5/n_{var}$. The factor 0.5 converts mean-squared errors into RMSEs, the division by n_{var} yields an average across the variables included in $\hat{\epsilon}_t$, and the factor 100 converts the differential into percentages. This scaling makes the log-determinant differentials comparable to the RMSE differentials depicted in Figure 5.2. The results are qualitatively consistent with the comparison of univariate RMSEs. Not surprisingly, for the group of quarterly variables (GDP, INVFIX, GOV) the gain from including within-quarter monthly information is smaller than for the group of monthly nonfinancial variables (UNR, HRS, CPI, IP, PCE). For quarterly variables the forecast accuracy gains relative to the QF-VAR

range from 11% ("+2 months" group) to 6% ("+0 months" group). For monthly variables the gains for the three forecast origin groups are 60%, 30% and 6% respectively. For $h \geq 3$ the QF-VAR catches up with the MF-VAR and the benefit from using monthly information vanishes. The only exception are the "+2" months forecasts of the monthly variables. Here the within-quarter monthly information remains even for forecast horizons exceeding one year. We exclude the financial variables (FF, TB, SP500) from the group of monthly variables because the financial variables are essentially known at the end of each quarter ("+2 months" group) which creates a near-singularity in forecast error covariance matrices that include one or more financial variables.

MF-VAR versus MIDAS. A popular alternative to the multivariate state-space framework used in this paper are MIDAS regressions. While there exist generalizations of the MIDAS approach to VAR settings, in most applications MIDAS regressions are used as univariate forecasting models. For a comparison of the two approaches we will focus on output growth. Our VAR models use 11 macroeconomic variables. If all of these variables are included in a MIDAS regression without any further restrictions, the MIDAS regression will perform very poorly. The distributed-lag restrictions on high-frequency regressors are designed to deal with many (high-frequency) observations of a single regressor but they are not designed to impose parsimony on a specification with many different right-hand-side variables. Thus, instead of comparing the 11-variable MF-VAR with MIDAS regressions, we will provide comparisons between bivariate MF-VARs and MIDAS regressions, estimated using the same set of variables.

Foroni, Marcellino, and Schumacher (2013) propose an unrestricted version of the MIDAS model (U-MIDAS) and show that when the mismatch of the frequency is low, like in macroeconomic applications that typically involve monthly and quarterly data only, this unrestricted version performs better in Monte Carlo experiments and provides a better

Figure 5.4: Relative RMSEs of Bivariate MF-VAR versus MIDAS



GDP nowcasting performance than a MIDAS regression with distributed-lag restrictions on the coefficients of the high-frequency variables. Thus, we consider U-MIDAS (instead of MIDAS) regressions in our comparison. The key aspect of our empirical analysis is the distinction between three groups of forecast origins, denoted by "+0," "+1," and "+2" (months). Each of these groups uses different within-quarter monthly information. Accord-

ingly, we use three separate U-MIDAS regressions, which, using the notation of Section 5.2, can be written as

$$\text{“+0”} \quad : \quad \tilde{y}_{q,t+3h} = \beta_0 + \beta_1 \tilde{y}_{q,t} + \beta_2 \tilde{y}_{q,t-3} + \sum_{s=1}^{6} \gamma_s x_{m,t-s+1} + resid_{t+3h} \tag{5.21}$$

$$\text{“+1”} \quad : \quad \tilde{y}_{q,t+3h} = \beta_0 + \beta_1 \tilde{y}_{q,t} + \beta_2 \tilde{y}_{q,t-3} + \sum_{s=1}^{6} \gamma_s x_{m,t-s+1} + \delta_1 x_{m,t+1} + resid_{t+3h}$$

$$\text{“+2”} \quad : \quad \tilde{y}_{q,t+3h} = \beta_0 + \beta_1 \tilde{y}_{q,t} + \beta_2 \tilde{y}_{q,t-3} + \sum_{s=1}^{6} \gamma_s x_{m,t-s+1} + \delta_1 x_{m,t+1} + \delta_2 x_{m,t+2} + resid_{t+3h},$$

where $t = 0, 3, 6, 9, \ldots$. The quarterly variable $\tilde{y}_{q,t+3h}$ was defined as average of unobserved monthly variables in (5.4) and corresponds to log GDP. The monthly variable $x_{m,t}$ is assumed to be scalar and we consider all eight of our monthly variables individually. The regression (5.21) is estimated by OLS for each group of forecast origins and for each forecast horizon separately. Thus, as in Foroni, Marcellino, and Schumacher (2013) we use direct estimation, i.e., the projection of $\tilde{y}_{q,t+3h}$ on the predictors available at the forecast origin, to determine the coefficients for the multi-step forecasting equation. Recall that under the Bayesian approach employed for the analysis of the MF-VAR multi-step forecasts are generated by iterating the VAR forward and using the posterior distribution to integrate out the unknown parameters.

Figure 5.4 illustrates the log GDP forecast performance of the bivariate MF-VARs relative to the MIDAS regressions. Each panel corresponds to a different monthly variable. Two results stand out. First, by and large, both the MF-VAR and MIDAS utilize the within-quarter monthly information equally well. The RMSE differentials are essentially the same for each of the three informational groups. For six out of the eight monthly variables the MF-VAR forecasts are more accurate than the MIDAS forecasts at some horizons, and no worse at the other horizons. For the unemployment rate, the gain from using the MF-VAR is highest for horizons of 2-5 quarters. For industrial production, the stock market index, hours, and the treasury bond rate the largest gain is realized at the long-horizon whereas

for PCE the improvement if fairly uniform for one- to eight-quarter ahead forecasts. Only for the federal funds rate and CPI inflation MIDAS forecasts appear to be marginally more accurate than the MF-VAR forecasts.

Other Comparisons. In the Online Appendix we also provide RMSE comparisons between the 11-variable MF-VAR and univariate QF-AR(2) models; and between a 4-variable MF-VAR (GDP, CPI, UNR, FF) and a 4-variable QF-VAR. The results are qualitatively very similar: there is a substantial gain from using the within-quarter-monthly information for nowcasting and short-horizon forecasting. This gain vanishes over one- to two-year horizons. Finally, the Online Appendix contains a careful comparison between MF-VAR forecasts and Greenbook (now Tealbook) forecasts, prepared by the staff of the Board of Governors for the meetings of the Federal Open Market Committee. At the nowcast horizon the unemployment forecasts of the MF-VAR are at par with the Greenbook forecasts, whereas the GDP growth and inflation forecasts are less accurate than the Greenbook forecasts. Over a four- to five-quarter horizon the MF-VAR generates more accurate GDP forecasts, whereas the Greenbook contains more precise inflation and unemployment rate forecasts.

5.4.3 Forecasting During the Great Recession

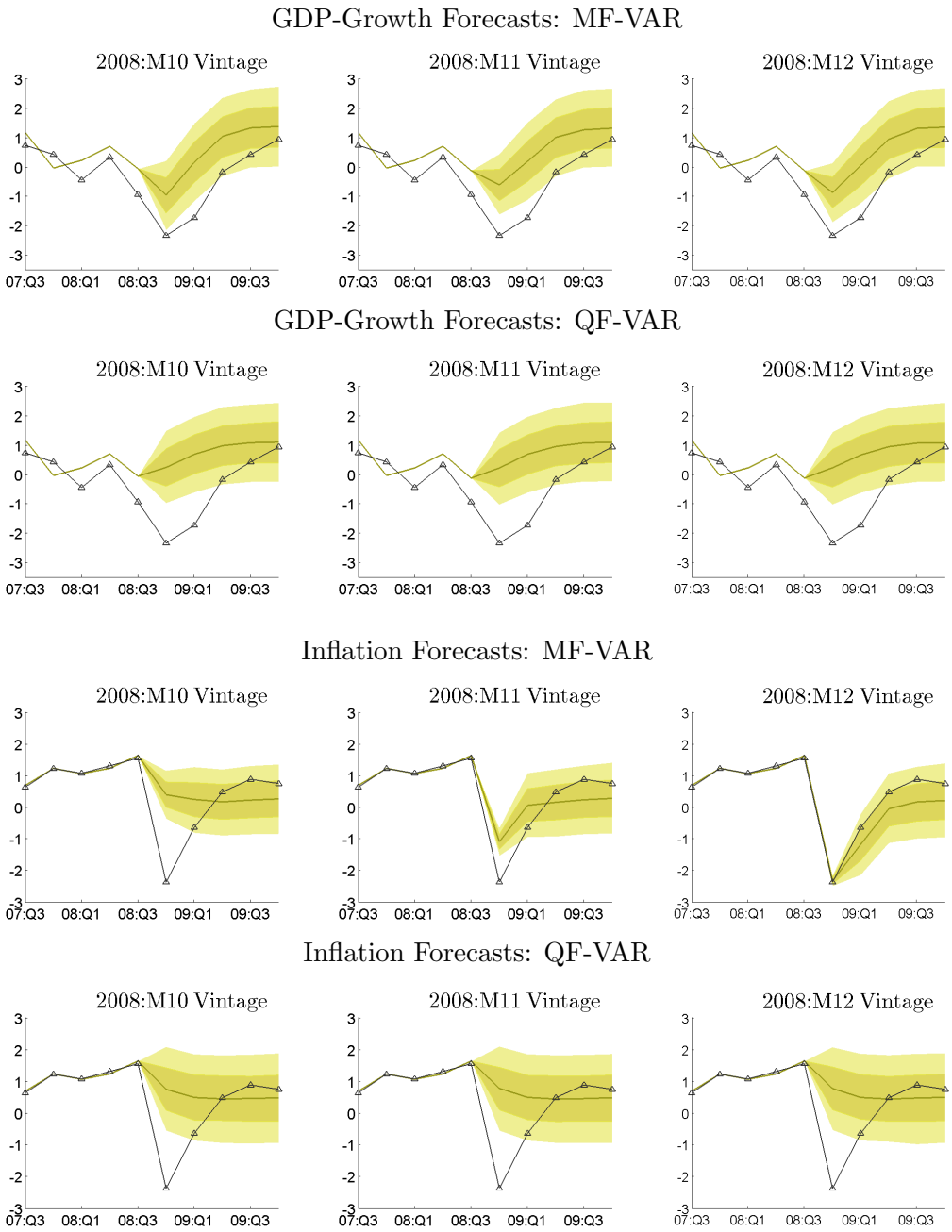
The pseudo-out-of-sample forecast performance of the previous section documented that the use of within-quarter monthly observations increases the precision of short-run forecast. We now examine how the use of monthly real-time information sharpened the VAR forecasts during the recent recession. We focus on the period from October to December 2008. Figure 5.5 depicts real-time interval forecasts from the MF-VAR and the QF-VAR. Moreover, we plot actual values using the 2011:M7 data vintage. We focus on real GDP growth and CPI inflation. The figure is divided into subpanels that correspond to particular estimation samples and forecast horizons. The first column of panels depicts October 2008 forecasts

("+0 months" group), and the second and third columns show November ("+1 month") and December ("+2 months") forecasts, respectively. A comparison between the first and second (third and forth) row of panels shows how monthly within-quarter information alters the density forecast for GDP (inflation).

The most striking feature of the top panels of Figure 5.5 is the -2% quarter-on-quarter growth rate of GDP in 2008:Q4. The magnitude of the drop in output growth in late 2008 is unexpected by the VAR models. It is, for all forecast origins, outside of the 90% predictive interval. The drop in GDP growth is equally unexpected by state-of-the-art dynamic stochastic general equilibrium (DSGE) models and the Blue Chip survey of professional forecasters as documented in Del Negro and Schorfheide (2013). A comparison of the MF-VAR and QF-VAR forecasts highlights how monthly information alters the within-quarter predictions. Notice that the QF-VAR forecasts do not stay constant within the quarter. The variation is caused by data revisions. As discussed in Section 5.3, each month new data releases for the previous quarter become available and change the lagged observations that determine the initial conditions for the VAR at the forecast origin. However, the within-quarter variation of the QF-VAR forecasts is fairly small.

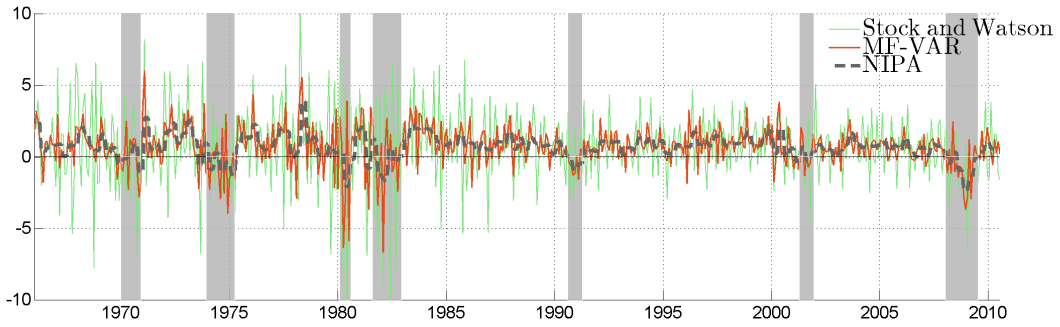
By December 2008 the QF-VAR nowcasts and forecasts show still no evidence of a severe downturn, because the latest information that is used to generate the predictions stems from 2008:Q3. The MF-VAR forecasts, on the other hand, do get revised more substantially during each quarter. In addition to the presence of data revisions, the forecasts are updated based on the information that is available at monthly frequency. Compared to the QF-VAR forecasts, the MF-VAR nowcasts during the fourth quarter of 2008 are a lot more pessimistic, which is in line with the actual realization of output growth. Over a one-year horizon the discrepancy between the MF-VAR and QF-VAR forecasts vanishes, which is consistent with the forecast error statistics presented in the previous section. According

Figure 5.5: Real-Time Forecasts During the Great Recession



Notes: Actual values are from the $T_* = 2012 : M1$ data vintage and are denoted as the black line with triangles. The title in each subplot indicates the forecast origin and the data vintage that are used in the estimation. We show the median, 60% bands, and 90% bands constructed from the predictive distribution.

Figure 5.6: Monthly GDP Growth (Scaled to a Quarterly Rate)



to both VARs the GDP growth forecasts are mean reverting. The models predict a GDP growth rate of about 1% for the second half of 2009. This prediction turned out to be accurate.

The bottom panels of Figure 5.5 depict the evolution of inflation forecasts in the last quarter of 2008. Since the CPI is published at a monthly frequency, the differences between within-quarter inflation forecasts from the MF-VAR and QF-VAR are much more pronounced than for GDP. Throughout 2008:Q4 the inflation forecasts from the QF-VAR stay essentially constant and miss the -2% deflation rate in 2008:Q4. The MF-VAR, on the other hand, detects the deflation by November 2008 as it is unfolding. At the longer horizon, the MF-VAR correctly predicts that the deflation episode is short-lived and that inflation rate will, with about 50% probability, be positive by the end of 2009. To summarize, these real-time forecasts during the Great Recession illustrate that the MF-VAR can transform within-quarter monthly information into more accurate nowcasts and forecasts of quarterly averages.

5.4.4 Monthly GDP

The estimation of the MF-VAR generates a monthly GDP series as a by-product. This series is implicitly extracted during the smoothing step of the Gibbs sampler (see Section 5.2.2)

from the eleven macroeconomic time series that enter the MF-VAR. A time series plot of monthly GDP growth is depicted in Figure 5.6. For each trajectory of log GDP generated with the Gibbs sampler, we compute month-on-month growth rates (scaled by a factor of 3 to make them comparable to quarter-on-quarter rates). For each month we then plot the median growth rate across the simulated trajectories. We overlay monthly GDP growth rates published by Stock and Watson (2010), who combine monthly information about GDP components to distribute quarterly GDP across the three months of the quarter.[1] Moreover, we plot growth rates computed from NIPA's quarterly GDP, implicitly assuming that GDP growth is constant within a quarter. Two observations stand out. First, at a monthly frequency GDP growth is much more volatile than at a quarterly level. Second, the monthly GDP growth series obtained from the MF-VAR estimation is somewhat smoother than the Stock-Watson series. While the two monthly measures are positively correlated, they are not perfectly synchronized, which is consistent with these measures being constructed from very different source data.

5.5 Conclusion

We have specified a VAR for observations that are observed at different frequencies, namely, monthly and quarterly. A Gibbs sampler was utilized to conduct Bayesian inference for model parameters and unobserved monthly variables. To cope with the dimensionality of the MF-VAR, we used a Minnesota prior that shrinks the VAR coefficients toward univariate random-walk representations. The degree of shrinkage is determined in a data-driven way, by maximizing the log MDD with respect to a low-dimensional vector of hyperparameters and we show how to approximate the MDD of a MF-VAR. Finally, we used the model to generate forecasts. The main finding is that within-quarter monthly information leads to

[1] Frale, Marcellino, Mazzi, and Proietti (2011) use a similar approach to construct a monthly GDP series for the Euro Area.

drastic improvements in the short-horizon forecasting performance. These improvements are increasing in the time that has passed since the beginning of the quarter. Over a one- to two-year horizon there are, however, no noticeable gains from using the monthly information.

5.6 Appendix

Section 5.6.1 of this appendix provides details of the implementation of the Bayesian computations for the MF-VAR presented in the main text. Section 5.6.2 discusses the construction of the real-time data set. Finally, Section 5.6.3 of this appendix provides tables and figures with additional empirical results. References to equations, tables, and figures without an A, B, or C prefix refer to equations, tables, and figures in the main text.

5.6.1 Implementation Details

Recall from the exposition in the main text (see equation (5.9)) that the Bayesian computations are implemented with a Gibbs sampler that iterates over the conditional distributions

$$p(\Phi, \Sigma | Z_{0:T}, Y_{-p+1:T}) \quad \text{and} \quad p(Z_{0:T} | \Phi, \Sigma, Y_{-p+1:T}).$$

Conditional on $Z_{0:T}$ the MF-VAR reduces to a standard linear Gaussian VAR with a conjugate prior. The reader is referred to Section 2 of the handbook chapter by Del Negro and Schorfheide (2011) for a detailed discussion of posterior inference for such a VAR.

We limit the exposition in this appendix to a brief presentation of the Minnesota prior and the hyperparameter selection (Section 5.6.1). The sampling from the conditional posterior of $Z_{0:T}|(\Phi, \Sigma, Y_{-p+1:T})$ is implemented with a standard simulation smoother, discussed in detail, for instance, in Carter and Kohn (1994b), the state-space model textbook of Durbin and Koopman (2001b), or the handbook chapter by Giordani, Pitt, and Kohn (2011). The only two aspects of our implementation that deserve further discussion are the initialization (Section 5.6.1) and the use of the more compact state-space representation for periods $t = 1, \ldots, T_b$ (Section 5.6.1).

Minnesota Prior and Its Hyperparameters

To simplify the exposition, suppose that $n = 2$ and $p = 2$. A transposed version of (5.1) can be written as

$$x_t' = [x_{t-1}', x_{t-2}', 1]'\Phi + u_t' = w_t'\Phi + u_t', \quad u_t \sim iidN(0, \Sigma). \tag{5.22}$$

We generate the Minnesota prior by dummy observations (x_*, w_*) that are indexed by a 5×1 vector of hyperparameters λ with elements λ_i. Using a pre-sample, let $\underline{x}$ and $\underline{s}$ be $n \times 1$ vectors of means and standard deviations. For time series that are observed at monthly frequency, the computation of pre-sample moments is straightforward. In order to obtain pre-sample means and standard deviations for those series that are observed at quarterly frequency, we simply equate $\underline{x_q}$ with the pre-sample mean of the observed quarterly values and set $\underline{s}$ equal to the pre-sample standard deviation of the observed quarterly series.

Dummy Observations for Φ_1.

$$\begin{bmatrix} \lambda_1 \underline{s}_1 & 0 \\ 0 & \lambda_1 \underline{s}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \underline{s}_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 \underline{s}_2 & 0 & 0 & 0 \end{bmatrix} \Phi + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}. \tag{5.23}$$

We can rewrite the first row of (5.23) as

$$\lambda_1 \underline{s}_1 = \lambda_1 \underline{s}_1 \phi_{11} + u_{11}, \quad 0 = \lambda_1 \underline{s}_1 \phi_{21} + u_{12}.$$

Since, according to (5.22) the u_t's are normally distributed, we can interpret the relationships as

$$\phi_{11} \sim \mathcal{N}(1, \Sigma_{11}/(\lambda_1^2 \underline{s}_1^2)), \quad \phi_{21} \sim \mathcal{N}(0, \Sigma_{22}/(\lambda_1^2, \underline{s}_1^2)).$$

where ϕ_{ij} denotes the element i, j of the matrix Φ, and Σ_{ij} corresponds to element i, j of Σ. The hyperparameter λ_1 controls the tightness of the prior.

Dummy Observations for Φ_2.

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \lambda_1 \underline{s}_1 2^{\lambda_2} & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 \underline{s}_2 2^{\lambda_2} & 0 \end{bmatrix} \Phi + U, \tag{5.24}$$

where the hyperparameter λ_2 is used to scale the prior standard deviations for coefficients associated with x_{t-l} according to $l^{-\lambda_2}$.

Dummy Observations for Σ. A prior for the covariance matrix Σ, centered at a matrix that is diagonal with elements equal to the pre-sample variance of x_t, is obtained by stacking the observations

$$\begin{bmatrix} \underline{s}_1 & 0 \\ 0 & \underline{s}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Phi + U \tag{5.25}$$

λ_3 times.

Sums-of-Coefficients Dummy Observations. When lagged values of a variable $x_{i,t}$ are at the level $\underline{x}_i$, the same value $\underline{x}_i$ is a priori likely to be a good forecast of $x_{i,t}$, regardless of the value of other variables:

$$\begin{bmatrix} \lambda_4\underline{x}_1 & 0 \\ 0 & \lambda_4\underline{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_4\underline{x}_1 & 0 & \lambda_4\underline{x}_1 & 0 & 0 \\ 0 & \lambda_4\underline{x}_2 & 0 & \lambda_4\underline{x}_2 & 0 \end{bmatrix} \Phi + U. \tag{5.26}$$

Co-persistence Dummy Observations. When all lagged x_t's are at the level $\underline{x}$, a priori x_t tends to persist at that level:

$$\begin{bmatrix} \lambda_5\underline{x}_1 & \lambda_5\underline{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_5\underline{x}_1 & \lambda_5\underline{x}_2 & \lambda_5\underline{x}_1 & \lambda_5\underline{x}_2 & \lambda_5 \end{bmatrix} \Phi + U. \tag{5.27}$$

Prior Distribution. After collecting the T^* dummy observations in matrices X^* and W^*, the likelihood function associated with (5.22) can be used to relate the dummy observations to the parameters Φ and Σ. If we combine the likelihood function with the improper prior $p(\Phi, \Sigma) \propto |\Sigma|^{-(n+1)/2}$, we can deduce that the product $p(X^*|\Phi, \Sigma) \cdot |\Sigma|^{-(n+1)/2}$ can be interpreted as

$$(\Phi, \Sigma) \sim MNIW(\underline{\Phi}, (W^{*'}W^*)^{-1}, \underline{S}, T^* - k), \tag{5.28}$$

where $\underline{\Phi}$ and $\underline{S}$ are

$$\underline{\Phi} = (W^{*\prime}W^*)^{-1}W^{*\prime}W^*, \quad \underline{S} = (X^* - W^*\underline{\Phi})'(X^* - W^*\underline{\Phi}).$$

Provided that $T^* > k + n$ and $W^{*'}W^*$ is invertible, the prior distribution is proper.

Hyperparameter Grid Search for MF-VAR: For the first recursive sample the grid search proceeds in three steps. Define:

$$\begin{aligned}
\Lambda_1^{(1)} &= \{0.01, 1.12, 2.23, 3.34, 4.45, 5.56, 6.67, 7.78, 8.89, 10\} \\
\Lambda_2^{(1)} &= \{0.01, 1.12, 2.23, 3.34, 4.45, 5.56, 6.67, 7.78, 8.89, 10\} \\
\Lambda_3^{(1)} &= \{1\} \\
\Lambda_4^{(1)} &= \{2.23, 2, 7, 3.34, 4.3, 4.45, 5.56\} \\
\Lambda_5^{(1)} &= \{2.23, 2, 7, 3.34, 4.3, 4.45, 5.56\}
\end{aligned}$$

The first grid is given by

$$\Lambda^{(1)} = \Lambda_1^{(1)} \otimes \Lambda_2^{(1)} \otimes \Lambda_3^{(1)} \otimes \Lambda_4^{(1)} \otimes \Lambda_5^{(1)},$$

where $\otimes$ denote the Cartesian product. Thus, we are fixing $\lambda_3 = 1$ throughout. We maximize $\ln \hat{p}(Y_{1:T}|Y_{-p+1:0,\lambda})$ with respect to $\lambda \in \Lambda^{(1)}$. By construction $\hat{\lambda}_3 = 1$. We retain the argmax values $\hat{\lambda}_4 = 2.7$ and $\hat{\lambda}_5 = 4.3$.

In the second step we refine the grids for λ_1 and λ_2 as follows:

$$\begin{aligned}
\Lambda_1^{(2)} &= \{0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.13, 0.15\} \\
\Lambda_2^{(2)} &= \{0.8, 1.3, 2.1, 2.8, 3.5, 4.3, 4.8, 5.2\}.
\end{aligned}$$

Maximization of the MDD with respect to $\Lambda^{(2)} = \Lambda_1^{(2)} \otimes \Lambda_2^{(2)} \otimes \{1.0\} \otimes \{2.70\} \otimes \{4.30\}$ yields $\hat{\lambda}$ for the first recursive sample.

In the third step we reoptimize the choice of λ_1 and λ_2 for recursive samples 40, 75, 110, and 151. In this step we use the following grids for λ_1 and λ_2:

$$\begin{aligned}
\Lambda_1^{(3)} &= \{0.05, 0.07, 0.09, 0.11, 0.13\} \\
\Lambda_2^{(3)} &= \{2.1, 2.8, 3.5, 4.3, 4.8\}.
\end{aligned}$$

Hyperparameter Grid Search for QF-VAR: For the QF-VAR we are also fixing $\lambda_3 = 1$. The grids for λ_1 and λ_2 are given by the 40 equally-spaced points on the interval $[0.01, 10]$. The grids for λ_4 and λ_5 are given by the 40 equally-spaced points on the interval $[0.1, 10]$.

Initial Distribution $p(z_0|Y_{-p+1:0})$

Recall that $t = 1$ corresponds to 1968:M1. Let $T_- = -11$ such that $t = T_-$ corresponds to 1967:M1. We then initialize z_{T_-} using actual observations. This is straightforward for x_{m,T_-}, x_{m,T_--1}, x_{m,T_--p} because they are observed. We set x_{q,T_-}, x_{q,T_--1}, x_{q,T_--p} equal to the observed quarterly values, assuming that during these periods the monthly within-quarter values simply equal the observed averages during the quarter. This provides us with a distribution for $p(z_{T_-})$ that is simply a point mass. We then set Φ and Σ equal to their respective prior means and apply the Kalman filter for $t = T_- + 1, \ldots, 0$ to the state-space system described in (5.2) and (5.7), updating the beliefs about the latent state z_t with pre-sample observations $Y_{T_-:0}$. In slight abuse of notation, we denote the distribution of z_t obtained after the period 0 updating by $p(z_0|Y_{-p+1})$. Note that this distribution does not depend on the "unknown" parameters Φ and Σ, because the Kalman filter iterations were implemented based on the prior means of these matrices.

Compact State-Space Representation

As discussed in the main text, the computational efficiency of the simulation-smoother step in the Gibbs sampler can be improved by eliminating, for $t = 1, \ldots, T_b$, the monthly observations $x_{m,t}$ from the state vector z_t that appears in the measurement equation (5.7). We begin by re-ordering the lags of x_t and the VAR coefficients in (5.1) to separate lags of $x_{m,t}$ from lags of $x_{q,t}$. Define the $pn_m \times 1$ vector $z_{m,t}$ and $pn_q \times 1$ vector $z_{q,t}$ as

$$z'_{m,t} = \left[x'_{m,t}, \ldots, x'_{m,t-p+1}\right], \quad z'_{q,t} = \left[x'_{q,t}, \ldots, x'_{q,t-p+1}\right].$$

In a similar manner, define the $n_m \times pn_m$ matrix Φ_{mm}, the $n_m \times pn_q$ matrix Φ_{mq}, the $n_q \times pn_m$ matrix Φ_{qm}, and the $n_q \times pn_q$ matrix Φ_{qq} such that (5.1) can be rewritten as

$$\begin{bmatrix} x_{m,t} \\ x_{q,t} \end{bmatrix} = \begin{bmatrix} \Phi_{mm} & \Phi_{mq} \\ \Phi_{qm} & \Phi_{qq} \end{bmatrix} \begin{bmatrix} z_{m,t-1} \\ z_{q,t-1} \end{bmatrix} + \begin{bmatrix} \Phi_{mc} \\ \Phi_{qc} \end{bmatrix} + \begin{bmatrix} u_{m,t} \\ u_{q,t} \end{bmatrix}. \tag{5.29}$$

Recall that for $t \leq T_b$, all the monthly series are observed. Thus, $y_{m,t} = x_{m,t}$ and, in slight abuse of notation, $z_{m,t-1} = y_{m,t-p:t-1}$. Now define $s_t = [x'_{q,t}, z'_{q,t-1}]'$ and notice that based on the second equation in (5.29), one can define matrices Γ_s, Γ_{zm}, Γ_c, and Γ_u such that we obtain a state-transition equation in companion form

$$s_t = \Gamma_s s_{t-1} + \Gamma_{zm} y_{m,t-p:t-1} + \Gamma_c + \Gamma_u u_{q,t}. \tag{5.30}$$

The measurement equation for the monthly series takes the form

$$y_{m,t} = \Lambda_{ms} s_t + \Phi_{mm} y_{m,t-p:t-1} + \Phi_{mc} + u_{m,t}. \tag{5.31}$$

Finally, the measurement equation for the quarterly series can be expressed as

$$y_{q,t} = M_{q,t} \Lambda_{qs} s_t, \tag{5.32}$$

where the matrix $\Lambda_{qs} s_t$ averages $x_{q,t}$, $x_{q,t-1}$, and $x_{q,t-2}$ and $M_{q,t}$ is a time-varying selection matrix that selects the elements of $\Lambda_{qs} s_t$ that are observed in period t. In sum, (5.30), (5.31), and (5.32) provide an alternative state-space representation of the MF-VAR that reduces the dimension of the state vector from np to $n_q(p+1)$. In this alternative representation, the "measurement errors" $u_{m,t}$ in (5.31) are correlated with the innovations $u_{q,t}$ in the state-transition equation (5.30). Moreover, the lagged observables $y_{m,t-p:t-1}$ directly enter the state-transition and measurement equations. Since these observables are part of the $t-1$ information, the modification of the Kalman filter and simulation smoother is straightforward.

At the end of period $t = T_b$, we switch from the state-space representation in terms of $s_t = [x'_{q,t}, \ldots, x'_{q,t-p}]'$ to a state-space representation in terms of $\tilde{z}_t = [z'_t, x'_{t-p}] =$

$[x_t', \ldots, x_{t-p}']'$.[2] In the forward pass of the Kalman filter, let $\hat{s}_{t|t} = \mathbb{E}[s_t|Y_{-p+1:t}]$ and $P^s_{t|t} = \mathbb{V}[s_t|Y_{-p+1:t}]$ (omitting (Φ, Σ) from the conditioning set). Since $x_{m,t}, \ldots, x_{m,t-p+1}$ is known conditional on the $Y_{-p+1:t}$, we can easily obtain $\hat{\tilde{z}}_{t|t} = \mathbb{E}[\tilde{z}_t|Y_{-p+1:t}]$ by augmenting $\hat{s}_{t|t}$ with $y_{m,t}, \ldots, y_{m,t-p}$. Moreover, $P^{\tilde{z}}_{t|t} = \mathbb{V}[\tilde{z}_t|Y_{-p+1:t}]$ can be obtained by augmenting $P^s_{t|t}$ by zeros, to reflect that $x_{m,t}, \ldots, x_{m,t-p}$ are known with certainty. In the backward pass of the simulation smoother we start out with a sequence of draws from $\tilde{z}_T|Y_{-p+1:T}$ and $\tilde{z}_t|(\tilde{Z}_{t+1:T}, Y_{-p+1:T})$ for $t = T-1, \ldots, T_b+1$. Let $\hat{\tilde{z}}_{t|T}$ and $P^{\tilde{z}}_{t|T}$ denote the mean and variance associated with this distribution. At $t = T_b$ we convert the conditional mean and variance of $\tilde{z}_{T_b}$ into a conditional mean and variance for s_{T_b}. This is done by eliminating all elements associated with $x_{m,t}, \ldots, x_{m,t-p}$.

5.6.2 Construction of Real-Time Data Set

The eleven real-time macroeconomic data series are obtained from the ALFRED database maintained by the Federal Reserve Bank of St. Louis. Table 5.6.2-1 summarizes how the series used in this paper are linked to the series provided by ALFRED.

We construct two sequences of dates that contain the set of forecast origins $(T_{min}, \ldots, T_{max})$. One sequence contains the last day of each month, and the other sequence will comprise the Greenbook forecast dates. ALFRED provides a publication date for each data vintage. We wrote a computer program that selects for every forecast origin, the most recent ALFRED vintage for each of the eleven variables and combines the series into a single data set. This leaves us with a real-time data set for each forecast origin. Based on the missing values in each real-time data set, we construct the selection matrices M_t, $t = T_b+1, \ldots, T$, that appear in (5.7). The patterns of missing values are summarized in Tables 5.1 and 5.4. Greenbook forecasts are also obtained from the ALFRED database.

[2] We augment the state vector z_t in (5.2) and (5.7) by an additional lag of x_t to ensure that s_t is a subvector of the resulting $\tilde{z}_t$. This augmentation requires a straightforward modification of the state-transition equation (5.2) and the measurement equations (5.7).

Table 5.3: ALFRED Series Used in Analysis

Time Series	ALFRED Name
Gross Domestic Product (GDP)	GDPC1
Fixed Investment (INVFIX)	FPIC1
Government Expenditures (GOV)	GCEC1
Unemployment Rate (UNR)	UNRATE
Hours Worked (HRS)	AWHI
Consumer Price Index (CPI)	CPIAUCSL
Industrial Production Index (IP)	INDPRO
Personal Consumption Expenditure (PCE)	PCEC96
Federal Fund Rate (FF)	FEDFUNDS
Treasury Bond Yield (TB)	GS10
S&P 500 (SP500)	SP500

Some of the vintages of PCE and INVFIX extracted from ALFRED were incomplete. The recent vintages of PCE and INVFIX from ALFRED do not include data prior to 1990 or 1995 (depending on the vintages). However, the most recent data for PCE and INVFIX can be obtained from BEA or NIPA, say, from 1/1/1967 to 1/1/2012. Let us consider PCE for illustration. For the vintages between 12/10/2003 and 6/25/2009, data start from 1/1/1990, and for the vintages between 7/31/2009 and the present, data start from 1/1/1995. First, we compute the growth rates from the most recent data. Based on the computed growth rates, we can backcast historical series up to 1/1/1967 using the 1/1/1990 (1/1/1995) data points as initializations. We think this is a reasonable way to construct the missing points. We eliminated 4 of the 151 samples (28, 29, 33, 145) because the vintages for PCE and INVFIX were incomplete. In principle, we could backcast as for the other vintages, but we took a shortcut.

Table 5.4 lists exceptions for the classification of information sets for specific forecast origins.

Table 5.4: Illustration of Information Sets: Exceptions

Exceptions E_0: January (+0 Months)

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M10	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q4	M11	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q4	M12	X	X	X	X	∅	X	X	X	QAv	QAv	QAv
Q1	M1	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

Exceptions E_1: February (+1 Month)

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M11	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q4	M12	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q1	M1	X	X	X	X	∅	X	X	X	∅	∅	∅
Q1	M2	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

Exceptions E_2: March (+2 Months)

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M12	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q1	M1	X	X	X	X	X	X	X	X	∅	∅	∅
Q1	M2	X	X	X	X	∅	X	X	X	∅	∅	∅
Q1	M3	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

Notes: ∅ indicates that the variable is missing. X denotes monthly observation and QAv denotes quarterly average. "+0 months" group: January, April, July, October; "+1 month" group: February, May, August, November; "+2 month" group: March, June, September, December. The table illustrates exceptions that arise due to an occasional two-month publication lag for PCE. Exception E_0 occurs for 28 out of 151 recursive samples (1, 4, 7, 10, 13, 16, 19, 22, 28, 37, 43, 52, 61, 64, 73, 79, 85, 88, 97, 106, 109, 115, 124, 130, 133, 139, 145, 151). Exception E_1 occurs for 14 out of 151 recursive samples (8, 20, 44, 53, 56, 68, 80, 89, 98, 101, 104, 116, 119, 140). Exception E_2 occurs for 5 out of 151 recursive samples (21, 27, 48, 51, 78).

5.6.3 Additional Empirical Results

11-Variable VAR, End-of-Month Forecasts

Table 5.5 provides numerical values for the RMSEs attained by the eleven-variable MF-VAR.

Figure 5.7 compares the 11-variable MF-VAR forecasts to quarterly-frequency AR(2) forecasts.

Figure 5.8 depicts recursive means of $h = 1$ and $h = 8$ step-ahead mean forecasts (setting future shocks equal to zero). Each hairline corresponds to a separate run of our MCMC algorithm. In each run, we generate 20,000 draws and discard the first 10,000 draws. We plot Monte Carlo averages based on the subsequent 500, 1,000, 1,500, ..., 10,000 draws. The units on the y-axis are percentages. With the exception of the eight-quarter-ahead federal funds rate forecast, the Monte Carlo variation is below one basis point and negligible compare to the overall forecast error.

Table 5.5: RMSEs for 11-Variable MF-VAR

Horizon	UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
					+0 Months						
1	0.21	0.50	0.57	0.99	0.55	0.21	0.17	3.05	0.57	1.73	0.80
2	0.47	0.80	0.61	1.45	0.69	0.69	0.43	7.93	0.79	2.45	0.75
3	0.80	0.98	0.64	1.71	0.74	1.10	0.61	8.04	0.86	2.86	0.74
4	1.12	1.01	0.62	1.72	0.72	1.45	0.70	7.97	0.88	2.85	0.77
5	1.40	0.96	0.64	1.66	0.69	1.78	0.79	7.72	0.86	2.77	0.78
6	1.63	0.91	0.64	1.59	0.68	2.08	0.86	7.72	0.83	2.66	0.74
7	1.84	0.87	0.63	1.56	0.65	2.31	0.89	7.98	0.79	2.54	0.69
8	2.00	0.85	0.63	1.54	0.64	2.50	0.94	7.82	0.79	2.59	0.79
					+1 Month						
1	0.15	0.39	0.33	0.98	0.49	0.07	0.08	1.24	0.57	1.53	0.81
2	0.44	0.79	0.62	1.44	0.71	0.55	0.34	7.97	0.79	2.39	0.75
3	0.75	0.97	0.64	1.71	0.75	0.93	0.53	7.98	0.86	2.86	0.75
4	1.07	1.01	0.62	1.72	0.73	1.29	0.63	7.98	0.88	2.87	0.75
5	1.36	0.98	0.63	1.69	0.70	1.64	0.74	7.77	0.86	2.79	0.72
6	1.61	0.93	0.62	1.61	0.67	1.95	0.80	7.75	0.83	2.70	0.75
7	1.81	0.88	0.64	1.59	0.66	2.20	0.82	7.84	0.82	2.59	0.74
8	1.98	0.86	0.63	1.56	0.66	2.40	0.86	7.84	0.80	2.59	0.77
					+2 Months						
1	0.08	0.30	0.20	0.73	0.38	0.00	0.00	0.00	0.50	1.41	0.81
2	0.30	0.60	0.60	1.15	0.67	0.39	0.38	7.05	0.68	2.06	0.77
3	0.59	0.90	0.62	1.63	0.75	0.76	0.61	8.02	0.84	2.77	0.76
4	0.92	1.01	0.62	1.72	0.74	1.12	0.72	7.84	0.89	2.89	0.76
5	1.23	0.99	0.62	1.67	0.70	1.50	0.81	7.79	0.86	2.82	0.71
6	1.50	0.93	0.63	1.60	0.68	1.81	0.90	7.84	0.85	2.72	0.74
7	1.72	0.89	0.64	1.58	0.66	2.08	0.87	7.74	0.82	2.63	0.77
8	1.90	0.86	0.62	1.58	0.65	2.27	0.87	7.94	0.81	2.57	0.76
					All Forecasts						
1	0.16	0.40	0.40	0.91	0.48	0.13	0.11	1.90	0.55	1.56	0.81
2	0.41	0.74	0.61	1.35	0.69	0.56	0.39	7.67	0.75	2.31	0.76
3	0.72	0.95	0.63	1.68	0.75	0.94	0.59	8.01	0.85	2.83	0.75
4	1.04	1.01	0.62	1.72	0.73	1.30	0.68	7.93	0.88	2.87	0.76
5	1.33	0.98	0.63	1.67	0.70	1.65	0.78	7.76	0.86	2.79	0.74
6	1.58	0.92	0.63	1.60	0.68	1.95	0.86	7.77	0.84	2.69	0.74
7	1.79	0.88	0.64	1.57	0.66	2.20	0.86	7.85	0.81	2.59	0.73
8	1.96	0.86	0.63	1.56	0.65	2.39	0.89	7.87	0.80	2.58	0.77

Notes: RMSEs for UNR (%), FF (annualized %), and TB (annualized %) refer to forecasts of levels. The remaining RMSEs refer to forecasts of quarter-on-quarter growth rates in percentages.

Figure 5.7: Relative RMSEs of MF-VAR versus QF-AR2

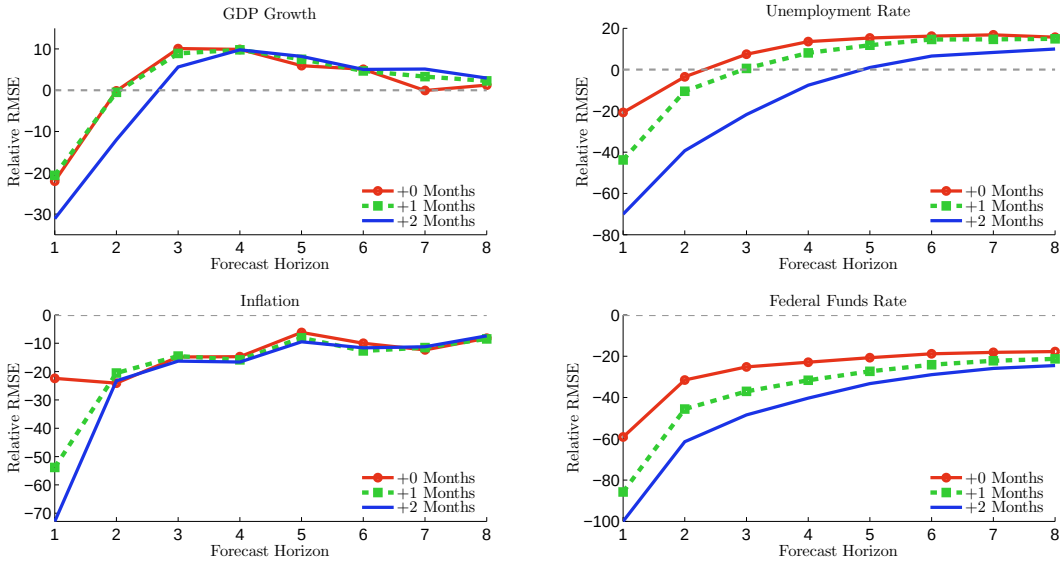
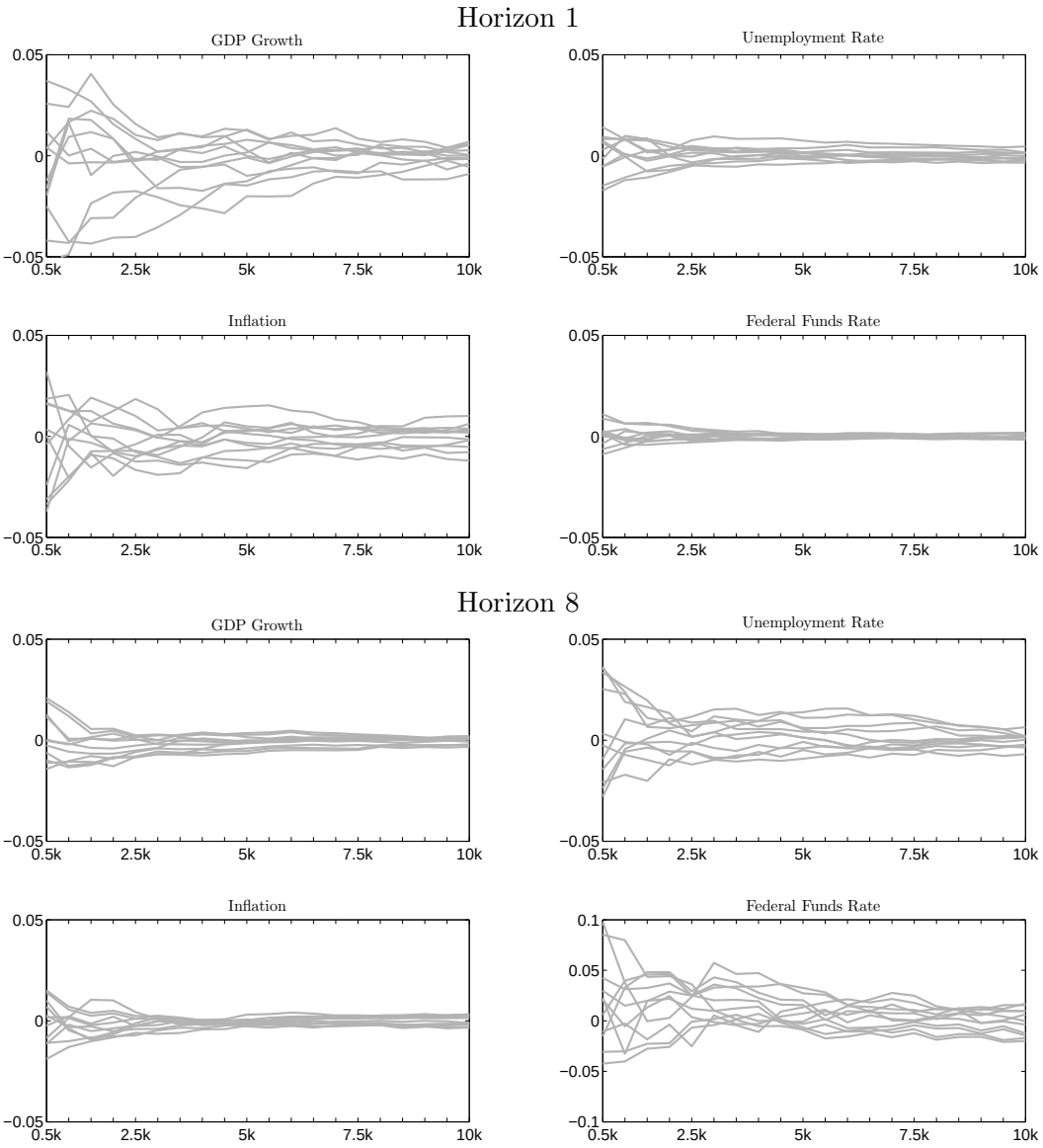
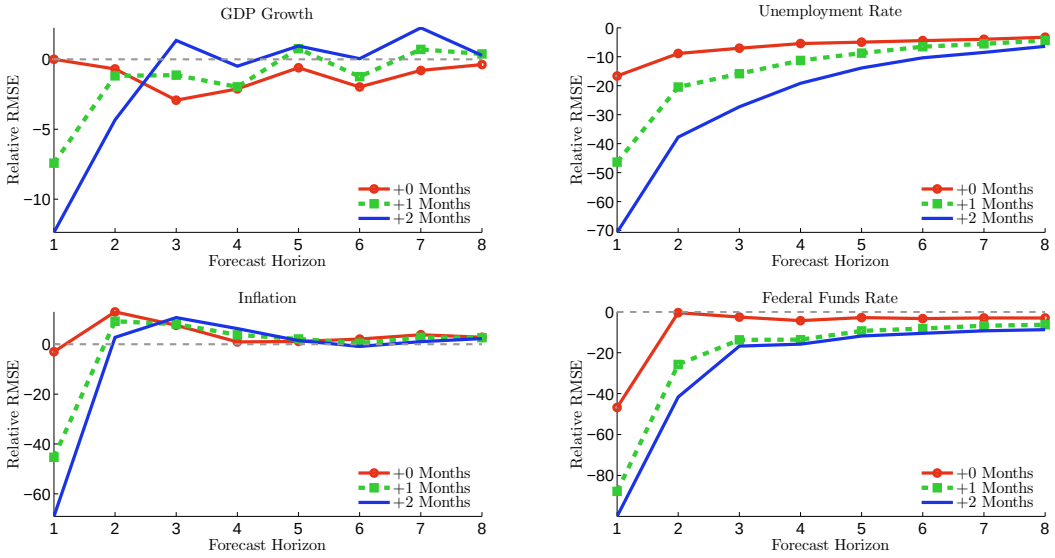


Figure 5.8: Convergence



Notes: The figure depicts recursive means of $h = 1$ and $h = 8$ step-ahead mean forecasts (setting future shocks equal to zero). Each hairline corresponds to a separate run of our MCMC algorithm. In each run, we generate 20,000 draws and discard the first 10,000 draws. We plot Monte Carlo averages based on the subsequent 500, 1,000, 1,500, ..., 10,000 draws. The units on the y-axis are percentages.

Figure 5.9: Relative RMSEs of 4-Variable MF-VAR versus QF-VAR



RMSEs for 4-Variable MF VAR

We also consider a four-variable MF-VAR based on one quarterly series and three monthly series. The three monthly series are the Consumer Price Index (CPI), Unemployment Rate (UNR), and Federal Funds Rate (FF). The quarterly series is Real GDP. Real GDP and CPI enter the MF-VAR in log levels, whereas UNR and FF are simply divided by 100 to make their scale comparable to the scale of the two other variables. As for the eleven-variable VAR, the number of lags is set to six.

Figure 5.9 reports RMSE ratios for the four-variable MF-VAR versus a four-variable QF-VAR using the end-of-month sample. The results are qualitatively similar to the ones reported in Figure 5.2. In general, the within-quarter monthly information of the MF-VAR increases the forecast accuracy compared to the QF-VAR. However, for GDP growth and the federal funds rate, these improvements are not as long-lived as in the eleven-variable setting.

11-Variable MF-VAR End-of-Month Density Forecasts

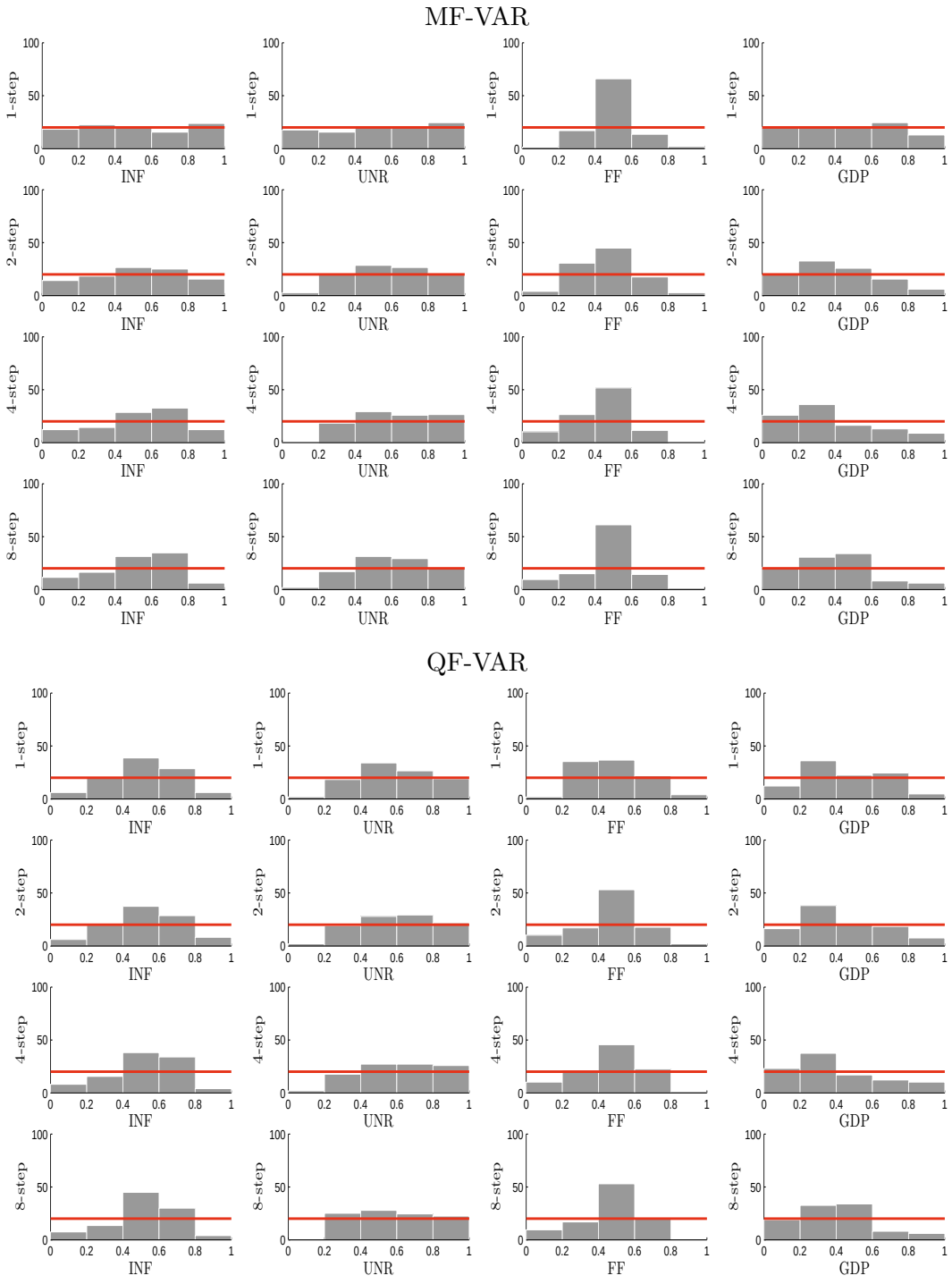
The MF-VAR generates an entire predictive distribution for the future trajectories of the eleven macroeconomic variables. While, strictly speaking, predictive distributions in a Bayesian framework are subjective, it is desirable that predicted probabilities are consistent with observed frequencies if the forecast procedure is applied in a sequential setting. To assess the MF-VAR density forecasts, we construct probability integral transformations (PITs) from (univariate) marginal predictive densities. The probability integral transformation of an h-step ahead forecast of $y_{i,t+h}$ based on time t information is defined as

$$z_{i,h,t} = \int_{-\infty}^{y_{i,t+h}} p(\widetilde{y}_{i,t+h}|Y_{1:t})d\widetilde{y}_{i,t+h}. \tag{5.33}$$

PITs, sometimes known as generalized residuals, are relatively easy to compute and facilitate comparisons among elements of a sequence of predictive distributions, each of which is distinct in that it conditions on the information available at the time of the prediction. For $h = 1$ the $z_{i,h,t}$'s are independent across time and uniformly distributed: $z_{i,h,t} \sim iid\mathcal{U}[0,1]$. For $h > 1$ the PITs remain uniformly distributed but are no longer independently distributed.

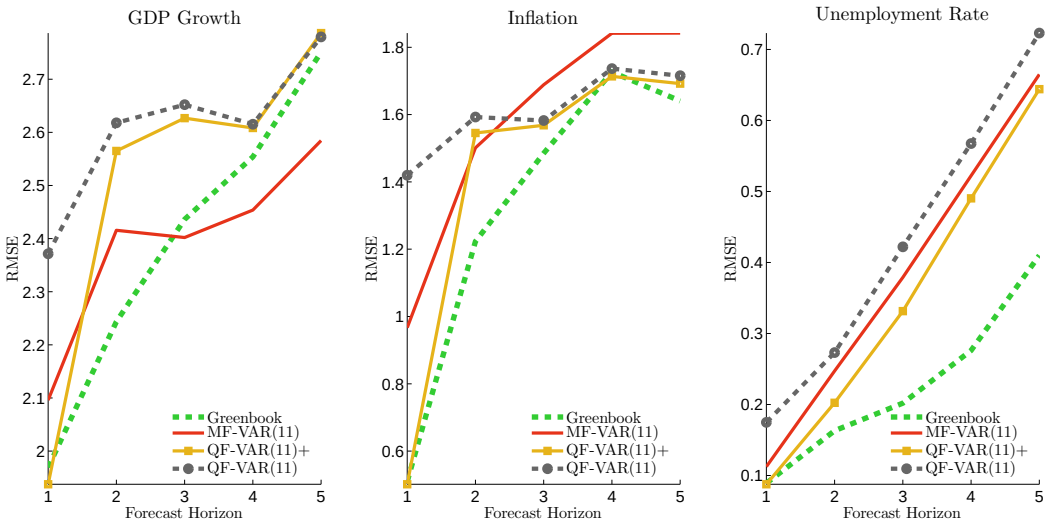
Figure 5.10 displays histograms for the PITs based on density forecasts from the MF-VAR and the QF-VAR using the end-of-month sample. The PITs are computed from the empirical distribution of the simulated trajectories $Y_{T+1:T+H}$. To generate the histogram plots, the unit interval is divided into $J = 5$ equally sized subintervals, and we depict the fraction of PITs (measured in percent) that fall in each bin. Since, under the predictive distribution, the PITs are uniformly distributed on the unit interval, we also plot the 20% line. For $h = 1$ (nowcast) and $h = 2$ (forecast for next quarter) the frequency of MF-VAR PITs falling in each of the five bins is close to 20% for inflation, unemployment, and output growth, indicating that the predictive densities are well calibrated. The federal funds rate density forecasts, on the other hand, appear to be too diffuse, because of the small number of PITs falling into the 0-0.2 and 0.8-1 bins. Over longer horizons, specifically for $h = 4$ and $h = 8$, the deviations from uniformity become more pronounced for all the series. The federal funds rate density forecasts remain too diffuse, and the MF-VAR tends to overpredict GDP growth and underpredict unemployment. For the QF-VAR the deviations from uniformity generally tend to be larger than for the MF-VAR forecasts.

Figure 5.10: PIT Histograms for 11-Variable VARs



Notes: Probability integral transformations for forecasts of inflation (INF), unemployment rate (UNR), federal funds rate (FF), and GDP growth (GDP). The bars represent the frequency of PITs falling in each bin. The solid line marks 20%.

Figure 5.11: RMSEs of 11-Variable MF-VAR versus Greenbook



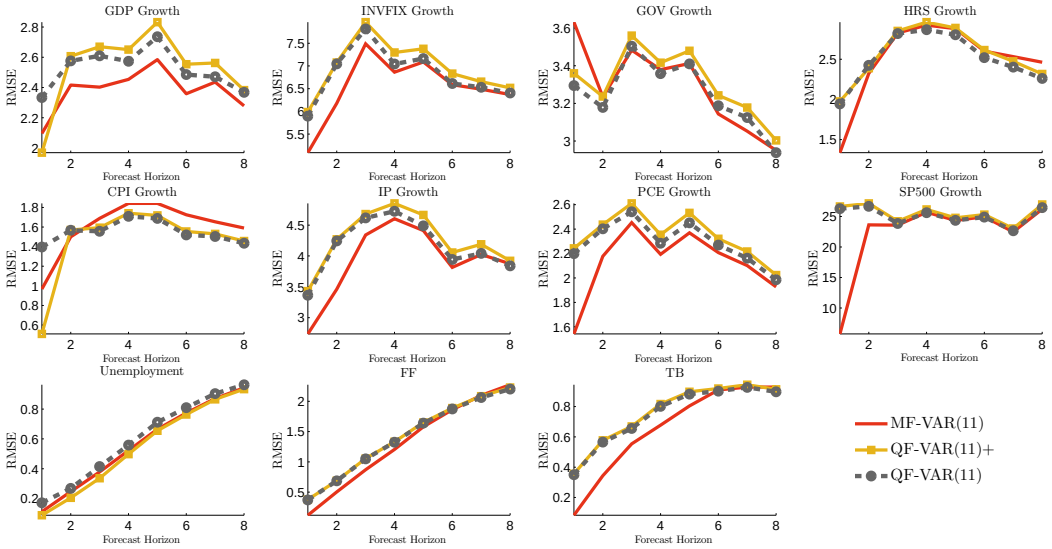
Notes: 22^{nd} and 38^{th} samples are eliminated because the vintages for PCE were incomplete.

11-Variable MF-VAR Forecasts: Comparison with Greenbook Forecasts

Data Set. We compare the MF-VAR forecasts to Greenbook forecasts, prepared by the staff of the Board of Governors for the FOMC meetings. Greenbook forecasts are publicly available with a five-year delay. The FOMC holds eight regularly scheduled meetings during the year and additional meetings as needed. Our comparison involves 63 Greenbook forecast dates from March 19, 1997, to December 8, 2004. Period $t = 1$ corresponds to 1968:M1. We construct the real-time data set for the Greenbook comparison as in Section 5.3.2 with one important exception. Financial variables are available in daily frequency, but typically their monthly averages are not yet available at the Greenbook publication dates. Since up-to-date information from the financial sector is potentially very important for short-run forecasting, we compute estimates for these variables based on weighted within-month averages of daily data up to the forecast origin. More specifically, we proceed as follows. Assume that there are four days in a month and denote the daily interest rate as r_τ. Imagine that at the forecast origin, only r_1 and r_2 are available. We replace the missing monthly interest rate by the expected monthly average $(r_1 + 3r_2)/4$ and include a measurement error with variance $5\hat{\sigma}_r^2/16$, where $\hat{\sigma}_r^2$ is the sample variance of past $r_\tau - r_{\tau-1}$'s. We do not group the Greenbook publication dates based on the availability of within-quarter monthly observations when computing forecast error statistics.

MF-VAR versus Greenbook Forecasts. We proceed by comparing the VAR forecasts

Figure 5.12: RMSEs of 11-Variable MF-VAR, QF-VAR, and QF-VAR+



Notes: The MF-VAR aligns the information that was available to the staff of the Board of Governors. The recursive estimation of the MF-VAR is repeated 62 times. The 22^{nd} sample is eliminated because the vintages for PCE were incomplete.

to Greenbook forecasts. Results are plotted in Figure 5.11, which depicts absolute RMSEs for quarter-on-quarter GDP growth (annualized), CPI inflation (annualized), and the unemployment rate. We are pooling the forecast errors from all estimation samples. At the nowcast horizon $h = 1$, the Greenbook forecasts and the MF-VAR forecasts for GDP growth and the unemployment rate attain roughly the same RMSE. For horizons $h \geq 3$, the MF-VAR produces more accurate output growth forecasts, while the Greenbook contains more precise unemployment rate predictions. In regard to inflation, the Greenbook forecasts dominate the MF-VAR forecasts at all horizons. As in the case of the end-of-month samples, the short-run forecasts from the MF-VAR attain a smaller RMSE than the QF-VAR forecasts. While the QF-VAR inflation forecasts slightly dominate the MF-VAR forecasts for horizons $h = 4$ and $h = 5$, the MF-VAR GDP growth and unemployment rate forecasts are more accurate than the QF-VAR forecasts at all horizons. A similar pattern also holds true for the remaining seven variables (not depicted in the figure). The MF-VAR forecasts are as good as the QF-VAR forecasts in the long run and substantially more accurate for short horizons.

As a low-brow alternative to the MF-VAR analysis, a forecaster with access to external nowcasts could simply condition the QF-VAR forecasts on these nowcasts to improve the

short-horizon forecast performance of the QF-VAR. In the following experiment, we assume that the forecaster is able to utilize the Greenbook nowcasts for quarterly GDP growth, inflation, and unemployment.[3] We refer to the resulting empirical model as QF-VAR+ and it is implemented as follows: when simulating $T+1$ draws from the predictive distribution of the QF-VAR, the forecaster uses one iteration of the Kalman filter to condition the simulated trajectories treating the nowcasts as actual observations. A detailed discussion of this procedure in the context of dynamic stochastic general equilibrium (DSGE) model forecasts is provided in Del Negro and Schorfheide (2013). The RMSEs for the QF-VAR+ are also plotted in Figure 5.11. With respect to GDP growth and inflation, the benefit of including the external nowcast into the QF-VAR is short-lived. While for $h = 1$ the QF-VAR+ attains the Greenbook RMSE, for horizons > 1 the performance resembles that of the QF-VAR. For the unemployment forecasts, the improvement in forecast performance extends to horizons $h > 1$. In fact, the RMSEs for the MF-VAR and the QF-VAR+ are quite similar. On balance, the MF-VAR compares well against a QF-VAR augmented by current-quarter nowcasts. A comparison for all 11 variables is provided in Figure 5.12.

[3]We thank Jonathan Wright for suggesting this experiment to us. We do not update the posterior distribution of the QF-VAR parameters in view of the additional information.

Bibliography

Albuquerque, R., M. Eichenbaum, and S. Rebelo (2012): "Valuation Risk and Asset Pricing," NBER Working Paper No. 18617.

Andrieu, C., A. Doucet, and R. Holenstein (2010): "Particle Markov Chain Monte Carlo Methods (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 72, 1–33.

Ang, A., J. Boivin, S. Dong, and R. Loo-Kung (2011): "Monetary Policy Shifts and the Term Structure," *Review of Economic Studies*, 78, 429–457.

Aruoba, B., and F. Schorfheide (2011): "Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-offs," *American Economic Journal: Macroeconomics*, 3, 60–90.

Aruoba, S. (2008): "Data Revisions Are Not Well Behaved," *Journal of Money, Credit and Banking*, 40(2-3), 319–340.

Aruoba, S., F. Diebold, J. Nalewaik, F. Schorfheide, and D. Song (2012): "Improving GDP Measurement: A Forecast Combination Perspective," in X. Chen and N. Swanson (eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis: Essays in Honour of Halbert L. White Jr.*, Springer, 1-26.

Aruoba, S., F. Diebold, and C. Scotti (2009a): "Real-Time Measurement of Business Conditions," *Journal of Business and Economic Statistics*, 27(4), 417–427.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009b): "Real-Time Measurement of Business Conditions," *Journal of Business & Economics Statistics*, 27(4), 417–427.

Ascari, G., and A. Sbordone (2013): "The Macroeconomics of Trend Inflation," Federal Reserve Bank of New York Staff Report No. 628.

Backus, D., M. Chernov, and S. Zin (2013): "Identifying Taylor Rules in Macro-finance Models," Manuscript, NYU and UCLA.

Baele, L., G. Bekaert, and K. Inghelbrecht (2010): "The Determinants of Stock and Bond Return Comovements," *The Review of Financial Studies*, 23, 2374–2428.

Bai, J., E. Ghysels, and Wright (2013): "State Space Models and MIDAS Regressions," *Econometric Reviews*, 32(7), 779–813.

Banbura, M., D. Giannone, and L. Reichlin (2010): "Large Bayesian VARs," *Journal of Applied Econometrics*, 25(1), 71–92.

Bansal, R., A. Gallant, and G. Tauchen (2007): "Rational Pessimism, Rational Exuberance, and Asset Pricing Models," *Review of Economic Studies*, 74, 1005–1033.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2013): "Volatility, the Macroeconomy and Asset Prices," *Journal of Finance*, Forthcoming.

Bansal, R., D. Kiku, and A. Yaron (2012a): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

——— (2012b): "Risks For the Long Run: Estimation with Time Aggregation," Manuscript, University of Pennsylvania and Duke University.

Bansal, R., and I. Shaliastovich (2013): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26, 1–33.

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

Bansal, R., and H. Zhou (2002): "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, 5, 1997–2043.

Barro, R. (2009): "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99, 243–264.

Beeler, J., and J. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bikbov, R., and M. Chernov (2013): "Monetary Policy Regimes and the Term Structure of Interest Rates," *Journal of Econometrics*, 174, 27–43.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Campbell, J., C. Pflueger, and L. Viceira (2013): "Monetary Policy Drivers of Bond and Equity Risks," Manuscript, Harvard University and University of British Columbia.

Campbell, J., and R. Shiller (1988a): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.

——— (1988b): "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

Campbell, J., and R. Shiller (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *The Review of Economic Studies*, pp. 495–514.

Campbell, J., A. Sunderam, and L. Viceira (2013): "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds," Manuscript, Harvard University.

Carriero, A., T. E. Clark, and M. Marcellino (2011): "Bayesian VARs: Specification Choices and Forecast Accuracy," *FRB Cleveland Working Paper 11-12*.

——— (2012): "Real-Time Nowcasting with a Bayesian Mixed Frequency Model with Stochastic Volatility," *FRB Cleveland Working Paper 12-27*.

Carter, C. K., and R. Kohn (1994a): "On Gibbs Sampling for State Space Models," *Biometrika*, 81(3), 541–553.

Carter, C. K., and R. Kohn (1994b): "On Gibbs Sampling for State Space Models," *Biometrika*, 81(3), 541–553.

Chauvet, M., and S. Potter (2013): "Forecasting Output," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2. Elsevier.

Chen, R., and J. Liu (2000): "Mixture Kalman Filters," *Journal of Royal Statistical Society Series B*, 62, 493–508.

Chiu, C. W. Eraker, B., A. Foerster, T. B. Kim, and H. Seoane (2012): "Estimating VAR's Sampled at Mixed or Irregular Spaced Frequencies: A Bayesian Approach," FRB Kansas City RWP 11-11.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 115(1), 147–180.

Cochrane, J. (2011): "Determinacy and Identification with Taylor Rules," *Journal of Political Economy*, 119, 565–615.

Cochrane, J., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–160.

Coibon, O., and Y. Gorodnichenko (2011): "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, 101, 341–370.

Dai, Q., and K. Singleton (2002): "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure," *Journal of Financial Economics*, 63(3), 415–441.

David, A., and P. Veronesi (2013): "What Ties Return Volatilities to Price Valuations and Fundamentals?," *Journal of Political Economy*, Forthcoming.

Davig, T., and E. Leeper (2007): "Generalizing the Taylor Principle," *American Economic Review*, 97(3), 607–635.

De Mol, C., D. Giannone, and L. Reichlin (2008): "Forecasting Using a Large Number of Predictors: Is Bayesian Shrinkage a Valid Alternative to Principal Components?," *Journal of Econometrics*, 146(2), 318–328.

Del Negro, M., and F. Schorfheide (2004): "Priors from General Equilibrium Models for VARs," *International Economic Review*, 45(2), 643 – 673.

——— (2010): "Bayesian Macroeconometrics," in *Handbook of Bayesian Econometrics*, ed. by H. K. van Dijk, G. Koop, and J. Geweke. Oxford University Press.

Del Negro, M., and F. Schorfheide (2011): "Bayesian Macroeconometrics," in *The Oxford Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. van Dijk, pp. 293–389. Oxford University Press.

Del Negro, M., and F. Schorfheide (2013): "DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2. Elsevier.

Diebold, F., and L. Kilian (2001): "Measuring Predictability: Theory and Macroeconomic Applications," *Journal of Applied Econometrics*, 16, 657–669.

Doan, T., R. Litterman, and C. Sims (1984): "Forecasting and conditional projection using realistic prior distributions," *Econometric Reviews*, 3(1), 1–100.

Doh, T. (2012): "Long-Run Risks in the Term Structure of Interest Rates: Estimation," *Journal of Applied Econometrics*, 28, 478–497.

Durbin, J., and S. Koopman (2001a): *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Durbin, J., and S. J. Koopman (2001b): *Time Series Analysis by State Space Methods*. Oxford University Press.

Edwards, C., and E. Howrey (1991): "A 'True' Time Series and Its Indicators: An Alternative Approach," *Journal of the American Statistical Association*, 86, 878–882.

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

Faust, J., J. Rogers, and J. Wright (2005): "News and Noise in G-7 GDP Announcements," *Journal of Money, Credit and Banking*, 37, 403–417.

Faust, J., and J. Wright (2013): "Forecasting Inflation," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2. Elsevier.

Fernández-Villaverde, J., P. Guerrón-Quintana, and J. Rubio-Ramírez (2010): "Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data," Manuscript, Duke University, Federal Reserve of Philadelphia, University of Pennsylvania.

Fernández-Villaverde, J., and J. Rubio-Ramírez (2011): "Macroeconomics and Volatility: Data, Models, and Estimation," in D. Acemoglu, M. Arellano and E. Deckel (eds.), *Advances in Economics and Econometrics: Theory and Applications*, Tenth World Congress of the Econometric Society, Cambridge University Press, Forthcoming.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74(4), 1059–1087.

Fixler, D., and J. Nalewaik (2009): "News, Noise, and Estimates of the "True" Unobserved State of the Economy," Manuscript, Bureau of Labor Statistics and Federal Reserve Board.

Fleischman, C., and J. Roberts (2011): "A Multivariate Estimate of Trends and Cycles," Manuscript, Federal Reserve Board.

Foroni, C., M. Marcellino, and C. Schumacher (2013): "U-MIDAS: MIDAS Regressions with Unrestricted Lag Polynomials," *Journal of the Royal Statistical Society, Series A*, Forthcoming.

Frale, C., M. Marcellino, G. L. Mazzi, and T. Proietti (2011): "EUROMIND: A Monthly Indicator of the Euro Area Economic Conditions," *Journal of the Royal Statistical Society, Series A*, 174(2), 439–470.

Gallmeyer, M., B. Hollifield, F. Palomino, and S. Zin (2007): "Arbitrage-Free Bond Pricing with Dynamic Macroeconomic Models," *The Federal Reserve Bank of St. Louis Review*, July/August 2007.

Gartaganis, A., and A. Goldberger (1955): "A Note on the Statistical Discrepancy in the National Accounts," *Econometrica*, 23, 166–173.

Geweke, J. (1977): "The Dynamic Factor Analysis of Economic Time Series Models," in D. Aigner and A. Goldberger (eds.), *Latent Variables in Socioeconomic Models*, North Holland, 365-383.

Geweke, J. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18(1), 1–126.

Geweke, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. New Jersey: John Wiley & Sons.

Ghysels, E. (2012): "Macroeconomics and the Reality of Mixed Frequency Data," Manuscript, University of North Carolina at Chapel Hill.

Ghysels, E., A. Sinko, and R. Valkanov (2007): "MIDAS Regressions: Further Results and New Directions," *Econometric Reviews*, 26(1), 53–90.

Giannone, D., M. Lenza, and G. Primiceri (2012): "Prior Selection for Vector Autoregressions," NBER Working Paper w18467.

Giannone, D., L. Reichlin, and D. Small (2008): "Nowcasting: The Real-Time Informational Content of Macroeconomic Data," *Journal of Monetary Economics*, 55(4), 665–676.

Giordani, P., M. K. Pitt, and R. Kohn (2011): "Bayesian Inference for Time Series State Space Models," in *The Oxford Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. K. van Dijk, pp. 61–124. Oxford University Press.

Greenaway-McGrevy, R. (2011): "Is GDP or GDI a Better Measure of Output? A Statistical Approach," *Manuscript, Bureau of Economic Analysis*.

Hall, R. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–357.

Harding, D., and R. Scutella (1996): "Efficient Estimates of GDP," Unpublished Seminar Notes, La Trobe University, Australia.

Harvey, A. (1989a): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

——— (1989b): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Hodrick, R. (1992): "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *The Review of Financial Studies*, 5, 357–386.

Jacobs, J., and S. van Norden (2011): "Modeling Data Revisions: Measurement Error and Dynamics of "True" Values," *Journal of Econometrics*, 161, 101–109.

Johannes, M., L. Lochstoer, and Y. Mou (2013): "Learning about Consumption Dynamics," Manuscript, Columbia Business School.

Karlsson, S. (2013): "Forecasting with Bayesian Vector Autoregressions," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2. Elsevier.

Kim, C.-J., and C. Nelson (1999): *State Space Models with Regime Switching*. MIT Press.

Kishor, N., and E. Koenig (2011): "VAR Estimation and Forecasting When Data Are Subject to Revision," *Journal of Business and Economic Statistics*, in press.

Komunjer, I., and S. Ng (2011): "Dynamic Identification of Dynamic Stochastic General Equilibrium Models," *Econometrica*, 79, 1995–2032.

Kuzin, V., M. Marcellino, and C. Schumacher (2011): "MIDAS vs. Mixed-Frequency VAR: Nowcasting GDP in the Euro Area," *International Journal of Forecasting*, 27(2), 529–542.

Landefeld, J., E. Seskin, and B. Fraumeni (2008): "Taking the Pulse of the Economy: Measuring GDP," *Journal of Economic Perspectives*, 22, 193–216.

Litterman, R. (1986): "Forecasting with Bayesian Vector Autoregressions - Five Years of Experience," *Journal of Business & Economic Statistics*, 4(1), 25–38.

Litterman, R. B. (1980): "Techniques for Forecasting with Vector Autoregressions," Ph.D. thesis, University of Minnesota.

Lubik, T. A., and F. Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94, 190–217.

Mankiw, N., D. Runkle, and M. Shapiro (1984): "Are Preliminary Announcements of the Money Stock Rational Forecasts?," *Journal of Monetary Economics*, 14, 15–27.

Mankiw, N., and M. Shapiro (1986): "News or Noise: An Analysis of GNP Revisions," *Survey of Current Business*, May, 20–25.

Mariano, R., and Y. Murasawa (2003a): "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18, 427–443.

Mariano, R. S., and Y. Murasawa (2003b): "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18(4), 427–443.

McCulla, S., and S. Smith (2007): "Measuring the Economy: A Primer on GDP and the National Income and Product Accounts," Washington, DC, Bureau of Economic Analysis, `http://www.bea.gov/national/pdf/nipa_primer.pdf`.

Nalewaik, J. (2010): "The Income- and Expenditure-Side Estimates of U.S. Output Growth," *Brookings Papers on Economic Activity*, 1, 71–127 (with discussion).

Nalewaik, J. (2012): "Estimating Probabilities of Recession in Real Time with GDP and GDI," *Journal of Money, Credit and Banking*, 44, 235–253.

Nalewaik, J., and E. Pinto (2012): "The Response of Capital Goods Shipments to Demand over the Business Cycle," Manuscript, Federal Reserve Board.

of Economic Analysis, B. (2006): "A Guide to the National Income and Product Accounts of the United States," Washington, DC, Bureau of Economic Analysis, `http://www.bea.gov/national/pdf/nipaguid.pdf`.

Phillips, P. C. B. (1996): "Econometric Model Determination," *Econometrica*, 64(4), 763–812.

Piazzesi, M., and M. Schneider (2006): "Equilibrium Yield Curves," *NBER Macro Annual*, pp. 389–442.

Pitt, M., and N. Shephard (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94, 590–599.

Rapach, D., J. Strauss, and G. Zhou (2010): "Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy," *Review of Financial Studies*, 23, 821–862.

Rassier, D. (2012): "The Role of Profits and Income in the Statistical Discrepancy," *Survey of Current Business*, pp. 8–22.

Rodriguez, A., and G. Puggioni (2010): "Mixed Frequency Models: Bayesian Approaches to Estimation and Prediction," *International Journal of Forecasting*, 26, 293–311.

Rudebusch, G. (2002): "Term structure evidence on interest rate smoothing and monetary policy inertia," *Journal of monetary economics*, 49(6), 1161–1187.

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.

Sargent, T., and C. Sims (1977): "Business Cycle Modeling Without Pretending to Have too Much a Priori Theory," in C.A. Sims (ed.), *New Methods in Business Cycle Research: Proceedings from a Conference, Federal Reserve Bank of Minneapolis, 45-109.*

Schorfheide, F. (2005): "Learning and Monetary Policy Shifts," *Review of Economic Dynamics*, 8, 392–419.

Schorfheide, F., and D. Song (2012): "Real-Time Forecasting with a Mixed-Frequency VAR," FRB Minneapolis Working Paper 701.

Schorfheide, F., D. Song, and A. Yaron (2013): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," Manuscript, University of Pennsylvania.

Shephard, N. (2013): "Martingale Unobserved Component Models," Manuscript, University of Oxford.

Sims, C., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96, 54–81.

Sims, C. A., and T. Zha (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39(4), 949–968.

Smith, R., M. Weale, and S. Satchell (1998): "Measurement Error with Accounting Constraints: Point and Interval Estimation for Latent Data with an Application to U.K. Gross Domestic Product," *The Review of Economic Studies*, 65(1), 109–134.

Stambaugh, R. F. (1999): "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

Stock, J. H., and M. W. Watson (2010): "Monthly Estimates of Real GDP/GDI: Research Memorandum," Manuscript, Harvard University and Princeton University.

Stone, R., D. Champernowne, and J. Meade (1942): "The Precision of National Income Estimates," *Review of Economic Studies*, 9, 111–125.

van Binsbergen, J., J. Fernández-Villaverde, R. Koijen, and J. Rubio-Ramírez (2012): "The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences," *Journal of Monetary Economics*, 59(7), 634–648.

Waggoner, D. F., and T. Zha (1999): "Conditional Forecasts in Dynamic Multivariate Models," *Review of Economics and Statistics*, 81(4), 639–651.

Watson, M., and R. Engle (1983): "Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC and Varying Coefficient Regression Models," *Journal of Econometrics*, 23, 385–400.

Wilcox, D. (1992): "The Construction of the U.S. Consumption Data: Some Facts and their Implications for Empirical Work," *American Economic Review*, 82, 922–941.