

# Core and ‘Crust’: Consumer Prices and the Term Structure of Interest Rates <sup>\*</sup>

Andrea Ajello,<sup>†</sup> Luca Benzoni,<sup>‡</sup> and Olena Chyruk<sup>§</sup>

June 7, 2018

---

<sup>\*</sup>We are grateful to Stijn Van Nieuwerburgh (Editor), Andrew Ang (NBER discussant), Anna Cieslak, and two anonymous referees for their comments and suggestions. We also thank Torben Andersen, Jeff Campbell, Larry Christiano, Charlie Evans, Spence Krane, Michael McCracken, Marcel Priebisch, Alejandro Justiniano, Giorgio Primiceri, Ken Singleton, and seminar participants at the Kellogg School of Management, the University of British Columbia, Essex University, the University of Lugano, Dartmouth College, the Board of Governors of the Federal Reserve System, the European Central Bank, the Bank of England, the 2013 Asset Pricing NBER SI, the 2012 WFA, the 2012 SoFiE, the 2012 SED, the 2014 EFA, the 2012 European Meeting of the Econometric Society, the 2012 Federal Reserve System conference on Business and Financial Analysis, the 2012 Midwest Macroeconomics Meetings, the 2014 HEC Applied Financial Time Series Workshop, the 2013 Computational and Financial Econometrics Conference, the 2013 Symposium on Control and Modeling of Social and Economic Behavior at UIUC, the 2016 RCEF, and the 2016 IAAE conference. All errors remain our sole responsibility. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System. The most recent version of this paper is at <http://ssrn.com/abstract=1851906>.

<sup>†</sup>Board of Governors of the Federal Reserve System, [andrea.ajello@frb.gov](mailto:andrea.ajello@frb.gov).

<sup>‡</sup>Corresponding author. Federal Reserve Bank of Chicago, [lbenzoni@frbchi.org](mailto:lbenzoni@frbchi.org).

<sup>§</sup>Federal Reserve Bank of Chicago, [ochyruk@frbchi.org](mailto:ochyruk@frbchi.org).

# Core and ‘Crust’: Consumer Prices and the Term Structure of Interest Rates

## Abstract

We propose a no-arbitrage model of the nominal and real term structures that accommodates the different persistence and volatility of distinct inflation components. Core, food, and energy inflation combine into a single total inflation measure that ties nominal and real risk-free bond prices together. The model is successful at extracting market participants’ expectations of future inflation from nominal yields and inflation data. Estimation uncovers a factor structure that is common to core inflation and interest rates and downplays the pass-through effect of short-lived food and energy shocks on inflation and interest rates. Model forecasts systematically outperform survey forecasts and other benchmarks.

# 1 Introduction

A general view in the empirical macro-finance literature is that financial variables do little to help forecast consumer prices. In particular, most empirical studies find that there is limited or no marginal information content in the nominal interest rate term structure for future inflation (Stock and Watson (2003)). The challenge to reconcile yield curve dynamics with inflation has become even harder during the recent financial crisis due to the wild fluctuations in consumer prices, largely driven by short-lived shocks to food and, especially, energy prices (Figure 1). There is hardly any trace of these oscillations in the term structure of interest rates. Core price indices, which exclude the volatile food and energy components, are more stable. Nonetheless, previous attempts to forecast core inflation using Treasury yields data have also had limited success.

We propose a dynamic term structure model (DTSM) that is successful at extracting market participants' expectations of future inflation from nominal yields, inflation, and real activity data. We price both the real and nominal Treasury yield curves using no-arbitrage restrictions. In the tradition of the affine DTSM literature (e.g., Duffie and Kan (1996), Piazzesi (2010), Duffie, Pan, and Singleton (2000)), we assume that the real spot rate is a linear combination of latent and observable macroeconomic factors. The macroeconomic factors are the three main determinants of consumer prices growth—core, food, and energy inflation—as well as real economic activity. We model them jointly with the latent factors in a vector autoregression.

This framework easily accommodates the properties of the different inflation components. Shocks to core inflation are much more persistent and less volatile compared to shocks to food and, especially, energy inflation (the 'crust' in the total consumer price index). The model fits these features by allowing for different persistence and volatility of the shocks to each of the three inflation measures, and for contemporaneous and lagged dependence

among the factors. In the model, the weighted average of the individual core, food, and energy components is the single measure of total inflation that ties nominal and real bond prices together.

Estimation on a panel data set of nominal Treasury yields, the three inflation series, and real activity delivers inflation forecasts that systematically outperform popular benchmarks. For instance, for total consumer price index (CPI) inflation the one-year-ahead root mean squared error (RMSE) produced by our DTSM over the 1999–2014 period is 26% lower than the RMSE for an autoregressive moving average model (ARMA).<sup>1</sup> At longer horizons DTSM forecasts improve further, with a 43% decrease in the five-year-ahead RMSE compared to the ARMA. The improvement in DTSM forecasts relative to other time-series models of inflation is even starker. Moreover, our DTSM yields RMSEs that are systematically lower than those of the forecasts from the Survey of Professional Forecasters (SPF) and Blue Chip Economic Indicators (BC). This is remarkable, as previous studies have documented that professional survey forecasts outperform all other forecasting models (e.g., Ang, Bekaert, and Wei (2007) and Faust and Wright (2009)). Our model does well at predicting core CPI inflation too, improving on the ARMA model and other benchmarks. The gain is most evident in long-run core inflation forecasts, where the DTSM cuts the ARMA RMSEs in half.

Our inflation forecasts not only reflect information from past price realizations, but also from yield curve dynamics. We allow the latent factors to shape the conditional mean of core inflation and model estimation finds support for such dependence. In contrast, models that include core, food, and energy inflation series but leave out interest rates data do much worse than our DTSM at forecasting inflation. These include weighted univariate ARMA models for the core, food, and energy series considered by Faust and Wright (2013) as well as a multivariate inflation vector autoregression (VAR) estimated on core, food, and energy

---

<sup>1</sup>This is the main benchmark favored by Ang, Bekaert, and Wei (2007) and Stock and Watson (1999).

inflation.

These results underscore the advantages of modeling the dynamics of the individual inflation components in a DTSM. Estimation finds shocks to food and, especially, energy inflation to be short-lived and to have limited impact on the yield curve and long-run inflation expectations. In contrast, a core inflation shock has a positive and lasting effect on short maturity yields that progressively declines with the yields' maturity. Most of the variation in yields and core inflation, however, is attributed to latent factors shocks. This shows that our DTSM distills information from the nominal yield curve that significantly improves our inflation forecasts.

Our framework offers other advantages:

- First, jointly modeling the three inflation factors (core, food, and energy) produces forecasts for total inflation as well as each of its components. In contrast, a DTSM that prices bonds out of a single measure of inflation delivers forecasts for the specific proxy of inflation used for estimation (e.g., total, core, or a principal component of several price series).
- Second, the distinct modelling of the inflation components allows us to explore which inflation shocks are priced in the term structure. We find shocks to energy inflation not to command a risk premium, while shocks to core and food inflation are priced.
- Third, we quantify the pass-through effect of energy shocks on core inflation and the term structure. This helps to inform policy makers' decision on how to react to an energy shock. Recursive model estimation shows that energy shocks have had limited impact on core inflation in recent times. Bond yields are largely unaffected by energy shocks as well.
- Fourth, the baseline DTSM does well at predicting personal consumption expenditures

(PCE) inflation too. This is of particular relevance for U.S. policy makers who pay close attention to PCE measures of inflation.

- Last, the model performs well at forecasting Treasury yields. Across maturities, the RMSEs for DTSM nominal yields' forecasts are lower than the RMSEs produced by several benchmarks, including the ARMA model and SPF yields forecasts.

Recent work on canonical (i.e., maximally identified) Gaussian affine term structure models shows that no-arbitrage restrictions do not affect out-of-sample forecasts of yields and macroeconomic variables relative to the forecasts produced by an “unconstrained factor-VAR model” (e.g., Joslin, Singleton, and Zhu (2011), Duffee (2011b), and Joslin, Le, Singleton (2013)). In this paper, we depart from the canonical Gaussian DTSM framework by imposing over-identifying conditions on the maximal specification that restrict the physical factor dynamics and the risk premia coefficients. Consequently, the irrelevance result of no-arbitrage restrictions need not hold<sup>2</sup> and, in fact, we show that our model outperforms an unconstrained factor-VAR out of sample.

A considerable improvement comes from the restrictions that we impose on the food and energy inflation dynamics. The presence of distinct core, food, and energy variables provides a flexible framework in which we find that food and, in particular, energy shocks have limited pass-through onto the yields and core inflation, while they remain an important source of variation for total inflation. Hence, we restrict their dynamics consistently with these findings. Economically, this is intuitive, as in past decades energy (and food) shocks have been volatile but short-lived. Moreover, our DTSM allows for a rich specification of the market prices of risk. This allows us to single out the sources of variation in risk premia with restrictions that lead to improved inflation forecasts. When we impose similar restrictions

---

<sup>2</sup>Joslin, Singleton, and Zhu (2011) conclude that improvements in the conditional forecasts of the pricing factors in Gaussian dynamic term structure models are due to the combined structure of no-arbitrage and  $\mathbb{P}$ -distribution restrictions. Duffee (2011b) and Joslin, Le, Singleton (2013) reach similar conclusions.

in a DTSM that includes a single measure of inflation (either total or core), we find a big deterioration in the inflation RMSEs.

We fit our baseline DTSM on quarterly nominal Treasury yields, core, food, and energy inflation, and real activity data. In several robustness checks, we also force it to match other observable quantities. We estimate the model on different yields' datasets and various measures of real activity, we expand the set of observables to include survey forecasts and market-based measures of real rate such as Treasury Inflation Protected Securities (TIPS) data, we estimate the model at the monthly rather than quarterly frequency, and explore the effect of macroeconomic spanning restrictions. When stretching the model along all these dimensions, we still find that our baseline DTSM fares well across this broad range of cases.

**Related Literature** Ang, Bekaert, and Wei (2007, 2008) estimate nominal and real term structures for U.S. Treasury rates with no-arbitrage models that include latent factors and one inflation factor (measured by either total or core realized inflation). The authors consider specifications with and without regime switches in the inflation dynamics. They find that term structure information does not generally lead to better inflation forecasts and often leads to inferior forecasts compared to those produced by models that use only aggregate activity measures. Their evidence confirms the results in Stock and Watson (1999), and extends them to a wide array of specifications that combine inflation, real activity, and yield dynamics. The relatively poor forecasting performance of term structure models applies to simple regression specifications, iterated long-horizon VAR forecasts, no-arbitrage affine models, and non-linear no-arbitrage models. They conclude that while inflation is very important for explaining the dynamics of the term structure (e.g., Ang, Bekaert, and Wei, 2008), yield curve information is less important for predicting future inflation. Yet, the yield curve should reflect market participants' expectations of future consumer price dynamics,

and our DTSM is successful at extracting such information.

Several studies incorporate market expectations in fitting real and nominal term structures of interest rates. For instance, Adrian and Wu (2010), Campbell, Sunderam, and Viceira (2013), Christensen, Lopez, and Rudebusch (2010), D’Amico, Kim, and Wei (2018), and Grishchenko and Huang (2013) combine nominal off-the-run yields constructed in Gürkaynak, Sack, and Wright (2007) with TIPS zero-coupon rates from Gürkaynak, Sack, and Wright (2010). Chen, Liu, and Cheng (2010) use raw U.S. TIPS data, while Barr and Campbell (1997) and Hördahl and Tristani (2012) focus on European index-linked bonds. Kim and Wright (2005) and Pennacchi (1991) rely on survey forecasts, while Haubrich, Pennacchi, and Ritchken (2014) introduce inflation swap rates to help identify real rates and expected inflation. In these studies, estimation typically forces the model to match survey- and market-based measures of real rates and expected inflation (TIPS data, survey inflation forecasts, or inflation swaps) up to a measurement error. Hence, model-implied real rates and inflation forecasts inherit the properties of these inputs by construction. In contrast, we propose a model that relies *exclusively* on nominal U.S. Treasury, inflation, and real activity data. Although we do not use surveys during estimation of our baseline DTSM, our inflation and interest rates forecasts systematically outperform the SPF, the BC, and the University of Michigan survey forecasts out of sample.

A vast related literature explores the relation between *nominal* interest rates and the macroeconomy. Early works directly relate current bond yields to past yields and macroeconomic variables using a vector auto-regression approach (e.g., Estrella and Mishkin (1997), and Evans and Marshall (1998, 2007)). More recently, several studies have explored similar questions using no-arbitrage dynamic term structure models (e.g., Ang and Piazzesi (2003), Ang, Piazzesi, and Wei (2006), Diebold, Rudebusch, and Aruoba (2006), Duffee (2006), Hördahl, Tristani, and Vestin (2006), Moench (2008), Diebold, Piazzesi, and Rudebusch



(2005), Piazzesi (2005), Rudebusch and Wu (2008)). Other contributions have extended these models to include market expectation in the form of survey forecasts (e.g., Chernov and Mueller (2012), Chun (2011), and Kim and Orphanides (2012)).

Another branch of the literature studies the link between bond risk premia and the macroeconomy (e.g., Cieslak and Povala (2015), Cochrane and Piazzesi (2005), Duffee (2011a), Joslin, Priebsch, and Singleton (2010)). These articles focus on the predictability of bond returns. We concentrate on no-arbitrage models of the nominal and real term structures, and explore their implications for expected inflation and the inflation risk premium.

The rest of the paper proceeds as follows. The model is described in Section 2. We discuss data and the estimation method in Section 3. The empirical results are in Section 4, while Section 5 concludes the paper.

## 2 A DTSM with Core and Crust

We assume that  $K_1$  latent factors  $L_t = [\ell_t^1, \dots, \ell_t^{K_1}]'$ ,  $K_2$  inflation factors  $\Pi_t = [\pi_t^1, \dots, \pi_t^{K_2}]'$ , and  $K_3$  real activity factors  $\Gamma_t = [\gamma_t^1, \dots, \gamma_t^{K_3}]'$  describe the time- $t$  state of the economy. Collecting the latent and macroeconomic factors into a state vector  $X_t = [L_t', \Pi_t', \Gamma_t']'$ , we define the state dynamics via a Gaussian VAR system,

$$X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t, \quad (1)$$

where  $\mu$  is a  $K \times 1$  vector of constants,  $K = K_1 + K_2 + K_3$ , and  $\Phi$  is a  $K \times K$  matrix with the autoregressive coefficients. The  $K \times 1$  vector of independent and identically distributed (i.i.d.) shocks  $\Sigma \varepsilon_t$  has Gaussian distribution  $N(0, V)$ , with  $V = \Sigma \Sigma'$ .

We assume that the the market price of risk  $\lambda_t$  is affine in the state vector  $X_t$ ,

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \quad (2)$$

for a  $K \times 1$  vector  $\lambda_0$  and the  $K \times K$  matrix  $\lambda_1$ . We then derive risk-adjusted state dynamics

$$X_t = \mu^Q + \Phi^Q X_{t-1} + \Sigma \varepsilon_t^Q, \quad (3)$$

where  $\mu^Q = \mu - \Sigma \lambda_0$  and  $\Phi^Q = \Phi - \Sigma \lambda_1$ .

## 2.1 Real Bond Prices

As in Ang, Bekaert, and Wei (2007, 2008), we assume that the one-period short real rate,  $r_t^*$ , is an affine function of the state vector  $X_t$ ,

$$r_t^* = \delta_0 + \delta_1' X_t. \quad (4)$$

where  $\delta_0$  is a scalar and  $\delta_1$  is a  $K \times 1$  vector. We specify the real pricing kernel  $m_{t+1}^*$ ,

$$m_{t+1}^* = \exp \left( -r_t^* - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1} \right), \quad (5)$$

and obtain the time- $t$  price of a real zero-coupon bond with  $(n+1)$  periods to maturity as:

$$p_t^{*n+1} = E_t [m_{t+1}^* p_{t+1}^{*n}]. \quad (6)$$

Since the model is affine, equation (6) has solution

$$p_t^{*n} = \exp \left( \bar{A}_n^* + \bar{B}_n^{*'} X_t \right), \quad (7)$$

where the coefficients  $\bar{A}_n^*$  and  $\bar{B}_n^{*'}$  solve the ordinary difference equations (ODEs):

$$\begin{aligned} \bar{A}_{n+1}^* &= -\delta_0 + \bar{A}_n^* + \bar{B}_n^{*'} \mu^Q + \frac{1}{2} \bar{B}_n^{*'} \Sigma \Sigma' \bar{B}_n^* \\ \bar{B}_{n+1}^{*'} &= -\delta_1' + \bar{B}_n^{*'} \Phi^Q. \end{aligned} \quad (8)$$

The real short rate equation (4) yields the initial conditions  $\bar{A}_1^* = -\delta_0$  and  $\bar{B}_1^{*'}$  =  $-\delta_1'$  for the ODEs (8). Thus, the real yield on an  $n$ -period zero-coupon bond is

$$y_t^{*n} = -\frac{\log(p_t^{*n})}{n} = A_n^* + B_n^{*'} X_t, \quad (9)$$

where  $A_n^* = -\frac{\bar{A}_n^*}{n}$  and  $B_n^{*'}$  =  $-\frac{\bar{B}_n^{*'}}{n}$ .

## 2.2 Nominal Bond Prices

If we define  $Q_t$  to be the price deflator, then the time  $t$  price of a nominal  $(n + 1)$ -period zero-coupon bond,  $p_t^{n+1}$ , is given by

$$p_t^{n+1} = p_t^{*n+1} Q_t = E_t \left[ m_{t+1}^* \frac{Q_t}{Q_{t+1}} p_{t+1}^{*n} Q_{t+1} \right] = E_t [m_{t+1} p_{t+1}^n] , \quad (10)$$

with the nominal pricing kernel  $m_{t+1}$  defined as

$$m_{t+1} = m_{t+1}^* \frac{Q_t}{Q_{t+1}} = m_{t+1}^* \exp(-\pi_{t+1}) = \exp \left( -r_t^* - \pi_{t+1} - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1} \right) , \quad (11)$$

where  $\pi_t \equiv \log(Q_t/Q_{t-1})$  is the inflation rate at which investors deflate nominal asset prices. All existing dynamic term structure models of the nominal and real yield curves specify  $\pi_t$  as a univariate process.<sup>3</sup> In contrast, we assume that the inflation rate is a weighted sum of the inflation factors in  $\Pi_t$ ,  $\pi_t = \sum_{j=1}^{K_2} \omega^j \pi_t^j$ , where the weights  $\omega^j$ ,  $j = 1, \dots, K_2$ , represent the relative importance of the distinct price series in the total consumer price index. In our preferred model,  $K_2 = 3$  and the three inflation series are the main components of the total consumer price index, i.e., core, food, and energy.

We define  $\mu^{Q,\pi} = \sum_{j=1}^{K_2} \omega^j \mu^{Q,\pi^j}$ , where  $\mu^{Q,\pi^j}$  is the element of the vector  $\mu^Q$  in the state dynamics (3) that corresponds to the inflation factor  $\pi^j$ ,  $j = 1, \dots, K_2$ . Similarly,  $\Phi^{Q,\pi} = \sum_{j=1}^{K_2} \omega^j \Phi^{Q,\pi^j}$  and  $\Sigma^\pi = \sum_{j=1}^{K_2} \omega^j \Sigma^{\pi^j}$  are the weighted averages of the rows  $\Phi^{Q,\pi^j}$  and  $\Sigma^{\pi^j}$  of the  $\Phi^Q$  and  $\Sigma$  matrices in equation (3) that correspond to the inflation factor  $\pi^j$ . Then, Appendix A.1 shows that nominal bond prices are an affine function of the state vector  $X$ :

$$p_t^n = \exp \left( \bar{A}_n + \bar{B}_n' X_t \right) , \quad (12)$$

---

<sup>3</sup>For instance, Ang, Bekaert, and Wei (2007, 2008) model the dynamics of a single inflation factor  $\pi_t$  measured by either total or core CPI and PCE inflation. Abrahams et al. (2013), D'Amico, Kim, and Wei (2018), Haubrich, Pennacchi, and Ritchken (2014) focus on total CPI inflation.

where the coefficients  $\bar{A}_n$  and  $\bar{B}'_n$  solve the ODEs:

$$\begin{aligned}\bar{A}_{n+1} &= -\delta_0 - \mu^{Q,\pi} + \frac{1}{2}\Sigma^\pi\Sigma^{\pi'} + \bar{A}_n + \bar{B}'_n\mu^Q + \frac{1}{2}\bar{B}'_n\Sigma\Sigma'\bar{B}_n - \bar{B}'_n\Sigma\Sigma^{\pi'} \\ \bar{B}'_{n+1} &= -\delta'_1 - \Phi^{Q,\pi} + \bar{B}'_n\Phi^Q,\end{aligned}\tag{13}$$

with initial conditions  $\bar{A}_1 = -\delta_0 - \mu^{Q,\pi} + \frac{1}{2}\Sigma^\pi\Sigma^{\pi'}$  and  $\bar{B}'_1 = -\delta'_1 - \Phi^{Q,\pi}$ . Thus, the yield on a nominal  $n$ -period zero-coupon bond is affine in the state vector,

$$y_t^n = -\frac{\log(p_t^n)}{n} = A_n + B'_n X_t,\tag{14}$$

where  $A_n = -\frac{\bar{A}_n}{n}$  and  $B_n = -\frac{\bar{B}_n}{n}$ .

### 2.3 Identifying Restrictions: The Maximal DTSM

We denote the vectors that contain the first  $K_1$  elements of  $\mu^Q$  and  $\delta_1$  by  $\mu_{L,1}^Q$  and  $\delta_{L,1}$ ; the upper-left sub-matrices of  $\Phi^Q$  and  $\Sigma$  by  $\Phi_{LL}^Q$  and  $\Sigma_{LL}$ ; and the upper-right sub-matrix of  $\Phi^Q$  by  $\Phi_{L,\Pi\Pi}^Q$ . Then, for identification, we impose the following restrictions: (i)  $\Phi_{LL}^Q$  is lower triangular with diagonal elements ordered in decreasing manner.<sup>4</sup> (ii) The elements of  $\Phi_{L,\Pi\Pi}^Q$  are fixed at zero. (iii) The covariance matrix  $\Sigma$  is lower triangular, and the sub-matrix  $\Sigma_{LL}$  is diagonal. (iv) The elements of the  $\delta_{L,1}$  vector are fixed at one. (v) Under risk-adjusted measure, the latent factors have zero unconditional mean,  $\mu_L^Q = 0$ .

With these identifying restrictions, the model is observationally equivalent to the canonical representation of Joslin, Le, and Singleton (2013).<sup>5</sup> We favor our normalization scheme since it lends itself naturally to incorporate the overidentifying restrictions that we specify below and help us improve our inflation forecasts.

<sup>4</sup>This identifying restriction applies to the case in which the autoregressive matrix  $\Phi^Q$  has real and distinct eigenvalues. Hamilton and Wu (2012) and Joslin, Singleton, and Zhu (2011) suggest normalizations that accommodate more general cases, such as  $\Phi^Q$  with complex, repeated, and zero eigenvalues.

<sup>5</sup>For instance, in the presence of  $K_1 = 3$  latent,  $K_2 = 3$  inflation, and  $K_3 = 1$  real activity factors, both representations allow for 124 identified coefficients.

## 2.4 Over-Identifying Restrictions: The Baseline DTSM\*

In the baseline model, we include  $K_1 = 3$  latent factors. Moreover, we specify the inflation vector to contain  $K_2 = 3$  factors,  $\Pi_t = [\pi_t^c, \pi_t^f, \pi_t^e]$ , where  $\pi_t^c$ ,  $\pi_t^f$ , and  $\pi_t^e$  are core, food, and energy inflation. Market participants deflate nominal asset prices in equation (11) at the total inflation rate, computed as the weighted sum of the three inflation series. That is,  $\pi_t = \pi_t^{tot} = \omega^c \pi_t^c + \omega^f \pi_t^f + \omega^e \pi_t^e$ , where  $\omega^c$ ,  $\omega^f$ , and  $\omega^e$  represent the relative importance of core, food, and energy prices in the total price index. Finally, we assume the model to have a single real activity factor,  $K_3 = 1$ .

These assumptions lead to a DTSM with seven state variables and 124 identified coefficients in its maximal representation. Prior to estimation we explore the effect of restricting some of these coefficients on the model fit and its out-of-sample performance. Due to the model complexity, we cannot conduct specification tests for all possible combinations of parameter restrictions. Hence, we use economic intuition and stylized empirical evidence to guide our model selection choices. Through this process, which we outline below, we impose restrictions on the matrices  $\Phi^Q$ ,  $\Sigma$ , and  $\Lambda$  to arrive at a baseline model, labeled DTSM\*, that contains 37 free parameters.

First, we assume that the autoregressive matrix  $\Phi_{LL}^Q$  has elements

$$\Phi_{LL}^Q = \begin{bmatrix} \phi_{\ell^1, \ell^1}^Q & 0 & 0 \\ (1 - \phi_{\ell^2, \ell^2}^Q) & \phi_{\ell^2, \ell^2}^Q & 0 \\ 0 & (1 - \phi_{\ell^3, \ell^3}^Q) & \phi_{\ell^3, \ell^3}^Q \end{bmatrix}. \quad (15)$$

Matrix (15) describes a central-tendency model in discrete time.<sup>6</sup> Central-tendency models have several advantages that make them popular in the term structure and, more broadly, financial economics literature.<sup>7</sup> They naturally span the yields' variation at different fre-

<sup>6</sup>This is more transparent in continuous time, where one can specify the  $k$ th-order factor to mean-revert towards the factor of order  $k - 1$ :  $d\ell_t^k = \beta^k (\ell_t^{k-1} - \ell_t^k) dt + \sigma_{\ell^k} dW_t^{\ell^k}$ . Discretization of the diffusion process for factors  $k = 1, 2$ , and  $3$  produces the VAR autoregressive matrix in equation (15).

<sup>7</sup>Within the Gaussian affine DTSM class, special cases include Vasicek (1977), in which the spot rate mean-reverts to a constant, and the model of Balduzzi, Das, and Foresi (1998), in which the conditional mean of the spot rate is described by two factors, one of which mean-reverts to a more persistent lower-order

quencies; the first factor impacts long-dated Treasuries the most, while higher-order factors exhibit a lower degree of persistence and mostly affect shorter-dated Treasuries. Furthermore, their autoregression matrix can be specified with very few coefficients. As in Calvet, Fisher, and Wu (2018), we impose a non-linear decay condition,  $\phi_{\ell^k, \ell^k}^Q = \exp\{-\beta_k\}$ ,  $\beta_k = \beta_1 b^{k-1}$ , with  $\beta_1 > 0$ ,  $b > 1$  and  $k = 1, \dots, K_1$ . This provides a flexible yet parsimonious representation that greatly facilitates model estimation. In addition, this structure ranks the factors in order of persistence, which avoids issues related to factors rotations (e.g., Collin-Dufresne, Goldstein, and Jones (2008), Dai and Singleton (2000), Hamilton and Wu (2012), Joslin, Priebsch, and Singleton (2010)), and guarantees very fast convergence to the likelihood function optimum.<sup>8</sup> Furthermore, we assume that the shocks to the latent factors are uncorrelated, with standard deviation  $\sigma_{\ell^k}$ ,  $k = 1, 2$ , and 3.

Second, to shape the food and energy dynamics, we turn to VAR models estimated on the yields' principal components along with individual measures of energy (or food) inflation series.<sup>9</sup> We find these models to do poorly out-of-sample in forecasting food and energy inflation, whereas a simple AR(1) for the univariate food/energy series outperforms them. Economically, this is intuitive, as in past decades energy and food shocks have been volatile but short-lived and have displayed very limited pass-through on core inflation and the yield curve. Hence, we implement similar restrictions in our DTSM by zeroing out the interactions between energy/food inflation and other factors. This greatly reduces the number of free DTSM coefficients and decreases the forecasting errors.

---

factor. Calvet, Fisher, and Wu (2018) extend this idea to a high-order cascade model that can approximate infinite-dimensional term structures. Central tendency specifications are also widely employed in stochastic-volatility (SV) models. For instance, Drechsler and Yaron (2011) use a central-tendency two-factor SV setup in an Epstein-Zin endowment economy with long-run risk, while Duffie, Pan and Singleton (2000) use it in reduced-form derivative pricing models.

<sup>8</sup>In Section 4.1.3, we document that imposing this central-tendency specification does not alter the out-of-sample performance of our DTSM. Hence, these restrictions are mostly for computational convenience.

<sup>9</sup>Joslin, Le, and Singleton (2013) show that a maximal Gaussian macro-finance term-structure model is nearly identical to a factor vector-autoregression. Hence, the out-of-sample performance of a VAR model estimated on the yields' principal components and food/energy inflation is similar to that of a maximal DTSM estimated on the same data

In regards to the core inflation dynamics, we find a VAR model estimated on the yields' principal components, real activity, and core inflation to do well out of sample. Hence, in the next step, we examine the interactions between core inflation, real activity, and the latent factors in our DTSM with the objective to improve the in- and out-of-sample model fit. To this end, we explore the effect of imposing restrictions on:

1. the dependence of core and activity on the lagged realizations of other variables;
2. the correlation between shocks to each latent factor and macroeconomic variables;
3. the correlations between shocks to the macroeconomic variables;
4. the market prices of risk associated with each variable in the state vector  $X$  and their dependence on all variables in  $X$ .

In all cases, we fix coefficients at zero based on the estimates' standard errors and model information criteria. We alternate and iterate over different sets of restrictions along the model dimensions 1 through 4 several times to alleviate the concern that the order in which they are imposed could lead to a sub-optimal outcome. Most importantly, we check the out-of-sample performance of the most promising cases to ensure that in-sample specification tests translate into better forecasting performance.

Through this process, we find core inflation to load on lagged values of the first and third latent factors,  $\ell^1$  and  $\ell^3$ , in addition to its own lag. Recall that the latent factors play a crucial role in a DTSM to capture the variation in bond yields. Hence, in our baseline model the term structure information condensed in  $\ell^1$  and  $\ell^3$  shapes the conditional mean of core inflation and thus inflation forecasts. In contrast, the pass-through effects of food and energy inflation on core inflation are small and insignificant, thus we omit these channels from the baseline DTSM. Term structure information enters the real activity dynamics too, but only through the third latent factor, which spans high-frequency fluctuations in short yields. The

coefficient is positive, consistent with a positive association between economic growth and increases in short rates, e.g., during a Fed tightening cycle. Moreover, we find real activity to load on its own lag and lagged food inflation. Food inflation follows a simple AR(1) process, while energy inflation only depends on food and its own lagged realizations. Among the correlations between shocks to the latent factors and the macroeconomic variables, only two are non-zero:  $\sigma_{\ell^3,core}$  and  $\sigma_{\ell^2,activity}$ . All covariances between macroeconomic variables are estimated in the baseline model, with the exception of the covariance between volatile energy shocks and either food or activity shocks.

Finally, we find shocks to the first latent factor, core and food inflation to be priced, with risk premia that depend on core inflation and economic activity. In contrast, we find shocks associated with energy, food, and activity not to command a risk premium. That is, coefficients other than  $\lambda_{\ell^1,core}$ ,  $\lambda_{core,core}$ ,  $\lambda_{core,activity}$  and  $\lambda_{food,core}$  are fixed at zero.

## 2.5 Benchmark Models

Here we list the various benchmarks against which we compare the out-of-sample performance of our baseline model. More details on these models and their estimation are in Appendix A.2.

### 2.5.1 Univariate Inflation Models: ARMA(1,1) and RW

The literature has proposed a wide array of models to forecast inflation (e.g., Stock and Watson (1999, 2003, and 2007)). Of these, the ARMA(1,1) and the random walk (RW) have proven particularly reliable in predicting consumer price dynamics over different sample periods. Thus, we consider both of these univariate models for comparison with our term structure specifications.



### 2.5.2 Multivariate Inflation Models: $\text{ARMA}_W(1,1)$ and $\text{VAR}_\Pi$

As in Faust and Wright (2013), we also consider various models that combine distinct core, food, and energy series, but leave out interest rates data. First, we construct  $\text{ARMA}_W$  forecasts for total inflation as a weighted sum of the  $\text{ARMA}(1,1)$  forecasts of the three inflation components using the relative importance weights. Second, we use an unconstrained  $\text{VAR}(1)$  estimated on core, food, and energy inflation data to forecast the three inflation components. Such forecasts recombine into a measure of total expected inflation, as in the  $\text{ARMA}_W$  case. We label this model  $\text{VAR}_\Pi$ .

### 2.5.3 Maximal DTSMs and Dynamic Factor Models

Joslin, Singleton, and Zhu (2011), Duffee (2011b), and Joslin, Le, Singleton (2013) show that no-arbitrage restrictions do not affect the forecasts of interest rates and macroeconomic variables in a canonical Gaussian affine term structure model. Hence, the model specified in Section 2.3, which is maximal in the number of identified coefficients, is virtually identical to an unconstrained VAR estimated on interest rates and macroeconomic data. We can therefore use such VAR as a benchmark for our baseline DTSM\*, in lieu of the maximal DTSM, to gauge the effect of the over-identifying restrictions imposed in Section 2.4. We label this case  $\text{DTSM}^{Max}$ .

Moreover, to illustrate the effect of the distinct modeling of the inflation series, we also consider two VARs estimated on the interest rates, real activity, and a single inflation series. These are equivalent to maximal DSTMs estimated on either total or core inflation, and therefore we label them  $\text{DTSM}^{Max,Tot}$  and  $\text{DTSM}^{Max,Core}$ .

Next, we consider a dynamic factor model that includes the same physical dynamics as our preferred DTSM but excludes the no-arbitrage restrictions; we label it  $\text{DFM}^{CT}$  to underscore that in this model the latent factors have a central-tendency specification. This benchmark

allows us to single out the effect of the over-identifying restrictions on the market prices of risk that we have imposed in Section 2.4. In a final check, we examine a more general DFM that excludes the restrictions on the autoregressive matrix in equation (15). A comparison between the  $DFM^{CT}$  and such DFM allows us to quantify the effect of the central-tendency structure on the latent factors on model forecasts.

#### 2.5.4 Surveys

Ang, Bekaert, and Wei (2007) show that inflation surveys outperform other popular forecasting methods (see also Faust and Wright (2009)). Hence, we include the following survey forecasts:

1. The median inflation forecast from the Michigan Survey of Consumers.
2. The median forecasts from the Survey of Professional Forecasters (SPF) for total CPI inflation; the three-month Treasury bill rate; and the 10-year Treasury bond rate.
3. The median forecasts from the Blue Chip Economic Indicators (BC) for total CPI inflation.

Appendix A.2.4 explains how we match these survey forecasts with the forecasts produced by our model.

## 3 Data and Estimation

We estimate the model on a quarterly sample of inflation, real activity, and nominal U.S.

Treasury yields with maturities of 3 months and 1, 3, 5, and 10 years from 1985 to 2015.<sup>10</sup>

The 3-month yield is from the Center for Research in Security Prices (CRSP) Fama-Bliss

---

<sup>10</sup>The sample period excludes the Fed's monetary experiment of the early 1980s and it is therefore less likely to include different regimes in inflation and interest rates. Ang, Bekaert, and Wei (2007) focus instead on a longer sample period with a model that includes regime shifts in a single inflation factor (either core or total); they find the model to perform very poorly out-of-sample. In contrast, we show in the Online Appendix that our baseline DTSM\* improves significantly over the benchmarks when estimated on a sample that starts in 1962.

Discount Bonds file. For other maturities, we construct zero-coupon yields from daily nominal constant-maturity par yields distributed by the Board of Governors in the H.15 data release. Prior to analysis, we interpolate the par yields into zero-coupon yields using a smoothed spline interpolation, as described in Section A.1 of the Online Appendix.<sup>11</sup> We then aggregate the daily series to the quarterly frequency.<sup>12</sup>

We focus on two widely-used measures of inflation:<sup>13</sup>

1. Monthly data on four Consumer Price Indices (CPI) constructed by the Bureau of Labor Statistics (BLS): (1) the total CPI for all Urban Consumers; (2) the core CPI (all items less food and energy); (3) the food CPI; and (4) the energy CPI.
2. Monthly price indices for Personal Consumption Expenditure (PCE) data released by the Bureau of Economic Analysis (BEA). Similar to the CPI series, we consider total, core, food, and energy PCE indices.

Table 1 contains CPI and PCE summary statistics.<sup>14</sup> The CPI- and PCE-weighted series are the total inflation series computed from their core, food, and energy components using the relative importance weights. Summary statistics for CPI- and PCE-weighted series are very close to those computed for the actual total CPI and PCE inflation series. Moreover, we find that the correlation between CPI and CPI-weighted total inflation series is 99.8%; for PCE data it is 99.9%. This confirms that weighted and actual measures of total inflation are virtually identical.

---

<sup>11</sup>The interpolated interest rates series are available at <https://www.chicagofed.org/~media/others/people/research-resources/benzoni/benzoni-abc-yields-csv.csv>. As a robustness check, we also estimate the model on the Gürkaynak, Sack, and Wright's (2007) U.S. Treasury yields data available from the Federal Reserve Board. Moreover, we confirm in unreported results that our findings are unchanged when we compute zero-coupon rates using a linear (rather than a smoothed spline) term-structure interpolation, similar to the unsmoothed Fama-Bliss method.

<sup>12</sup>As a robustness check, we also estimate the model at the monthly frequency.

<sup>13</sup>Section A.2 in the Online Appendix describes the main constituents of the core, food, and energy indices and explains the differences between the CPI and PCE series. Appendix A.3 explains how we measure the weights  $\omega^c$ ,  $\omega^f$ , and  $\omega^e$  associated with the core, food, and energy components.

<sup>14</sup>All price series are seasonally adjusted. We compute quarterly price indices by averaging over the monthly observations. Growth rates are quarter over quarter logarithmic differences in the index levels.

Table 1 also illustrates the difference in persistence across inflation series. The first-order auto-correlation for CPI-core inflation exceeds 0.8; higher-order auto-correlations remain high. The CPI-food series is much less persistent, with a first-order auto-correlation of 0.49 that declines at longer lags. In contrast, the shocks to the CPI-energy series are short lived, with a first-order auto-correlations of 0.22. Shocks die away quickly, resulting in second- and third-order correlations that are small and negative. Consequently, total CPI inflation is less persistent than core inflation. This is also evident from Figure 1, which plots the four inflation series over the 1985Q1-2015Q4 sample period. PCE inflation shares similar properties with the CPI series.

For both CPI and PCE series, the core component has a predominant weight in the total inflation index. The average relative importance of CPI core, food, and energy are 0.77, 0.15, and 0.08, respectively. In the PCE series, core prices have a higher average weight of 0.86, while the food and energy weights are lower at 0.09 and 0.05, respectively.

In our DTSM, we assume the inflation weights to be constant. Hence, in the baseline case we fix the weights  $\omega_c$ ,  $\omega_f$ , and  $\omega_e$  at the sample averages of the relative importance series. As a robustness check, we let the weights vary over the sample period and find the results to be similar to those obtained with fixed weights. This is not surprising, as in the data the weights show little time variation, with a standard deviation that is nearly zero across series and auto-correlations that are high at all lags (Table 1).

We focus on aggregate real consumption growth as a measure of the real activity factor. This series is commonly used in consumption-based asset pricing models. Moreover, we can compute it at both the monthly and quarterly frequencies using real personal consumption expenditures data released by the BEA.<sup>15</sup>

---

<sup>15</sup>Furthermore, we document below that the out-of-sample performance of the model is robust to using other activity series, including quarterly gross domestic product (GDP) growth and the Chicago Fed National Activity Index (CFNAI).

We apply the Kalman filter to estimate the model via maximum likelihood.<sup>16</sup> In the baseline case, the observable variables are the inflation and real activity factors,  $\Pi_t$  and  $\Gamma_t$ , and all five principal components (PCs) extracted from the panel of yields. We assume the inflation and real activity factors to be measured without error, while the yields' PCs have i.i.d. zero-mean Gaussian errors with a common variance coefficient for the first three PCs<sup>17</sup> and distinct variance coefficients for the remaining PCs. As a robustness check, we also estimate the model directly on the panel of yields, rather than their PCs, and find similar out-of-sample results.<sup>18</sup>

## 4 Empirical Results

The yields' factor loading, i.e., the  $B_n$  coefficients in pricing equation (14), determine the effect of each factor on the yield curve. Figure 2 shows estimates of these coefficients scaled to correspond to one-standard-deviation movements of the factors. Due to their recursive structure, the latent factors in our model behave in a way that is consistent with a low-dimensional version of the cascade model of Calvet, Fisher, and Wu (2018). They span the variation of the yields at different maturities, with the first latent factor impacting long-dated Treasuries the most, while higher-order factors exhibit faster mean reversion and mostly

---

<sup>16</sup>In unreported results we confirm that imposing restrictions on the conditional maximum Sharpe ratios (e.g., Duffee (2010)) produces results that are similar to those based on unconstrained maximum-likelihood estimation.

<sup>17</sup>A DTSM with three latent factors can fit the first three PCs of the yields extremely well. Consistent with this observation, we estimate the variance coefficients for the first three errors to be very small, and find that imposing an equality restriction on these parameters does not alter the results.

<sup>18</sup>Absent measurement errors, there would be *no difference* in the likelihood function when rotating the measurement equations from yields to the entire set of PCs. The same is true if one were to assume Gaussian measurement errors on the yields, and then rotate the covariance matrix of the errors consistent with the rotation of the measurement equations from the yields to their PCs (e.g., Joslin, Singleton, and Le 2013). Some small difference arises, however, when restrictions are imposed on the covariance matrix of the measurement errors. If the yields are the observables, a diagonal covariance matrix implies that the errors on the yields are uncorrelated, while the error vector on the PCs has a full covariance matrix. In contrast, imposing a diagonal covariance matrix for the PCs themselves, as in our baseline case, produces a full covariance matrix for the yields. We find that treating the PCs as observable and imposing a diagonal covariance matrix on the associated error vector produces a slightly higher likelihood function value than what we obtain when we assume that the yields are observed with uncorrelated errors. This is intuitive, as by design the PCs are orthogonal.

impact shorter-dated Treasuries (the black lines). The impact of a one-standard-deviation core inflation shock (the red line in Figure 2) is highest for short maturity yields and it progressively declines with the yields' maturity. At the ten-year maturity, the immediate reaction to a core inflation shock is about one fifth of the reaction to a one-standard-deviation  $\ell^1$  shock. Food and activity shocks have a moderate impact on short-maturity yields that fades away at longer maturities (the blue and purple lines), while the factor loadings on energy inflation (the green line) are very small across the yield curve.

Next, we explore how shocks to the state variables propagate over time. To this end, Table 2 shows a variance decomposition of the forecast errors for nominal yields, CPI inflation, and real activity. Panels A-C report results for yields with maturity of one quarter, five and ten years. In all cases, the latent factors account for the majority of the variation in yields' dynamics. The variation in the one-quarter yield forecasts at the one-year horizon is primarily driven by innovations to the second latent factor and core inflation; real activity and food inflation combined together explain about 7% of the short-run variation in the spot yield, while energy shocks are largely unimportant. At longer horizons, however, the first latent factor takes over, while the effect of core inflation and other shocks dissipates. Together,  $\ell^1$  and  $\ell^2$  account for 95% of the unconditional variation of the one-quarter yield forecast error. Shocks to core inflation only explain a small fraction of the variation in the five- and ten-year yield forecast error, while shocks to real activity, food and energy inflation account for virtually no variation across the entire term structure.

Table 2, Panel D, shows a variance decomposition for the CPI core inflation forecast error. At the one year horizon, shocks to the third latent factor and core inflation jointly explain more than 90% of the error variation.<sup>19</sup> At longer horizons the importance of  $\ell^3$  and

---

<sup>19</sup>In the DTSM\*, shocks to the inflation factors are correlated; moreover, core inflation shocks correlate with shocks to the third latent factor. Thus, as customary we identify the shocks via a Cholesky factorization of the covariance matrix with the latent factors entering first in the state vector, then core, food, and energy inflation and, lastly, real activity. Alternative identification assumptions produce similar results with one exception: when ordering core inflation before  $\ell^3$ , shocks to core display a larger impact on the variance of

core shocks decreases, while shocks to the other latent factors become dominant. Together,  $\ell^1$  and  $\ell^2$  explain around 70% of the unconditional variation, with the remaining portion split between  $\ell^3$  and core inflation. Across horizons, the contribution of shocks to real activity, food and energy inflation is very small.

Panels E and F decompose the variance of the forecast errors for food and energy inflation. The model attributes most of the variation to food and energy shocks themselves, respectively, consistent with these two variables displaying strong autoregressive dynamics. Moreover, core shocks propagate through the contemporaneous correlation with food and energy shocks and account for about 16% and 9% of the forecasting error. While both food and energy inflation do not directly depend on the latent factors, shocks to the latent factors pass through core inflation and real activity and can therefore impact food and energy inflation indirectly. However, the importance of this indirect channel is limited and mostly visible at long horizons.

In our model, core, food, and energy dynamics recombine to produce a total inflation forecast. While energy goods constitute only about 8% of the basket of goods and services that make up the consumer price index, the variation in energy inflation dwarfs that of food and, especially, core inflation (Table 1). Since energy shocks explain most of the energy forecasting errors (Table 2, Panel F), it is no surprise that they drive a large portion (66%) of the variation in total inflation forecasting errors as well (Table 2, Panel G). Shocks to core inflation and the latent factors explain the remaining 30%.

Panel H in Table 2 documents the presence of significant interactions between activity, inflation, and the latent factors. Across horizons, shocks to core and food inflation,  $\ell^2$  and  $\ell^3$  explain approximately 35% of the variation in the activity forecasting error, while the rest is attributed to activity shocks. In our model, real activity is largely unimportant to explain

---

the core inflation forecast error at the one year horizon.

other variables, except for its contribution to the variance of the one-quarter yield forecast error.

Taken together, these results suggest that core inflation and interest rates share a common factor structure. Shocks to the latent factors drive both core inflation and the yield curve, while shocks to core inflation mostly affect inflation and the spot rate at short horizons. Moreover, while core shocks are directly linked to nominal yields, especially at short yield maturities, food and energy shocks are unimportant towards explaining the term structure of interest rates. In contrast, energy shocks dwarf innovations to other variables as a driver of the variation of the forecast error for total inflation.

## 4.1 Inflation Forecasts

We repeatedly estimate the DTSM\* using quarterly yields, inflation, and real activity data over the period beginning in 1985Q1 and ending on date  $t$ , where  $t$  ranges from 1999Q4 through 2014Q4.<sup>20</sup> For each set of coefficients obtained with data up to and including quarter  $t$ , we forecast core, food, and energy inflation at quarter  $t + j$ ,  $j = 1, \dots, J$ . As in Ang, Bekaert, and Wei (2007), for each series  $i$  we construct an inflation forecast at horizon  $J$  by summing the  $J$  quarterly forecasts,  $E_t[\pi_{t+J,J}^i] = \sum_{j=1}^J E_t[\pi_{t+j}^i]$ , where  $\pi_{t+J,J}^i = \sum_{j=1}^J \pi_{t+j}^i$  and  $\pi_{t+j}^i = \log(Q_{t+j}^i/Q_{t+j-1}^i)$ . Moreover, we use the weights  $\omega_t^c$ ,  $\omega_t^f$ , and  $\omega_t^e$  to compute a forecast of total inflation,  $E_t[\pi_{t+J,J}^{tot}] \equiv \omega^c E_t[\pi_{t+J,J}^c] + \omega^f E_t[\pi_{t+J,J}^f] + \omega^e E_t[\pi_{t+J,J}^e]$ . We assess the forecast error against realized inflation based on the root mean squared error criterion,

$$\text{RMSE} = \sqrt{E[(E_t(\pi_{t+J,J}^i) - \pi_{t+J,J}^i)^2]} = \sqrt{\frac{1}{N} \sum_{t=1}^N (E_t(\pi_{t+J,J}^i) - \pi_{t+J,J}^i)^2}, \quad (16)$$

where  $N$  is the number of predictions in the out-of-sample window.

Table 3 reports RMSEs in percent per year for inflation forecasts at the one-year horizon ( $J = 4$  quarters). Panels A and B show results for CPI inflation, while Panel C focuses on

<sup>20</sup>In the Online Appendix, we document that the DTSM\* outperforms the inflation forecasts of the benchmark models when estimated over a longer sample period that starts in 1962.



PCE inflation. We choose the ARMA(1,1) to be the main benchmark (Stock and Watson (1999) and Ang, Bekaert, and Wei (2007)) against which we compare the relative performance of all models, including the DTSM\* and the other benchmark models in Section 2.5. In particular, the table shows West (1996)  $p$ -values for a test of equal forecast accuracy computed under the null that the RMSE for the ARMA equals the RMSE for the competing model and against the alternative that the RMSE for the ARMA exceeds the DTSM\* RMSE.

#### 4.1.1 DTSM\* vs. Time-Series Models of Inflation

The baseline DTSM\* outperforms each of the time-series models of inflation in predicting total and core CPI inflation. In particular, the DTSM\* RMSE for total inflation is 1.19%; this is a 26% and 34% improvement over the univariate ARMA and RW models, respectively. Time series models that include distinct core, food, and energy components such as the ARMA<sub>W</sub> and VAR<sub>II</sub> fare better than univariate ARMA and RW models of total inflation. Still, their RMSEs are 20-24% higher than the DTSM\* RMSE. That is, the improvement in forecasting performance of the DTSM\* relative to the ARMA benchmark cannot be attributed solely to the distinct modeling of the inflation series, nor to the averaging of the individual core, food, and energy forecasts.

The DTSM\* produces a 0.50% RMSE for CPI core inflation, which is 6% and 11% lower than the ARMA and RW RMSEs, respectively. The VAR<sub>II</sub>, which is jointly estimated on core, food, and energy inflation, does much worse than all other models on core inflation; in particular, the VAR<sub>II</sub> RMSE is 26% higher than that for the DTSM\*. This further underscores that the distinct modeling of the inflation components is, by itself, insufficient to decrease the forecasting errors. The DTSM\* outperforms all time series models on CPI food inflation too, and is roughly at par with the ARMA in predicting energy inflation.

The West (1996) test for equal forecast accuracy rejects the null that the DTSM\* and ARMA RMSEs for total CPI inflation are identical (the  $p$ -value is 3%) against the alternative

that the DTSM\* does better than the ARMA. For core CPI inflation we fail to reject the null at conventional significance levels. However, at longer forecasting horizons the West (1996) test rejects the null of equal predictive accuracy for core inflation too. Indeed, Table 4 shows out-of-sample RMSEs associated with CPI inflation forecasts at horizons ranging from one quarter to five years. The DTSM\* RMSE on core inflation declines from 0.58% at the one-quarter forecasting horizon to 0.33% for forecasts five years out. Compared to the ARMA, the DTSM\* long-run forecasting errors are twice as low, with a West (1996)  $p$ -value that is essentially zero. Long-run DTSM\* forecasts of total inflation are also dramatically better than ARMA forecasts; the five-year RMSE is 0.49%, which is 43% lower than the ARMA RMSE (the West  $p$ -value is 4%). The DTSM\* improves upon the ARMA on food inflation too, while it is at par with the ARMA for energy.

The results we discussed so far are based on a specific 1999-2014 out-of-sample window. To illustrate the robustness of our findings to the choice of the testing period we compute RMSEs over a grid of out-of-sample windows with start date ranging from 1997Q4 to 2000Q4 and end date from 2004Q4 to 2014Q4. For each window in the grid, we compute core and total CPI inflation RMSEs for the baseline DTSM\* and ARMA. Figure 3 plots their percentage ratio,  $100 \times (\text{RMSE DTSM}^*/\text{RMSE ARMA} - 1)$ . That is, negative numbers in the plot signal that the DTSM\* does better than the ARMA.

It is evident that the DTSM\* greatly outperforms the ARMA on both core and total inflation. In particular, the DTSM\* beats the ARMA at forecasting core inflation in virtually all of the out-of-sample windows, with an improvement upon the ARMA RMSEs of up to 40%. For total inflation the improvement occurs in 96% of the cases with a reduction in RMSEs that ranges from 18% to 26%, except for a few out-of-sample windows that have an early end date.

### 4.1.2 Survey Forecasts

Due to the timing of data releases, it is not possible to perfectly match the information sets of survey participants at the time they make a forecast with the information set available to the econometrician when she estimates the DTSM. However, we can choose the release date of the survey forecasts to minimize the difference in the two information sets. To this end, we consider two alternative approaches, described in more detail in Appendix A.2.4. In one case, we use forecasts released at the end of the estimation quarter, we exclude the “nowcast” for the current quarter that corresponds to the last estimation period, and collect the forecasts for the following four quarters. In this first case, professional forecasters will not have seen the inflation data corresponding to the end of the last estimation quarter that are instead known to the econometrician when she estimates the model. In the second case, we use survey forecasts released in the quarter that follows the end of the estimation period. In this case, professional forecasters will have seen all the data that go into the estimation of the model, but will also have access to additional data that are realized beyond the end of the sample period. Hence, survey participants will have a significant advantage compared to the econometrician.

Here, we present RMSEs for both of these cases in an attempt to establish lower and upper bounds within which the information set of the econometrician is enclosed. We label the first case “surveys without nowcast,” since the nowcast for the first quarter of the survey release, computed with partially observed data, is excluded. In the second case, professional forecasters have partially observed the variable of interest in the first quarter of the forecasting period, hence we label this latter case “surveys with nowcast.”

In either of these two cases, professional forecasters do quite well at predicting inflation (Ang, Bekaert, and Wei (2007), Faust and Wright (2009)). For SPF forecasts, the total-inflation RMSEs over the 1999-2014 out-of-sample window are 1.30% when the nowcast is

excluded and 1.22% when it is included (Table 3, Panel A). BC forecasts produce similar RMSEs of 1.28% and 1.21%, respectively. In all cases, BC and SPF RMSEs are lower than those for the time-series models of inflation; a test of equal forecast accuracy rejects the null hypothesis that the BC/SPF and ARMA perform identically. Remarkably, the DTSM\* RMSE is even lower albeit by a small amount. We find a much bigger improvement over the University of Michigan survey forecasts.

Given the prominence of inflation surveys by professional forecasters in the literature, in Table 5 we directly compare the DTSM\* and the SPF/BC forecasts. In Panel A we exclude the survey nowcast released in the last quarter of the estimation period and compare survey forecasts over the next year. The DTSM\* RMSE is significantly lower than those of both BC and SPF forecasts, with West (1996)  $p$ -values of 1%.<sup>21</sup>

Panel B in Table 5 shows a similar comparison for the second case in which we give survey participants an information advantage compared to the econometrician. The survey *nowcast* is clearly superior to the one-quarter-ahead forecast of our DTSM\*. This is to be expected, as the survey nowcast incorporates information from Treasury yields, inflation swaps, energy futures, as well as plain observation of consumer prices (e.g., gas prices) realized during the first forecasting horizon that help survey participants refine their measure of inflation for that quarter (e.g., Faust and Wright 2013). However, our baseline model outperforms the survey *forecasts*: over horizons from 2-5 quarters ahead, the DTSM\* produces a 1.20% RMSE, compared to 1.31% and 1.28% for the SPF and BC surveys, with West (1996)  $p$ -values of 2%.

Figures 4 and 5 extend these results to a grid of out-of-sample windows with start date ranging from 1997Q4 to 2000Q4 and end date from 2004Q4 to 2014Q4. At the 1-4 quarters horizon, the DTSM\* outperforms the SPF and BC survey forecasts 98% and 95% of the

---

<sup>21</sup>We obtain similar  $p$ -values when using the bootstrapping method of Chernov and Mueller (2013). The results are in the Online Appendix.

times, respectively, with percentage ratios  $100 \times (\text{RMSE DTSM}^*/\text{RMSE survey} - 1)$  that are as small as -10%. At the 2-5 quarters horizon, the RMSE ratios favor the baseline model over the surveys in 90% and 89% of the out-of-sample windows, with improvements of up to 9%.

#### 4.1.3 The Role of the Core and Crust Decomposition and the Overidentifying Restrictions

In this section, we document that the improvement in our inflation forecasts stems from the combination of our decomposition of inflation into its core, food, and energy components and the over-identifying restrictions on the individual inflation dynamics and the market prices of risk. We start by examining the performance of the maximal version of our model, the  $\text{DTSM}^{Max}$  specification described in Section 2.5.3. Compared to our  $\text{DTSM}^*$ , the  $\text{DTSM}^{Max}$  also includes the individual inflation components but only features the minimal set of coefficient restrictions necessary for identification. We then progressively add constraints on the model coefficients to elicit their impact on inflation forecasts. In particular, we separate the effect of restrictions on the food and energy dynamics, the central tendency structure for the latent factors, and the specification of the market prices of risk. Finally, we explore maximal DTSMs that include a single measure of inflation (either total or core) and compare their performance to our core and crust models.

Table 3, Panel B, examines the out-of-sample performance of the maximal version of our baseline model, the  $\text{DTSM}^{Max}$ . On core inflation, the  $\text{DTSM}^{Max}$  does as well as the  $\text{DTSM}^*$ . In contrast, for total, food, and energy inflation it delivers RMSEs that are higher than those for our baseline model and professional survey forecasts. The results however improve when we constrain the food and energy dynamics as in our  $\text{DTSM}^*$ , in which the dependence of food and energy inflation on the latent factors and other macroeconomic factors is reduced via over-identifying restrictions. This is also consistent with the experience of the past

decades, in which food and especially energy shocks have been short-lived and partially decoupled from yields' and core fluctuations. In particular, we estimate the DFM described in Section 2.5.3, in which the food and energy processes are modeled as in our DTSM<sup>\*</sup>. We obtain a considerable reduction in the energy inflation RMSE, which in turn lowers the total inflation RMSE to 1.32%. The DFM also includes the DTSM<sup>\*</sup> restrictions in the core equation. The core RMSE is however largely unaffected compared to the DTSM<sup>Max</sup>.

Note that the DFM does not restrict the latent factor dynamics to follow a central tendency structure. Hence, in DFM<sup>CT</sup> we impose a recursive specification on the latent factors that is identical to the one in our baseline model. We find this restriction to facilitate model estimation without affecting the out-of-sample performance of the model. Across all inflation series, the DFM and DFM<sup>CT</sup> RMSEs are identical.

Compared to the DFM<sup>CT</sup>, the DTSM<sup>\*</sup> produces similar results for core inflation while it delivers a significant improvement in the total inflation RMSE, which declines from 1.33% to 1.19%. This is due to the effect of the overidentifying restrictions on the market prices of risk discussed in Section 2.4.

Next, we illustrate the effect of modeling distinct inflation components relative to a DTSM that uses a single measure of either total or core inflation. The DTSM<sup>Max,Core</sup>, which includes a single core inflation series, does well at forecasting core inflation, with an RMSE that is at par with our baseline model.<sup>22</sup> However, the DTSM<sup>Max,Core</sup> is silent about total inflation dynamics. Hence, we turn to the DTSM<sup>Max,Tot</sup>, which includes a single total inflation series. In this case, the total inflation RMSE is 1.33%, which is in line with the DFM model. Indeed, there are similarities in the way the DFM and DTSM<sup>Max,Tot</sup> deal with the volatile energy shocks. In the DFM, energy has its own dynamics but its dependence on the other variables is restricted. These restrictions improve the energy inflation forecasts relative

---

<sup>22</sup>??We should acknowledge the difference in results compared to ABW due to the estimation period.

to the DTSM<sup>Max</sup>, in which energy inflation depends on all other factors, and therefore result in lower total inflation RMSEs. Instead, in the DTSM<sup>Max,Tot</sup> the volatile energy and food shocks are bundled together with the smooth core series in a total inflation measure. In the DTSM<sup>Max,Tot</sup>, total inflation also depends on all other variables, but the series is smoother relative to energy inflation and therefore we do not observe the deterioration of total inflation forecasts that we find in the maximal DTSM<sup>Max</sup>.

However, the maximal DTSM<sup>Max,Total</sup> and DTSM<sup>Max,Core</sup> lack the flexibility of the baseline model when it comes to incorporating overidentifying restrictions. With the necessary adjustments, we impose the restrictions described in Section 2.4 in a DTSM<sup>Total</sup> that includes a single measure of total inflation and compare its performance to that of the maximal DTSM<sup>Max,Total</sup>. The RMSE for total inflation is 1.61%, which is higher than the 1.33% for DTSM<sup>Max,Total</sup>. Similarly, the core RMSE for DTSM<sup>Core</sup> is 0.66%, compared to 0.48% for DTSM<sup>Max,Core</sup>.

Overall these results show that time-series and no-arbitrage over-identifying restrictions combine in our baseline DTSM to improve its performance relative to the maximal DTSM specification in which it is nested. A considerable improvement comes from the restrictions on the food and energy inflation dynamics. The central tendency specification for the latent factors greatly facilitates estimation without changing model forecasting performance. Finally, the distinct modeling of the inflation series provides a flexible specification of the market prices of risk. This allows us to single out the shocks that are being priced in the term structure and the variables that drive the variation in their risk premia, resulting in better inflation forecasts. In contrast, similar restrictions applied in DTSMs that contain a single measure of inflation do not work as well.

#### 4.1.4 Robustness Checks

Panel C in Table 3 contains various robustness checks. We consider estimation on different yields' datasets and alternative measures of real activity, we expand the set of observables to include TIPS data and survey forecasts, we re-estimate the model at the monthly rather than quarterly frequency, and explore the effect of macroeconomic spanning restrictions. In all cases, which we describe in more detail below, we rely on the same baseline model rather than choose new overidentifying restrictions for the specific purpose of each robustness check. When stretching the model along all these dimensions, we still find the DTSM\* to fare well across this broad range of cases.

**Yields Data** In DTSM<sup>yields</sup> we estimate the model on the panel of the nominal zero-coupon yields themselves, instead of their principal components. This approach produces a 1.19% RMSE for total inflation, identical to the DTSM\* case. The core inflation RMSE is 0.53%, which is in line with the baseline model.

The DTSM<sup>GSW</sup> uses instead a panel of yields released by the Federal Reserve Board based on the Gürkaynak, Sack, and Wright (GSW, 2007) term structure interpolation. The main difference between our and the GSW yields data is that GSW construct their term structure from off-the-run Treasuries, while we use U.S. Treasury yields' data from the H.15 release that are mainly on-the-run.<sup>23</sup> While the datasets are similar, there is a visible off-on-the-run spread between long-maturity zero-coupon GSW and our yields. At the 10-year maturity, the spread averages 17 basis points over the 1985-2015 sample period and rises during times of market stress; for instance, it peaks at around 70 basis points during the financial crisis (see Figure 1 in the Online Appendix). With a 1.26% RMSE, the DTSM<sup>GSW</sup> still outperforms the ARMA on total inflation (the West (1996)  $p$ -value is 7%). However,

---

<sup>23</sup>Another smaller difference is that GSW use an extended Nelson-Siegel yield curve, while we rely on a constrained spline interpolation.



compared to the DTSM\* its total-inflation RMSEs is 6% higher. This suggests that on-the-run Treasuries, like those that we use in our baseline case, contain a slightly better inflation signal than the more illiquid off-the-run securities.

**Real Activity Measures** We consider two alternative real activity measures: real GDP growth and the Chicago Fed National Activity Index. The DTSM<sup>GDP</sup> produces a 1.19% RMSE for total inflation, identical to the estimate obtained for DTSM\*, and a 0.52% RMSE for core inflation, which is similar to what we have found with real consumption growth. Switching to the CFNAI series in DTSM<sup>CFNAI</sup> we obtain RMSEs of 1.23% and 0.56% for total and core inflation, respectively.

**TIPS and Surveys** Our model produces estimates of the term structure of real rates. Hence, it is natural to augment the system of observation equations to include market-based measures of real yields and match them with the corresponding model-implied quantities. To this end, we include TIPS yields with 2-, 5-, and 10-year maturity released by the Federal Reserve Board based on the Gürkaynak, Sack, and Wright (2010) method. Similar to Chernov and Mueller (2012), we include TIPS data starting from 2004, as prior to that date the TIPS market was still in its infancy and suffered from significant liquidity problems (e.g., D’Amico, Kim, and Wei 2018). We label this case DTSM<sup>TIPS</sup>.

In the presence of TIPS, the total inflation RMSE increases to 1.28%; while somewhat worse than our baseline’s, this value is lower than that of the ARMA, with a 9%  $p$ -value for the West (1996) test of equal forecasting accuracy. At 0.51%, the core RMSE is indistinguishable from the one produced by the DTSM\*.

Next, we augment the baseline model to include survey forecasts. In DTSM<sup>SPF</sup> we add the one-year-ahead SPF surveys in the measurement equation, while DTSM<sup>BC</sup> incorporates Blue-Chip survey forecasts. In the DTSM<sup>SPF</sup> case, the total and core inflation RMSEs

are 1.25% and 0.53%; the  $\text{DTSM}^{BC}$  results are similar. Finally, we jointly use TIPS and survey data during estimation of the  $\text{DTSM}^{TIPS+SPF}$  and  $\text{DTSM}^{TIPS+BC}$  models. The total inflation RMSEs are mostly unchanged, while the core inflation RMSE improves slightly to 0.47%.

**Monthly vs. Quarterly Frequency** Here we show that our results are not very sensitive to the choice of the quarterly estimation frequency. We repeatedly estimate the  $\text{DTSM}^{Monthly}$  on monthly term structure, inflation, and real activity data over the period beginning in January 1985 and ending on date  $t$ , where  $t$  ranges from December 1999 through December 2014. For ease of comparison with the quarterly baseline results, we only focus on the forecasts formulated in the last month of each quarter. Similar to the baseline case, we aggregate those forecasts at the one-year horizon. We find a 1.18% total inflation RMSE, which is nearly identical to that of the  $\text{DTSM}^*$ . For core inflation the  $\text{DTSM}^{Monthly}$  produces a 0.59% RMSEs, compared to 0.50% in the baseline case.

**Unspanned Inflation Risk** Figure 2 shows that the factor loadings of nominal yields on energy inflation are tiny. Here we show that fixing them at zero does not improve the forecasting performance of the baseline  $\text{DTSM}^*$ . In Section A.3 of the Online Appendix we derive restrictions along the lines of Joslin, Pribsch, and Singleton (2014) and Wright (2011) such that a subset of the macroeconomic shocks is unspanned by the yield curve. We label the case with unspanned energy risk  $\text{DTSM}^{UMRE}$ . The results in Table 3 are mostly similar to those of our baseline model.

In contrast, the factor loadings on core inflation and, to a lesser extent, food inflation and real activity are different from zero. In the next robustness check, we assume that all macro risks are unspanned and label this case  $\text{DTSM}^{UMR}$ . The RMSE on total inflation increases to 1.28% relative to the 1.19% estimate for the baseline (Table 3), while core

inflation forecasts improve somewhat over the 1999-2014 out-of-sample window. Moreover, in unreported results we compare the DTSM<sup>UMR</sup> to survey forecasts of total inflation. In virtually all out-of-sample windows, we find the DTSM<sup>UMR</sup> to underperform both SPF and BC surveys. This is in contrast to the baseline DTSM<sup>\*</sup>, which systematically beats the surveys (Figures 4 and 5). The most dramatic difference between the DTSM<sup>UMR</sup> and DTSM<sup>\*</sup>, however, is in the nominal yields' forecasts. We show in Section 4.4 below that UMR restrictions on core inflation produce much higher RMSEs across yields' maturities.

## 4.2 PCE Inflation Forecasts

Table 3, Panel D, shows that the results for PCE inflation series are largely consistent with the evidence on CPI inflation. Namely, for total PCE inflation the DTSM<sup>\*</sup> outperforms all time-series models of inflation and produces a 17% decrease in RMSE compared to the ARMA case; the associated West (1996)  $p$ -value is 18%. There is a 9% improvement in the food inflation RMSE, while the RMSE for energy is slightly higher than the ARMA RMSE. On core inflation, the DTSM<sup>\*</sup> underperforms the ARMA, although the test for equal forecast accuracy does not reject the null that the two models perform identically.

## 4.3 Energy Pass-Through

In the baseline DTSM<sup>\*</sup>, we have fixed the  $\phi_{\pi^c, \pi^e}$  coefficient that links lagged realizations of energy inflation to core inflation at zero. Such coefficient is a measure of the pass-through effect of energy inflation shocks onto core inflation. To illustrate the extent of the pass-through in our setting, here we consider a flavor of the model in which  $\phi_{\pi^c, \pi^e}$  is a free parameter. We estimate this model over samples with start date of 1985Q1 and end dates ranging from 1995Q1 to 2015Q4. For each sample period, in Figure 6 we report the estimate for the  $\phi_{\pi^c, \pi^e}$  coefficient along with 90% confidence bands. Over the entire 1995-2015 period, the  $\phi_{\pi^c, \pi^e}$  estimate is close to zero and statistically insignificant. This shows that the pass-

through effect of energy shocks has been negligible in recent years. These results extend the analysis of, e.g., Clark and Terry (2010), Hooker (2002), and Stock and Watson (2010) to a DTSM setting. Stock and Watson (2010) attribute the decline in the energy pass-through since the 1980s to multiple factors, e.g., they argue that energy is a smaller share of expenditures than it was during the oil price shocks of the 70s, labor union membership has declined sharply over the past forty years, and there has been a shift from production of goods to production of services.

#### 4.4 Nominal Yields Forecasts

While the main focus of our work is on predicting inflation, it is worth noting that our DTSM does very well at forecasting nominal yields too.<sup>24</sup> Table 6 shows the RMSEs for one-year-ahead forecasts of Treasury rates over the 1999-2014 out-of-sample window. Across yields' maturities, the baseline DTSM\* estimated on CPI data outperforms the ARMA benchmark. The  $p$ -values for a West (1996) test of equal forecast accuracy are 6%, 4%, and 2% for the one-quarter, five- and ten-year yields, respectively. The DTSM\* RMSEs are much lower than those for the random walk on short and medium maturity yields: the DTSM\* improves by 32% and 11% on the RW for one-quarter and five-year yields, while it is at par with the RW on the ten-year yield.

Further, Table 6 shows that the DTSM\* outperforms the SPF forecasts of the one-quarter yield with a 20% decline in RMSE. This is remarkable given the role that professional survey forecasts of interest rates play in the term structure literature. For instance, it is common to include survey forecasts of the spot rate to anchor the expectation component of interest rates in the estimation of a DTSM (e.g., Chun (2011), D'Amico, Kim, and Wei (2018), Chernov and Mueller (2012), Kim and Orphanides (2012)). However, our results show that

---

<sup>24</sup>As for the in-sample fit, the root mean squared errors on nominal Treasury yields range from 4.2 to 7.1 basis points across maturities over the 1985–2015 period.

interest rate surveys are not particularly accurate measures of future yields.

The UMR restrictions on core, food, energy inflation, and real activity worsen the DTSM yields' forecasts considerably, especially for short-maturity yields. For instance, the DTSM\* RMSE is 19% lower than the DTSM<sup>UMR</sup> RMSE for the one-quarter yield. In contrast, the DTSM<sup>UMRE</sup>, in which macroeconomic variables are spanned by the yields with the exception of energy inflation, performs on par with the baseline DTSM\*.

Table 6, Panel C, reports similar results obtained by estimating the baseline DTSM\* on PCE inflation data.

## 5 Conclusions

Much of the empirical macro-finance literature finds that financial variables contain little predictive content for consumer price inflation. Nonetheless, this conclusion is at odds with the intuition that the yield curve reflects market participants' expectations of future price dynamics. We address this puzzle with a novel DTSM that includes separate dynamics for core, food, and energy inflation, which then combine into a measure of total inflation that we use to price Treasury bonds. This framework captures the different degrees of persistence and volatility in shocks to the three inflation components. In particular, it downplays the role of short-lived fluctuations in energy prices in determining expectations of future inflation and bond yields.

The model does very well at predicting inflation and nominal yields, with out-of-sample errors that are smaller than the errors produced by popular benchmarks, including survey forecasts. Variance decomposition of the forecasting errors shows that a common set of latent factors shapes the dynamics of nominal yields and core inflation. Thus, our forecasts embody information from yield curve dynamics as well as past price realizations, and the latent factors explain most of the variation in core inflation and bond yields. Taken together, these results

suggest that we find predictive content in the yield curve to forecast future inflation.

Bond yields are the sum of (1) expected inflation and real spot rate paths and (2) real and inflation risk premium components. Our results suggest that the baseline DTSM provides a good estimate of the expected paths and a close fit of the yield curve. Hence, we expect the model to also pin down the residual risk premium components well. In particular, our core and crust framework accommodates a rich specification of the market prices of risk in which distinct inflation variables can command different risk premia. This allows us to elicit which inflation shocks are priced in the term structure and to determine the sources of variation in their risk premia. This setting extends naturally to other macroeconomic aggregates. For instance, we can use it to study the risk premia associated with the various components of consumption and GDP growth. We investigate these issues in a companion article, Ajello, Benzoni, and Chyruk (2018).

## Appendix

### A.1 Nominal Bond Prices

The price of a one-period nominal zero-coupon bond is:

$$\begin{aligned}
p_t^1 &= E_t[m_{t+1}] = E_t \left[ \exp \left( -r_t^* - \pi_{t+1} - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1} \right) \right] \\
&= E_t \left[ \exp \left( -\delta_0 - \delta_1' X_t - \sum_{j=1}^{K_2} \omega^j \pi_{t+1}^j - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1} \right) \right] \\
&= \exp \left( -\delta_0 - \delta_1' X_t - \mu^\pi - \Phi^\pi X_t - \frac{1}{2} \lambda_t' \lambda_t \right) E_t [\exp(-(\lambda_t' + \Sigma^\pi) \varepsilon_{t+1})] . \quad (\text{A.1})
\end{aligned}$$

Since  $\varepsilon_{t+1} \sim N(0, I)$ , then  $E_t[\exp(-(\lambda_t' + \Sigma^\pi) \varepsilon_{t+1})] = \exp(\frac{1}{2}(\lambda_t' + \Sigma^\pi)(\lambda_t' + \Sigma^\pi)')$ . Substituting in equation (A.1) and rearranging terms we obtain

$$p_t^1 = \exp \left( -\delta_0 - \delta_1' X_t - \mu^\pi - \Phi^\pi X_t + \frac{1}{2} \Sigma^\pi \Sigma^{\pi'} + \Sigma^\pi (\lambda_0 + \lambda_1 X_t) \right) = \exp(\bar{A}_1 + \bar{B}_1' X_t) , \quad (\text{A.2})$$

where  $\bar{A}_1 = -\delta_0 - \mu^{Q,\pi} + \frac{1}{2} \Sigma^\pi \Sigma^{\pi'}$  and  $\bar{B}_1 = -\delta_1' - \Phi^{Q,\pi}$ .

Assume now that equation (12) prices a nominal  $n$ -period zero-coupon bond. Then, the same formula prices an  $(n + 1)$ -period bond. To verify this claim, combine equations (10)-(12):

$$\begin{aligned}
p_t^{n+1} &= E_t \left[ \exp \left( -r_t^* - \sum_{j=1}^{K_2} \omega^j \pi_{t+1}^j - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1} + \bar{A}_n + \bar{B}_n' X_{t+1} \right) \right] \\
&= \exp \left( -\delta_0 - \mu^\pi + \bar{B}_n' \mu - \frac{1}{2} \lambda_t' \lambda_t + \bar{A}_n + (-\delta_1' - \Phi^\pi + \bar{B}_n' \Phi) X_t \right) \\
&\quad \times E_t \left[ \exp \left( (-\lambda_t' - \Sigma^\pi + \bar{B}_n' \Sigma) \varepsilon_{t+1} \right) \right] \\
&= \exp \left( -\delta_0 + \bar{A}_n + \bar{B}_n' \mu^Q - \mu^{Q,\pi} + \frac{1}{2} \bar{B}_n' \Sigma \Sigma' \bar{B}_n + \frac{1}{2} \Sigma^\pi \Sigma^{\pi'} - \bar{B}_n' \Sigma \Sigma^{\pi'} \right. \\
&\quad \left. + (-\delta_1' - \Phi^{Q,\pi} + \bar{B}_n' \Phi^Q) X_t \right). \tag{A.3}
\end{aligned}$$

We collect terms linear in  $X_t$  and independent of  $X_t$  to obtain the ODEs (13).

## A.2 Benchmark Models

Here we provide details on the specification and estimation of the benchmark models in Section 2.5.

### A.2.1 Univariate Inflation Models: ARMA(1,1) and RW

The ARMA(1,1) model for an inflation series  $\pi^i$  is

$$\pi_t^i = \mu + \rho \pi_{t-1}^i + \varepsilon_t + \theta \varepsilon_{t-1}. \tag{A.4}$$

Estimation is by maximum likelihood.

As in Atkeson and Ohanian (2001), the random walk forecast for an inflation series at any future horizon  $J$  is the average of the realizations during the past four quarters:

$$E_t[\pi_{t+J,J}^{tot}] = \sum_{j=1}^4 \pi_{t-j+1}^i, \tag{A.5}$$

where  $\pi_{t+J,J}^i = \sum_{j=1}^J \pi_{t+j}^i$ .

### A.2.2 Multivariate Inflation Models: $\text{ARMA}_W(1,1)$ and $\text{VAR}_\Pi$

The  $\text{ARMA}_W(1,1)$  forecast of total inflation is

$$E_t[\pi_{t+J,J}^{tot}] = \omega^c E_t[\pi_{t+J,J}^c] + \omega^f E_t[\pi_{t+J,J}^f] + \omega^e E_t[\pi_{t+J,J}^e], \quad (\text{A.6})$$

where  $E_t[\pi_{t+J,J}^c]$ ,  $E_t[\pi_{t+J,J}^f]$ , and  $E_t[\pi_{t+J,J}^e]$  are the  $\text{ARMA}(1,1)$  forecasts of core, food, and energy inflation at horizon  $J$ . The weights  $\omega^c$ ,  $\omega^f$ , and  $\omega^e$  are the relative importance of the core, food, and energy indices in the total consumer price index (see Appendix A.3 for more details). Also in this case, we estimate each  $\text{ARMA}(1,1)$  model by maximum likelihood.

Similarly, the  $\text{VAR}_\Pi$  forecast of total inflation is also computed according to equation (A.6), except that  $E_t[\pi_{t+J,J}^c]$ ,  $E_t[\pi_{t+J,J}^f]$ , and  $E_t[\pi_{t+J,J}^e]$  are the core, food, and energy inflation forecasts derived from the OLS estimates of an unconstrained VAR for the three inflation factors with one lag.

### A.2.3 Maximal DTSMs and Dynamic Factor Models

The  $\text{DTSM}^{Max}$  model extends the  $\text{VAR}_\Pi$  by including the first three principal components of the nominal yields and real activity in addition to the three inflation series. The  $\text{DTSM}^{Max,Tot}$  and  $\text{DTSM}^{Max,Core}$  follow the same approach of the  $\text{DTSM}^{Max}$ , except that we use a single measure of inflation (either total or core) instead of the three individual components.

The  $\text{DFM}^{CT}$  is a state-space model with state dynamics identical to those of our preferred DTSM. Similar to our DTSM, the observation equation includes the first five principal components of the nominal Treasury yields. However, in the  $\text{DFM}^{CT}$  the factor loadings are free coefficients, while in the DTSM they are determined by the solution of the ODEs in equation (13). Estimation proceeds via maximum likelihood with the Kalman filter. The more general DFM follows the same approach of the  $\text{DFM}^{CT}$ , except that we remove the restrictions that the latent factors have a central tendency specification.



### A.2.4 Surveys

**The University of Michigan Surveys of Consumers (UMSC)** Each month, the University of Michigan surveys a sample of households representative of all American households asking them approximately 50 core questions, each of which tracks a different aspect of consumer attitudes and expectations. Among these, we focus on the median answers to the question “By about what percent do you expect prices to go (up/down), on the average, during the next 12 months?” reported from January 1978 in Table 32 on the UMSC Internet site at <https://data.sca.isr.umich.edu/>

**Blue Chip Economic Indicators** The Blue Chip Economic Indicators surveys business economists monthly and collects their forecasts of U.S. macroeconomic variables and other indicators of future business activity. Among these, we focus on the forecasts of total CPI inflation, available to us since 1980.

**The Survey of Professional Forecasters** The Survey of Professional Forecasters is administered quarterly by the Federal Reserve Bank of Philadelphia. We obtain median forecasts for total CPI inflation and the nominal three-month and ten-year U.S. Treasury yields at the Internet site <https://www.philadelphiafed.org/research-and-data>. The series start in 1981Q3 (inflation and three-month yield) and 1992Q1 (ten-year yield).

**Forecast Timing** In comparing our model to survey forecasts, we need to match the information set of the forecasters at the time they formulate a forecast with the information set available to the econometrician when estimating the DTSM. To fix ideas, consider the case of one-year-ahead DTSM forecasts computed with data through December of a given year. While December Treasury yields are available in (nearly) real time, December inflation data is released in the second half of the following January.

In the case of the BC survey, we have two options:

1. We can use BC inflation forecasts for quarters Q1-Q4 of the following year released in December. In this case, we put BC forecasters at a disadvantage compared to the model, since they will not have seen the December CPI release and some of the December yields at the time they turned in their forecasts.
2. To make sure that the BC forecasters incorporate the December CPI release in their forecast, we could instead use the February BC release. However, such forecasts reflect much other information that is excluded from the DTSM information set. For instance, by the end of January professional forecasters have observed at least one additional month of Treasury yields, spot oil prices, energy futures, breakeven inflation rates, inflation swaps, etc., all of which help them improve their forecasts for the first quarter well beyond what our DTSM could do. In particular, the Q1 inflation forecast reported in the February BC release is a “nowcast,” since it refers to a period for which the variable of interest is partially observed.<sup>25</sup> Not surprisingly, it is well known that inflation nowcasts are very accurate (e.g., Faust and Wright (2013)). Also, since they are not “pure forecasts,” many authors do not use them at all (e.g., Aruoba (2016)).

To address these issues, we draw three comparisons between our model and the BC survey forecasts:

- First, we compare one-year-ahead DTSM forecasts to BC forecasts released in the last month of model estimation (e.g., December for a sample period ending in Q4).
- Second, we construct DTSM forecasts for the one-year period that spans quarters from 2 to 5 after the end of the estimation period. We compare them to the corresponding

---

<sup>25</sup>A third alternative would be to pair the December DTSM forecasts with the January release of the BC forecasts. There are two problems that lead us to dismiss this option. First, similar to the December BC release, at the time of the January release the BC forecasters have not yet seen the December CPI number. Second, the Q1 forecast in the January BC release is a nowcast and is therefore subject to the same caveats we face with the February release.

BC forecasts from the release in the second month after the end of the estimation period (e.g., February for a sample period ending in Q4). That is, we exclude the BC nowcast from the comparison.

- Third, for completeness and full transparency, we also compare the one-quarter-ahead DTSM forecasts to the Blue-Chip one-quarter nowcast released in the second month after the end of the estimation period (e.g., February for a sample period ending in Q4).

All comparisons have merit, but the first and the second ones are the most relevant in our case. In both cases, they show that our model outperforms survey forecasts, even when, in the second case, forecasters are given 1-2 months of additional information compared to our model. The third comparison confirms that forecasters produce very informed *nowcasts*. However, when we look at *forecasts*, our model outperforms professional forecasters.

Similar issues arise with the SPF; we address them in the same way.

### A.3 Core, food, and energy weights

Market participants deflate nominal asset prices in equation (11) at the total inflation rate,  $\pi_t$ . In the model that has three inflation factors, we compute  $\pi_t$  as the weighted sum of the core, food, and energy inflation series. That is,  $\pi_t = \pi_t^{tot} = \omega_t^c \pi_t^c + \omega_t^f \pi_t^f + \omega_t^e \pi_t^e$ , where the weights  $\omega_t^c$ ,  $\omega_t^f$ , and  $\omega_t^e$  represent the relative importance of core, food, and energy prices in the total price index at time  $t$ . This appendix describes how we construct such weights.

#### A.3.1 Consumer price index weights

For the CPI weights we use the relative importance of core, food, and energy in the CPI reported by the Bureau of Labor Statistics (BLS). The relative importance of a component is the percentage share of the expenditure on that component relative to the expenditure on all items within an area. The BLS conducts a Consumer Expenditure Survey to deter-

mine how these shares change over time to reflect fluctuations in the consumption patterns of the population. Each year since 1987, the BLS releases the December value of these series based on the core, food, and energy consumption baskets for that year. Monthly fluctuations in prices result in changes in the relative importance shares for these baskets compared to the values reported the previous December. To account for this pattern, we update the value of the December shares to obtain monthly series that reflect the changes in the cost to purchase the same food, core, and energy baskets. The BLS Internet site at [http://www.bls.gov/cpi/cpi\\_riar.htm](http://www.bls.gov/cpi/cpi_riar.htm) explains in details how to do that. The BLS does not make relative importance shares broadly available for years prior to 1987. We thank the BLS for sharing such data with us.

### **A.3.2 Personal consumption expenditures weights**

Similar to the CPI weights, the PCE weights are the shares of the expenditures on the core, food, and energy baskets relative to total personal consumption expenditures. To compute these shares, we use data from the national income and product account (NIPA) Table 2.3.5U, Personal Consumption Expenditures by Major Type of Product and by Major Function. The variables are (1) Personal consumption expenditures; (2) Personal consumption expenditures excluding food and energy; (3) Food and beverages purchased for off-premises consumption; and (4) Energy goods and services.

## Figures and Tables

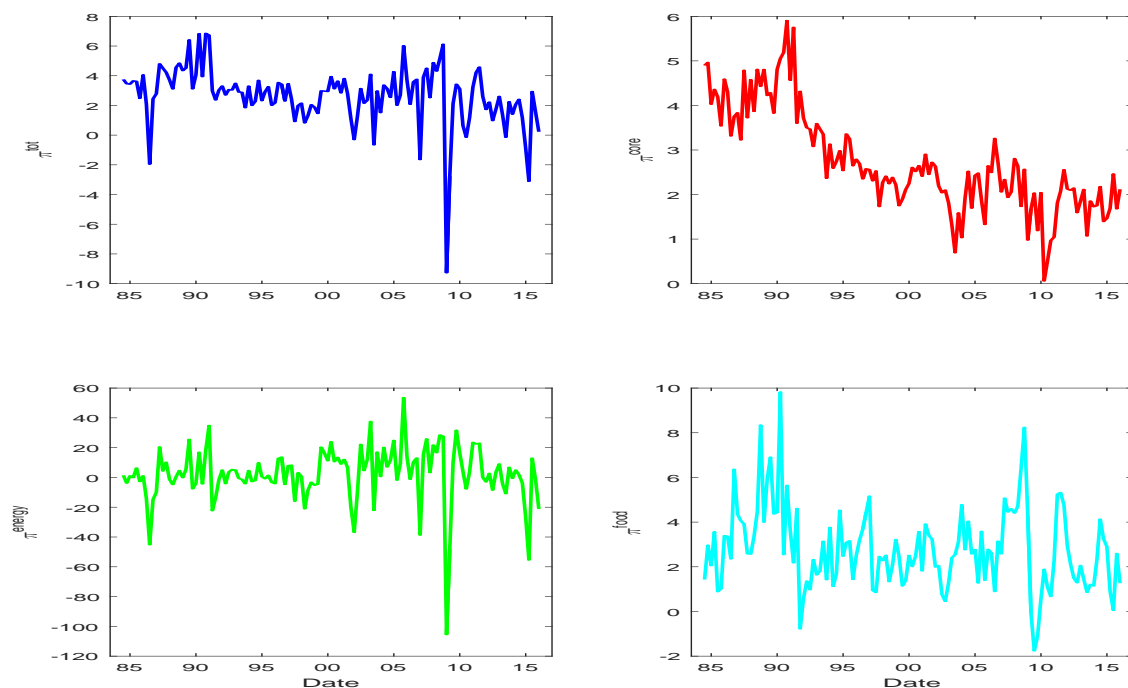


Figure 1: **CPI Inflation Series.** The plots depict total, core, food, and energy quarterly CPI inflation series. The sample period is 1985Q1-2015Q4.

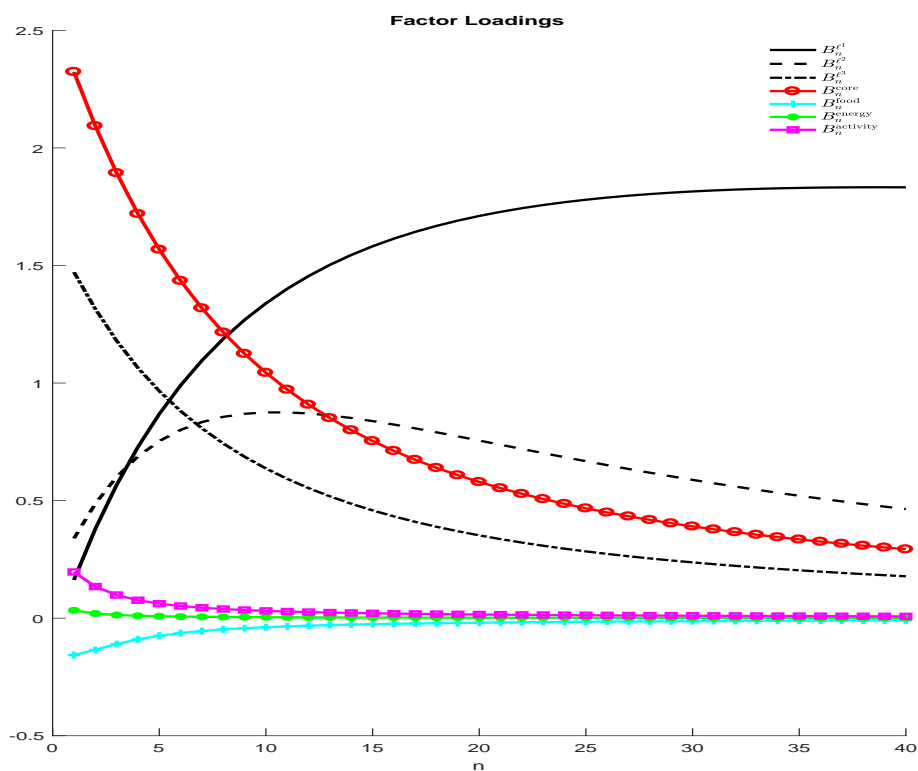


Figure 2: **DTSM\* Factor Loadings.** The plot depicts the factor loadings for nominal yields on the latent factors ( $B_n^{\ell^1}$ ,  $B_n^{\ell^2}$ , and  $B_n^{\ell^3}$ ), inflation factors ( $B_n^{core}$ ,  $B_n^{food}$ , and  $B_n^{energy}$ ), and real activity  $B_n^{activity}$  where  $n$  denotes quarters to maturity. Factor loadings are scaled to correspond to one standard deviation movement in the factors. The sample period is 1985Q1-2015Q4.

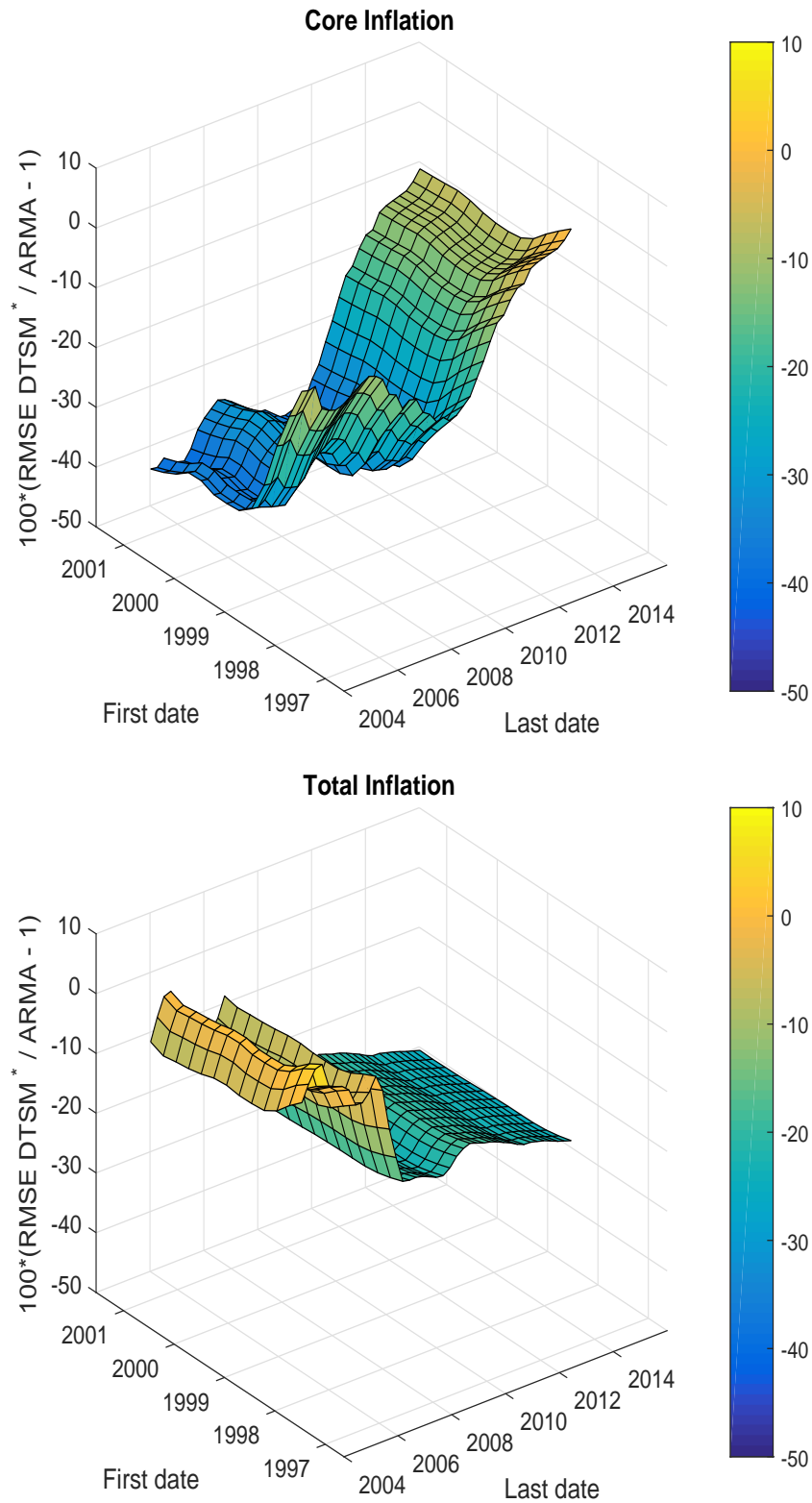


Figure 3: **RMSE Percentage Ratios: DTSM\* vs. ARMA(1,1)**. For the DTSM\* and ARMA models, we compute one-year-ahead RMSEs over a grid of out-of-sample windows with start date ranging from 1997Q4 to 2000Q4 and end date from 2004Q4 to 2014Q4. The figure displays their percentage ratio,  $100 \times (\text{RMSE DTSM}^* / \text{RMSE ARMA} - 1)$ . Negative numbers in the plot signal that the DTSM\* outperforms the ARMA. The top panel shows results for core CPI inflation, while results for total CPI inflation are in the bottom panel.

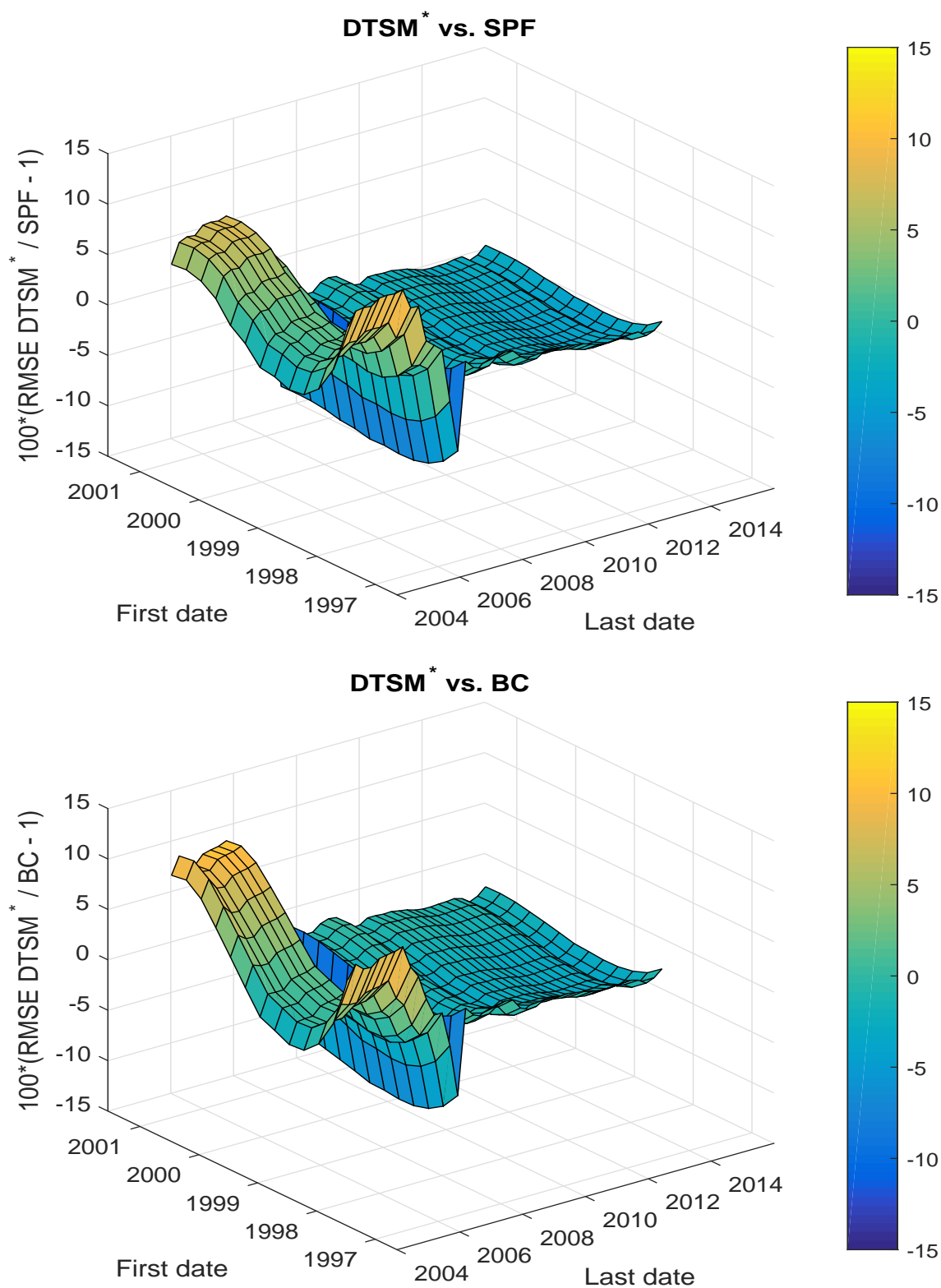


Figure 4: **RMSE Percentage Ratios: DTSM\* vs. Survey Forecasts.** We compute RMSEs for the DTSM\* and SPF / BC total inflation forecasts one year out over a grid of out-of-sample windows with start date ranging from 1997Q4 to 2000Q4 and end date from 2004Q4 to 2014Q4. The figure displays the percentage ratios  $100 \times (\text{RMSE DTSM}^* / \text{RMSE Surveys} - 1)$ . Negative numbers in the plot signal that the DTSM\* outperforms the survey forecasts.

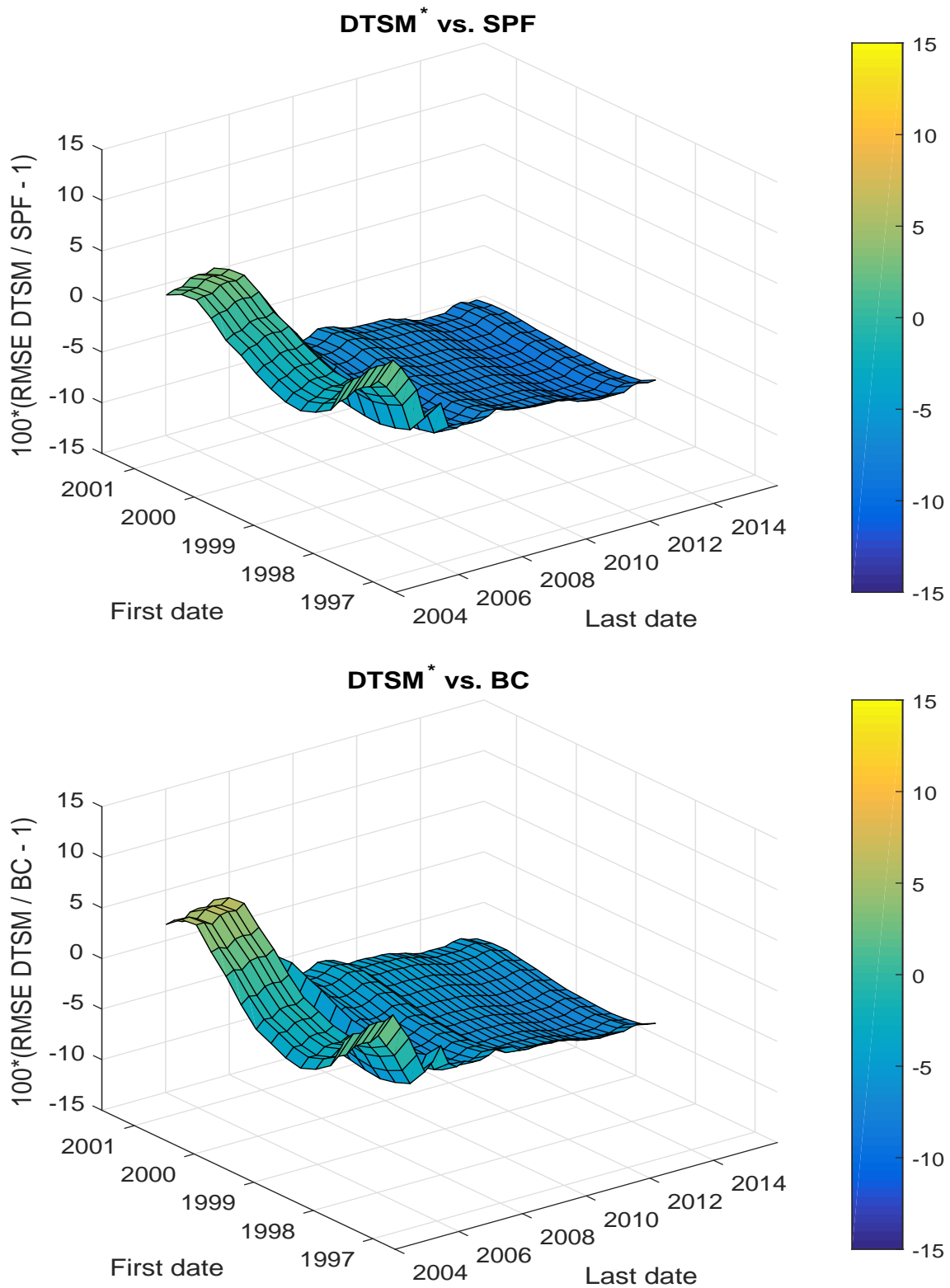


Figure 5: **RMSE Percentage Ratios, 2-5 Quarters Ahead: DTSM\* vs. Survey Forecasts.** We compute RMSEs for the DTSM\* and SPF / BC total inflation forecasts 2-to-5 quarters ahead over a grid of out-of-sample windows with start date ranging from 1997Q4 to 2000Q4 and end date from 2004Q4 to 2014Q3. The figure displays the percentage ratios  $100 \times (\text{RMSE DTSM}^* / \text{RMSE Surveys} - 1)$ . Negative numbers in the plot signal that the DTSM\* outperforms the survey forecasts.



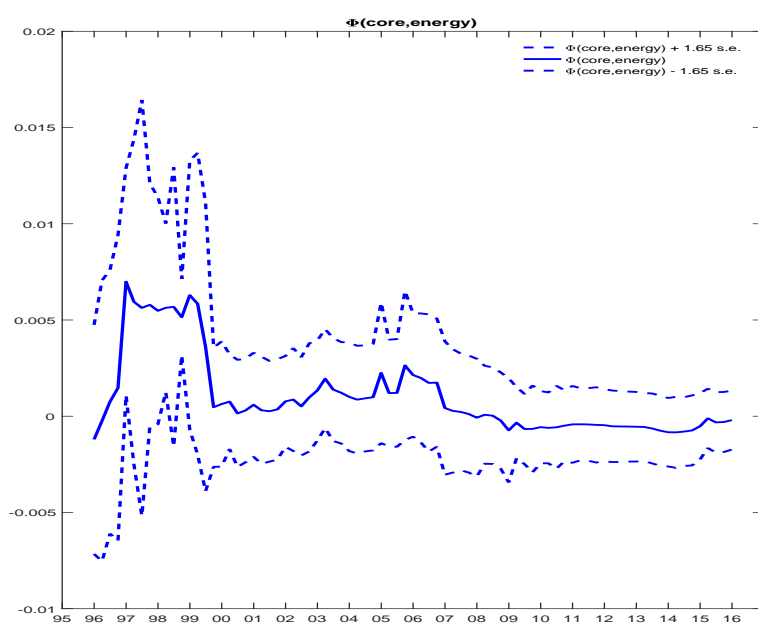


Figure 6: **Energy Pass Through.** We repeatedly estimate the model using data that start in 1985Q1 and end on dates that range from 1995Q4 to 2015Q4. For each sample period, the plot shows the estimate of the  $\phi_{\pi^c, \pi^e}$  coefficient and its 90% confidence bands.

Table 1: **Summary Data Statistics.** The table reports summary statistics for CPI and PCE inflation series on core, food, energy and total consumer price indices; as well as CPI and PCE measures of relative importance weights for the core, food, and energy price indices. CPI- and PCE-weighted are the total inflation series computed from their core, food, and energy components using the relative importance weights. The sample period is 1985Q1-2015Q4.

	Central moments				Autocorrelations		
	Mean	Std. Dev.	Skewness	Kurtosis	AC(1)	AC(2)	AC(3)
Panel A: CPI inflation and weights							
CPI	2.65	2.01	-1.92	12.35	0.32	0.07	0.13
CPI-weighted	2.69	1.92	-1.48	9.59	0.34	0.08	0.15
CPI-core	2.72	1.15	0.56	2.93	0.84	0.84	0.79
CPI-food	2.79	1.83	0.82	4.76	0.49	0.35	0.22
CPI-energy	2.09	18.88	-1.76	11.22	0.22	-0.13	-0.03
Weight-core	0.77	0.02	-2.26	7.80	0.92	0.83	0.76
Weight-food	0.15	0.01	1.08	3.82	0.96	0.92	0.88
Weight-energy	0.08	0.01	0.84	2.74	0.92	0.82	0.78
Panel B: PCE inflation and weights							
PCE	2.23	1.52	-1.30	8.68	0.44	0.19	0.23
PCE-weighted	2.25	1.48	-1.06	7.37	0.45	0.20	0.25
PCE-core	2.25	1.03	0.85	3.20	0.76	0.75	0.72
PCE-food	2.30	2.16	0.59	4.03	0.42	0.30	0.09
PCE-energy	2.17	19.59	-1.74	10.96	0.21	-0.13	-0.02
Weight-core	0.86	0.02	-1.10	3.55	0.95	0.91	0.86
Weight-food	0.09	0.01	0.82	2.31	0.98	0.95	0.93
Weight-energy	0.05	0.01	0.80	4.23	0.91	0.81	0.72

Table 2: **Variance Decomposition.** We use the baseline DTSM\* to decompose the variation in nominal yields, inflation, and real activity into proportions associated with shocks to the latent factors  $\ell^1$ ,  $\ell^2$ , and  $\ell^3$ ; core, food, and energy inflation; and real activity. In Panels A-C, we decompose the variance of the one-quarter, five- and ten-year yield; Panels D-H show a similar decomposition for the variance of core, food, energy, and total CPI inflation, and real activity. The sample period is 1985Q1-2015Q4.

Horizon (quarters):	4	8	12	20	$\infty$
Panel A: 1-quarter yield					
$\ell^1$	15.17	28.31	37.06	48.79	65.54
$\ell^2$	52.49	56.32	52.34	43.52	29.26
$\ell^3$	2.92	1.73	1.27	0.96	0.69
CPI-core	21.84	10.08	6.89	4.97	3.35
CPI-food	2.07	0.97	0.66	0.47	0.31
CPI-energy	0.12	0.05	0.03	0.02	0.02
Activity	5.39	2.55	1.74	1.25	0.83
Panel B: 5-year yield					
$\ell^1$	63.16	69.12	74.03	80.28	86.95
$\ell^2$	35.11	29.81	25.10	18.97	12.39
$\ell^3$	0.34	0.33	0.30	0.26	0.23
CPI-core	1.25	0.66	0.51	0.44	0.38
CPI-food	0.02	0.02	0.01	0.01	0.01
CPI-energy	0.00	0.00	0.00	0.00	0.00
Activity	0.12	0.07	0.05	0.04	0.04
Panel C: 10-year yield					
$\ell^1$	83.66	86.75	89.28	92.18	94.27
$\ell^2$	15.67	12.73	10.24	7.34	5.25
$\ell^3$	0.27	0.28	0.26	0.23	0.21
CPI-core	0.35	0.21	0.20	0.22	0.25
CPI-food	0.01	0.00	0.00	0.00	0.00
CPI-energy	0.00	0.00	0.00	0.00	0.00
Activity	0.03	0.02	0.02	0.02	0.02

Table 2, continued

Horizon (quarters):	4	8	12	20	$\infty$
Panel D: Core CPI Inflation					
$\ell^1$	4.84	18.21	29.11	42.47	60.20
$\ell^2$	2.30	9.79	13.02	12.59	9.12
$\ell^3$	75.37	57.83	46.48	36.08	24.56
CPI-core	16.42	13.12	10.54	8.20	5.66
CPI-food	0.03	0.07	0.06	0.05	0.03
CPI-energy	0.00	0.00	0.00	0.00	0.00
Activity	1.03	0.99	0.80	0.62	0.43
Panel E: Food CPI Inflation					
$\ell^1$	0.03	0.41	1.04	2.23	4.96
$\ell^2$	0.23	0.54	0.87	1.11	1.14
$\ell^3$	0.85	0.86	0.86	0.85	0.83
CPI-core	16.31	16.24	16.08	15.85	15.40
CPI-food	81.62	80.95	80.16	78.99	76.72
CPI-energy	0.00	0.00	0.00	0.00	0.00
Activity	0.96	1.00	0.99	0.97	0.95
Panel F: Energy CPI Inflation					
$\ell^1$	0.01	0.13	0.34	0.75	1.72
$\ell^2$	0.07	0.13	0.22	0.28	0.30
$\ell^3$	0.42	0.42	0.42	0.42	0.42
CPI-core	8.67	8.69	8.67	8.63	8.55
CPI-food	2.35	2.37	2.37	2.35	2.33
CPI-energy	88.13	87.90	87.64	87.22	86.35
Activity	0.34	0.35	0.35	0.35	0.34
Panel G: Total CPI Inflation					
$\ell^1$	0.35	1.46	2.70	4.76	9.18
$\ell^2$	0.36	1.02	1.46	1.68	1.68
$\ell^3$	6.41	6.34	6.23	6.09	5.81
CPI-core	15.59	15.31	15.05	14.69	14.01
CPI-food	2.52	2.51	2.47	2.41	2.30
CPI-energy	74.00	72.61	71.35	69.65	66.34
Activity	0.76	0.75	0.74	0.72	0.69

Table 2, continued

Horizon (quarters):	4	8	12	20	$\infty$
Panel H: Real Activity					
$\ell^1$	0.02	0.14	0.28	0.52	1.08
$\ell^2$	14.40	14.77	14.99	15.09	15.03
$\ell^3$	1.25	1.28	1.28	1.27	1.26
CPI-core	13.26	13.17	13.11	13.06	12.98
CPI-food	6.47	6.50	6.47	6.45	6.41
CPI-energy	0.00	0.00	0.00	0.00	0.00
Activity	64.60	64.14	63.87	63.61	63.24

Table 3: **Forecasts of Annual Inflation Series.** We repeatedly estimate each model using data that start in 1985Q1 and end on dates ranging from 1999Q4 to 2014Q4; for each of these sample periods we forecast inflation one year out. For each model, the table shows RMSEs in percent per year and  $p$ -values for a test of equal forecast accuracy (West (1996)) computed under the null that the RMSE for that model equals the RMSE for the ARMA(1,1), when the alternative is that the RMSE for the ARMA(1,1) exceeds the RMSE for that model.

Panel A: Time series models and survey forecasts of CPI inflation

	Total		Core		Food		Energy	
	RMSE	$p$ -val.	RMSE	$p$ -val.	RMSE	$p$ -val.	RMSE	$p$ -val.
Time series models of inflation								
ARMA	1.60		0.53		1.49		12.95	
ARMA <sub>W</sub>	1.43	0.14						
VAR <sub>II</sub>	1.47	0.06	0.63	0.93	1.43	0.32	13.11	0.64
RW	1.81	0.87	0.56	0.85	1.98	0.99	18.75	0.97
Surveys without nowcast								
U. of M.	1.80	0.95						
SPF	1.30	0.05						
BC	1.28	0.05						
Surveys with nowcast								
U. of M.	1.74	0.89						
SPF	1.22	0.03						
BC	1.21	0.03						

Panel B: DTSMs estimated on CPI data

	Total		Core		Food		Energy	
	RMSE	<i>p</i> -val.	RMSE	<i>p</i> -val.	RMSE	<i>p</i> -val.	RMSE	<i>p</i> -val.
Baseline DTSM								
DTSM <sup>*</sup>	1.19	0.03	0.50	0.35	1.31	0.15	12.96	0.50
Over-identifying restrictions								
DTSM <sup>Max</sup>	1.41	0.10	0.48	0.15	1.46	0.45	14.16	0.91
DFM	1.32	0.06	0.49	0.28	1.42	0.31	13.28	0.72
DFM <sup>CT</sup>	1.33	0.07	0.49	0.26	1.42	0.32	13.31	0.74
DTSM <sup>Max,Tot</sup>	1.33	0.06						
DTSM <sup>Max,Core</sup>			0.48	0.16				
DTSM <sup>Tot</sup>	1.61	0.51						
DTSM <sup>Core</sup>			0.66	0.84				

Panel C: Robustness checks

Yields series								
DTSM <sup>Yields</sup>	1.19	0.03	0.53	0.49	1.30	0.15	12.87	0.45
DTSM <sup>G<sub>SW</sub></sup>	1.26	0.07	0.64	0.84	1.49	0.49	13.00	0.53
Real activity measures								
DTSM <sup>GDP</sup>	1.19	0.03	0.52	0.46	1.30	0.17	12.83	0.42
DTSM <sup>CFNAI</sup>	1.23	0.05	0.56	0.64	1.25	0.19	12.82	0.40
TIPS and surveys								
DTSM <sup>TIPS</sup>	1.28	0.09	0.51	0.37	1.32	0.23	12.49	0.19
DTSM <sup>BC</sup>	1.26	0.05	0.54	0.51	1.32	0.20	13.20	0.76
DTSM <sup>SPF</sup>	1.25	0.04	0.53	0.48	1.32	0.18	13.26	0.76
DTSM <sup>TIPS+BC</sup>	1.27	0.05	0.47	0.15	1.33	0.25	13.07	0.64
DTSM <sup>TIPS+SPF</sup>	1.27	0.05	0.47	0.13	1.32	0.23	13.12	0.66
Monthly frequency								
DTSM <sup>Monthly</sup>	1.18	0.04	0.59	0.74	1.27	0.15	12.75	0.31
Unspanned macroeconomic factors								
DTSM <sup>UMRE</sup>	1.20	0.03	0.52	0.43	1.31	0.15	13.03	0.55
DTSM <sup>UMR</sup>	1.27	0.05	0.47	0.11	1.36	0.13	13.01	0.61

Panel D: Time series models and DTSMs estimated on PCE data

	Total		Core		Food		Energy	
	RMSE	<i>p</i> -val.	RMSE	<i>p</i> -val.	RMSE	<i>p</i> -val.	RMSE	<i>p</i> -val.
Time series models of inflation								
ARMA	1.19		0.41		1.87		13.49	
ARMA <sub>W</sub>	1.03	0.12						
VAR <sub>II</sub>	1.07	0.15	0.51	1.00	1.69	0.19	13.83	0.90
RW	1.33	0.82	0.42	0.56	2.48	0.98	19.52	0.97
Baseline DTSM								
DTSM*	0.99	0.18	0.56	0.91	1.71	0.20	14.05	0.89

Table 4: **Long-Run Inflation Forecasts.** We repeatedly estimate each model using data that start in 1985Q1 and end on dates ranging from 1999Q4 to 2014Q4; for each of these sample periods we forecast inflation from one quarter to five years out. For each model, the table shows RMSEs in percent per year and  $p$ -values for a test of equal forecast accuracy (West (1996)) computed under the null that the RMSE for that model equals the RMSE for the ARMA(1,1), when the alternative is that the RMSE for the ARMA(1,1) exceeds the RMSE for that model.

	CPI		CPI-core		CPI-food		CPI-energy	
	ARMA	DTSM*	ARMA	DTSM*	ARMA	DTSM*	ARMA	DTSM*
1Q	2.51	2.30 (0.08)	0.56	0.58 (0.60)	1.51	1.47 (0.37)	23.37	23.52 (0.60)
1Y	1.60	1.19 (0.03)	0.53	0.50 (0.35)	1.49	1.31 (0.15)	12.95	12.96 (0.50)
2Y	1.19	0.80 (0.04)	0.57	0.49 (0.20)	1.23	0.96 (0.09)	8.12	8.28 (0.57)
3Y	1.03	0.67 (0.04)	0.61	0.49 (0.08)	0.86	0.64 (0.10)	6.80	6.84 (0.52)
4Y	0.94	0.58 (0.03)	0.61	0.42 (0.01)	0.70	0.54 (0.19)	6.16	6.00 (0.42)
5Y	0.86	0.49 (0.04)	0.61	0.33 (0.00)	0.67	0.57 (0.29)	5.35	4.98 (0.32)

Table 5: **Total CPI Inflation Forecasts: DTSM\* vs. Surveys.** We repeatedly estimate the DTSM\* using data that start in 1985Q1 and end on dates ranging from 1999Q4 to 2014Q4; for each of these sample periods we forecast inflation from one to five quarters out. The table shows RMSEs for the DTSM\* and the corresponding SPF/BC survey forecasts in percent per year. In parentheses are the  $p$ -values for a test of equal forecast accuracy (West (1996)) computed under the null that the RMSE for the DTSM\* equals the SPF/BC RMSE, when the alternative is that the SPF/BC RMSE exceeds the RMSE for the DTSM\*.

	SPF	DTSM*	BC	DTSM*
Panel A: Without Nowcast				
1-4Q	1.30	1.19 (0.01)	1.28	1.19 (0.01)
Panel B: With Nowcast				
1Q (nowcast)	1.42	2.24 (0.94)	1.46	2.24 (0.95)
2-5Q	1.31	1.20 (0.02)	1.28	1.20 (0.02)



Table 6: **Treasury Yields Forecasts.** We repeatedly estimate each model using data that start in 1985Q1 and end on dates ranging from 1999Q4 to 2014Q4; for each of these sample periods we forecast nominal Treasury yields one year out. For each model, the table shows RMSEs in percentage per year and  $p$ -values for a test of equal forecast accuracy (West (1996)) computed under the null that the RMSE for that model equals the RMSE for the ARMA(1,1), when the alternative is that the RMSE for the ARMA(1,1) exceeds the RMSE for that model.

	1Q Yield		5Y Yield		10Y Yield	
	RMSE	$p$ -val.	RMSE	$p$ -val.	RMSE	$p$ -val.
Panel A: Univariate interest rates models and survey forecasts						
ARMA	1.38		1.10		0.89	
RW	1.77	0.99	1.00	0.12	0.73	0.00
SPF	1.52	0.88			1.02	0.99
Panel B: DTSMs, estimation on CPI data						
DTSM*	1.21	0.06	0.89	0.04	0.74	0.02
DTSM <sup>UMRE</sup>	1.21	0.06	0.90	0.04	0.74	0.02
DTSM <sup>UMR</sup>	1.49	0.84	1.01	0.32	0.85	0.43
Panel C: DTSM, estimation on PCE data						
DTSM*	1.23	0.08	0.88	0.02	0.72	0.01

## References

- Abrahams, Michael, Tobias Adrian, Richard K. Crump, and Emanuel Moench, 2013, Pricing TIPS and Treasuries with Linear Regressions, Working Paper, Federal Reserve Bank of New York.
- Adrian, Tobias, and Hao Wu, 2010, The Term Structure of Inflation Expectations, Working Paper, Federal Reserve Bank of New York.
- Ajello, Andrea, Luca Benzoni, and Olena Chyruk, 2018, Bond Risk Premia and the Macroeconomy, Working Paper, Federal Reserve Bank of Chicago.
- Ang, Andrew, Geert Bekaert, and Min Wei, 2008, The Term Structure of Real Rates and Expected Inflation, *Journal of Finance* 63, 797-849.
- Ang, Andrew, Geert Bekaert, and Min Wei, 2007, Do macro variables, asset markets, or surveys forecast inflation better? Original Research Article, *Journal of Monetary Economics* 54, 1163-1212.
- Ang, Andrew, and Monika Piazzesi, 2003, A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, *Journal of Monetary Economics* 50, 745-787.
- Ang, Andrew, Monika Piazzesi, and Min Wei, 2006, What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131, 359-403.
- Atkeson, Andrew, Ohanian, Lee E., 2001, Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review* 25, 2-11.
- Barr, David G., and John Y. Campbell, 1997, Inflation, real interest rates, and the bond market: A study of UK nominal and index-linked government bond prices, *Journal of Monetary Economics* 39, 361-383.

- Bauer, Michael D., and Glenn D. Rudebusch, 2015, Resolving the Spanning Puzzle in Macro-Finance Term Structure Models, Working Paper, Federal Reserve Bank of San Francisco.
- Buraschi, Andrea, and Alexei Jiltsov, 2005, Inflation risk premia and the expectations hypothesis, *Journal of Financial Economics* 75, 429-490.
- Calvet, Laurent E., Adlai J. Fisher, and Liuren Wu, 2018, Staying on Top of the Curve: A Cascade Model of Term Structure Dynamics, *Journal of Financial and Quantitative Analysis*, 1-27.
- Campbell, John Y., Robert Shiller, and Luis M. Viceira, 2009, Understanding Inflation-Indexed Bond Markets, *Brookings Papers on Economic Activity*, 79-138.
- Campbell, John Y., Adi Sunderam, and Luis M. Viceira, 2013, Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds, Working Paper, Harvard University.
- Chen, Ren-Raw, Bo Liu, and Xiaolin Cheng, 2010, Inflation, Fisher Equation, and The Term Structure of Inflation Risk Premia: Theory and Evidence from TIPS, *Journal of Empirical Finance* 17(5), 702-721.
- Chernov, Mikhail, and Philippe Mueller, 2012, The Term Structure of Inflation Forecasts, *Journal of Financial Economics* 106(2), 367-394.
- Christensen, Jens H., Jose A. Lopez, and Glenn D. Rudebusch, 2010, Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields, *Journal of Money, Credit, and Banking* 42, 143-178.
- Chun, Albert L, 2011, Expectations, Bond Yields and Monetary Policy, *Review of Financial Studies* 24, 208-247.
- Cieslak, Anna, and Pavol Povala, 2015, Expected returns in Treasury bonds, *Review of Financial Studies* 28(10), 2859-2901.

- Clark, Todd E., and Stephen J. Terry, 2010, Time Variation in the Inflation Passthrough of Energy Prices, *Journal of Money, Credit and Banking* 42, 1419-1433.
- Cochrane, John H., and Monika Piazzesi, 2005, Bond Risk Premia, *American Economic Review* 95, 138-160.
- Cochrane, John H., and Monika Piazzesi, 2008, Decomposing the Yield Curve, Working Paper, University of Chicago and Stanford University.
- Collin-Dufresne, Pierre, Robert S. Goldstein, and Christopher S. Jones, 2008, Identification of Maximal Affine Term Structure Models, *Journal of Finance* 63, 743-795.
- Dai, Qiang and Kenneth J. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance* 55, 1943-1978.
- D'Amico, Stefania, Don H. Kim, and Min Wei, 2018, Tips from TIPS: The Informational Content of Treasury Inflation-Protected Security Prices, *Journal of Financial and Quantitative Analysis* 53(1), 395-436.
- Diebold, Francis X., Monika Piazzesi, and Glenn Rudebusch, 2005, Modeling bond yields in finance and macroeconomics, *American Economic Review: Papers and Proceedings* 95, 415-420.
- Diebold, Francis X., Glenn D. Rudebusch, S. Borağan Aruoba, 2006, The macroeconomy and the yield curve: a dynamic latent factor approach, *Journal of Econometrics* 131, 309-338.
- Drton, Mathias, and Martyn Plummer, 2016, A Bayesian information criterion for singular models, Working Paper, University of Washington
- Duffee, Gregory R., 2006, Term structure estimation without using latent factors, *Journal of Financial Economics* 79, 507-536.
- Duffee, Gregory R., 2010, Sharpe ratios in term structure models, Working Paper, Johns Hopkins University.

- Duffee, Gregory R., 2011a, Information in (and not in) the term structure, *Review of Financial Studies* 24, 2895-2934.
- Duffee, Gregory R., 2011b, Forecasting with the term structure: the role of no-arbitrage restrictions, Working Paper, Johns Hopkins University.
- Duffie, Darrell, and Rui Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379-406.
- Duffie, Darrell, Jun Pan, and Kenneth J. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump Diffusions, *Econometrica* 68, 1343-1376.
- Estrella, Arturo, and Frederic S. Mishkin, 1997, The predictive power of the term structure of interest rates in Europe and the United States: implications for the European central bank. *European Economic Review* 41, 1375-1401.
- Evans, Charles L., and David A. Marshall, 1998, Monetary policy and the term structure of nominal interest rates: evidence and theory. *Carnegie-Rochester Conference Series on Public Policy* 49, 53-111.
- Evans, Charles L., and David A. Marshall, 2007, Economic determinants of the nominal treasury yield curve, *Journal of Monetary Economics* 54, 1986-2003.
- Faust, Jon, and Jonathan H. Wright, 2009, Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset, *Journal of Business and Economic Statistics* 27, 468-479.
- Faust, Jon, and Jonathan H. Wright, 2013, Forecasting Inflation, In: G. Elliott and A. Timmermann (Eds), *Handbook of Economic Forecasting*, Volume 2A, Chapter 1, 2-56, North Holland, Elsevier.
- Grishchenko, Olesya V., and Jing-zhi Huang, 2013, Inflation Risk Premium: Evidence from the TIPS Market, *Journal of Fixed Income* 22(4), 5-30.

- Gürkaynak, Refet S., Brian Sack, Jonathan H. Wright, 2007, The U.S. Treasury yield curve: 1961 to the present, *Journal of Monetary Economics* 54, 2291-2304.
- Gürkaynak, Refet S., Brian Sack, Jonathan H. Wright, 2010, The TIPS yield curve and Inflation Compensation, *American Economic Journal: Macroeconomics* 2, 70-92.
- Hamilton, James D., and Jing Cynthia Wu, 2012, Identification and Estimation of Gaussian Affine Term Structure Models, *Journal of Econometrics* 168(2), 315-331.
- Haubrich, Joseph G., Pennacchi George G., and Peter H. Ritchken, 2014, Inflation Expectations, Real Rates, and Risk Premia: Evidence from Inflation Swaps, *Review of Financial Studies* 25(5), 1588-1629.
- Hooker, Mark A., 2002, Are Oil Shocks Inflationary? A symmetric and Nonlinear Specifications versus Changes in Regime, *Journal of Money, Credit, and Banking* 34, 540-561.
- Hördahl, Peter, and Oreste Tristani, 2012, Inflation risk premia in the term structure of interest rates, *Journal of the European Economic Association* 10(3), 634-657.
- Hördahl, Peter, and Oreste Tristani, David Vestin, 2006, A joint econometric model of macroeconomic and term-structure dynamics, *Journal of Econometrics* 131, 405-444.
- Joslin, Scott, Anh Le, and Kenneth J. Singleton, 2013, Why Gaussian Macro-Finance Term Structure Models Are (Nearly) Unconstrained Factor-VARs, *Journal of Financial Economics* 109(3), 604-622.
- Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu, 2011, A New Perspective on Gaussian Dynamic Term Structure Models, *Review of Financial Studies* 24(3), 926-970.
- Joslin, Scott, Marcel Pribsch and Kenneth J. Singleton, 2014, Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks. *Journal of Finance* 69(3), 1197-1233.

- Kim, Don H., and Athanasios Orphanides, 2012, Term Structure Estimation with Survey Data on Interest Rate Forecasts, *Journal of Financial and Quantitative Analysis* 47(01), 241-272
- Kim, Don H., and Jonathan H. Wright, 2005, An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates, Working paper, Board of Governors of the Federal Reserve System.
- Mehra, Yash P., 2002, Survey measures of expected inflation: revisiting the issues of predictive content and rationality, *Federal Reserve Bank of Richmond Economic Quarterly* 88, 17-36.
- Moench, Emanuel, 2008, Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach, *Journal of Econometrics* 146, 26-43.
- Pennacchi, George G., 1991, Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data, *The Review of Financial Studies* 4, 53-86.
- Piazzesi, Monika, 2005, Bond yields and the Federal Reserve, *Journal of Political Economy*, 113, 311-344.
- Piazzesi, Monika, 2010, Affine Term Structure Models. In Yacine Aït-Sahalia and Lars Peter Hansen (Eds), *Handbook of Financial Econometrics, Volume 1*, Chapter 12, pp. 691-766, North Holland, Elsevier.
- Rudebusch, Glenn D., and Tao Wu, 2008, A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy, *Economic Journal* 118, 906-926.
- Souleles, Nicholas S., 2004, Expectations, heterogeneous forecast errors and consumption: micro evidence from the Michigan consumer sentiment surveys, *Journal of Money, Credit and Banking* 36, 39-72.
- Stock, James H., Mark W. Watson, 1999, Forecasting inflation, *Journal of Monetary Economics* 44, 293-335.

- Stock, James H., Mark W. Watson, 2003, Forecasting Output and Inflation: The Role of Asset Prices, *Journal of Economic Literature* 41, 788-829.
- Stock, James H., Mark W. Watson, 2007, Why Has U.S. Inflation Become Harder to Forecast?, *Journal of Money, Credit and Banking*, Supplement to Vol. 39, 3-33.
- Stock, James H., Mark W. Watson, 2010, Modeling Inflation After the Crisis, in *Macroeconomic Policy: Post-Crisis and Risks Ahead*, Proceedings of the Federal Reserve Bank of Kansas City, 2010 Jackson Hole Symposium.
- Thomas, Lloyd B., 1999, Survey measures of expected U.S. inflation, *Journal of Economic Perspectives* 13, 125-144.
- West, Kenneth D. 1996, Asymptotic Inference about Predictive Ability, *Econometrica* 64, 1067-1084.
- Wright, Jonathan H., 2011, Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset, *American Economic Review* 101(4): 1514-1534.