A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations[*]

Joshua C.C. Chan
University of Technology Sydney

Todd Clark
Federal Reserve Bank of Cleveland

Gary Koop
University of Strathclyde

May 10, 2017

Abstract: A knowledge of the level of trend inflation is key to many current policy decisions and several methods of estimating trend inflation exist. This paper adds to the growing literature which uses survey-based long-run forecasts of inflation to estimate trend inflation. We develop a bivariate model of inflation and long-run forecasts of inflation which allows for the estimation of the link between trend inflation and the long-run forecast. Thus, our model allows for the possibilities that long-run forecasts taken from surveys can be equated with trend inflation, that the two are completely unrelated, or anything in between. By including stochastic volatility and time-variation in coefficients, it extends existing methods in empirically important ways. We use our model with a variety of inflation measures and survey-based forecasts for several countries. We find that long-run forecasts can provide substantial help in refining estimates and fitting and forecasting inflation. The same evidence indicates it is less helpful to simply equate trend inflation with the long-run forecasts.

Keywords: trend inflation, inflation expectations, state space model, stochastic volatility

JEL Classification: C11, C32, E31

1 Introduction

As is evident in public commentary (see, e.g., Bernanke 2007 and Mishkin 2007), central bankers and other policymakers pay considerable attention to measures of long-run inflation expectations. These expectations are viewed as shedding light on the credibility of monetary policy. Monetary policy tools work differently if long-run inflation expectations are firmly anchored than if they are not. In general, monetary policy is thought to be most effective when long-run inflation expectations are stable.

These considerations have contributed to the development of a large literature on the measurement of long-run inflation expectations. One simple approach is to rely on direct estimates of inflation expectations from surveys of professionals or consumers.[1] For example, Federal Reserve commentary such as Mishkin (2007) includes long-run expectations based on the Survey of Professional Forecasters' (SPF) projection of average inflation 1 to 10 years ahead.

Other approaches focus on econometric estimates of trend inflation. A large literature uses econometric methods to estimate inflation trends and forecast inflation (see, among many others, Stock and Watson 2007, Chan, Koop and Potter 2013, and Clark and Doh 2014).[2] One portion of this literature combines econometric models of trend with the information in surveys (see, among others, Kozicki and Tinsley 2012, Wright 2013, Nason and Smith 2014, Mertens 2016, and Del Negro, et al. 2017).[3]

In recent years, some countries have experienced extended periods of inflation running below survey-based estimates of long-run inflation expectations. For example, Fuhrer, Olivei, and Tootell (2012) show that actual inflation in Japan consistently ran below (survey-based) long-run inflation expectations in their sample, from the early 1990s to 2010. More recently, in the United States, for each year between 2008 and 2016, inflation in the core PCE price index ran below the SPF long-run forecast of roughly 2 percent (which coincides with the Federal Reserve's official goal for inflation).[4] Even though survey-based inflation expectations have been stable, actual inflation has been low enough for long enough to pull some common econometric estimates of trend inflation well below 2 percent (see, e.g., Bednar and Clark 2014). These experiences raise the question of whether it is possible for survey-based inflation expectations to become disconnected from actual inflation. Such a disconnect (if irrational) would make such expectations less useful for gauging the credibility of monetary policy and for forecasting inflation.

---

[1]Direct estimates of inflation expectations can also be obtained based on the relationship between real and nominal bonds. However, estimates of break-even inflation calculated using these are usually available only for a short time span. And there are reasons to expect that break-even inflation might reflect factors other than just long-run inflation expectations (e.g., if the risk premium is time-varying). Faust and Wright (2013) find it too volatile to be a sensible forecast for long-run expected inflation. For these reasons, we do not use break-even inflation data in this paper.

[2]The reader is referred to Faust and Wright (2013) for a recent survey on inflation forecasting, including a discussion of inflation surveys and methods for estimating trend inflation.

[3]Some DSGE models — developed in Del Negro and Schorfheide (2013) and references therein — treat the inflation target of the central bank as a random walk process and include survey measures of long-run inflation expectations as indicators of the target in model estimation. In a different vein, Aruoba (2016) develops an econometric, three-factor model of the term structure of inflation expectations.

[4]This statement is based on Q4/Q4 inflation rates for each year. The statement also applies to headline inflation, except that headline inflation rose above two percent for one year, 2011.

In this paper we develop a new model to examine the relationship between inflation, long-run inflation expectations, and trend inflation. We build on papers such as Kozicki and Tinsley (2012) by using models which are more flexible in empirically important directions, extending recent work with unobserved components models with stochastic volatility (UCSV) such as Stock and Watson (2007, 2015), Chan, Koop and Potter (2013), Clark and Doh (2014), Garnier, Mertens, and Nelson (2015), and Mertens (2016). Papers such as Kozicki and Tinsley (2012) equate long-run forecasts with trend inflation. Similarly, econometric estimates of trend inflation are sometimes calibrated to be the same as surveys. We also build on work by Nason and Smith (2014, 2016) that considers the possible disconnect between inflation and short-run inflation expectations in the context of a simple unobserved components model.

Our model permits us to assess the evidence for the links between trend inflation and long-run inflation expectations that have been assumed in some of the aforementioned literature. For example, the model of Mertens (2016) assumes that trend inflation moves one-for-one with long-run inflation expectations but allows a constant difference in the levels of trend inflation and long-run inflation expectations. Our approach allows us to assess the evidence in favor of such restrictions. We are able to estimate the relationship to investigate whether equating trend inflation with inflation expectations based on surveys improves the model of inflation. Our model permits the relationship to vary over time, such that trend inflation can be equal to the forecasts provided in the surveys at some points in time, but at other points in time forecasts can provide biased or inefficient estimates of trend inflation. We include comparisons to other, restricted versions of the model to assess the importance of such time variation to the trend estimate, model fit, and forecasting. Another point of departure from the existing literature is that, in our baseline model (although not all our models), we only use survey data on long-run inflation forecasts, allowing us to avoid the use of a subsidiary (possibly mis-specified) model linking short-run forecasts to long-run inflation expectations.

In our empirical work, we compare the fit and forecasting performance of our model to more restricted alternatives and some other models from the literature, using data for both the U.S. and a few other countries. We focus on results for CPI inflation and inflation expectations from Blue Chip and show our key results to be robust to two other data choices for the U.S. We present evidence that extensions over simpler approaches such as the addition of stochastic volatility and time-varying coefficients are important in practice. Survey-based measures of inflation expectations are found to be useful for estimating trend inflation, producing smoother and more precise estimates than a UCSV model. However, we also present evidence that the survey-based measures should not simply be equated with trend inflation; the relationship between the two is more complicated and, in some cases, time-varying. We include results from a pseudo-out-of-sample forecasting exercise, which shows point and density forecasts from our model to be at least as good as those from other models that have been found successful in the inflation forecasting literature. After establishing these results in U.S. data, we consider model estimates based on inflation and long-run survey expectations for Italy, Japan, and the UK. For these countries, it continues to be the case that the evidence indicates long-run survey expectations to be helpful to trend estimation, model fit, and forecasting. Although for Italy the data indicate the survey and trend inflation move one-for-one with no bias, for

Japan and the UK the data support a more flexible relationship.

Although our main empirical work does not directly address the question of why long-run surveys may differ from trend inflation, the final section of this paper includes some discussion of this issue in light of recent work on various topics, including work on informational rigidities in the professionals' forecasts by Coibion and Gorodnichenko (2015) and Mertens and Nason (2015).

2 Econometric Modeling of Trend Inflation

As discussed in sources such as Mertens (2016), an unobserved components framework is commonly used to model inflation, π_t, as being composed of trend (or underlying) inflation, π_t^*, and a deviation from trend, the inflation gap, c_t:

$$\pi_t = \pi_t^* + c_t. \tag{1}$$

The trend in inflation is defined (consistent with the Beveridge-Nelson decomposition) as the infinite-horizon forecast of inflation conditional on the information set available in period t, denoted Ω_t:

$$\lim_{j \to \infty} E\left[\pi_{t+j} | \Omega_t\right] = \pi_t^*, \tag{2}$$

which implies a random walk process for the trend π_t^* and a stationary, mean-zero inflation gap, c_t.

There are many possible econometric models consistent with this simple decomposition, and we will argue for a particular modeling framework soon. But the basic justification for using surveys of long-run forecasts can be clearly seen from (2). If those surveyed at time t about what inflation will be in period $t + j$ are rational forecasters, they can be expected to be reporting $E\left[\pi_{t+j} | \Omega_t\right]$. Thus, using (2), forecasts of long-run inflation will correspond to trend inflation, π_t^*. There are several ways that this relationship plus data on long-run forecasts made at time t (z_t) can be used to produce estimates of current trend inflation, with Kozicki and Tinsley (2012) being an influential recent approach.

However, there are reasons to be cautious about simply equating long-run forecasts from surveys with inflation trends, partly in light of the simple observations on the recent experiences in the U.S. and Japan noted in the introduction. For instance, surveys may produce forecasts that are biased, at least at some points in time. Survey forecasts at long horizons might also not move one-for-one with trend inflation. Surveys might also contain some noise, due to factors such as changes in participants from one survey date to another. In addition, papers such as Coibion and Gorodnichenko (2015) and Mertens and Nason (2015) find evidence of informational rigidities such that professional forecasters are slow to adjust their expectations. Accordingly, we desire an econometric specification that allows us to estimate the relationship between trend inflation and the long-run expectation of forecasters rather than imposing a particular form. In our model, a finding that long-run forecasts taken from surveys can be equated with trend inflation is possible, but not assumed a priori.

Earlier work also suggests many other desirable features we want our econometric model to have. First, Faust and Wright (2013) find improvements in forecast performance

by using the inflation gap (as opposed to inflation itself) as a dependent variable and modeling the inflation gap as the deviation of actual inflation from a slowly evolving trend. Many of the other studies mentioned above with time-varying inflation trends focus on an inflation gap. Our econometric specification follows this practice.

Second, the inflation gap, c_t, should be stationary but may exhibit persistence. For instance, a central bank may tolerate deviations of inflation from a trend or target for a certain period of time, provided such deviations are temporary. Furthermore, the central bank's toleration for such deviations may change over time. For instance, Chan, Koop and Potter (2013) discuss how the high inflation in the 1970s may have been partly due to the combination of a large inflation gap (with only a small increase in trend inflation) with a Federal Reserve tolerant of a high degree of inflation gap persistence. When Paul Volcker subsequently became the Fed chair, this tolerance decreased and inflation gap persistence dropped. We want our model to be able to accommodate such shifts in persistence.

Third, a large number of papers, such as Stock and Watson (2007), have found the importance of allowing for stochastic volatility, not only in the inflation equation but also in the state equations which describe the evolution of trend inflation. We include this feature in all of our models.

Finally, a general theme of many papers on inflation modeling, including Faust and Wright (2013) and Stella and Stock (2013), is time-varying predictability. The time-varying persistence and stochastic volatility features mentioned above are two such sources of time-varying predictability, accommodated by the model features mentioned above. The work of D'Agostino, Gambetti, and Giannone (2013) also indicates time-varying parameters to be helpful to forecast accuracy. Accordingly, we want a model with not only stochastic volatility but also time-varying parameters (TVP).

2.1   Baseline Model

All of these features are built into the following extremely flexible model, which should be able to accommodate any relevant empirical properties of the data on inflation (π_t) and the survey-based inflation expectation (z_t). (Note that all of the errors defined in the model below are independent over time and with each other.) We refer to this specification as model M1:

$$\pi_t - \pi_t^* = b_t(\pi_{t-1} - \pi_{t-1}^*) + v_t, \tag{3}$$

$$z_t = d_{0t} + d_{1t}\pi_t^* + \varepsilon_{z,t} + \psi\varepsilon_{z,t-1}, \ \varepsilon_{z,t} \sim \ N(0, \sigma_z^2) \tag{4}$$

$$\pi_t^* = \pi_{t-1}^* + n_t, \tag{5}$$

$$b_t = b_{t-1} + \varepsilon_{b,t}, \ \varepsilon_{b,t} \sim TN(0, \sigma_b^2), \tag{6}$$

$$d_{it} - \mu_{di} = \rho_{di}(d_{i,t-1} - \mu_{di}) + \varepsilon_{di,t}, \ \varepsilon_{di,t} \sim \ N(0, \sigma_{di}^2), \ i = 0, 1, \tag{7}$$

$$v_t = \lambda_{v,t}^{0.5}\varepsilon_{v,t}, \ \varepsilon_{v,t} \sim N(0,1), \tag{8}$$

$$n_t = \lambda_{n,t}^{0.5}\varepsilon_{n,t}, \ \varepsilon_{n,t} \sim N(0,1), \tag{9}$$

$$\log(\lambda_{i,t}) = \log(\lambda_{i,t-1}) + \nu_{i,t}, \ \nu_{i,t} \sim \ N(0, \phi_i), \ i = v, n. \tag{10}$$

In this model, the inflation gap $\pi_t - \pi_t^*$ follows an AR(1) process with a time-varying coefficient and stochastic volatility. Allowing b_t to be time-varying accommodates potential changes in the degree of persistence in the inflation gap. Note that we truncate the innovations to the AR(1) coefficient in (6) so as to ensure the inflation gap is stationary at every point in time ($TN(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2 truncated to ensure $0 < b_t < 1$). Trend inflation π_t^* follows a random walk with stochastic volatility in its innovations.

The long-run inflation expectation z_t is dependent on trend inflation, with a time-varying intercept d_{0t} and slope coefficient d_{1t} and an MA(1) error term. Accordingly, our model captures three dimensions along with the survey expectation can provide what we call a "biased" — a deliberate simplification of terms — measure of trend, through: (1) a non-zero intercept, d_{0t}; (2) a non-unity slope, d_{1t}; and (3) an MA component in the error term, reflected in ψ. We focus on the first two forms of "bias," in either a constant differential between trend inflation and the survey forecast or a failure of the survey to move one-for-one with trend.[5] Since d_{0t} and d_{1t} are time varying, we have the potential to estimate changes in the relationship between long-run forecasts and trend inflation. For instance, it is possible that long-run forecasts are unbiased estimates of trend inflation at some points in time, but not others. Our model allows for this possibility, but a constant coefficient model would not. Thus, investigating restrictions relating to d_{0t} and d_{1t} is of economic interest. To allow for persistence in a long-term inflation forecast that may not be adequately picked up by persistence in trend inflation, we add an MA(1) error term to (4). Although the empirical evidence for the need for this MA error term is weak in one of our U.S. data combinations (PCE inflation with PTR), in our baseline results for the U.S. and in the results for other countries, the MA term is empirically important to model fit and we include it in our general specification.

Variants of the model described above, excluding z_t, involving only (possibly restricted versions of) (3), (5), (6), (8), (9) and (10) have been used to estimate trend inflation by several authors. For instance, the popular UCSV model of Stock and Watson (2007) is this model with $b_t = 0$, and Chan, et al. (2013) use this model with bounded trend inflation but without stochastic volatility in $\varepsilon_{n,t}$. We stress that stochastic volatility is often found to be important in models of trend inflation such as these.[6] This feature allows for the possibility that the volatility of trend inflation or deviations of inflation from trend vary over time.

By adding the additional equations (4) and (7) to a conventional unobserved components model such as the one defined by (3), (5), (6), (8), (9) and (10), we can potentially improve the model's ability to fit historical inflation data and its estimates of trend infla-

---

[5]Conceptually, the distinction between the infinite-horizon forecast that constitutes trend inflation and the 10-year horizon of the survey expectation could cause $d_{0,t}$ to differ from 0 and $d_{1,t}$ to differ from 1. In practice, though, for professional forecasters, it seems likely that the 10-year ahead survey forecast is equivalent to an infinite-horizon forecast. For example, since the Federal Reserve established its longer-run inflation objective of 2 percent, the 10 year-ahead forecast of PCE inflation from the Survey of Professional Forecasters has stayed close to 2 percent. Moreover, in a cross-country analysis, Mehrotra and Yetman (2014) find that survey forecasts at just a 24-month ahead horizon tend to cluster around central bank inflation targets.

[6]For the errors in other equations, preliminary estimates suggest that an assumption of homoskedasticity is reasonable.

tion. That is, adding the relationship between z_t and π_t^* should provide extra information for estimating trend inflation beyond that provided in a univariate model involving inflation only. This information could improve precision of trend estimates, the model's ability to fit inflation, and forecast accuracy. Our model is less restrictive than those used in some other studies that relate inflation and survey measures of inflation expectations, and our specification can be seen as consistent with the cointegration restrictions imposed in these other studies (e.g., Mertens 2016, Mertens and Nason 2015, and Nason and Smith 2014). These other studies impose stationarity of the difference between actual inflation and survey expectations. Our model is consistent with cointegration of the survey expectation z_t with trend inflation π_t^*: the innovation term of the z_t equation is a stationary MA(1) process. Although the posterior of $d_{0,t}$ and $d_{1,t}$ need not be close to 0 or 1, respectively, our prior centers the initial values of these coefficients at 0 and 1, respectively. So our prior implies cointegration of z_t with trend inflation π_t^* with a slope coefficient of 1. With π_t^* the source of integration in π_t, it follows that we can think of π_t and z_t as cointegrated as well.

In sum, our model is a structural time series model (e.g., the wish to directly construct a direct measure of trend inflation implies a particular structure in our state space model) with additional features added by empirical necessity (e.g., the necessity of allowing for stochastic volatility, the possibility of persistence in the inflation process through addition of MA errors, but also that short lag lengths suffice with inflation data). It builds on, and shares, many similarities with other models in this literature. In our empirical work, we will investigate some reduced form time series approaches. These will include, in our forecast comparisons, a vector autoregression and vector error correction model, both with time-varying parameters (TVP-VAR and TVP-VECM, respectively).

We use Bayesian methods to estimate all the unknown parameters of our models, including latent variables such as trend inflation. The Markov Chain Monte Carlo (MCMC) algorithm used for estimation is similar to that used in previous work (e.g., Chan et al. 2016) and, hence, we say no more of it here. The priors used in this paper are informative, but not dogmatically so. In models such as ours, involving many unobserved latent variables, use of informative priors is typically necessary.[7] An earlier version of this paper, Federal Reserve Bank of Cleveland Working Paper 15-20, presented results from a prior sensitivity analysis of our baseline model, showing our results are fairly robust to changes in our prior. Complete details of the MCMC algorithm and prior are given in the Technical Appendix.

2.2 Extensions of the Baseline Model

Our baseline model excludes an economic activity indicator from the inflation gap equation (4). We do so in the interest of parsimony, motivated in part by evidence in the forecasting literature (see Faust and Wright 2013 and references therein) of the difficulty of using economic activity variables to improve predictions of inflation. However, in our

---

[7]Indeed, in the UC-SV model of Stock and Watson (2007), the stochastic volatility equations equivalent to our (10) are assumed to have a common error variance and this common variance is fixed at a specific value. Our prior is much less restrictive than this.

analysis for the U.S., we also consider a specification (denoted M7) augmented to include in the inflation equation an unemployment rate gap with a time-varying coefficient. Our specification with the unemployment gap has precedents in other recent studies, including: Stella and Stock (2013), which generalizes the UCSV formulation of Stock and Watson to relate the inflation gap to an unemployment gap; Jarocinski and Lenza (2015), which considers a specification involving a factor model of economic activity, for the purpose of estimating the output gap, with a structure for inflation, trend inflation, and inflation expectations that corresponds to a restricted, constant parameter version of our formulation; and Morley, Piger, and Rasche (2015), which considers a bivariate, constant parameter model relating inflation less a random walk trend to an unemployment gap.

Our baseline model includes only long-run inflation expectations since they should most directly reflect trend inflation. From Blue Chip, we have data on short-run expectations. To assess the potential value of short-horizon expectations, we also consider a version of our model (denoted M6) augmented to include these expectations, using an additional state equation which is the same as (4) except that a measure of short-run inflation expectations is the dependent variable.

In our results below, we also include for comparison a model (denoted M8) which adds one lag of the inflation gap in the equation for z_t and drops the MA component:

$$z_t = d_{0t} + d_{1t}\pi_t^* + d_{2t}(\pi_{t-1} - \pi_{t-1}^*) + \epsilon_{z,t} \quad \epsilon_{z,t} \sim N(0, \sigma_z^2). \tag{11}$$

The motivation for this specification can be found in papers such as Erceg and Levin (2003), which argue that, in the absence of a credible monetary regime, long-run inflation forecasts may respond to short-run movements in inflation as well as changes in trend inflation.

2.3   Restrictions on the Baseline Model

To help assess the ability of our model to improve the precision of trend estimates, the fit of inflation, and forecasts of inflation, we will also consider some more restricted models. The first of these additional models, M2, restricts d_{0t} and d_{1t} to be constants, d_0 and d_1. Model M3 imposes $d_0 = 0$ and $d_1 = 1$, which is the restriction that long-run inflation forecasts are unbiased estimates of trend inflation. These two models will shed light on the value of time variation in the coefficients and the value of allowing some bias in the relationship between the survey expectation and trend inflation (using the broad definition of bias indicated above).

Model M4 restricts our baseline model M1 by making no use of inflation expectations — which will shed light on the value of those expectations to inflation modeling. As such, it is a UCSV model like that of Stock and Watson (2007) but extended to allow an

autoregressive component:[8]

$$
\begin{align}
\pi_t - \pi_t^* &= b_t(\pi_{t-1} - \pi_{t-1}^*) + v_t, \tag{12} \\
\pi_t^* &= \pi_{t-1}^* + n_t, \tag{13} \\
b_t &= b_{t-1} + \varepsilon_{b,t}, \ \ \varepsilon_{b,t} \sim TN(0, \sigma_b^2), \tag{14} \\
v_t &= \lambda_{v,t}^{0.5}\varepsilon_{v,t}, \ \ \varepsilon_{v,t} \sim N(0,1), \tag{15} \\
n_t &= \lambda_{n,t}^{0.5}\varepsilon_{n,t}, \ \ \varepsilon_{n,t} \sim N(0,1), \tag{16} \\
\log(\lambda_{i,t}) &= \log(\lambda_{i,t-1}) + \nu_{i,t}, \ \ \nu_{i,t} \sim N(0, \phi_i), \ i = v, n. \tag{17}
\end{align}
$$

Model M5 is an AR(1) model in "gap form" similar to that used in Faust and Wright (2013), which they describe as "amazingly hard to beat by much." We call this the Faust and Wright model below.[9] We add stochastic volatility to this model to aid in comparability with our own. Specifically, we define the gap as $g_t = \pi_t - z_t$ and use the model:

$$
\begin{align}
g_t &= \beta g_{t-1} + \epsilon_{g,t}, \quad \epsilon_{g,t} \sim N(0, \lambda_{g,t}), \tag{18} \\
\log(\lambda_{g,t}) &= \log(\lambda_{g,t-1}) + \nu_{g,t}, \quad \nu_{g,t} \sim N(0, \phi_g), \tag{19}
\end{align}
$$

where we assume $|\beta| < 1$. The forecast for π_{t+k} given data until time t is computed by adding z_t to a forecast for g_{t+k}.

Finally, in the out-of-sample forecast comparison we consider bivariate TVP-VAR and TVP-VECM specifications, featuring stochastic volatility. Both models use a data vector containing inflation and long-run inflation expectations, such that $y_t = (\pi_t, z_t)'$. The TVP-VAR is given by

$$
B_{0t}y_t = b_{0t} + B_{1t}y_{t-1} + B_{2t}y_{t-2} + \epsilon_t^y, \quad \epsilon_t^y \sim N(0, \Sigma_t), \tag{20}
$$

where b_{0t} is a 2×1 vector of time-varying intercepts, B_{1t}, B_{2t} are 2×2 VAR coefficient matrices, B_{0t} is a 2×2 lower triangular matrix with ones on the diagonal and $\Sigma_t = \mathrm{diag}(\exp(h_{1t}), \exp(h_{2t}))$. The log-volatilities and the VAR coefficients evolve according to independent random walks. The TVP-VECM takes the following form:

$$
C_{0t}\Delta y_t = c_{0t} + c_{1t}(\pi_{t-1} - z_{t-1}) + C_{1t}\Delta y_{t-1} + \epsilon_t^{\Delta y}, \quad \epsilon_t^{\Delta y} \sim N(0, \Sigma_t), \tag{21}
$$

where c_{0t} and c_{1t} are 2×1 vectors of time-varying intercepts and coefficients, C_{1t} is a 2×2 coefficient matrix, and C_{0t} is a 2×2 lower triangular matrix with ones on the diagonal. Again the VAR coefficients and log-volatilities evolve according to independent random walks.

The models considered in this paper are summarized in the following table. Note, however, that not every model is used with every dataset, partly due to data availability and partly out of consideration for brevity. For example, for some datasets we have no short-run inflations expectations data so that M6 is not estimated, and some models (the TVP-VAR and TVP-VECM models) are only used in the forecast comparison.

---

[8]The supplemental appendix of Cogley, Primiceri and Sargent (2010) makes use of a similar model.

[9]Our specification generalizes their "fixed ρ" model by estimating coefficients. Accordingly, our model takes the same form as their "AR-gap" model, except that, at all horizons, we use the 1-step ahead form of the model and iterated forecasts, whereas they use a direct multi-step form of the model.

Model	Brief Description
M1	Our model as defined in equations (3) through (10)
M2	Restricts M1 such that d_{0t} and d_{1t} are constant
M3	Restricts M1 such that $d_{0t} = 0$ and $d_{1t} = 1$
M4	UCSV (M1 without inflation expectations data) of equations (12) through (17)
M5	Faust-Wright model (AR(1) in gap form) in equations (18) and (19)
M6	M1 augmented with short-run inflation expectations
M7	M1 augmented with unemployment rate gap in inflation equation
M8	M1 augmented with one lag of inflation gap in equation for z_t
M9	TVP-VAR
M10	TVP-VECM

3 Data

Policymakers are interested in a range of different measures of inflation, and the research literature considers a range of measures. Accordingly, for the U.S., we provide results for several combinations of measures of inflation and inflation expectations. Subsequently, we present an international comparison using data from Italy, Japan, and the UK. We chose these countries in part because the forecast data go back as far as 1990 and in part because the survey-based long-run forecasts show some noticeable time variation.

For the U.S., we use three different measures of quarterly inflation (π_t in the model): i) inflation based on the consumer price index (CPI inflation), ii) inflation based on the consumer price index excluding food and energy (core CPI inflation), and iii) inflation based on the price index for personal consumption expenditures (PCE inflation). Inflation rates are computed as annualized log percent changes ($\pi_t = 400 \ln\left(P_t/P_{t-1}\right)$, where P_t is a price index). The CPI has the advantage of being widely familiar to the public, and for much of our sample, the available inflation expectations data refer to it. However, changes over time in the methodology used to construct the CPI — such as the 1983 change in the treatment of housing costs to use rental equivalence — may create structural instabilities, because the historical data are not revised to reflect methodology changes. One reason we also consider PCE inflation is that its historical data has been revised to reflect methodology changes, reducing concerns with instabilities created by methodology changes. Another reason is that the Federal Reserve's preferred inflation measure is PCE inflation; its longer-run inflation objective is stated in terms of PCE inflation.

Reflecting data availability, our results draw on a few different sources of long-run inflation expectations. In most of our results for the U.S., we use the Blue Chip Consensus (the mean of respondents' forecasts, from Blue Chip Economic Indicators) to measure long-run inflation expectations (z_t in the model). Blue Chip has been publishing long-run (6-10 year) forecasts of CPI inflation and GNP or GDP deflator inflation since 1979 in the latter case and 1983 in the former case. To extend the CPI forecast survey back to 1979, we fill in data for 1979 to 1983 using deflator forecasts from Blue Chip.[10] The forecasts

---

[10]For the next several years following 1983, Blue Chip's long-run forecasts of CPI and GDP inflation are very similar.

are only published twice a year; we construct quarterly values using interpolation.

Partly for the purpose of using a longer sample, in some of our results we instead use the long-run inflation expectation series included (as the series denoted PTR) in the Federal Reserve Board of Governor's FRB/US econometric model. Defined in CPI terms, the PTR series in the Board's model splices (1) econometric estimates of inflation expectations from Kozicki and Tinsley (2001) early in the sample to (2) 5- to 10-year-ahead survey measures compiled by Richard Hoey to (3) 1- to 10-year ahead expectations from the Survey of Professional Forecasters.[11] Defined in the PCE terms actually used in the FRB/US model, the series uses the same sources, but from 1960 through 2006, the source data are adjusted (by Board staff, for use in the FRB/US model) to a PCE basis by subtracting 50 basis points from the inflation expectations measured in CPI terms. Although some readers may be concerned by the econometric component to the PTR time series and the approximations used to translate from CPI to PCE terms, we only use the series in a relatively small set of results.

We present results for three combinations of inflation with corresponding inflation expectations: i) CPI inflation plus Blue Chip forecasts, ii) core CPI inflation plus Blue Chip forecasts and iii) PCE inflation plus PTR long-run forecasts. This set addresses robustness to different inflation measures and to different measures of inflation expectations.[12] In results based on Blue Chip expectations, the estimation sample period is 1980:Q1 to 2016:Q1. In results based on the PTR measure of inflation expectations, we estimate the model using data from 1960:Q2 to 2016:Q1.

As detailed above, one model we consider as a robustness check includes a short-run inflation expectation (in addition to the long-run expectation). We measure the short-run expectation with the three-quarter ahead forecast of CPI inflation from the Blue Chip Consensus. Out of concern for data consistency, we only estimate this model with CPI inflation and the long-run expectation from Blue Chip.

A second model we consider as a robustness check includes economic activity as a predictor of inflation with a time-varying coefficient. In this model, we follow common practice (e.g., Morley, Piger, and Rasche 2015, Stella and Stock 2013) and define the relevant activity variable as an unemployment gap, defined as the actual unemployment rate less the Congressional Budget Office's estimate of the natural rate of unemployment.[13]

For our international analysis, we use CPI inflation rates and long-run forecasts of CPI inflation from Consensus Economics (hereafter, CE). The exception is the UK, for which we use the retail price index excluding indirect taxes (RPI) and the CE forecasts of RPI inflation. We obtained CPI data from Haver Analytics and the UK's RPI from the website of the Office of National Statistics. The long-run forecasts obtained from

---

[11]Surveys of professional forecasters have long included projections of CPI inflation or the GNP/GDP price deflator/price index, but only recently has any survey included PCE inflation. The Blue Chip consensus tracks expectations of inflation in both the CPI and GDP price index. The Survey of Professional Forecasters tracks expectations of CPI inflation and, since 2007, PCE inflation.

[12]An earlier version of this paper, released as Federal Reserve Bank of Cleveland Working Paper 15-20, contains results for a wider range of combinations, including for GDP deflator inflation.

[13]Following studies such as Rudd and Peneva (2015), we use the measure the CBO refers to as its short-term estimate of the natural rate, which incorporates a temporary, substantial rise in the natural rate in the period following the start of the Great Recession, attributable to structural factors such as extended unemployment insurance benefits.

CE are conceptually comparable to the U.S. forecasts published by Blue Chip; they are projections of average inflation 6 to 10 years ahead, reported as the average across private forecasters who participate in the survey. Since mid-2014, the CE forecasts have been published on a quarterly basis (in the first month of each quarter). Prior to that, the forecasts were only published twice a year (April and October), and we construct quarterly values using interpolation. For Italy, Japan, and the UK, data runs from 1990:Q2 through 2016:Q2. In light of the shorter samples of expectations data available for these other countries, in the international assessment we only report full-sample estimates and omit out-of-sample forecast comparisons.

4 Empirical Results using U.S. Data

In this section, we present results for three different combinations of inflation and expectations measures for the U.S. In addition to our baseline model, we present selected results from six to seven other models, detailed above. The primary purpose of this paper is to develop an appropriate model for investigating the relationship between inflation, trend inflation and inflation expectations. However, it is also of interest to see whether it forecasts better than plausible alternatives. To this end, we carry out a pseudo-out-of-sample forecasting exercise. In our results based on long-run expectations from Blue Chip, the evaluation sample begins with 1995Q1. In results based on the PTR measure of inflation expectations, for which a longer history is available, the forecast evaluation period begins in 1975Q1.[14]

Empirical results are mostly presented using figures. In each case, the first set of figures focusses on M1. It plots posterior means (along with an interval estimate) of all the latent variables in the model (i.e., $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}, d_{1t}$). The figure for π_t^* also plots actual inflation (π_t) along with long-run forecasts taken from the surveys (z_t). The next set of figures presents comparisons of these latent variables across our models. For the baseline case of CPI inflation with long-run inflation expectations measured by 6-10 year ahead forecasts of Blue Chip (for brevity, we omit the same for the other data combinations), we include some additional charts to compare the precision of trend estimates and pseudo-real time estimates of trend. Finally, tables of marginal likelihoods and measures of forecast performance are provided. For the latter, we present root mean squared forecast errors (RMSFEs) and sums of log predictive likelihoods, both taken relative to the UCSV-AR model (M4). When computing forecasts for model M7, we assume an AR(4) model for the unemployment gap.

4.1 Results Using CPI Inflation and Blue Chip Forecasts

4.1.1 Estimation results using the full sample

We begin by presenting evidence on how well our baseline model performs relative to alternative models at estimating trend inflation and other features of interest using the

---

[14]We repeated the analysis with a shorter forecast evaluation period beginning in 1985Q1 (after the Great Moderation) and found results to be qualitatively similar.

full sample.

Figure 1 presents estimates of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for our baseline model. Trend inflation estimates can be seen to be much smoother than actual inflation. In a general sense, they track long-run survey-based forecasts fairly well. However, trend inflation lies consistently below survey forecasts and this difference is large in a statistical sense. That is, z_t consistently lies above the upper bound of the credible interval for π_t^* and the professionals were forecasting long-run inflation to be somewhat higher than our estimate of trend inflation. A finding that the professionals' forecasts are often slightly above our estimates of trend inflation can also be seen in the results for d_{0t} and d_{1t}. Remember that $d_{0t} = 0$ and $d_{1t} = 1$ implies long-run forecasts are unbiased estimates of trend inflation. In Figure 1, most of the posterior probability of d_{0t} lies in the positive region and (with high posterior probability) d_{1t} is above one, particularly early in our sample. These values jointly imply that our trend inflation estimates are slightly below those of the professionals.

Estimates of b_t tend to be consistent with a fair amount of inflation persistence (at roughly 0.5), with slight evidence of some decrease over time. There is also strong evidence of stochastic volatility, both in the inflation equation and in the one for trend inflation. This is consistent with the findings of Stock and Watson (2007) in their univariate model for inflation. It is interesting to note that, as in Stock and Watson (2007), both types of stochastic volatility were high around 1980 and fell subsequently. The recent financial crisis was associated with a large increase in the volatility of shocks to the inflation gap, but no increase in the volatility of shocks to trend inflation. Insofar as low volatility in trend inflation reflects a firm anchoring of inflation expectations, then our results suggest the Fed has succeeded in anchoring inflation expectations since the 1980s and that these expectations were not shaken by the financial crisis.

Figure 2 compares parameter and trend inflation estimates across models (except for a trend estimate from the Faust-Wright model (M5), which does not produce such an estimate). These results indicate that, relative to our baseline model, estimates are only modestly changed (early in the sample) by the addition of short-run inflation expectations (M6) or an unemployment gap (M7). Restricting the baseline model by making the coefficients d_0 and d_1 of the inflation expectations equation constant or restricting them to specific values (0 and 1, respectively) has somewhat more noticeable effects on the time-varying volatility of innovations to trend inflation ($\lambda_{n,t}$), the coefficients d_0 and d_1, and trend inflation. For example, restricting d_0 and d_1 to be constant in model M2 lowers the estimate of the slope d_1 from more than 1 in M1 to a little more than 0.8 in M2 and raises the intercept d_0 from 0.3 or less in model M1 to about 0.8 in model M2.

For both M2 and M3, the estimated trend is well above the estimate from model M1 for about the first 10 years of the sample. Perhaps not surprisingly, with d_0 and d_1 restricted to 0 and 1, respectively, the trend estimate from model M3 is essentially the same as the survey expectation z_t (so much so as to obscure the line for z_t in the top panel's chart). Broadly, the estimates from the various models covered in Figure 2 increase the weight of evidence against $d_{0t} = 0$ and $d_{1t} = 1$. For example, M6 and M7 roughly line up with model 1 in their estimates of these time-varying coefficients, with d_0 above 0 and d_1 above 1. The estimates of model M2 shows that, even with a constant coefficient model, estimates of these coefficients differ from the (0,1) case. The
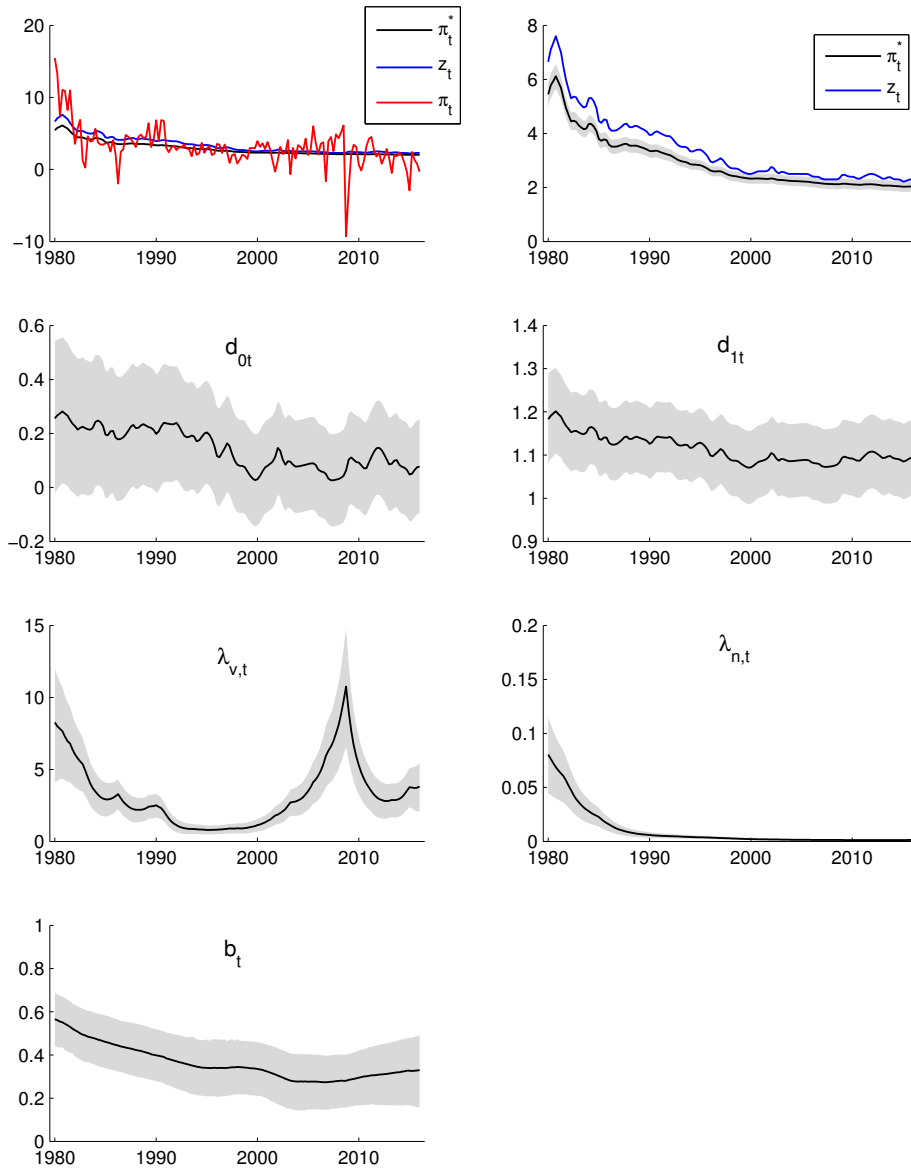
Figure 1: Posterior Means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for M1 (CPI+Blue Chip). Shaded bands are 16th-84th percentiles
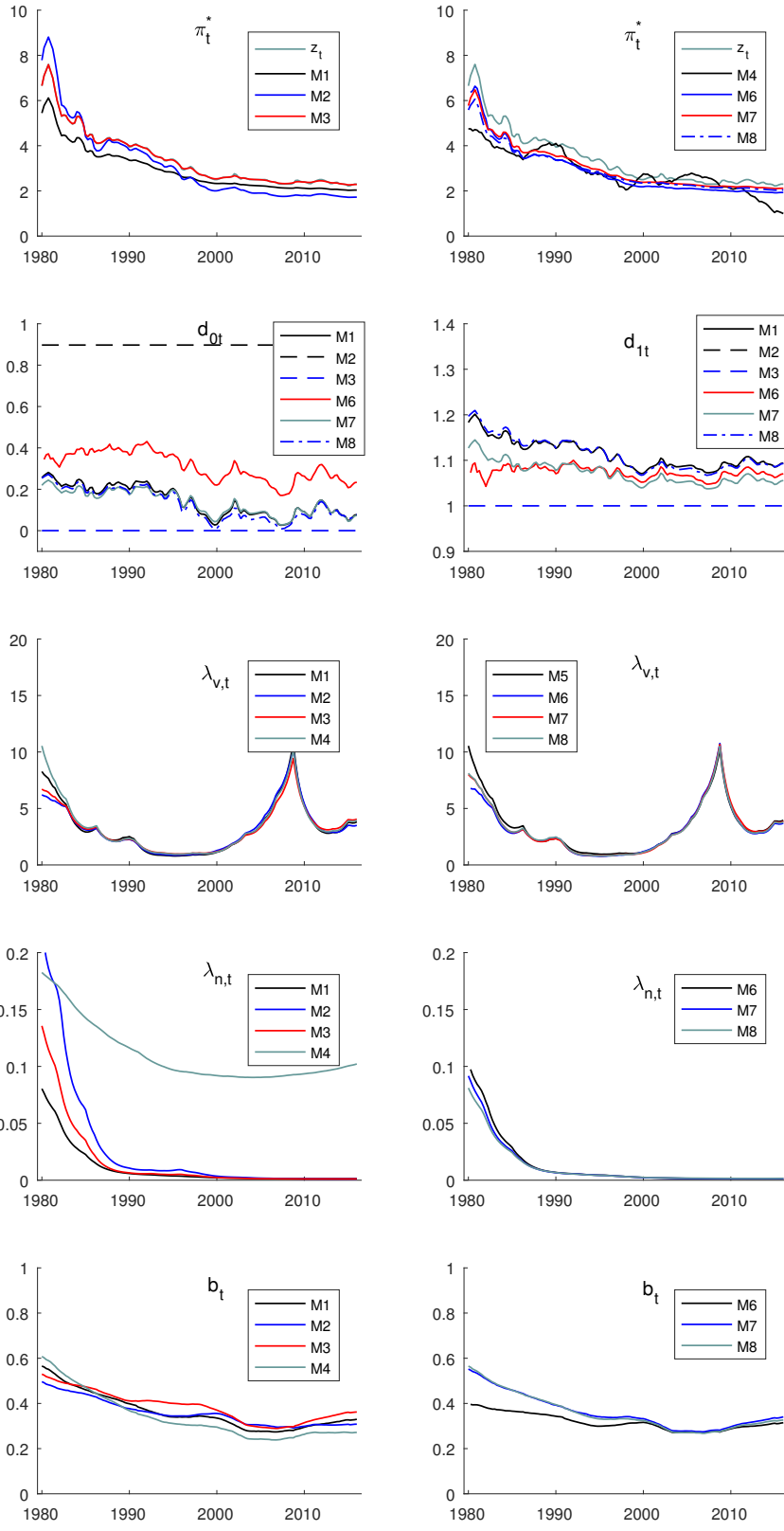
Figure 2: Comparison of posterior means of π_t^*, b_t, $\lambda_{v,t}$, $\lambda_{n,t}$, d_{0t} and d_{1t} for different models (CPI+Blue Chip)
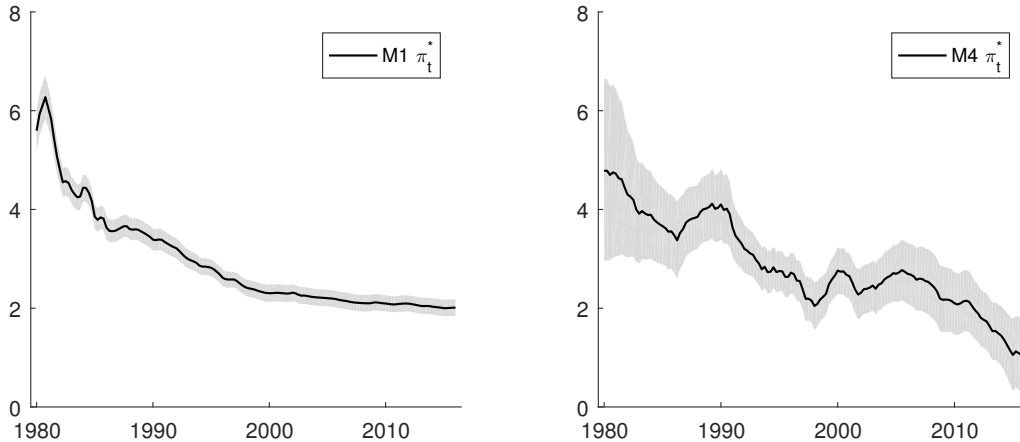
Figure 3: Estimates of π_t^* for M1 and M4 with 16th-84th percentiles as shaded bands (CPI + Blue Chip)

estimates for model M8 indicate that adding a lagged inflation gap to the equation for inflation expectations has little effect on the estimate of trend inflation or coefficients and volatilities of the model. This finding reflects the result — not shown in the interest of brevity — that the coefficient d_{2t} is estimated to be small.

The estimate of trend inflation produced by M1 lies consistently below z_t. To shed some light on this, note that the sample means of z_t, π_t and our point estimates of π_t^* are 3.5, 3.1 and 2.9 percent, respectively. Thus, on average, the professionals' forecasts lie above average inflation whereas our estimates of trend inflation are, sensibly, pulled to be much closer to it.

Dropping long-run inflation expectations out of the model, as does the UCSV-AR specification of M4, creates larger differences in estimates compared to the baseline model. The estimate of the time-varying volatility to trend inflation ($\lambda_{n,t}$) is noticeably higher for M4 than the baseline specification. In addition, the estimate of trend inflation from M4 differs from the baseline in some important respects. As evident from the top row of Figure 2, M4's trend inflation estimate tends to be more variable and substantially lower around 1980 than any of the other approaches which include long-run inflation expectations. In addition, as shown in Figure 3, the credible set around the estimate of trend inflation is much narrower with M1 than M4. Using a survey-based measure of inflation expectations to inform the estimate greatly increases the precision of the trend estimate.

To assess the ability of our model to fit inflation data, Table 1 provides marginal likelihoods for the eight models under consideration.[15] We start by comparing our baseline model to the UCSV-AR (M4) and Faust-Wright models (M5) and then consider the effects on model fit of restrictions on the d coefficients and of model extensions. By the classic recommendations of Jeffreys for interpreting Bayes factors (see, e.g., page 777 of Kass and Raftery, 1995), the evidence in favor of our model against models M4 and M5 is strong (decisive for M4 and substantial for M5). Restricting the d coefficients in models M2

---

[15]The Technical Appendix details the computation of the marginal likelihood. These are constructed using the predictive likelihood associated solely with inflation so as to ensure comparability across models.

Table 1: Log marginal likelihood estimates (CPI + Blue Chip)

M1	M2	M3	M4	M5	M6	M7	M8
-277.29	-278.60	-278.60	-284.41	-279.33	-275.33	-283.10	-277.45

and M3 modestly reduces model fit compared to the baseline M1. By the standards of Jeffreys, the evidence in favor of our time-varying d coefficients over constant coefficients is substantial, but not strong. Finally, extending our model to include short-horizon forecasts yields a substantial improvement in model fit, whereas extending it to include the unemployment gap makes model fit much worse. Extending our model to include the lagged inflation gap in the process for long-run expectations with model M8 has little effect on model fit relative to our baseline specification M1.

4.1.2   Using our model in real time

Up to this point, we have focused on full-sample estimates of the models and smoothed estimates of trend. However, models like these are often used in real time for forecasting or policy purposes. For instance, a central banker might use a model like this to regularly assess inflation trends. Hence, in this sub-section we present: i) historical time series of pseudo-real time estimates of trend inflation, ii) pseudo-real time forecasts and iii) a pure out-of-sample forecasting exercise up to 2021.

Pseudo-real time estimates of trend inflation are given in Figure 4. Starting in 1990:Q1, in each quarter t, we use the historical data up to that point in time to estimate the models and their inflation trends, saving the trend as of period t as the pseudo-real time estimate, and repeating the estimation at each subsequent quarter. As expected, these pseudo-real time trend estimates are noisier than their full-sample smoothed counterparts. The estimate from model M8 is very similar to the baseline from M1. The estimates from models M1, M2, M3, and M6 are broadly similar to one another, although there certainly can be sizable differences across models. The estimate from M7 (which includes short-run expectations as well as long-run expectations) has a similar contour to these other models, but tends to be higher. The estimate from M4 is much more noticeably different from the other estimates, particularly in its much higher volatility. In pseudo-real time estimates, including the survey-based measure of long-run inflation expectations greatly reduces the variability of trend inflation estimates.

Overall, these findings support the view that including information from survey forecasts and adding time-variation in parameters is useful in helping refine estimates of trend inflation, in dimensions including the capture of features that seem to exist in estimates of our relatively flexible model, the precision of trend estimates ex post, and the variability of pseudo-real time estimates of trend inflation. But simply assuming survey forecasts to be unbiased measures of trend inflation appears unduly restrictive.

To assess the value of long-run inflation expectations for forecasting future inflation, Table 2 reports the accuracy of point and density forecasts, as ratios of RMSFEs of each model relative to the UCSV-AR specification (M4) and as differences in log predictive likelihoods relative to the M4 model baseline (a RMSFE ratio less than 1 denotes im-
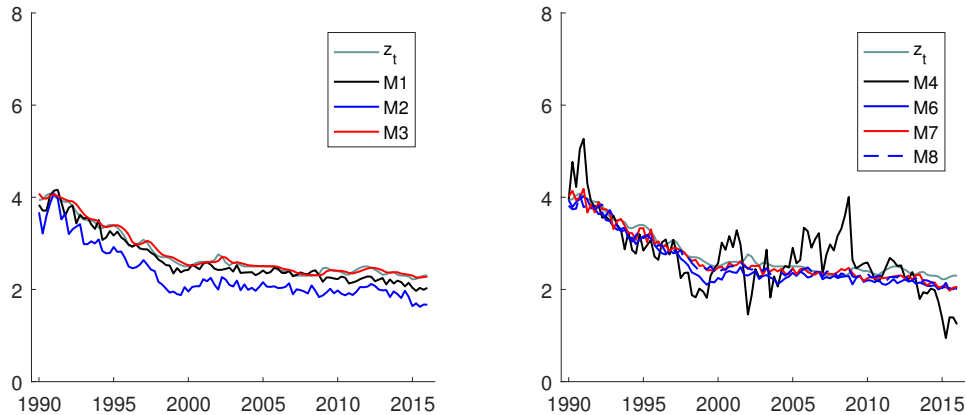
Figure 4: Posterior Means of pseudo-real time estimates of π_t^* for M1 (CPI+Blue Chip).

provement on the baseline, as does a positive relative log predictive likelihood). The horizons range from 1 through 20 quarters ahead, as well as 6-10 years ahead. These 6-10 years ahead forecasts refer to the average rate of inflation 6-10 years ahead. Note that, at this very long horizon, some caution is required in drawing strong conclusions from the results in that the number of fully independent observations is limited in the available data sample.

At most horizons, all of the models that include long-run inflation expectations (abstracting for the moment from models M9 and M10) improve on the accuracy of the UCSV-AR model. At short horizons, the gains are admittedly small to modest; practically speaking, there is little to distinguish the models in forecast accuracy. At longer horizons, up to 20 quarters ahead, the gains increase to as much as about 16 percent for point forecasts and more than 20 points in log predictive likelihood. The more restricted models M3 (which sets d_0 to 0 and d_1 to 1 for all time) and M5 (the Faust-Wright model) are slightly less accurate than the less restrictive models M1 and M6, but not meaningfully so. The picture is mostly similar at the very long horizon of 6-10 years ahead, with the main difference being that the performances of models M3 and M5 modestly deteriorate at this horizon. The performance of the less restrictive models M9 (TVP-VAR) and M10 (TVP-VECM) is more mixed. At short horizons, these models are comparable in accuracy to our proposed models. However, at longer horizons, the performance of models M9 and M10 can be driven by explosive forecasts. We have taken steps in the results to essentially eliminate this behavior for model M9 forecasts and sharply reduce it for model M10 forecasts.[16] At longer horizons, the relatively less restricted model M9 performs comparably to our preferred models in both point and density forecasts. But model M10 fares less well, such that it is dominated by other specifications, especially in density forecasts.

Finally, to illustrate some of the practical differences with our preferred model com-

---

[16]For model M9, we followed Cogley and Sargent's (2005) approach to imposing stationarity on the VAR estimates, and in simulating the forecast distribution, we shut down time variation in the latent states by holding them constant at end of sample values. For model M10, in simulating the forecast distribution, we shut down time variation in the latent states by holding them constant at end of sample values, and we use the posterior median rather than the mean as the point forecast.

Table 2: RMSFEs and log predictive likelihood for forecasting CPI inflation relative to UCSV-AR

|	Relative RMSFE							
	1Q	2Q	4Q	8Q	12Q	16Q	20Q	6-10Y
M1	0.97	0.94	0.88	0.90	0.90	0.89	0.90	0.90
M2	0.98	0.95	0.89	0.88	0.87	0.84	0.84	0.72
M3	0.98	0.95	0.90	0.92	0.94	0.93	0.94	1.02
M5	0.98	0.97	0.92	0.92	0.93	0.93	0.94	1.01
M6	0.97	0.94	0.88	0.90	0.89	0.88	0.88	0.84
M7	0.98	0.95	0.91	0.91	0.91	0.90	0.91	0.78
M8	0.97	0.94	0.88	0.90	0.90	0.89	0.90	0.89
M9	0.99	0.98	0.96	0.94	0.93	0.91	0.89	0.76
M10	0.98	0.95	0.90	0.90	0.91	0.95	0.99	1.60
	Relative log predictive likelihood							
	1Q	2Q	4Q	8Q	12Q	16Q	20Q	6-10Y
M1	2.86	4.40	7.84	10.71	14.10	17.52	18.70	53.60
M2	2.34	4.61	8.67	12.64	17.56	20.47	20.68	64.37
M3	1.08	2.33	5.52	8.44	10.54	13.35	14.24	32.05
M5	1.43	2.82	6.96	10.54	11.67	15.35	14.78	13.79
M6	2.44	3.85	7.98	10.77	15.84	18.57	20.34	58.68
M7	0.78	1.10	3.83	10.10	14.60	16.80	17.95	55.16
M8	1.75	3.82	8.49	11.95	14.20	18.35	19.26	53.42
M9	-0.24	0.44	2.40	7.34	13.30	19.41	21.92	54.74
M10	0.47	3.04	9.41	12.08	7.60	-7.42	-26.49	< -100
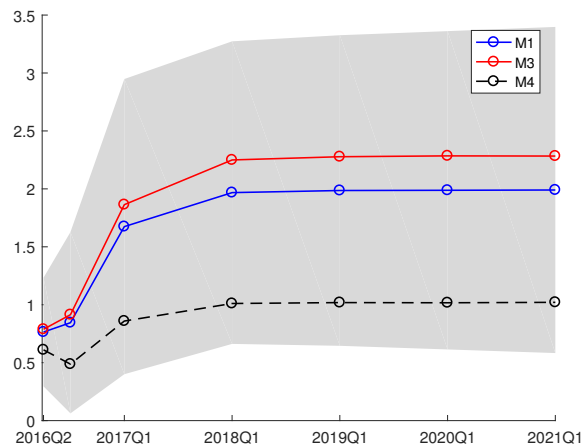
Figure 5: Point forecasts for CPI inflation (four-quarter average rates) from M1 with the 16th-84th percentiles as shaded bands.

pared to some basic alternatives, we present "true" out-of-sample forecasts using the end of our estimation sample as the jumping-off point. These forecasts (aggregated to represent four-quarter average rates of inflation, in keeping with common central bank practice) cover a 20 quarter period, from 2016:Q2 through 2021:Q1. We include forecasts for M1 along with two key comparison models, M3 and M4. Figure 5 shows that forecasts from M3 lay slightly above forecasts from M1. This reflects the estimated trends of the model, in which, at the end of the estimation sample shown in Figure 2, the trend estimate of model M3 — which imposes the restriction that the long-run survey expectation is an unbiased measure of trend inflation — slightly exceeds the trend estimate of the baseline model M1. M4, the UCSV-AR model which does not contain information from inflation surveys, produces forecasts which are lower still, reflecting the low level of recent inflation and the absence of any influence from long-run inflation expectations. These differences in forecasts highlight practical differences in the models that could, at a given moment in time, have important policy implications. For example, a monetary policymaker relying on the forecast of model M4 would likely prefer a much more accommodative monetary policy than would a policymaker relying on the forecast of our preferred model M1.

4.2 Results Using Core CPI inflation and Blue Chip Forecasts

4.2.1 Estimation results using the full sample

Results using core CPI inflation, given in Figures 6 and 7 and Table 3, are broadly similar to those using CPI inflation. In particular, we are still finding that our estimate of trend inflation lies below z_t and that d_{0t} and d_{1t} differ from the (0,1) values which imply that professionals are producing unbiased forecasts of trend inflation. The UCSV-AR model produces trend inflation estimates which are more erratic than those produced using models which incorporate inflation expectations (although we omit the results in the interest of brevity, this model also yields trends estimates that are less precise and

Table 3: Log marginal likelihood estimates (core CPI + Blue Chip)

M1	M2	M3	M4	M5	M7	M8
-148.70	-146.91	-151.14	-154.67	-152.26	-151.98	-147.99

much more variable in pseudo-real time). However, there are some interesting differences. There is less evidence of time-variation in d_{0t} and d_{1t} than with CPI inflation. Another point worth noting is that the volatilities, $\lambda_{v,t}, \lambda_{n,t}$, are large in 1980 but both continually fall over the sample period. This contrasts with the CPI inflation results where $\lambda_{v,t}$ shoots up at the time of the financial crisis.

In terms of model fit as captured by the marginal likelihoods of Table 3, our baseline model (M1) yields considerable gains relative to the UCSV-AR (M4) and Faust-Wright (M5) models. In contrast to the results for headline CPI inflation, for core CPI inflation, restricting the d_0 and d_1 coefficients to be constants improves model fit, yielding the best-fitting model. However, restricting these coefficients to 0 and 1, respectively, significantly harms model fit. These findings indicate that survey-based long-run inflation expectations are closely related to the trend in core CPI inflation but not an unbiased measure. Once again, extending the model to include the unemployment gap (M7) makes model fit much worse.

4.2.2 Using our model in real time

Figure 8 presents filtered, real-time, estimates of trend inflation. Our previous finding, that incorporation of survey-based forecasts can reduce the volatility in real-time estimates relative to models such as UCSV-AR which do not incorporate them, is also found for core CPI inflation.

The pseudo-out-of-sample forecasting results in Table 4 show that, with core CPI inflation, not all of the models incorporating long-run inflation expectations improve on the accuracy of the UCSV-AR model. Models M3 and M5 — the models that equate the long-run expectation with trend inflation — are generally less accurate than the UCSV-AR model, although in some cases only by small margins. Our proposed model yields forecasts slightly more accurate than those of the UCSV-AR baseline. Restricting the d_0 and d_1 coefficients to be constant as in model M2 yields more sizable improvements in forecast accuracy, especially at longer horizons. For core CPI inflation, model M2 forecasts best.

Finally, the purely out-of-sample forecasts out to 2021 are given in Figure 9. Once again the forecast from model M3 lies above — more sizably with core inflation than headline inflation — that from our baseline model M1, reflecting the same pattern in their estimates of trend inflation, in turn reflecting M3's restriction that the long-run survey expectation is an unbiased measure of trend inflation. In contrast to the findings for CPI inflation, we are now finding that model M4 (which does not use any information from the professional forecasts) produces a forecast which lies above that of the baseline model M1.
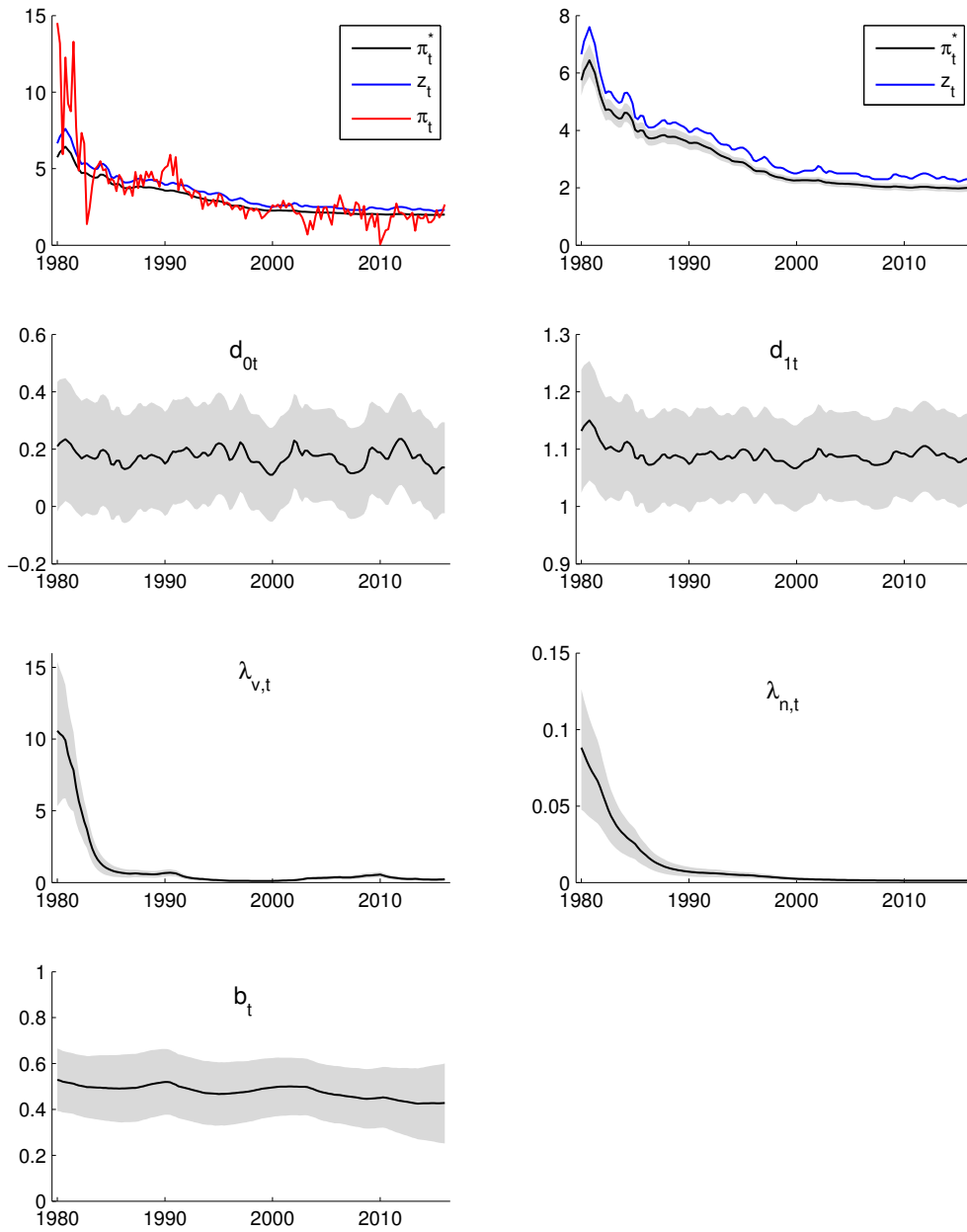
Figure 6: Posterior Means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for M1 (core CPI+Blue Chip)
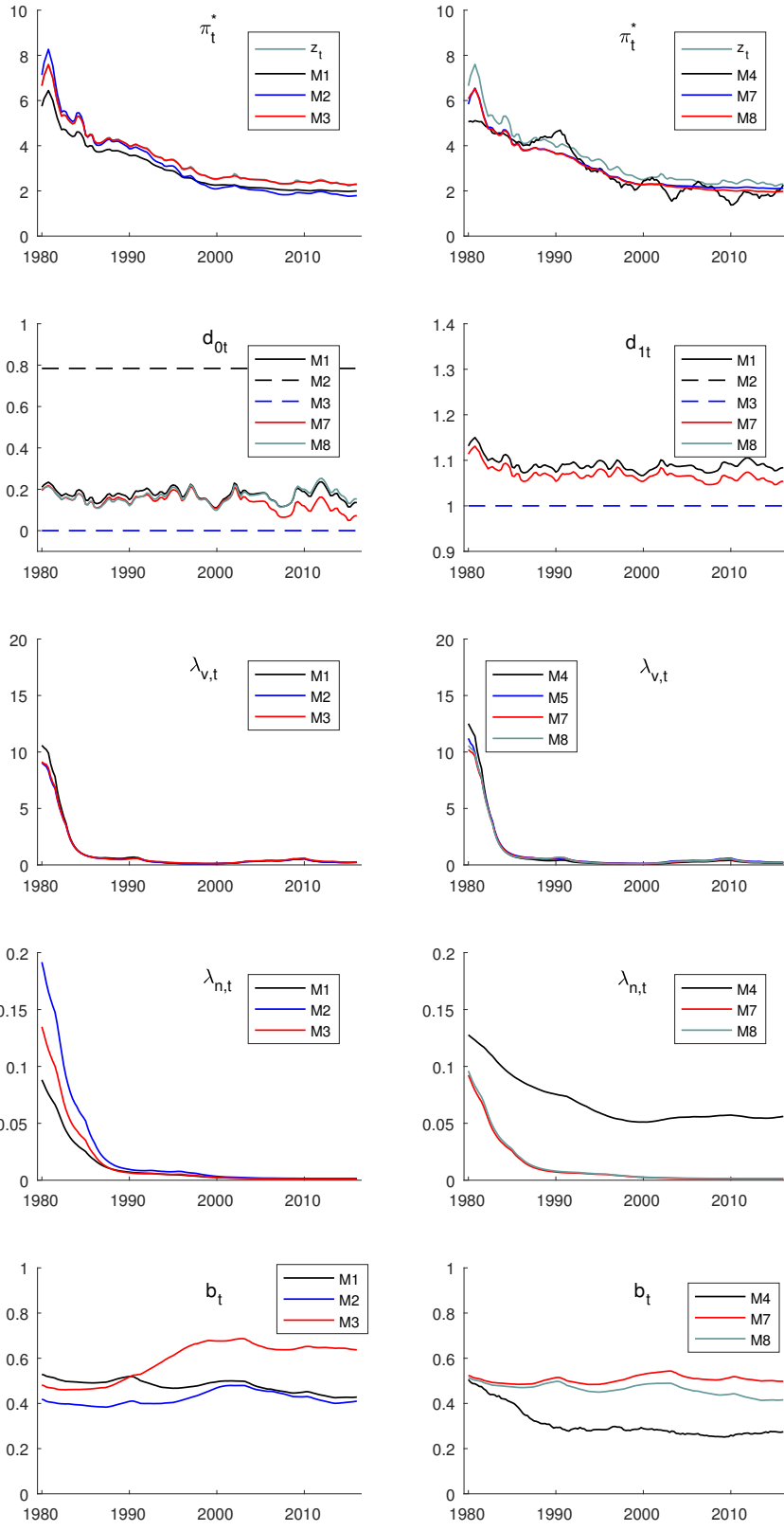
Figure 7: Comparison of posterior means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for different models (core CPI+Blue Chip)
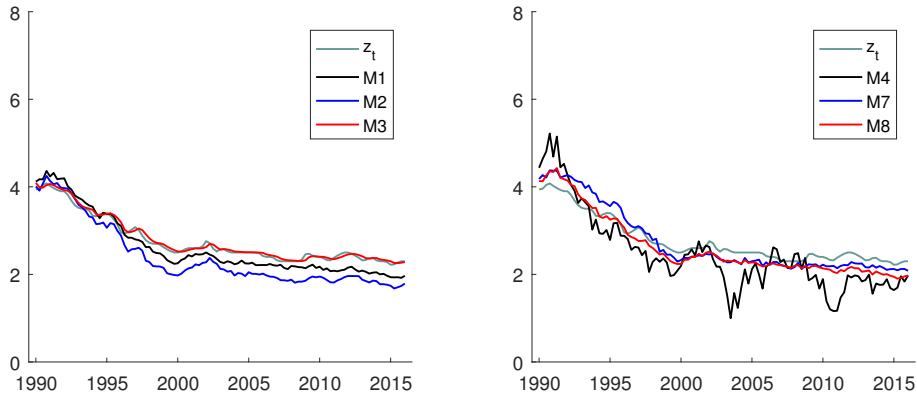
Figure 8: Posterior means of pseudo-real time estimates of π_t^* (core CPI+Blue Chip).

Table 4: RMSFEs and log predictive likelihood for forecasting core CPI inflation relative to UCSV-AR

|	1Q	2Q	4Q	8Q	12Q	16Q	20Q	6-10Y
				Relative RMSFE				
M1	0.99	0.99	0.98	0.95	0.94	0.95	0.95	1.00
M2	0.97	0.94	0.88	0.79	0.76	0.78	0.79	0.87
M3	1.04	1.08	1.12	1.11	1.06	1.04	1.00	1.03
M5	1.04	1.10	1.16	1.14	1.06	1.04	1.00	1.02
M7	0.98	0.95	0.95	0.98	1.06	1.07	1.07	1.15
M8	0.99	0.99	0.97	0.94	0.92	0.94	0.94	1.00
M9	0.98	0.98	1.00	1.00	0.95	0.89	0.88	0.88
M10	0.98	0.97	1.00	1.09	1.12	1.08	1.10	1.24
			Relative log predictive likelihood					
	1Q	2Q	4Q	8Q	12Q	16Q	20Q	6-10Y
M1	1.31	0.74	0.52	5.69	11.24	14.05	14.43	0.08
M2	2.95	4.24	8.52	22.13	31.52	33.70	33.07	25.57
M3	-3.04	-6.98	-12.70	-17.75	-15.87	-10.56	-8.09	-35.78
M5	-3.21	-9.28	-18.66	-28.61	-29.65	-27.88	-27.99	< -100
M7	1.49	2.63	5.50	6.37	0.59	3.46	3.17	-16.01
M8	1.42	0.84	0.86	5.75	12.26	15.70	16.11	3.28
M9	0.91	1.84	-2.29	-2.63	1.45	12.51	23.77	13.06
M10	1.69	2.93	0.61	-3.08	-5.53	-2.26	-7.45	-29.88
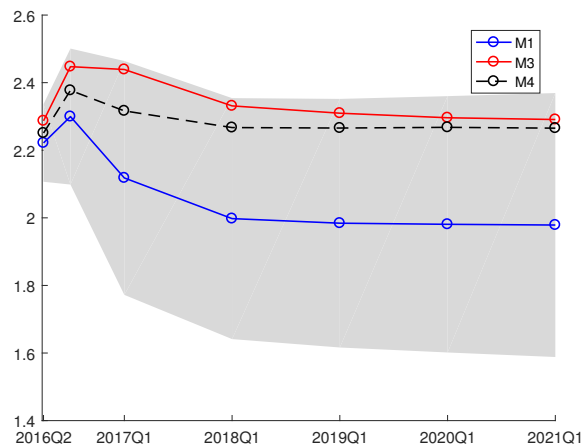
Figure 9: Point forecasts for core CPI inflation (four-quarter average rates) from M1 with the 16th-84th percentiles as shaded bands.

4.3 Results Using PCE Inflation and PTR Forecasts

4.3.1 Estimation results using the full sample

In this sub-section, the inflation measure is PCE inflation, and the long-run inflation expectations measure is PTR. For this data combination, our sample goes back to 1960 and so we are able to examine the performance of our model over a longer time period. Figures 10 and 11 and Table 5 provides the results.

For much of the sample, especially in the late 1970s and early 1980s, we are again finding strong evidence that our estimate of trend inflation lies modestly below the professionals' long-run forecast. Our estimates of d_0 and d_1 are relatively high from the late 1970s through roughly 1995, with d_1 trending up through 1980 and then down for some years. Post-1980, results for $\lambda_{v,t}$ and $\lambda_{n,t}$ are similar to those for CPI inflation. Pre-1980, $\lambda_{v,t}$ (the volatility in the inflation gap equation) follows the expected pattern in the mid- to late- 1970s before falling with the Great Moderation. But it is interesting to note that this pattern is not replicated for $\lambda_{n,t}$ (the volatility in trend inflation), which slowly rises throughout the 1970s before reaching a peak in the early 1980s and falling thereafter.

Turning to our other models, we are again finding that the UCSV-AR model is producing more erratic estimates of trend inflation (a pattern more evident in the 1960-2016 sample used in these results than in the 1980-2016 sample of our CPI results).

The marginal likelihoods of Table 5 yield some differences with respect to the baseline results we obtained with CPI inflation measures. In model fit, our baseline model (M1) continues to yield considerable gains relative to the Faust-Wright (M5) model, but not relative to the UCSV-AR (M4) specification. With PCE inflation, restricting the d_0 and d_1 to be constants slightly improves model fit, such that M2 and M3 are not really different from the UCSV-AR (M4) specification in model fit. Once again, extending the model to include the unemployment gap makes model fit much worse.
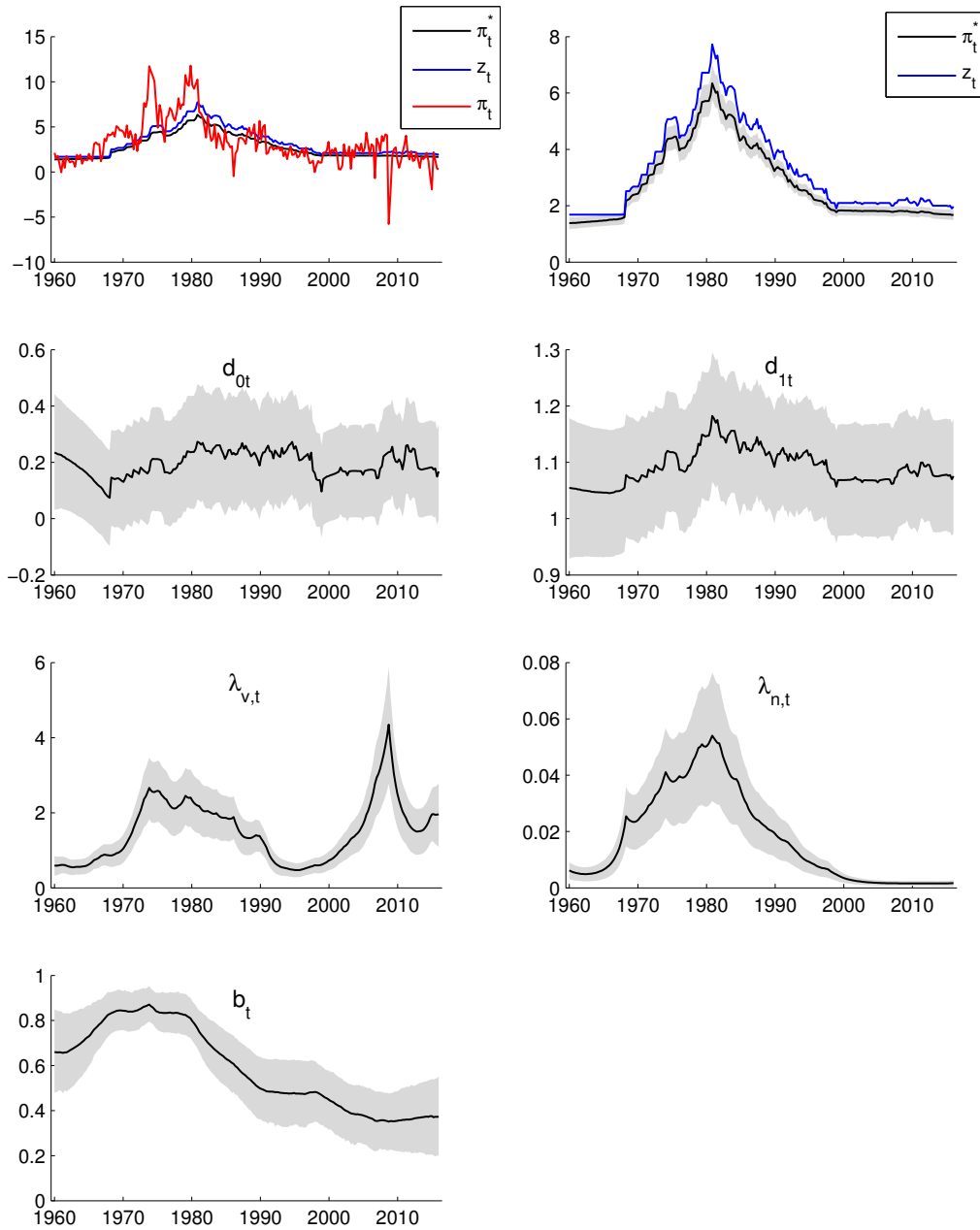
Figure 10: Posterior Means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for M1 (PCE+PTR)

Table 5: Log marginal likelihood estimates (PCE+PTR)

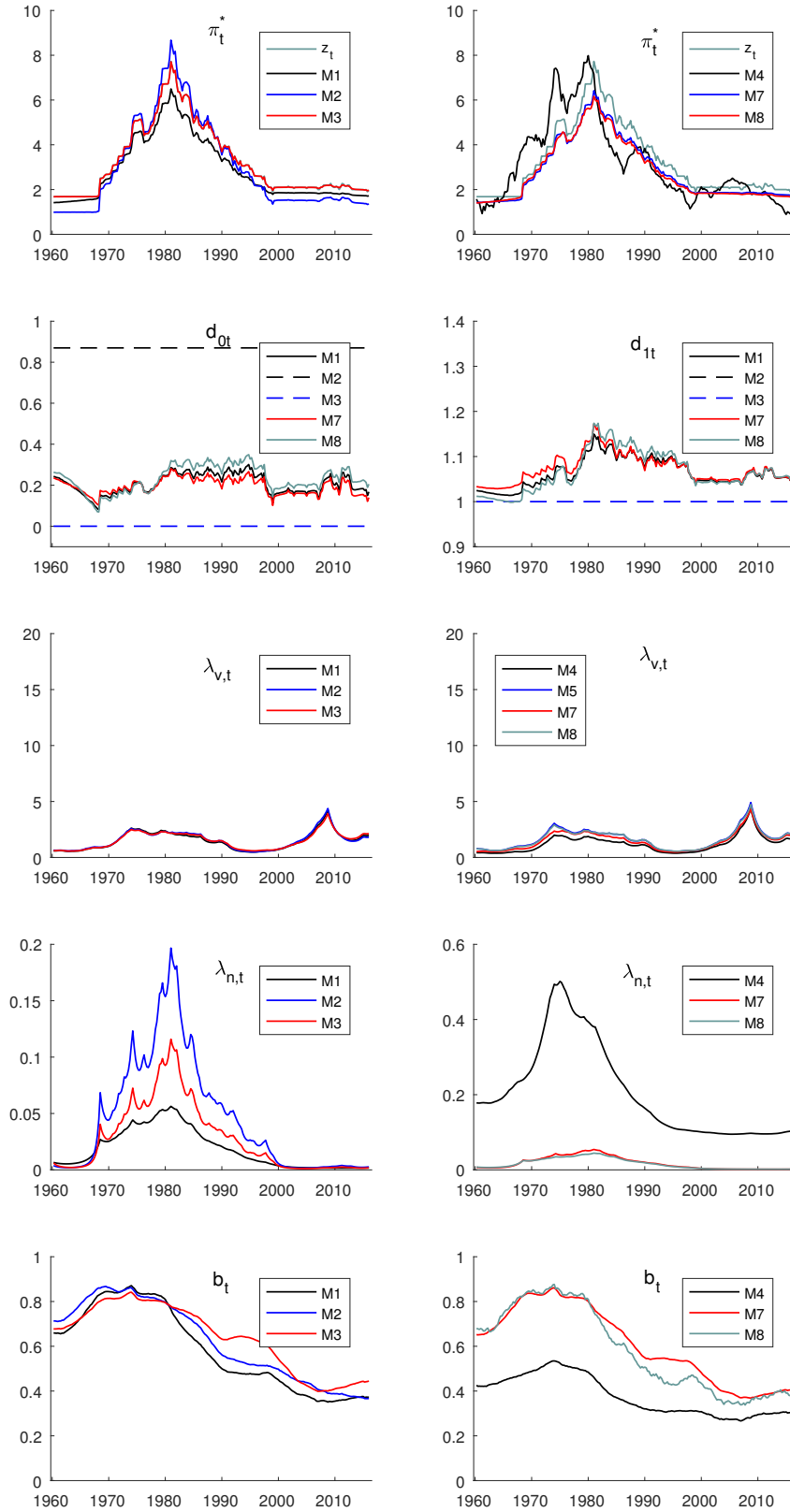| M1	M2	M3	M4	M5	M7	M8
-367.28	-366.26	-366.80	-366.35	-372.61	-373.89	-368.34

Figure 11: Comparison of posterior means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for different models (PCE+PTR)
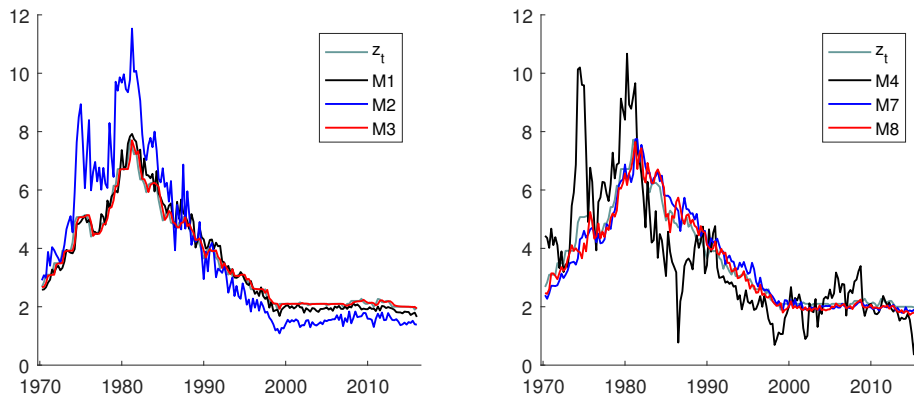
Figure 12: Posterior means of pseudo-real time estimates of π_t^* (PCE + PTR).

4.3.2   Using our model in real time

The real-time estimates of trend inflation are given in Figure 12. M1 is again producing more reasonable — in our assessment, less variable — estimates of trend inflation than the more volatile ones produced by UCSV-AR.

In Table 6's pseudo out-of-sample forecasting results for PCE inflation, the forecast performance of models incorporating long-run inflation expectations is broadly similar to the performance of the UCSV-AR model. Our proposed model M1 often improves on the accuracy of the baseline model, but only slightly. Restricting the model's d_0 and d_1 coefficients to be constant at 0 and 1, respectively, very slightly improves the accuracy of the model, but not to a notable degree.

Finally, the out-of-sample forecasts to 2021 for PCE inflation in Figure 13 show a pattern similar to those obtained for CPI inflation (reported in Figure 5). Again reflecting differences in trend estimates, the forecast of PCE inflation from M1 is modestly higher than the forecast from the restricted model M3 and notably above the forecast from the UCSV-AR model, which has inflation below 1% for the next five years. That said, all of the point forecasts are below the level of long-run inflation expectations (not shown in the chart), which is about 2 percent.

5   An International Comparison

The CE data allow us to use methods developed in this paper with survey forecasts constructed in an internationally comparable way. In this section, we present results for Italy, Japan, and the UK using the CE long-run forecasts as measures of expected inflation. In the interest of brevity, we focus on models M1 through M5 and M8 (i.e., the models which use only data on inflation and a long-run survey forecast). Note that these datasets have a shorter sample span, so our estimates begin in 1990. Since the period from 1990 to the Great Recession and financial crisis was a relatively stable time in most advanced economies, in this section we are missing some of the variability which was present in the U.S. datasets of the preceding section.

28

Table 6: RMSFEs and log predictive likelihood for forecasting PCE inflation relative to UCSV-AR

|	1Q	2Q	4Q	8Q	12Q	16Q	20Q	6-10Y
				Relative RMSFE				
M1	0.98	0.97	0.96	0.98	0.99	1.01	1.06	1.07
M2	0.98	0.98	0.96	0.98	0.99	1.02	1.15	1.24
M3	0.98	0.98	0.95	0.96	0.96	0.98	1.02	1.02
M5	1.02	1.04	1.05	1.02	1.03	1.05	1.09	1.02
M7	0.99	0.99	0.98	0.99	1.00	1.03	1.07	1.04
M8	0.98	0.97	0.95	0.97	0.97	0.99	1.04	1.06
M9	1.00	1.02	1.10	1.23	1.36	1.55	1.82	3.85
M10	1.03	1.06	1.15	1.18	1.26	1.35	1.45	> 100
				Relative log predictive likelihood				
	1Q	2Q	4Q	8Q	12Q	16Q	20Q	6-10Y
M1	1.45	2.73	4.25	1.23	8.02	5.61	0.32	-73.73
M2	1.24	2.69	6.29	5.60	11.49	9.34	5.03	3.71
M3	0.80	1.88	3.88	4.35	11.07	10.57	8.25	-45.78
M5	-3.57	-4.55	-4.78	-12.28	-12.18	-16.04	-20.02	< -100
M7	-1.15	-1.85	-0.15	2.17	5.42	0.67	-4.44	-59.23
M8	0.31	1.15	4.31	5.51	9.45	5.22	2.84	-77.62
M9	-0.03	-1.94	-5.68	-19.99	-44.30	-66.03	< -100	< -100
M10	-3.90	-6.43	-12.35	-27.60	-57.13	-46.85	-57.68	< -100
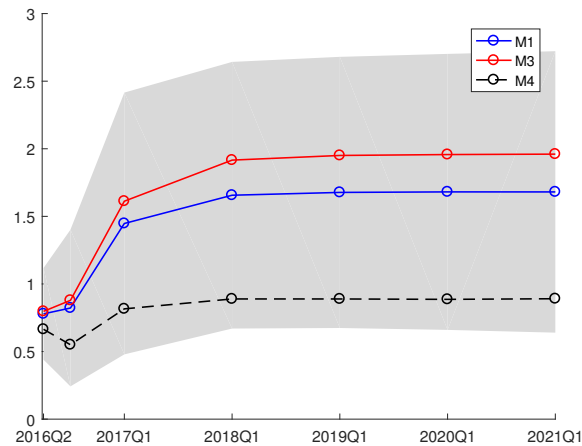
Figure 13: Point forecasts for PCE inflation (four-quarter average rates) from M1 with the 16th-84th percentiles as shaded bands.

Table 7: Log marginal likelihood estimates, other countries

country	M1	M2	M3	M4	M5	M8
Italy	-134.34	-134.96	-134.02	-137.92	-135.14	-134.70
Japan	-196.79	-193.74	-202.53	-196.72	-203.10	-197.09
UK	-164.79	-165.88	-165.41	-166.56	-167.37	-164.50

Figures 14, 15, and 16 provide estimates of our baseline model (M1) for Italy, Japan, and the UK, respectively. Figures 17, 18, and 19 provide comparisons of estimates across models M1 through M5 and M8, for Italy, Japan, and the UK, respectively.

Consider first the estimates of trend inflation. In the preceding section, we found our model produced estimates which were consistently slightly less than the professionals' forecasts. This finding also holds true for Japan (Figure 15). For the UK it holds much of the time. But for Italy, our estimate of trend inflation is very close to the professionals' survey (Figure 14). In the U.S. results, we also found the trend estimates from the UCSV-AR to be more erratic than those from our baseline model. In the shorter sample for other countries, this same finding applies to Italy (Figure 17) but not Japan (Figure 18) or the UK (Figure 19).

With the U.S. data, we found considerable evidence against the $d_{0t} = 0$ and $d_{1t} = 1$ restrictions. This also holds true in our estimates for Japan and the UK but not Italy.[17] However, the way each country departs from this restriction is a bit different. For Japan, there is support for the restriction $d_{1t} = 1$, but d_{0t} is positive and quite large, indicating that professionals' forecasts are consistently above trend inflation. A similar pattern holds in the UK, but only from the late 1990s until the financial crisis. There is substantial time variation in the UK estimates of d_{0t} and d_{1t}. All in all, we are finding a range of patterns but, apart from Italy, we are never finding strong support that the long-run surveys provide unbiased estimates of trend inflation.

With regards to stochastic volatility, we are finding somewhat less evidence of its presence in the shorter samples for Italy, Japan, and the UK than in the longer samples of U.S. data. As noted above, with our sample for the CPI in the U.S., $\lambda_{v,t}$ (inflation gap volatility) and $\lambda_{n,t}$ (trend inflation volatility) trended down in the 1980s and then were little-changed, with the notable exception of a spike in $\lambda_{v,t}$ around the Great Recession. With the other countries, there is some decline in $\lambda_{n,t}$ in the 1990s and some time variation in $\lambda_{v,t}$ for Japan, but otherwise, volatility is relatively stable. It is also interesting to note that, especially for Italy and the UK, the estimates of $\lambda_{n,t}$ are much higher for the UCSV-AR model (M4) than our baseline model (M1), which explains why this model produces more erratic estimates of trend inflation.

Table 7 presents marginal likelihood comparisons for the models applied to each country. In all cases, our proposed model fits inflation data as well as or better than the UCSV-AR (M4) and Faust-Wright (M5) models. For Italy, M1 is significantly better than M4 but only modestly better than M5. For Japan, M1 is significantly better than

---

[17]In addition, our international results corroborate our U.S. results that indicate estimates of model M8, which allows z_t to depend on a lagged inflation gap, are very similar to those of the baseline model M1.
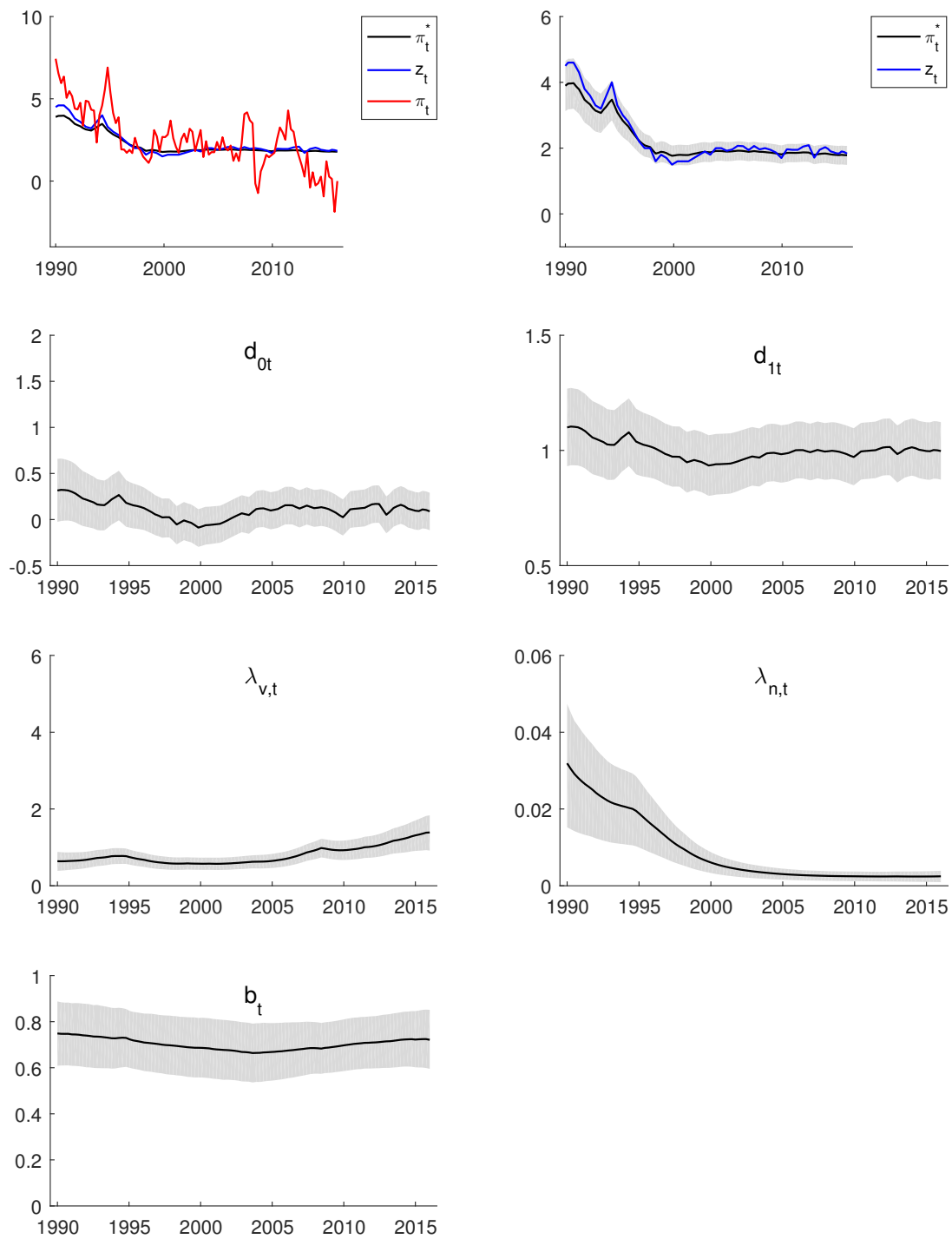
Figure 14: Posterior Means of π_t^*, b_t, $\lambda_{v,t}$, $\lambda_{n,t}$, d_{0t} and d_{1t} for Italy. Shaded bands are 16th-84th percentiles
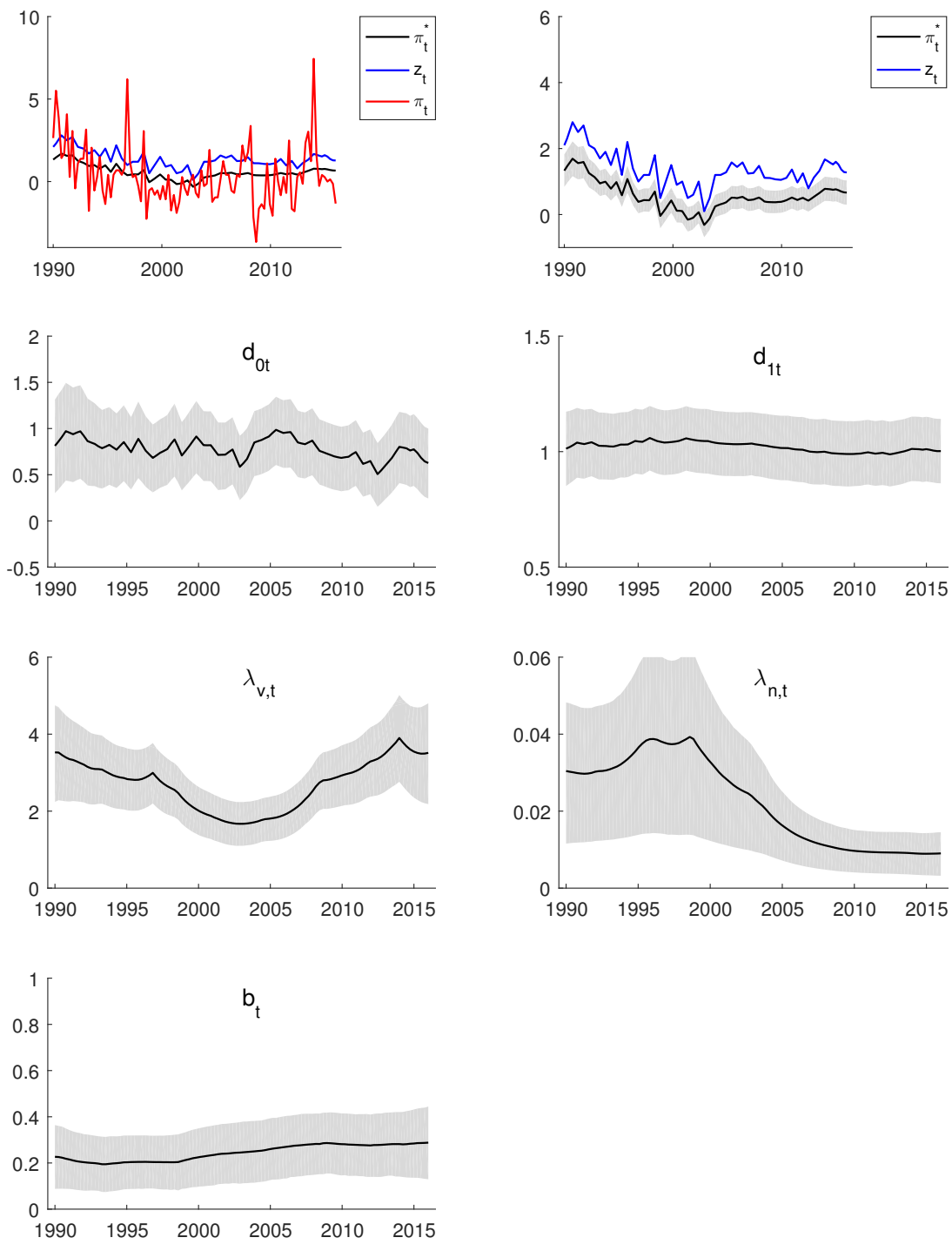
Figure 15: Posterior Means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for Japan. Shaded bands are 16th-84th percentiles
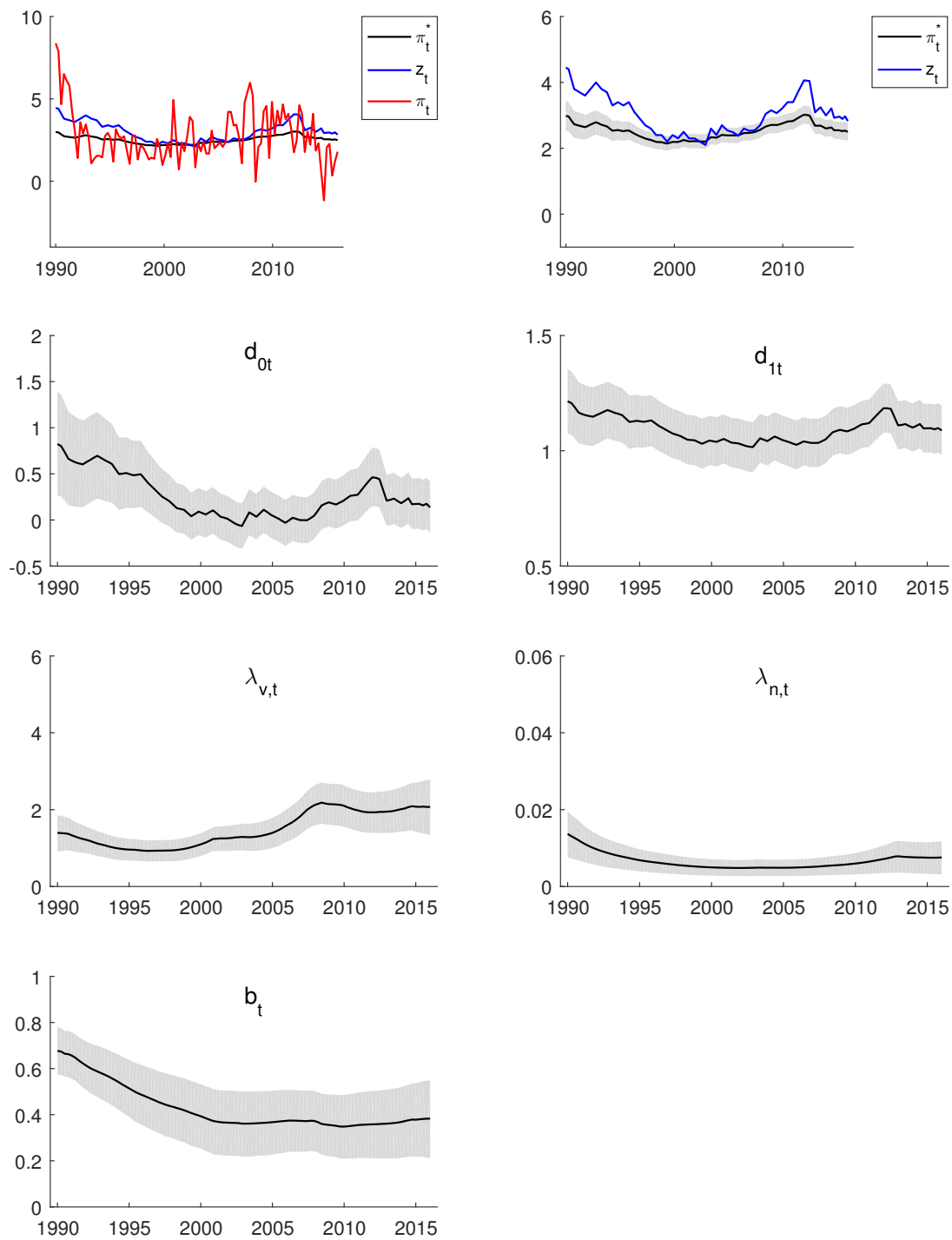
Figure 16: Posterior Means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for the UK. Shaded bands are 16th-84th percentiles
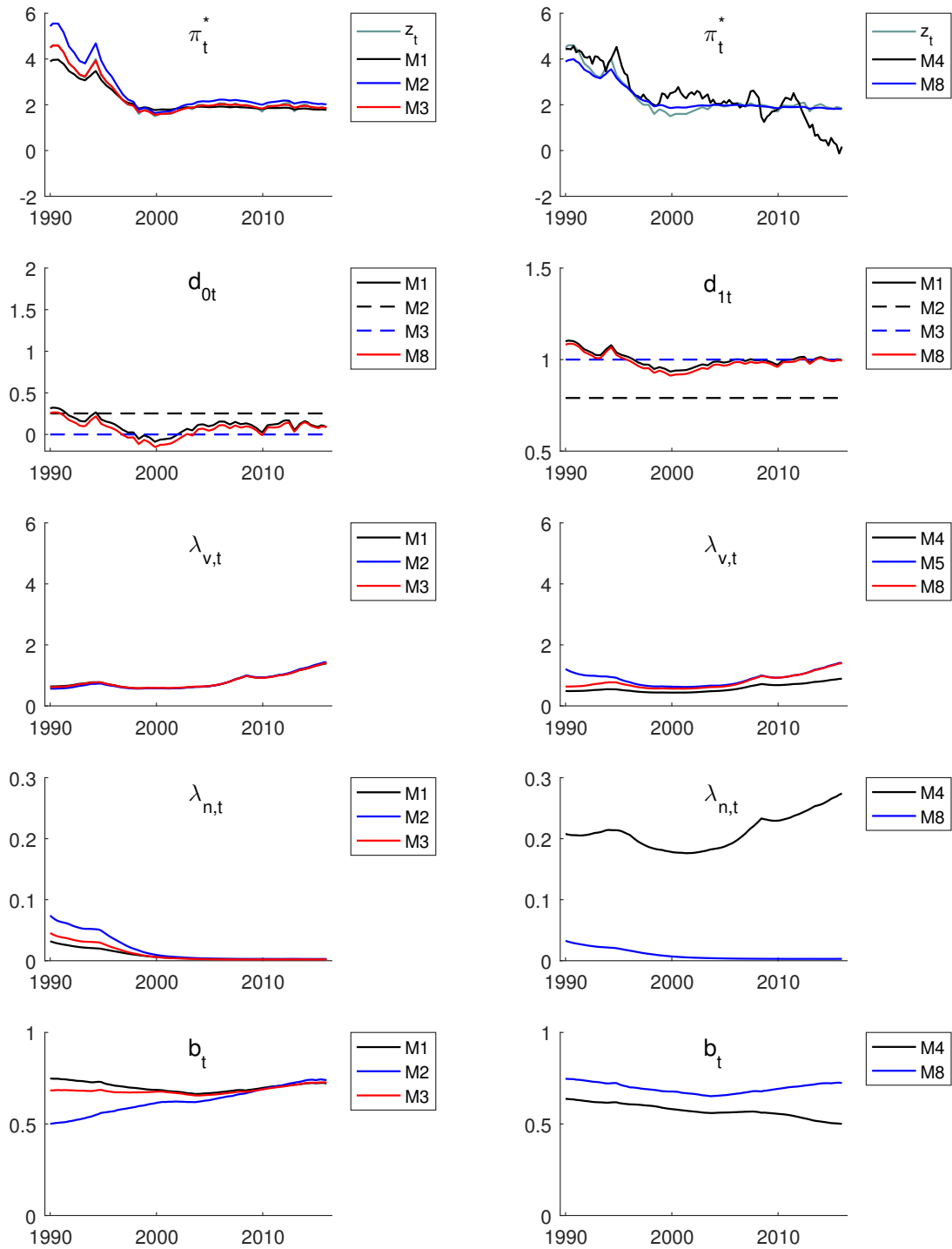
Figure 17: Comparison of posterior means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for different models, for Italy
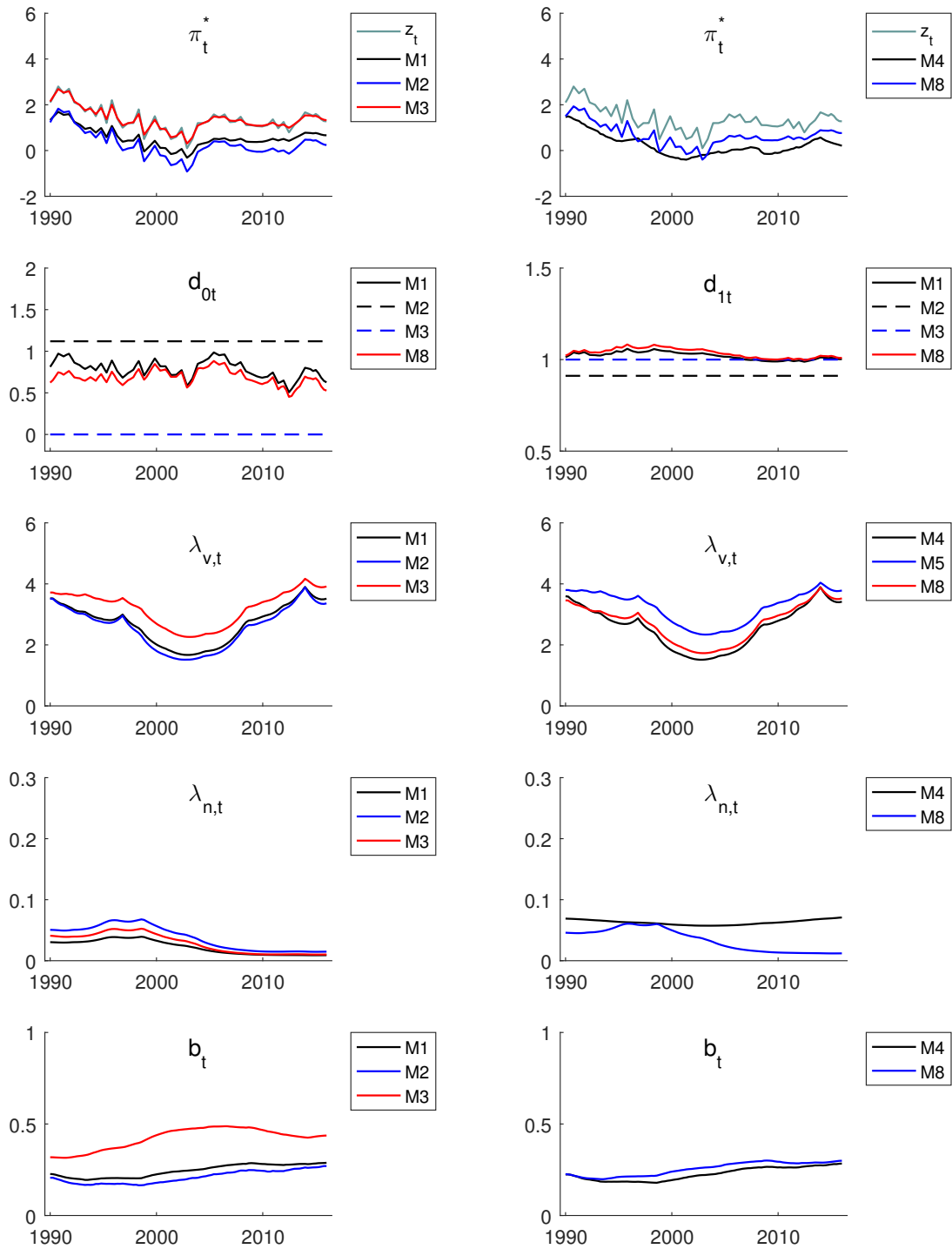
Figure 18: Comparison of posterior means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for different models, for Japan
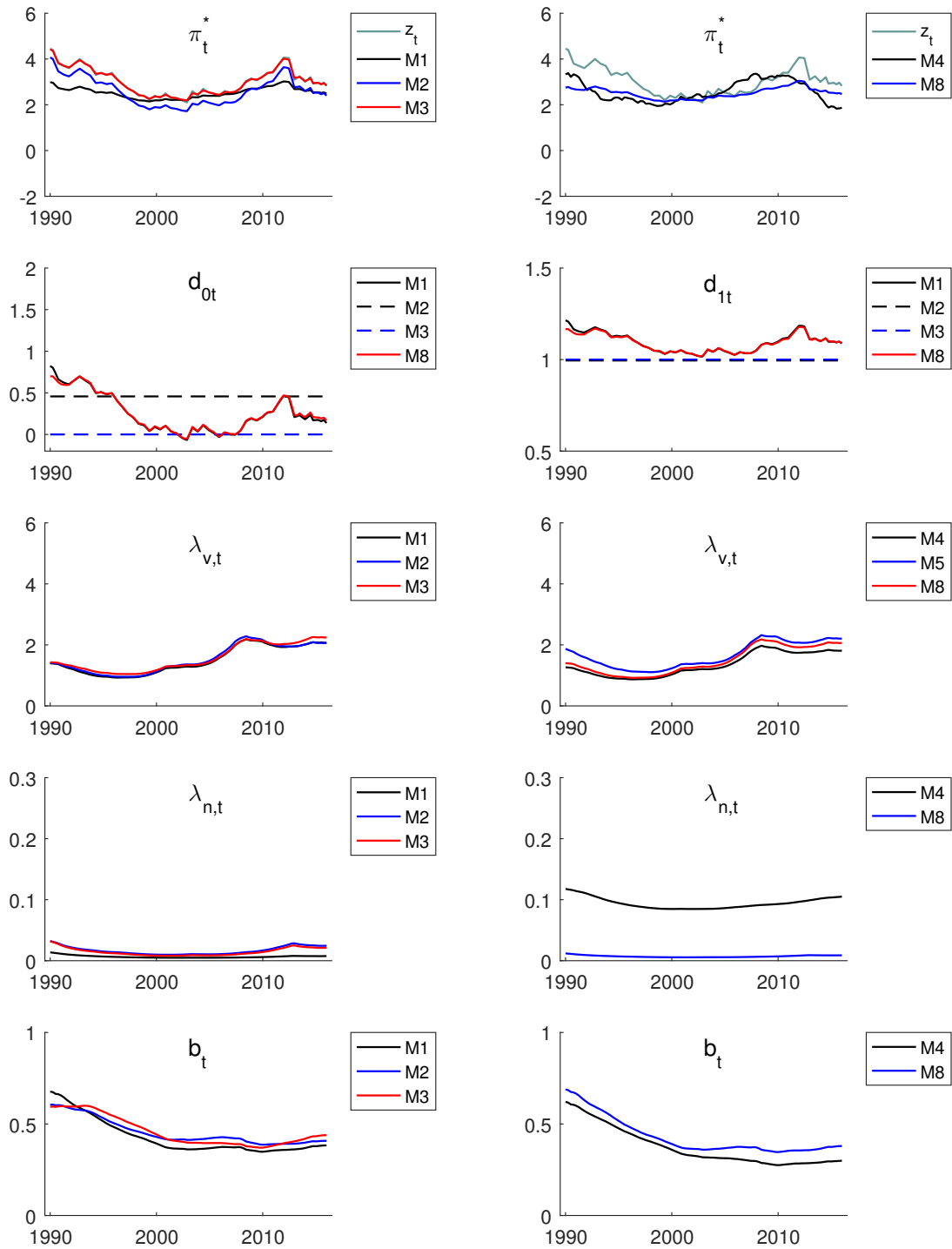
Figure 19: Comparison of posterior means of $\pi_t^*, b_t, \lambda_{v,t}, \lambda_{n,t}, d_{0t}$ and d_{1t} for different models, for the UK

M5 but about the same in fit as model M1. For the UK, M1 fits the data significantly better than both models M4 and M5. Across countries, the evidence regarding restrictions on the d_0 and d_1 coefficients is mixed. For Italy, the restrictions of models M2 and M3 neither harm nor help model fit. For Japan, imposing 0,1 restrictions significantly reduces model fit, but making the coefficients constants to be estimated significantly improves fit, making M2 the best-fitting specification. For the UK, imposing the restrictions of models M2 and M3 slightly reduces model fit. Extending the baseline model to allow the inflation expectation z_t to depend on the lagged inflation gap — as does model M8 — does not appear to improve model fit.

On balance, we interpret these international results as indicating that our model, which allows information about professionals' forecasts to help estimate trend inflation without imposing the restrictions that effectively equates such forecasts with trend inflation, is working successfully in a variety of countries with different inflationary experiences. Put another way, the evidence points to the value of using long-run expectations to help estimate trend inflation without imposing restrictions (constant coefficients of 0 and 1 in our model) that effectively equate the two.

6   Discussion

We have proposed our new model for several reasons. First, as a way of improving estimates of trend inflation by drawing strength from surveys of professional forecasters. Second, as a way of investigating whether these surveys are unbiased in the sense used in this paper (i.e., that they provide unbiased estimates of an econometric estimate of trend inflation or, equivalently, that $d_{0t} = 0$ and $d_{1t} = 1$). Third, as a model that might improve on existing specifications in fitting historical inflation data. Finally, as a simple model that might improve inflation forecasts over other simple models such as UCSV. Recently, there have been several influential papers which attempt to address the question of why survey-based forecasts might be biased. This is not the main focus of our paper, but some short discussion of this issue is warranted.

Our model allows for the possibility that survey expectations may become disconnected from the longer-run trend in inflation. That disconnection could take the relatively modest form of a systematic bias, or it could take the form of a more dramatic departure from rational expectations, with the survey expectation showing little connection to the longer-run trend in inflation. Studies such as Coibion and Gorodnichenko (2015) have presented evidence that survey forecasts depart from rationality in that they are subject to sluggish adjustment consistent with information rigidities. Mertens and Nason (2015) develop a joint model of inflation and inflation forecasts that permits time variation in the strength of the information rigidities. In light of this evidence, we have used our sample of Blue Chip forecasts of CPI inflation to produce estimates of the Coibion-Gorodnichenko stickiness regression. In this data, these estimates do not point to strong evidence of such information rigidities (however, this does not rule out bias or other manifestations of irrationality in the forecasts).[18] Although this evidence can be seen as supporting our

---

[18]We ran the Coibion-Gorodnichenko regression using both long-horizon and short-horizon Blue Chip forecasts of CPI inflation over various samples.

development of a model that does not impose parametric restrictions consistent with the Coibion-Gorodnichenko framework, our model can be seen as incorporating features that could capture the effects of information rigidities in a flexible way. In broad terms, our model can be seen as similar to that of Mertens and Nason (2015); the differences reflect a deliberate choice to impose fewer parametric restrictions and permit greater flexibility, particularly in the representation of survey forecasts of inflation. More specifically, if we abstract from time-varying parameters and volatilities for simplicity, our process for actual inflation is quite similar to that in Mertens and Nason. In the process for the survey forecast (recall also that we differ in our specification of the forecast horizon), Mertens and Nason incorporate a hierarchical structure with an additional latent state for the inflation forecast on which the observed survey forecast depends, with the latent state incorporating autoregressive dynamics and trend inflation, whereas we instead relate the observed survey forecast to a time-varying intercept, trend inflation with a time-varying coefficient, and an MA error term. To the extent the survey forecasts feature stickiness, this stickiness can be subsumed in our time-varying intercept and MA error term.

More broadly, some recent research develops frameworks in which long-run inflation expectations might depart from measures of trend inflation. Cecchetti, et al. (2017) argue that inflation expectations need not be very closely tied to inflation. Based on their empirical results, they characterize inflation as fluctuating around a time-varying local mean (which they also capture with a random walk process like our trend). They suggest that survey-based inflation expectations have value as indicators of the time-varying local mean but otherwise have little information content for inflation. At a high level, their interpretation of the data would seem to be one in which survey measures of long-run inflation could be biased measures of trend inflation (their local mean). More formally, some other recent work by Hills, Nakata, and Schmidt (2016) on complexities created by monetary policy constrained by the zero lower bound on interest rates indicates that average inflation can consistently fall short of central bank targets. Results in Kiley and Roberts (2017) and other studies cited therein suggest the need for a risk adjustment to policy interest rates, to make policy more accommodative than it might otherwise be, in order for central banks to achieve inflation targets. Until the public comes to fully appreciate such implications, it seems plausible that survey-based measures of inflation expectations might exceed average (trend) inflation. Admittedly, such concerns with zero lower bound constraints only apply to fairly recent data for most countries, except Japan, for which the relevant history is longer. However, they point to the fact that much remains to be done to fully understand the dynamics of inflation and inflation expectations, as described in sources such as Bernanke (2007).

7 Summary and Conclusion

In this paper, we have developed a bivariate model of inflation and inflation expectations that incorporates empirically-important features such as time-varying parameters and stochastic volatility. In a broad sense, we have used our model to investigate the relationship between these two variables. In a narrower sense, we have investigated the degree to which survey-based long-run inflation forecasts can be used to inform estimates

of trend inflation (e.g., by increasing precision), improve the fit of historical inflation data, and improve the accuracy of out-of-sample forecasts. In an extensive empirical exercise involving three combinations of measures of U.S. inflation and long-run inflation forecasts, we find a consistent story: Long-run inflation forecasts do provide useful additional information in informing estimates of trend inflation and in improving the fit of inflation models. However, the forecasts themselves cannot simply be equated with trend inflation. In out-of-sample forecasting, our model yields point and density forecasts that are at least as good as those from other models that have been found successful in the inflation forecasting literature. In estimates for Italy, Japan, and the UK, we find a similar story in most cases. However, for Italy we find simply equating trend inflation with long-run forecasts by the professionals' may be sufficient. However, it is reassuring that we are uncovering this result in the context of a flexible econometric model instead of simply imposing it a priori.

The history captured by our estimates indicates the distinction between trend inflation and long-run inflation expectations captured by surveys in the U.S. is practically important. For example, as noted in the introduction, for most of the period since 2008, inflation in the PCE price index has run below the Federal Reserve's longer-run inflation objective of 2 percent. Over the past couple of years, inflation has declined to very low levels. Yet, for several years before the recession that began in 2007, inflation ran steadily above target. Some estimates of trend inflation based entirely on inflation — as in the UCSV specification of Stock and Watson (2007) — have moved around with inflation, rising in the early to mid-2000s and declining markedly from the mid-2000s through 2016. At the other extreme, long-run inflation expectations measured from the Survey of Professional Forecasters have remained steady around 2 percent (with occasional up-ticks and down-ticks). Drawing on the information in both inflation and the survey's long-run expectation, our model's estimate of trend is much smoother than the estimate from a univariate UCSV specification, implying the trend to be stable in the face of both the rise of inflation in the years before the recession and the fall since the recession. In fact, our model estimates show trend inflation to be even more stable than the survey expectation (containing a little less noise than the survey). However, in keeping with a historical bias in the survey forecast, our estimate of trend inflation has for some time been stable, slightly below the survey expectation. At times, the differences in models could have important practical implications for policy. In particular, as we show in out of sample forecasts through 2021, the forecast of headline inflation from our preferred model is much more consistent with the Federal Reserve's longer-run inflation objective than is the forecast from a UCSV specification similar to that of Stock and Watson (2007).

References

Aruoba, B. (2016). "Term Structures of Inflation Expectations and Real Interest Rates." Manuscript, University of Maryland.

Bednar, W., and T. Clark. (2014). "Methods for Evaluating Recent Trend Inflation." Economic Trends, Federal Reserve Bank of Cleveland, March 28.

Bernanke, B. (2007). "Inflation Expectations and Inflation Forecasting." Speech at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts, July 10.

Cecchetti, S., Feroli, M., Hooper, P., Kashyap, A. and Schoenholtz, K. (2017). "Deflating Inflation Expectations: The Implications of Inflation's Simple Dynamics." Manuscript prepared for the 2017 U.S. Monetary Policy Forum.

Chan, J. (2013). "Moving Average Stochastic Volatility Models with Application to Inflation Forecast." Journal of Econometrics 176, 162–172.

Chan, J. and E. Eisenstat. (2015). "Bayesian Model Comparison for Time-varying Parameter VARs with Stochastic Volatility."CAMA Working Paper 32/2015.

Chan, J., and I. Jeliazkov. (2009). "Efficient Simulation and Integrated Likelihood Estimation in State Space Models." International Journal of Mathematical Modelling and Numerical Optimisation 1, 101-120.

Chan, J., G. Koop, and S.M. Potter. (2013). "A New Model of Trend Inflation." Journal of Business and Economic Statistics 31, 94-106.

Chan, J., G. Koop, and S.M. Potter. (2016). "A Bounded Model of Time Variation in Trend Inflation, NAIRU and the Phillips Curve." Journal of Applied Econometrics 31, 551-565.

Clark, T., and T. Doh. (2014). "Evaluating Alternative Models of Trend Inflation." International Journal of Forecasting 30, 426-448.

Coibion, O., and Y. Gorodnichenko. (2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." American Economic Review 105, 2644-2678.

Cogley, T., G. Primiceri, and T. Sargent. (2010). "Inflation-gap Persistence in the US." American Economic Journal: Macroeconomics 2, 43-69.

D'Agostino, A., L. Gambetti, and D. Giannone. (2013). "Macroeconomic Forecasting and Structural Change." Journal of Applied Econometrics 28, 82-101.

Del Negro, M., D. Giannone, M. Giannoni, and A. Tambalotti. (2017). "Safety, Liquidity, and the Natural Rate of Interest." Brookings Papers on Economic Activity, forthcoming.

Del Negro, M., and F. Schorfheide. (2013). "DSGE Model-based Forecasting." In G. Elliott & A. Timmermann (Eds.), Handbook of Economic Forecasting, Volume 2. Amsterdam: North Holland.

Erceg, C. and A. Levin. (2003). "Imperfect Credibility and Inflation Persistence," 50, 915–944.

Faust, J., and J. Wright. (2013). "Forecasting Inflation." In G. Elliott & A. Timmermann (Eds.), Handbook of Economic Forecasting, Volume 2. Amsterdam: North Holland.

Fuhrer, J., G. Olivei, and G. Tootell. (2012). "Inflation Dynamics When Inflation is Near Zero." *Journal of Money, Credit and Banking,* 44, 83-122.

Garnier, C., E. Mertens, and E. Nelson. (2015). "Trend Inflation in Advanced Economies." International Journal of Central Banking 11 (Supplement 1), 65-136.

Jarocinski, M., and M. Lenza. (2015). "Output Gap and Inflation Forecasts in a Bayesian Dynamic Factor Model of the Euro Area." Manuscript, European Central Bank.

Kass, R., and A. Raftery. (1995). "Bayes Factors." Journal of the American Statistical Association 90, 773-795.

Kiley, M., and J. Roberts. (2017). "Monetary Policy in a Low Interest Rate World." Brookings Papers on Economic Activity, forthcoming.

Kim, S., N. Shepherd and S. Chib. (1998). "Stochastic Volatility: Likelihood Inference and Comparison with ARCH models." Review of Economic Studies 65, 361-393.

Koop, G. (2003). Bayesian Econometrics. New York: John Wiley & Sons.

Koop, G., and D. Korobilis. (2010). "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics." Foundations and Trends in Econometrics 3, 267-358.

Kozicki, S., and P. Tinsley. (2001). "Shifting Endpoints in the Term Structure of Interest Rates." Journal of Monetary Economics 47, 613-652.

Kozicki, S., and P. Tinsley. (2012). "Effective Use of Survey Information in Estimating the Evolution of Expected Inflation." Journal of Money, Credit and Banking 44, 145-169.

Mertens, E. (2016). "Measuring the Level and Uncertainty of Trend Inflation." Review of Economics and Statistics, 98, 950-967.

Mertens, E. and Nason, J. (2015). "Inflation and Professional Forecast Dynamics: An Evaluation of Stickiness, Persistence and Volatility." Manuscript available at http://www.elmarmertens.com/research/workingpapers.

Mehrotra, A., and J. Yetman. (2014). "Decaying Expectations: What Inflation Forecasts Tell us About the Anchoring of Inflation Expectations." BIS Working Papers No. 464.

Mishkin, F. (2007). "Inflation Dynamics." Speech at the Annual Macro Conference, Federal Reserve Bank of San Francisco, March 23.

Morley, J., J. Piger, and R. Rasche. (2015). "Inflation in the G7: Mind the Gap(s)?" Macroeconomic Dynamics 19, 883-912.

Hills, T., T. Nakata, and S. Schmidt. (2016). "The Risky Steady State and the Interest Rate Lower Bound." manuscript.

Nason, J., and G. Smith. (2014). "Measuring the Slowly Evolving Trend in US Inflation with Professional Forecasts." Centre for Applied Macroeconomic Analysis, Australian National University, Working Paper 7/2014.

Nason, J., and G. Smith. (2016). "Sticky Professional Forecasts and the Unobserved Components Model of US Inflation." Manuscript, North Carolina State University.

Rudd, J., and E. Peneva. (2015). "The Passthrough of Labor Costs to Price Inflation." Finance and Economics Discussion Series Working Paper 2015-042, Board of Governors of the Federal Reserve System.

Stella, A., and J. Stock. (2013). "A State-dependent Model for Inflation Forecasting." Board of Governors of the Federal Reserve System, International Finance Discussion Papers, Number 1062.

Stock, J. and M. Watson. (2007). "Why has U.S. Inflation Become Harder to Forecast?" Journal of Money, Credit and Banking 39, 3-33.

Stock, J. and M. Watson. (2015). "Core Inflation and Trend Inflation." Review of Economics and Statistics, forthcoming.

Wright, J. (2013). "Evaluating Real Time VAR Forecasts with an Informative Democratic Prior." Journal of Applied Econometrics 28, 762-776.

Technical Appendix

This appendix provides the prior and MCMC algorithm used in this paper for our baseline model (model M1) and its variants, details the estimation algorithm, explains the computation of marginal likelihoods, and provides the priors for the TVP-VAR and TVP-VECM used in forecast comparison.

A   Priors for Model M1

The model is given in equations (3) through (10). We initialize the state equations (5), (6), (7), and (10) by $\pi_1^* \sim N(\pi_0^*, \lambda_{n,1} V_{\pi^*})$, $b_1 \sim N(b_0, V_b)$, $d_{i1} \sim N(\mu_{di}, \sigma_{di}^2/(1 - \rho_{di}^2))$, $i = 0, 1$, and $\log(\lambda_{i,1}) \sim N(\log(\lambda_{i,0}), V_{\lambda_i})$, $i = v, n$, with $\lambda_{i,0} = 1$, $b_0 = \pi_0^* = 0$ and $V_{\lambda_i} = V_b = V_{\pi^*} = 100$.

Our prior choices are motivated by a desire to use relatively non-informative choices, but to reflect the prior beliefs about the parameters in our structural time-series model. We note that, for some parameters, sensible prior information can easily be elicited. For instance, if the professionals are providing unbiased forecasts of trend inflation, then $d_{0t} = 0$ and $d_{1t} = 1$ and we would expect relatively little deviation from these values. Our prior reflects such information. In a prior sensitivity analysis (not reported here), we found a fair degree of robustness to prior choices (e.g., doubling or halving prior standard deviations had little effect). However, using an extremely non-informative prior in some cases leads to extremely erratic estimates of some of the states in our state space model. The avoidance of extremely non-informative priors is common with structural time series models such as ours. Indeed papers with simpler models than ours, such as Stock and Watson (2007), do not even attempt to estimate some parameters, instead setting them to specific values. More recent applications of the UCSV model often still calibrate key parameters as opposed to estimating them (see, e.g., Cecchetti, Feroli, Hooper, Kashyap and Schoenholtz, 2017).

For later reference, let $\pi = (\pi_1, \ldots, \pi_T)'$ and $d = (d_{01}, d_{11}, \ldots, d_{0T}, d_{1T})'$, and similarly define z, π^*, b, λ_v and λ_n. In addition, let θ denote the model parameters, i.e., $\theta = (\psi, \mu_{d0}, \mu_{d1}, \rho_{d0}, \rho_{d1}, \sigma_{d0}^2, \sigma_{d1}^2, \sigma_b^2, \sigma_z^2, \phi_v, \phi_n)'$.

We assume independent priors for elements of the parameter vector θ which are proper and weakly informative. The priors for μ_{di} and ρ_{di} are:

$$\mu_{d0} \sim N(a_0, V_\mu), \quad \mu_{d1} \sim N(a_1, V_\mu), \quad \rho_{di} \sim TN_{(c_1, c_2)}(a_2, V_\rho),$$

where $TN_{(c_1, c_2)}(\mu, \sigma)$ denotes the $N(\mu, \sigma)$ distribution truncated to the interval (c_1, c_2), and we set $a_0 = 0$, $a_1 = 1$, $a_2 = 0.95$, $V_\mu = 0.1^2$ and $V_\rho = 0.1^2$. These choices imply relatively informative priors centered at the values which imply trend inflation is equal to long-run inflation forecasts (apart from a mean zero error).

For the MA(1) coefficient, we consider the relatively non-informative prior which restricts the MA process to be invertible: $\psi \sim TN_{(-1,1)}(0, V_\psi)$ with $V_\psi = 0.25^2$.

Finally, we assume independent inverse gamma priors for the variance parameters. In particular, the degree of freedom parameters are all set to the relatively non-informative value of 5, and the scale parameters are set such that $E(\sigma_{d0}^2) = E(\sigma_w^2) = E(\phi_v) = E(\phi_n) = 0.01$ and $E(\sigma_{d1}^2) = E(\sigma_b^2) = 0.001$. The prior hyperparameter choices reflect

a desire to allow for appreciable parameter change in these models, but such change to occur smoothly. The prior mean for σ_{d0}^2 implies that with high probability the difference between consecutive d_{0t} lies within the values -0.2 and 0.2. This means that, in each period, d_{0t} is not expected to change by a large amount. However, if such changes were to occur several periods in a row, appreciable change in d_{0t} could be accommodated.

For those models that are variants of the baseline model M1, we use the same prior for all parameters which they hold in common with M1. For the model M2 that restricts the d_0 and d_1 coefficients to be constant over time, we use prior means of 0 and 1, respectively, with prior variances of $V_\mu = 0.5^2$ and $V_\rho = 0.5^2$. For the additional coefficient d_{2t} of model M8, we use a prior on the initial value that takes the same form and values used for the coefficient d_{1t}, with a mean of 0 and standard deviation of 0.1.

B   Estimation Algorithm for Model M1

To estimate the model in equations (3) through (10), we extend the MCMC sampler developed in Chan, Koop and Potter (2013) which was used for a univariate bounded inflation trend model. Moreover, we also incorporate the sampler in Chan (2013) for handling the MA innovations with stochastic volatility. Specifically, we sequentially draw from the following densities:

1. $p(\pi^* \,|\, Data, b, d, \lambda_v, \lambda_n, \theta)$;

2. $p(b \,|\, Data, \pi^*, d, \lambda_v, \lambda_n, \theta)$;

3. $p(d \,|\, Data, \pi^*, b, \lambda_v, \lambda_n, \theta)$;

4. $p(\lambda_v, \lambda_n \,|\, Data, \pi^*, b, d, \theta)$;

5. $p(\mu_{d0}, \mu_{d1} \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\mu_{d0}, \mu_{d1}\}})$;

6. $p(\sigma_{d0}^2, \sigma_{d1}^2 \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\sigma_{d0}^2, \sigma_{d1}^2\}})$;

7. $p(\rho_{d0}, \rho_{d1} \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\rho_{d0}, \rho_{d1}\}})$;

8. $p(\psi \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\psi\}})$;

9. $p(\sigma_b^2, \sigma_z^2, \phi_v, \phi_n \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\sigma_b^2, \sigma_z^2, \phi_v, \phi_n\}})$.

Step 1: To implement Step 1, note that information about π^* comes from three sources: the two measurement equations (3) and (4), and the state equation (5). We derive an expression for each component in turn. First, write (3) as

$$H_b \pi = H_b \pi^* + \tilde{\alpha}_{\pi^*} + v, \quad v \sim N(0, \Lambda_v),$$

where $\tilde{\alpha}_{\pi^*} = (b_1(\pi_0 - \pi_0^*), 0, \ldots, 0)'$, $\Lambda_v = \mathrm{diag}(\lambda_{v,1}, \ldots, \lambda_{v,T})$ and

$$
H_b = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
-b_2 & 1 & 0 & \cdots & 0 \\
0 & -b_3 & 1 & \cdots & 0 \\
\vdots & & & \ddots & \vdots \\
0 & 0 & \cdots & -b_T & 1
\end{pmatrix}.
$$

Since $|H_b| = 1$ for any b, H_b is invertible. Therefore, we have

$$
(\pi \mid \pi^*, b, \lambda_v) \sim N(\pi^* + \alpha_{\pi^*}, (H_b' \Lambda_v^{-1} H_b)^{-1}),
$$

with log density

$$
\log p(\pi \mid \pi^*, b, \lambda_v) \propto -\frac{1}{2}(\pi - \pi^* - \alpha_{\pi^*})' H_b' \Lambda_v^{-1} H_b (\pi - \pi^* - \alpha_{\pi^*}), \tag{22}
$$

where $\alpha_{\pi^*} = H_b^{-1} \tilde{\alpha}_{\pi^*}$. Note that H_b is a band matrix and α_{π^*} can be obtained quickly by solving the band system $H_b x = \tilde{\alpha}_{\pi^*}$ for x without computing the inverse H_b^{-1}.

The second component comes from (4) which can be written as:

$$
z = d_0 + X_{\pi^*} \pi^* + H_\psi \epsilon_z, \quad \epsilon_z \sim N(0, \sigma_w^2 I_T),
$$

where $d_0 = (d_{01}, \ldots, d_{0T})'$, $X_{\pi^*} = \mathrm{diag}(d_{11}, \ldots, d_{1T})$ and

$$
H_\psi = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
\psi & 1 & 0 & \cdots & 0 \\
0 & \psi & 1 & \cdots & 0 \\
\vdots & & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & \psi & 1
\end{pmatrix}.
$$

Thus, ignoring any terms not involving π^*, we have

$$
\log p(z \mid \pi^*, d, \sigma_w^2) \propto -\frac{1}{2\sigma_w^2}(z - d_0 - X_{\pi^*}\pi^*)'(H_\psi H_\psi)'^{-1}(z - d_0 - X_{\pi^*}\pi^*),
$$

$$
= -\frac{1}{2\sigma_w^2}(\tilde{z} - \tilde{X}_{\pi^*}\pi^*)'(\tilde{z} - \tilde{X}_{\pi^*}\pi^*), \tag{23}
$$

where $\tilde{z} = H_\psi^{-1}(z - d_0)$ and $\tilde{X}_{\pi^*} = H_\psi^{-1} X_{\pi^*}$. Since H_ψ is a band matrix, $\tilde{z}$ can be computed quickly by solving a linear system of equations without finding the inverse H_ψ^{-1}. The matrix $\tilde{X}_{\pi^*}$ is lower triangular that is in general not banded. However, most of the elements away from the main diagonal band are close to zero. In our implementation we construct a band approximation by replacing all elements below 10^{-6} with 0. Since the cut-off point is so small, it has no impact on the results, but it substantially speeds up the computation.

The third component is contributed by the state equation (5):

$$
\log p(\pi^* \mid \lambda_n) \propto -\frac{1}{2}(\pi^* - \delta_{\pi^*})' H' \Lambda_n^{-1} H(\pi^* - \delta_{\pi^*}), \tag{24}
$$

45

where H is the $T \times T$ first difference matrix, $\Lambda_n = \text{diag}(\lambda_{n,1} V_{\pi^*}, \lambda_{n,2}, \ldots, \lambda_{n,T})$ and $\delta_{\pi^*} = H^{-1}(\pi_0^*, 0, \ldots, 0)'$. Then, combining (22), (23) and (24), we finally obtain

$$\log p(\pi^* \mid Data, b, d, \lambda_v, \lambda_n, \theta)$$

$$\propto -\frac{1}{2}(\pi - \pi^* - \alpha_{\pi^*})' H_b'^{-1}{}_v H_b (\pi - \pi^* - \alpha_{\pi^*}) - \frac{1}{2\sigma_w^2}(\tilde{z} - \tilde{X}_{\pi^*}\pi^*)'(\tilde{z} - X_{\pi^*}\pi^*)$$

$$- \frac{1}{2}(\pi^* - \delta_{\pi^*})' H' \Lambda_n^{-1} H (\pi^* - \delta_{\pi^*}),$$

$$\propto -\frac{1}{2}(\pi^* - \hat{\pi}^*)' K_{\pi^*}(\pi^* - \hat{\pi}^*),$$

which is the kernel of the $N(\hat{\pi}^*, K_{\pi^*}^{-1})$ distribution, where

$$K_{\pi^*} = \left(H_b'^{-1}{}_v H_b + \frac{1}{\sigma_w^2}\tilde{X}_{\pi^*}'\tilde{X}_{\pi^*} + H'\Lambda_n^{-1} H \right)^{-1},$$

$$\hat{\pi}^* = K_{\pi^*}^{-1}\left(H_b'^{-1} H_b(\pi - \alpha_{\pi^*}) + \frac{1}{\sigma_w^2}\tilde{X}_{\pi^*}'\tilde{z} + H'\Lambda_n^{-1} H\delta_{\pi^*} \right).$$

If we use the band approximation of $\tilde{X}_{\pi^*}$ as described above, the precision K_{π^*} is also a band matrix. Then, we use the precision sampler in Chan and Jeliazkov (2009) to sample π^* from the conditional distribution $(\pi^* \mid Data, b, d, \lambda_v, \lambda_n, \theta)$.

Step 2: Next, we derive the conditional density $p(b \mid Data, \pi^*, d, \lambda_v, \lambda_n, \theta)$. Due to the inequality restriction $0 < b_t < 1$, this joint density is non-normal. We first rewrite (3) as:

$$\tilde{\pi} = X_b b + v, \quad v \sim N(0, \Lambda_v),$$

where $\tilde{\pi} = (\pi_1 - \pi_1^*, \ldots, \pi_T - \pi_T^*)'$ and $X_b = \text{diag}(\pi_0 - \pi_0^*, \ldots, \pi_{T-1} - \pi_{T-1}^*)$. It follows that the log density of $(\pi \mid \pi^*, b, \lambda_v)$ can also be written as follows:

$$\log p(\pi \mid \pi^*, b, \lambda_v) \propto -\frac{1}{2}(\tilde{\pi} - X_b b)'\Lambda_v^{-1}(\tilde{\pi} - X_b b). \tag{25}$$

Next, write (6) as

$$Hb = \tilde{\delta}_b + \epsilon_b,$$

where $\tilde{\delta}_b = (b_0, 0, \ldots, 0)'$ and the elements of ϵ_b are independent truncated normal variables. Note that $\Pr(0 < b_1 < 1) = \Phi((1 - b_0)/\sqrt{V_b}) - \Phi(b_0/\sqrt{V_b})$ and

$$\Pr(0 < b_t < 1) = \Phi\left(\frac{1 - b_{t-1}}{\sigma_b}\right) - \Phi\left(\frac{-b_{t-1}}{\sigma_b}\right),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Hence, the prior density for b is given by

$$\log p(b \mid \sigma_b^2) \propto \frac{1}{2}(b - \delta_b)' H' \Sigma_b^{-1} H (b - \delta_b) + g_b(b, \sigma_b^2), \tag{26}$$

where $\Sigma_b = \text{diag}(V_b, \sigma_b^2, \ldots, \sigma_b^2)$, $\delta_b = H^{-1}\tilde{\delta}_b$ and

$$g_b(b, \sigma_b^2) = -\sum_{t=2}^{T} \log\left(\Phi\left(\frac{1 - b_{t-1}}{\sigma_b}\right) - \Phi\left(\frac{-b_{t-1}}{\sigma_b}\right) \right).$$

Combining (25) and (26), we obtain

$$\log p(b \,|\, Data, \pi^*, d, \lambda_v, \lambda_n, \theta) \propto -\frac{1}{2}(b - \hat{b})' K_b^{-1}(b - \hat{b}) + g_b(b, \sigma_b^2),$$

where

$$K_b = \left(H'\Sigma_b^{-1}H + X_b'\Lambda_v^{-1}X_b\right)^{-1}, \quad \hat{\tau}^\pi = K_b^{-1}(H_b'^{-1}\tilde{\delta}_b + X_b'\Lambda_v^{-1}\tilde{\pi}).$$

We follow Chan, Koop and Potter (2013) to sample b. Specifically, candidate draws are first obtained from the $N(\hat{b}, K_b^{-1})$ distribution using the precision sampler in Chan and Jeliazkov (2009), and they are accepted or rejected via an acceptance-rejection Metropolis-Hastings step.

Step 3: To sample from $p(d \,|\, Data, \pi^*, b, \lambda_v, \lambda_n, \theta)$, we first rewrite (4) and (7) as

$$z = X_d d + H_\psi \varepsilon_z, \qquad\qquad \varepsilon_z \sim N(0, \sigma_w^2 I_T),$$
$$H_{\rho_d} d = \tilde{\delta}_d + \varepsilon_d \qquad\qquad \varepsilon_d \sim N(0, \Sigma_d),$$

where $\tilde{\delta}_d = (\mu_{d0}, \mu_{d1}, (1-\rho_{d0})\mu_{d0}, (1-\rho_{d1})\mu_{d1}, \ldots, (1-\rho_{d0})\mu_{d0}, (1-\rho_{d1})\mu_{d1})'$, $\Sigma_d = \mathrm{diag}(\sigma_{d0}^2/(1-\rho_{d0}^2), \sigma_{d1}^2/(1-\rho_{d1}^2), \sigma_{d0}^2, \sigma_{d1}^2, \ldots, \sigma_{d0}^2, \sigma_{d1}^2)$,

$$X_d = \begin{pmatrix} 1 & \pi_1^* & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \pi_2^* & 0 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 1 & \pi_T^* \end{pmatrix}, \quad H_{\rho_d} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ -\rho_{d0} & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -\rho_{d1} & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & -\rho_{d0} & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & -\rho_{d1} & 0 & 1 \end{pmatrix}.$$

Using standard linear regression results (see, e.g., Koop, 2003, pp. 60-61), we have $(d \,|\, Data, \pi^*, b, \lambda_v, \lambda_n, \theta) \sim N(\hat{d}, K_d^{-1})$, where

$$K_d = \left(H_{\rho_d}' \Sigma_d^{-1} H_{\rho_d} + \frac{1}{\sigma_w^2}\tilde{X}_d'\tilde{X}_d\right)^{-1}, \quad \hat{d} = K_d^{-1}\left(H_{\rho_d}'\Sigma_d^{-1}\tilde{\delta}_d + \frac{1}{\sigma_w^2}\tilde{X}_d'(H_\psi^{-1}z)\right)$$

with $\tilde{X}_d = H_\psi^{-1}X_d$. As before, we construct a band approximation of $\tilde{X}_d$ by replacing all elements less than 10^{-6} with 0. Then, the precision K_d is a band matrix and the precision sampler in Chan and Jeliazkov (2009) is used to sample d.

Step 4: To implement Step 4, note that λ_v and λ_n are conditionally independent given the parameters and other states. Hence, we can draw them sequentially using the auxiliary mixture sampler of Kim, Shepherd and Chib (1998). See also Koop and Korobilis (2010), p. 308–310, for a textbook treatment. Note that in conventional implementations, a forward-filtering-backward-smoothing algorithm is used; here it is replaced by the more efficient precision sampler of Chan and Jeliazkov (2009).

Steps 5 and 6: Both the densities of (μ_{d0}, μ_{d1}) and $(\sigma_{d0}^2, \sigma_{d1}^2)$ are standard. In fact, we have

$$(\mu_{di} \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\mu_{d0}, \mu_{d1}\}}) \sim N(\hat{\mu}_{di}, K_{di}^{-1}),$$
$$(\sigma_{di}^2 \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\sigma_{d0}^2, \sigma_{d1}^2\}}) \sim IG(\nu_{di} + T/2, \tilde{S}_{di}),$$

47

where $K_{di} = 1/V_\mu + (1-\rho_{di}^2)/\sigma_{di}^2 + (T-1)(1-\rho_{di})^2/\sigma_{di}^2$, $\hat{\mu}_{di} = K_{di}^{-1}(a_i/V_\mu + (1-\rho_{di}^2)d_{i1}/\sigma_{di}^2 + \sum_{t=2}^{T}(1-\rho_{di})(d_{it} - \rho_{di}d_{i,t-1})/\sigma_{di}^2)$ and $\tilde{S}_{di} = S_{di} + ((1-\rho_{di}^2)(d_{i1} - \mu_{di})^2 + \sum_{t=2}^{T}(d_{it} - \mu_{di}(1-\rho_{di}) - \rho_{di}d_{i,t-1})^2)/2$.

Step 7: It follows from (7) that

$$p(\rho_{di} \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\rho_{d0}, \rho_{d1}\}}) \propto p(\rho_{di})g_{\rho_{di}}(\rho_{di})e^{-\frac{1}{2\sigma_{di}^2}\sum_{t=2}^{T}(d_{it} - \mu_{di} - \rho_{di}(d_{i,t-1} - \mu_{di}))^2},$$

where $p(\rho_{di})$ is the truncated normal prior for ρ_{di} and $g(\rho_{di}) = (1-\rho_{di}^2)^{1/2}\exp(-\frac{1}{2\sigma_{di}^2}(1-\rho_{di}^2)(d_{i1} - \mu_{di})^2)$. This conditional density is non-standard, and we implement an independence-chain Metropolis-Hastings step with proposal distribution $N(\hat{\rho}_{di}, K_{\rho_{di}}^{-1})$, where $K_{\rho_{di}} = 1/V_\rho + X_{\rho_{di}}'X_{\rho_{di}}/\sigma_{di}^2$ and $\hat{\rho}_{di} = K_{\rho_{di}}^{-1}(a_2/V_\rho + X_{\rho_{di}}'y_{\rho_{di}}/\sigma_{di}^2)$, with $X_{\rho_{di}} = (d_{i1} - \mu_{di}, \ldots, d_{i,T-1} - \mu_{di})'$ and $y_{\rho_{di}} = (d_{i2} - \mu_{di}, \ldots, d_{iT} - \mu_{di})'$. Then, given the current draw ρ_{di}, a proposal ρ_{di}^* is accepted with probability $\min(1, g_{\rho_{di}}(\rho_{di}^*)/g_{\rho_{di}}(\rho_{di}))$; otherwise the Markov chain stays at the current state ρ_{di}.

Step 8: To sample ψ, note that

$$\log p(\psi \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\psi\}}) \propto \log p(z \,|\, \pi^*, d, \sigma_w^2) + \log p(\psi)$$

$$\propto -\frac{1}{2\sigma_w^2}(z - d_0 - X_{\pi^*}\pi^*)'(H_\psi H_\psi)'^{-1}(z - d_0 - X_{\pi^*}\pi^*) + \log p(\psi),$$

where $p(\psi)$ is the prior density of ψ. Following Chan (2013), we sample ψ via an independence-chain Metropolis-Hastings step. Specifically, since this log density can be quickly evaluated using band matrix routines, we maximize it numerically to obtain the mode and negative Hessian, denoted as $\hat{\psi}$ and K_ψ, respectively. Then, we generate candidate draws from the $N(\hat{\psi}, K_\psi^{-1})$ distribution.

Step 9: To sample $\sigma_b^2, \sigma_w^2, \phi_v$ and ϕ_n, first note that these parameters are conditionally independent given the data and the states. Hence, we can sample each element one by one. The variance parameters σ_w^2, ϕ_v and ϕ_n follow inverse-Gamma distributions:

$$(\sigma_w^2 \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\sigma_b^2, \sigma_w^2, \phi_v, \phi_n\}}) \sim IG\left(\nu_{\sigma_w^2} + \frac{T}{2}, S_{\sigma_w^2} + \frac{1}{2}\sum_{t=1}^{T}\epsilon_{z,t}^2\right),$$

$$(\phi_i \,|\, Data, \pi^*, b, d, \lambda_v, \lambda_n, \theta_{-\{\sigma_b^2, \sigma_w^2, \phi_v, \phi_n\}})$$

$$\sim IG\left(\nu_{\phi_i} + \frac{T-1}{2}, S_{\phi_i} + \frac{1}{2}\sum_{t=2}^{T}(\log(\lambda_{it}) - \log(\lambda_{i,t-1}))^2\right), i = v, n,$$

where the elements of ϵ_z can be computed as $\epsilon_z = H_\psi^{-1}(z - X_d d)$. Next, the log conditional density for σ_b^2 is given by

$$\log(\sigma_b^2 \,|\, Data, \pi^*, b, d, \lambda_n, \theta_{-\{\sigma_b^2, \sigma_w^2, \phi_v, \phi_n\}})$$

$$\propto -(\nu_{\sigma_b^2} + 1)\log\sigma_b^2 - \frac{S_{\sigma_b^2}}{\sigma_b^2} - \frac{T-1}{2}\log\sigma_b^2 - \frac{1}{2\sigma_b^2}\sum_{t=2}^{T}(b_t - b_{t-1})^2 + g_b(b, \sigma_b^2),$$

which is a nonstandard density. To proceed, we implement an MH step with the proposal density

$$IG\left(\nu_{\sigma_b^2} + \frac{T-1}{2}, S_{\sigma_b^2} + \frac{1}{2}\sum_{t=2}^{T}(b_t - b_{t-1})^2\right).$$

C   Computation of Marginal Likelihoods

Marginal likelihoods are calculated by decomposing the marginal density of the inflation data as the product of predictive likelihoods. Specifically, let $\pi_{1:t} = (\pi_1, \ldots, \pi_t)'$ denote the inflation data up to time t. Then, we can factor the marginal likelihood for model M_k as follows:

$$p(\pi \,|\, X_k, M_k) = \prod_{t=5}^{T} p(\pi_t \,|\, \pi_{1:t-1}, X_{1:t,k}, M_k),$$

where $p(\pi_{t+1} \,|\, \pi_{1:t}, M_k)$ is the predictive likelihood and X_k is a set of covariates used in model M_k such as the survey data or the unemployment gap.[19]

D   Priors for TVP-VAR and TVP-VECM Models

Regarding priors for the TVP-VAR and TVP-VECM models, let Ω_B denote the error covariance matrix in the state equation of the VAR coefficients. Similarly, define Ω_h to be the error covariance matrix corresponding to the log-volatilities. We assume the diagonal elements of Ω_B and Ω_h are independently distributed as

$$\omega_{B,i}^2 \sim IG(\nu_{B,i}, S_{B,i}), \qquad \omega_{h,j}^2 \sim IG(\nu_{h,j}, S_{h,j}), \qquad i = 1, \ldots, k_B,\ j = 1, \ldots, n,$$

where $k_B = (np+1)n + n(n-1)/2 = 11$. We set the hyperparameters as follows. For the degree of freedom parameters, they are assumed to be small: $\nu_{B,i} = \nu_{B,j} = 5$. The scale parameters are set so that the prior mean of $\omega_{h,j}^2$ is 0.1^2. In other words, the difference between consecutive log-volatilities is within 0.2 with probability of about 0.95. Similarly, the implied prior mean of $\omega_{B,i}^2$ is 0.01^2 if it is associated with a VAR coefficient and 0.1^2 for an intercept.

    Although we omit details in the interest of brevity, the models can be fitted using standard Gibbs samplers, such as the one described in Chan and Eisenstat (2015). Since the TVP-VECM can be written as a TVP-VAR, it can be estimated with the same methods used for the TVP-VAR.

---

[19]We discard an initial four predictive likelihoods to reduce the sensitivity to priors.