



Norwegian
Business School

This file was downloaded from BI Open, the institutional repository (open access) at BI Norwegian Business School <http://biopen.bi.no>

It contains the accepted and peer reviewed manuscript to the article cited below. It may contain minor differences from the journal's pdf version.

Guo, N., Zhang, B., & Cross, J. L. (2022). Time-varying trend models for forecasting inflation in Australia. *Journal of Forecasting*, 41(2), 316-330.
<https://onlinelibrary.wiley.com/doi/10.1002/for.2814>

Copyright policy of *Wiley*, the publisher of this journal:

Authors are permitted to self-archive the peer-reviewed (but not final) version of a contribution on the contributor's personal website, in the contributor's institutional repository or archive, subject to an embargo period of 24 months for social science and humanities (SSH) journals and 12 months for scientific, technical, and medical (STM) journals following publication of the final contribution.

<https://authorservices.wiley.com/author-resources/Journal-Authors/licensing/self-archiving.html>

Time-varying Trend Models for Forecasting Inflation in Australia*

Na Guo[†] Bo Zhang[‡] Jamie Cross[§]

February 18, 2021

Abstract

We investigate whether a class of trend models, which decompose a time series into an underlying trend and transitory component, with various error term structures can improve upon the forecast performance of commonly used time series models when forecasting CPI inflation in Australia. The main result is that trend models tend to provide more accurate point and density forecasts at medium to long forecasting horizons compared to conventional autoregressive and Phillips curve models. The best medium-term point forecasts come from a trend model with stochastic volatility in the transitory component and that with a moving average component, while long-run point forecasts are better made by trend models with stochastic volatilities and a moving average component. In a full sample study, we also find that trend models can capture various dynamics in periods of significance to the Australian economy which conventional models cannot. This includes the dramatic reduction in inflation when the RBA adopted inflation targeting, a one-off 10 per cent Goods and Services Tax inflationary episode in 2000, and then gradually decline in inflation since 2014.

Keywords: trend model, inflation forecast, Bayesian analysis, stochastic volatility

JEL classification codes: C11, C52, E31, E37

*This work is supported by the National Natural Science Foundation of China (71903142), and it is part of the research activities at the Centre for Applied Macroeconomics and Commodity Prices (CAMP) at the BI Norwegian Business School.

[†]School of Finance, Tianjin University of Finance and Economics, Tianjin, 300222 P. R. China, Laboratory for Fintech and Risk Management in Tianjin.

[‡]*Corresponding author:* School of Economics, Shanghai University, Shanghai, 200444 P. R. China, Centre for Applied Macroeconomic Analysis (CAMA), Australian National University, Australia. Email: bozhangyc@gmail.com

[§]BI Norwegian Business School, Center of Applied Macroeconomics and Commodity Prices (CAMP), Centre for Applied Macroeconomic Analysis (CAMA).

1 Introduction

Increases in the general price level of goods and services, known as inflation, affect the decisions of households, firms and governments making it one of the most important macroeconomic indicators. Since inflation can be affected by monetary policy, a central objective of many central banks around the world is to maintain a stable rate of inflation thereby achieving stable macroeconomic outcomes. For instance, in the early 1990s, the Reserve Bank of Australia (RBA) formulated an explicit inflation target of 2–3 per cent, on average, over the medium term, making it one of the early adopters of the inflation targeting framework (Dixon and Lim, 2004). Since then, Australian CPI inflation has typically ranged between 0-6 per cent in any given quarter. The Australian experience has shown that inflation is difficult to fine-tune within a narrow band. As a result, obtaining accurate inflation forecasts is critical in making correct policy decisions.

Despite being an early adopter of inflation targeting, the literature on forecasting Australian inflation is relatively sparse (see, e.g., Beechey and Österholm, 2010; Garnier et al., 2015; Cross and Poon, 2016; Zhang and Nguyen, 2020, and references therein). Moreover, none of these papers has investigated whether explicit modeling of the underlying trend can enhance forecast performance. This is surprising since fluctuations in the trend, as opposed to period-to-period fluctuations, are more in line with the RBA’s definition of a medium-term inflation target. With this potential shortcoming in mind, our objective in this paper is to determine whether a class of trend models with various error term structures can improve upon commonly used models in the literature.

To this end, we provide the first systematic study on forecasting Australia inflation using time-varying trend models with various specifications for flexible error structures.¹ In trend models, the time series is decomposed into an underlying trend and transitory component. In the literature on forecasting US CPI inflation, it has been shown that allowing for time-varying volatility in both components enhances overall forecast performance (Stock and Watson, 2007). Zhang et al. (2020) have also shown that models with stochastic volatility and various flexible innovations in the error terms can provide competitive forecasts for G7 economies. Since this is the first paper to consider such models for the Australian economy, we consider a trend model with stochastic volatility in the measurement equation (Trend-SV), a trend model with a moving average and stochastic volatility (Trend-SV-MA) (Chan, 2013), and also revisit the trend model with stochastic volatility in both the measurement equation and the state equation (Trend-2SV) (Stock and Watson, 2007). Set in this manner we can learn which model features best forecast inflation in Australia. In addition to these models, we also consider more conventionally used autoregressive (AR) and Phillips curve (PC) models, along with a combination based forecast. AR models have been shown to provide competitive forecasts of Australian inflation Cross and Poon (2016) and are therefore a natural benchmark for our analysis. PC models, which incorporate additional information from the unemployment

¹Such models have also been referred to as unobserved components models or trend-cycle models in the broader literature on inflation modeling (e.g., Stock and Watson, 2007; Chan, 2013).

rate, achieved remarkable forecasting performance in the U.S. from 1984 to 1996, and remain an important candidate model for US inflation forecasts (e.g., Staiger, Stock, and Watson, 1997; Brayton, Roberts, and Williams, 1999; Garratt, Mitchell, Vahey, and Wakerly, 2011). Such models have also been shown to be useful when forecasting inflation (Robinson et al., 2003) and modeling of other macroeconomic variables in Australia (Gruen et al., 2005).²

To remain consistent with Australia’s adoption of inflation targeting, our forecast evaluation sample ranges from 1993Q3-2019Q4.³ Our results show that trend models with stochastic volatility consistently performs well across medium to long forecast horizons in both point and density forecasts. More specifically, we find that Trend-SV forecasts well at medium to long horizons in point forecast and almost all forecast horizons in density forecasts, while Trend-2SV-MA and Trend-SV-MA have the best forecasting performance on longer horizons in point forecasts. In both point and density forecast analysis, we also find that the trend models can better predict underlying changes in the inflation dynamics as compared to the AR and PC models. In fact, in the parameter estimation section, trend models have already shown that they can capture the dramatic decrease in the underlying trend of inflation when the inflation targeting policy was implemented, a one-off 10 per cent Goods and Services Tax inflationary episode in 2000, and then gradually decline in inflation since 2014.

The remainder of the paper proceeds as follows. Section 2 describes the specifications of the trend models. Section 3 presents the simulation methods of the parameters, followed by the studies of parameter estimation. Section 4 discusses the forecast results of all the competing models for the recursive forecast, combination forecast, and rolling window forecast. Section 5 concludes.

2 Trend Models and Other Competing Models

The models used in this paper for forecasting Australia’s inflation can be divided into three groups. In the first instance, we introduce the trend model group, which is specified by the underlying trend of inflation and time-varying parameters. There are then two groups of competing models: The autoregressive model (AR) group and the Phillips curve (PC) group. Since PCs are often estimated using levels or first differences, we further split the PC group into two subgroups. The first subgroup specifies PCs in terms

²Gruen et al. (2005) examine Australian GDP data, and Phillips curve models present good forecasting performance on the output gap of Australia by using the Australian Real-Time Macroeconomics Database at the University of Melbourne. They conclude that reasonably reliable output gap estimates can be obtained in real-time despite the well-known problems of data revisions and endpoint problems in real-time data.

³As discussed in Cross (2019), the formal announcement of an explicit inflation target was made in 1996, however, reference to a target was made in official RBA speeches in August of 1992 and 1993. With this in mind, we start the forecast evaluation period in 1993Q3, however, the results are also robust to starting the forecast evaluation sample in 1992Q3.

of the inflation and the unemployment rate (PC group), and the second subgroup uses their first differences (PCd group). The following sections provide details of the model specifications in each group.

2.1 The Trend Model Group

2.1.1 Trend

The first model is a trend model with a Gaussian distributed error term (Trend) and constant variance, which is defined as:

$$y_t = \tau_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \sigma_y^2), \quad (1)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \quad (2)$$

where the error terms ε_t^y and ε_t^τ are respectively serially uncorrelated and assumed to follow independent and identically distributed (iid) Gaussian distributions. Equation (1) is known as the measurement equation, in which τ_t is the underlying trend. Equation (2) is the state equation that we specify as a random walk.

2.1.2 Trend-SV

The trend model can be extended by allowing the measurement equation to have stochastic volatility (Trend-SV):

$$y_t = \tau_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, e^{ht}), \quad (3)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \quad (4)$$

$$h_t = h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \quad (5)$$

where ε_t^y , ε_t^τ and ε_t^h are respectively serially uncorrelated iid Gaussian distributed error terms. The difference between the Trend model and the Trend-SV model is that the latter specifies stochastic volatility in the measurement equation (3). This has the effect of allowing the magnitude of the error variance to change over time. As in the case of the Trend model, the state equation for the Trend-SV model (4) is a random walk with a constant error term.

2.1.3 Trend-2SV

Following Stock and Watson (2007), the Trend-SV model can be further generalized to have stochastic volatility in the state equation (Trend-2SV):

$$y_t = \tau_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, e^{ht}), \quad (6)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim \mathcal{N}(0, e^{gt}), \quad (7)$$

$$h_t = h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \quad (8)$$

$$g_t = g_{t-1} + \varepsilon_t^g, \quad \varepsilon_t^g \sim \mathcal{N}(0, \sigma_g^2), \quad (9)$$

where the error terms $\varepsilon_t^y, \varepsilon_t^\tau, \varepsilon_t^h$ and ε_t^g are respectively serially uncorrelated iid Gaussian distributed error terms. The difference between the Trend-SV model and the Trend-2SV model is that the latter specifies stochastic volatility in the state equation (7). This has the effect of allowing the trend error variance to change over time.

2.1.4 Trend-SV-MA

Following Chan (2013), we also consider a version of the Trend-SV model in which the measurement equation has moving average (MA) errors (Trend-SV-MA):

$$y_t = \tau_t + \varepsilon_t^y, \quad (10)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \quad (11)$$

$$\varepsilon_t^y = u_t + \psi_1 u_{t-1} + \cdots + \psi_q u_{t-q}, \quad u_t \sim \mathcal{N}(0, e^{ht}), \quad (12)$$

$$h_t = h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \quad (13)$$

where ε_t^y has an MA process, and σ_τ^2 is the variance of the underlying trend of inflation. The inclusion of the MA term is designed to capture any serial dependence in the observations. Following Chan (2013) and Zhang et al. (2020), we set the order of moving average component to be one for simplicity.

2.1.5 Trend-2SV-MA

The last trend model allows for stochastic volatility in the state equation of the Trend-SV-MA model (Trend-2SV-MA):

$$y_t = \tau_t + \varepsilon_t^y, \quad (14)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim \mathcal{N}(0, e^{gt}), \quad (15)$$

$$\varepsilon_t^y = u_t + \psi_1 u_{t-1} + \cdots + \psi_q u_{t-q}, \quad u_t \sim \mathcal{N}(0, e^{ht}), \quad (16)$$

$$h_t = h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2). \quad (17)$$

where the error terms $\varepsilon_t^y, \varepsilon_t^\tau, \varepsilon_t^h$ and ε_t^g are serially uncorrelated and independent of each other, and ε_t^y has an MA process. In practice, the order of moving average component is again set to be one.

2.2 The AR Group

In addition to trend-models, we consider four AR models. This includes the standard AR with homoscedastic Gaussian distributed errors (AR), an AR with stochastic volatility (AR-SV), an AR with an MA process (AR-MA), and an AR with both stochastic volatility and an MA process (AR-SV-MA).

2.2.1 The Benchmark AR model

The AR model is defined as:

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2).$$

When conducting the forecasting exercise, the k -step ahead forecast inflation of the AR model can be calculated by

$$y_{t+k}^k - y_t = a_0^k + a^p(A)\Delta y_t + \varepsilon_t^k,$$

where $a^p(A)$ denotes the polynomials in lag operator A and ε_t^k is the k -step ahead forecast error. Following Stock and Watson (2007), we use the Akaike Information Criterion (AIC), and also the Hannan-Quinn information criterion (HQC) to determine the values of p for benchmark models in the forecasting exercises. Based on the results of AIC and HQC, the optimal lag lengths of AR model is one, and the results are reported in the first row of Table 1.

Table 1: Akaike Information Criterion (AIC) and Hannan-Quinn Information Criterion (HQC) for AR(p) Models and PC(p,q) and PCd(p,q) Models in 1993Q3-2019Q4.

	AIC	HQC
AR	1	1
PC	(4,4)	(4,4)
PCd	(3,4)	(3,4)

2.3 Phillips Curve Models

2.3.1 Phillips Curve I

Following Stock and Watson (1999, 2007) we use a non-accelerating inflation rate of the unemployment (NAIRU) Phillips curve (PC), which is defined by:

$$y_t = \sum_{i=1}^p b_{1i} \Delta y_{t-i} + \sum_{j=0}^q (b_{2j} (U_{t-j} - \bar{U}) + \varepsilon_t), \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

where U_t denotes the unemployment rate at date t and \bar{U} is the non-accelerating inflation rate of unemployment (NAIRU) which is assumed to be time invariant. The parameters p and q respectively denote the lag lengths of inflation and the unemployment rate in the PC. The k -step ahead forecast inflation of the PC is obtained as follows:

$$y_{t+k}^k - y_t = b_1^k(B)y_t + b_2^k(B)(U_t - \bar{U}) + \varepsilon_t^k, \quad (18)$$

where $b_1^k(B)$ and $b_2^k(B)$ denote the polynomials in lag operator B , respectively.

Let $b_0^k = -b_2^k(B)\bar{U}$; then, Equation (18) can be rewritten as Equation (19) with a constant term b_0^k :

$$y_{t+k}^k - y_t = b_0^k + b_1^k(B)\Delta y_t + b_2^k(B)U_t + \varepsilon_t^k. \quad (19)$$

This Phillips curve with an unemployment rate level is a conventional Phillips curve model. The transformed Equation (19) can describe the unemployment rate directly for the inflation forecast. It is denoted as PC(p, q), where p is the lag length of the first differential of inflation and q is the lag length of the unemployment rate. As with the AR model, the lag lengths p and q are determined by the AIC and HQC. The results are $p = 4$ and $q = 4$ and are reported in the second row of Table 1.

Along with the conventional PC model, we also consider PC models with SV, with MA, and with SV-MA (PC-SV, PC-MA, and PC-SV-MA) as competing models. The specifications of these models are the same as those in the trend model group. For simplicity, the lag lengths of inflation and the unemployment rate in these models are both four in the forecasting time period to consistent with PC.

2.3.2 Phillips Curve II

Phillips curve II is specified by an assumption that autoregressive distributed lags (ADL) exist in both the inflation and unemployment rates, as also explored by Stock and Watson (2007). This type of Phillips curve is represented as PCd(p, q). The expression of PCd model with Gaussian distributions in error terms is below:

$$y_t = \sum_{i=1}^p c_{1i}y_{t-i} + \sum_{j=1}^q c_{2j}\Delta U_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

where both the inflation and unemployment rates are assumed to be integrated of order one I(1) and the stationary predictors Δy_t and ΔU_t are included in the model. The k -step ahead forecast inflation of Phillips curve II is as follows:

$$y_{t+k}^k - y_t = c_1^k + c_2^k(B)\Delta y_t + c_3^k(B)\Delta U_t + \varepsilon_t^k,$$

Consistent with the AR and PC models, the lag lengths p and q of PCd are determined by the AIC and HQC in the third row of Table 1. The optimal lag lengths of PCd are $p = 3$ and $q = 4$ in 1993Q3-2019Q4.

As in the case of the PC models, we also consider PCd model variants with SV, with MA, and with SV-MA (PCd-SV, PCd-MA, and PCd-SV-MA) as competing models. The specifications of these models are the same as the corresponding trend models. For simplicity, the lag lengths of inflation and the unemployment rate in these models are consistent with PCd.

3 Estimation

3.1 Data

In the present paper, we use the quarterly consumer price index (CPI) and the unemployment rates (the percentage of the labor force 15 years and over) of Australia from 1978Q3 to 2019Q4 for parameter estimation, which are released by the Reserve Bank of Australia (RBA). Both of these macroeconomic variables are seasonally adjusted. The CPI inflation rate is calculated by the following formula:

$$y_t = 400 * \log(\text{CPI}_t / \text{CPI}_{t-1}).$$

The quarterly CPI inflation, unemployment rate, and their first differences are respectively plotted in Figures 1 and 2. Figure 1 shows that while inflation tends to be more volatile than the unemployment rate, both series have gradually decreased since the RBA adopted inflation targeting in 1993. This latter result has been attributed to good monetary policy by the RBA (Cross, 2019).

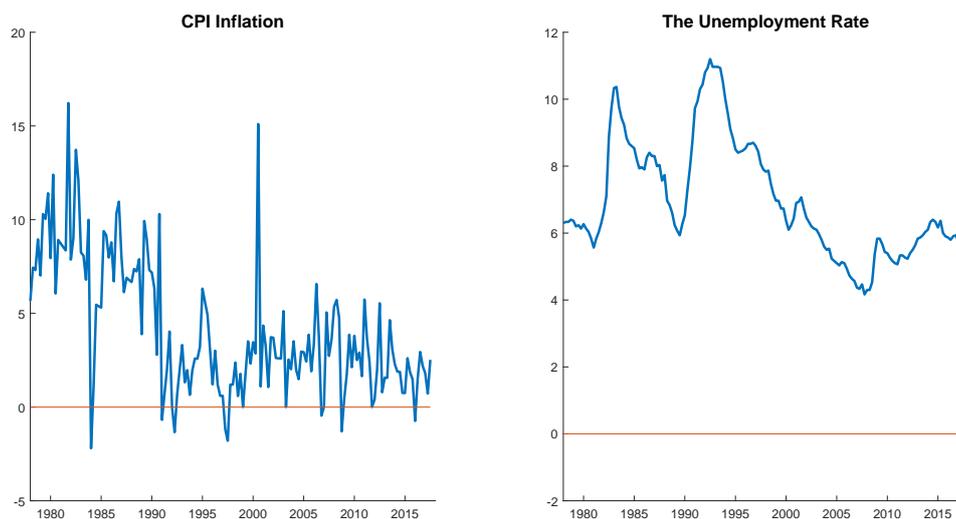


Figure 1: CPI inflation and the unemployment rates.

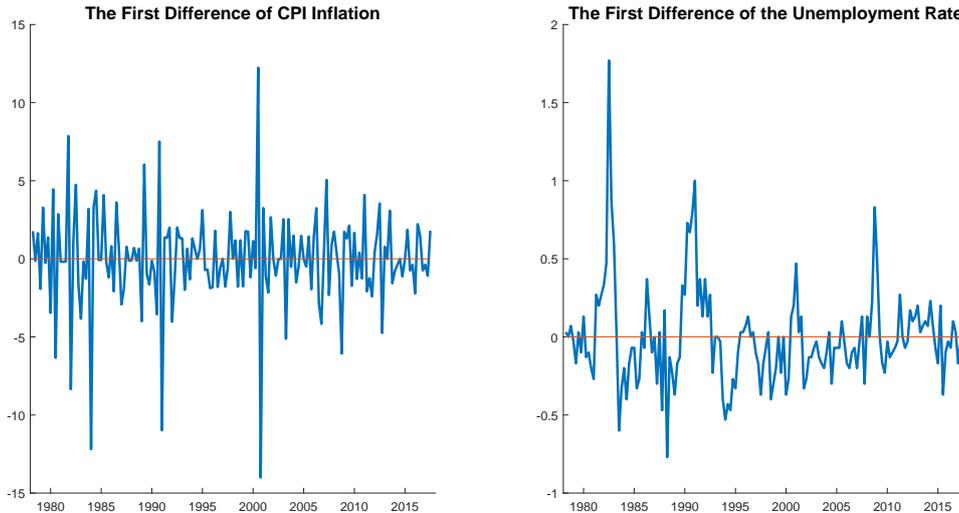


Figure 2: The first differences of CPI inflation and the unemployment rates.

3.2 Priors and Simulation Method

All models are estimated using a Bayesian paradigm with Markov chain Monte Carlo (MCMC) algorithms: Gibbs sampling and Metropolis-Hastings. Following Chan and Jeliazkov (2009), the latent states are efficiently sampled using a precision-based sampler as opposed to conventional Kalman or particle filters. The key idea is that the states can be efficiently drawn from the conditional posterior distribution by applying block-banded and sparse matrix algorithms on the precision matrix of that distribution. In this section, we provide details of parameter estimation setup for trend models.

3.2.1 Initial Values and Priors

Where appropriate, the initial values of τ , h and g in the trend models are assumed to follow Gaussian distributions:

$$\tau_1 \sim \mathcal{N}(\tau_0, \sigma_{0\tau}^2), \quad h_1 \sim \mathcal{N}(h_0, \sigma_{0h}^2), \quad g_1 \sim \mathcal{N}(g_0, \sigma_{0g}^2),$$

where $\tau_0 = h_0 = g_0 = 0$ and $\sigma_{0\tau}^2 = \sigma_{0h}^2 = \sigma_{0g}^2 = 5$ are set following (e.g., Chan, 2013; Stock and Watson, 2007; Zhang, Chan, and Cross, 2020). Therefore, the initial values of these two parameters are distributed with mean 0 and variance 5. Considering the properties of the growth rates for macroeconomic time series, the prior distribution is around 0 and large prior variances are used within $(-5, 5)$, thus, the initial values are reasonable and relatively non-informative.

The priors of σ_τ^2 , σ_h^2 , and σ_g^2 are assumed to be independent and follow inverse-gamma distributions:

$$\sigma_y^2 \sim \mathcal{IG}(\nu_y, S_y), \quad \sigma_\tau^2 \sim \mathcal{IG}(\nu_\tau, S_\tau), \quad \sigma_h^2 \sim \mathcal{IG}(\nu_h, S_h), \quad \sigma_g^2 \sim \mathcal{IG}(\nu_g, S_g).$$

Based on suggestions from previous studies on inflation by trend models (e.g., Chan, 2013; Stock and Watson, 2007), we set the hyperparameters to $\nu_y = \nu_\tau = \nu_h = \nu_g = 10$, $S_y = 9$, $S_\tau = 0.18$, and $S_h = S_g = 0.45$. These prior values imply relatively noninformative values for the shape parameters ν of the inverse-gamma distribution. The scale parameters mean that $\mathbb{E}\sigma_y^2 = 1$, $\mathbb{E}\sigma_\tau^2 = 0.141^2$ and $\mathbb{E}\sigma_h^2 = \mathbb{E}\sigma_g^2 = 0.224^2$; thus, the state transition is reasonably smooth and the results are comparable to those in the literature.

Finally, the moving average order in the trend models with MA variants is set to be one for simplicity (Chan, 2013). The MA coefficient is assumed to be a normal prior which is constrained within $(-1, 1)$ for the MA process invertibility:

$$\psi \sim \mathcal{N}(\psi_0, \sigma_{0\psi}^2),$$

where we set $\psi_0 = 0$ and $\sigma_{0\psi}^2 = 1$.

3.2.2 Posterior Simulation Method

The simulation of draws from the posterior distribution is conducted using MCMC methods, and the related Bayesian inference is similar to that of Chan (2013) and Zhang (2019). Specifically, the posteriors of the trend models are sampled cyclically in the following sequence:

For the Trend model:

1. $p(\boldsymbol{\tau} \mid \mathbf{y}, \sigma_y^2, \sigma_\tau^2)$,
2. $p(\sigma_y^2, \sigma_\tau^2 \mid \boldsymbol{\tau}) = p(\sigma_y^2 \mid \boldsymbol{\tau})p(\sigma_\tau^2 \mid \boldsymbol{\tau})$.

For the Trend-SV model:

1. $p(\boldsymbol{\tau} \mid \mathbf{y}, \mathbf{h}, \sigma_h^2, \sigma_\tau^2)$,
2. $p(\mathbf{h} \mid \mathbf{y}, \boldsymbol{\tau}, \sigma_h^2, \sigma_\tau^2)$,
3. $p(\sigma_h^2, \sigma_\tau^2 \mid \boldsymbol{\tau}, \mathbf{h}) = p(\sigma_h^2 \mid \mathbf{h})p(\sigma_\tau^2 \mid \boldsymbol{\tau})$.

For the Trend-2SV model:

1. $p(\boldsymbol{\tau} \mid \mathbf{y}, \mathbf{h}, \mathbf{g}, \sigma_h^2, \sigma_g^2)$,

2. $p(\mathbf{h}, \mathbf{g} | \mathbf{y}, \boldsymbol{\tau}, \sigma_h^2, \sigma_g^2) = p(\mathbf{h} | \mathbf{y}, \boldsymbol{\tau}, \sigma_h^2)p(\mathbf{g} | \mathbf{y}, \boldsymbol{\tau}, \sigma_g^2)$,
3. $p(\sigma_h^2, \sigma_g^2 | \mathbf{h}, \mathbf{g}) = p(\sigma_h^2 | \mathbf{h})p(\sigma_g^2 | \mathbf{g})$.

For the Trend-SV-MA model:

1. $p(\boldsymbol{\tau} | \mathbf{y}, \mathbf{h}, \psi, \sigma_\tau^2)$,
2. $p(\mathbf{h} | \mathbf{y}, \boldsymbol{\tau}, \psi, \sigma_h^2)$,
3. $p(\psi, \sigma_\tau^2, \sigma_h^2 | \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}) = p(\psi | \mathbf{y}, \boldsymbol{\tau}, \mathbf{h})p(\sigma_h^2 | \mathbf{h})p(\sigma_\tau^2 | \boldsymbol{\tau})$.

For the Trend-2SV-MA model:

1. $p(\boldsymbol{\tau} | \mathbf{y}, \mathbf{h}, \mathbf{g}, \psi, \sigma_h^2, \sigma_g^2)$,
2. $p(\mathbf{h}, \mathbf{g} | \mathbf{y}, \boldsymbol{\tau}, \psi, \sigma_h^2, \sigma_g^2) = p(\mathbf{h} | \mathbf{y}, \boldsymbol{\tau}, \psi, \sigma_h^2)p(\mathbf{g} | \mathbf{y}, \boldsymbol{\tau}, \psi, \sigma_g^2)$,
3. $p(\psi, \sigma_g^2, \sigma_h^2 | \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \mathbf{g}) = p(\psi | \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \mathbf{g})p(\sigma_h^2 | \mathbf{h})p(\sigma_g^2 | \mathbf{g})$.

3.3 Posterior Estimation

Before forecasting, we present estimates of the stochastic volatility parameters and moving average coefficient over the full sample: 1978Q2-2019Q4. All of the estimates are based on 50,000 draws from the posterior distribution after discarding the first 5,000 draws as a burn-in.

3.3.1 Posterior Estimates of SV Parameters

Figure 3 presents posterior means and credible intervals of the stochastic volatility parameters h and g obtained under Trend-SV and Trend-2SV, while Figure 4 presents a similar figure under Trend-SV-MA and Trend-2SV-MA. In both figures, we observe that the latter models produce smoother estimates of h , which allow for an extra time-varying variance parameter g to explain the volatility of inflation in Australia. Given that h and g are from exponential functions, which are monotonically increasing functions, the values of h and g change substantially over the sample period. This indicates that the SV models capture the presence of volatility clustering within the inflation series and that the inclusion of SV in the trend models may lead to improved forecast performance over a model variant with constant variance.

It is also interesting to note that the curves of h and g vary over time. During the early 1990s, the peak of h under Trend-2SV-MA is lower than that under Trend-SV-MA, while the value of g channeling the volatility of underlying trend from the state equation of

Trend-2SV-MA also reaches a peak in that time period. However, h begins to increase from 2005 under both models, while g is stable around those years. As discussed in Cross (2019), the increase in volatility in the early 2000s is likely due to the commencement of Australia’s resource boom, while the reduction in volatility in the 2010s follows the decline of the boom.⁴

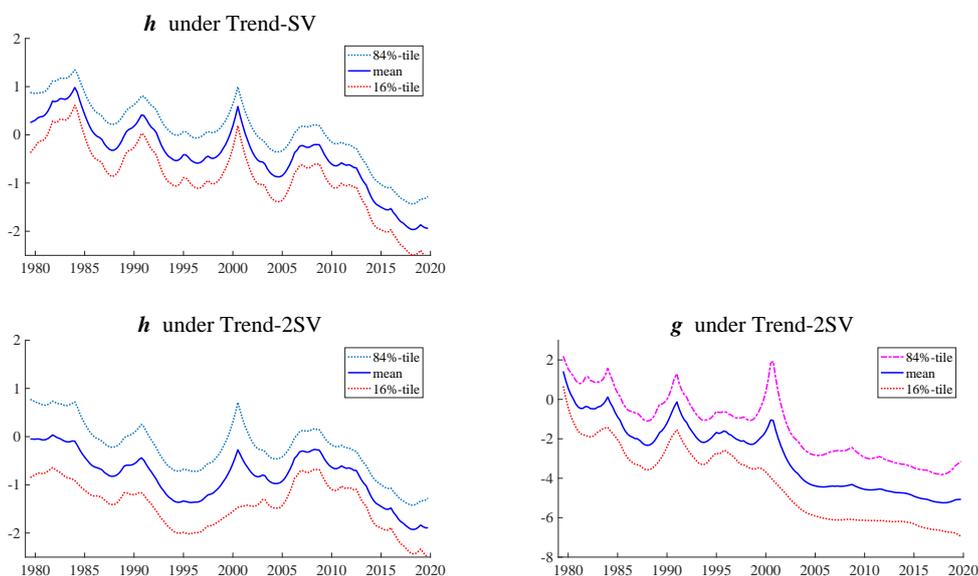


Figure 3: Posterior estimates and quantiles for SV parameter h and g under Trend-SV and Trend-2SV.

⁴For a discussion of the resource boom and its effects on the Australian economy see Sheehan and Gregory (2013).

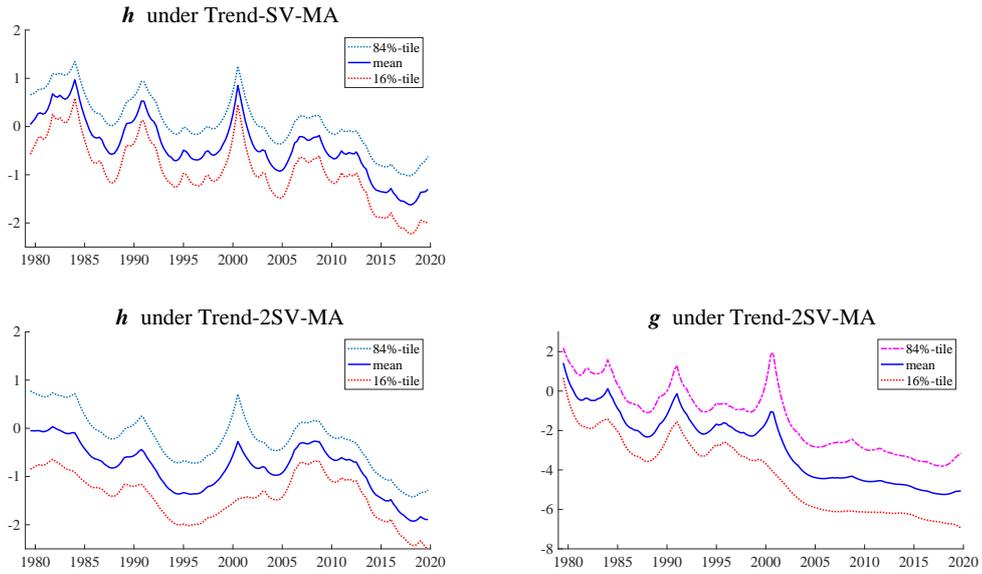


Figure 4: Posterior estimates and quantiles of SV parameter h and g under Trend-SV-MA and Trend-2SV-MA.

3.3.2 Posterior distributions of MA coefficients

The marginal density estimates for the MA coefficient $p(\psi|y)$ under Trend-SV-MA and Trend-2SV-MA are shown in Figure 5. Both values of ψ are concentrated around 0.2 and the mass is away from 0, which suggests that the MA specification is unlikely to be 0 in these two trend models. Specifically, under Trend-SV-MA, ψ has a higher probability and more concentrated around its mean than that under Trend-2SV-MA. It indicates that moving average components share more weight in channeling the volatility of the inflation when there is no stochastic volatility parameter in the state equation.

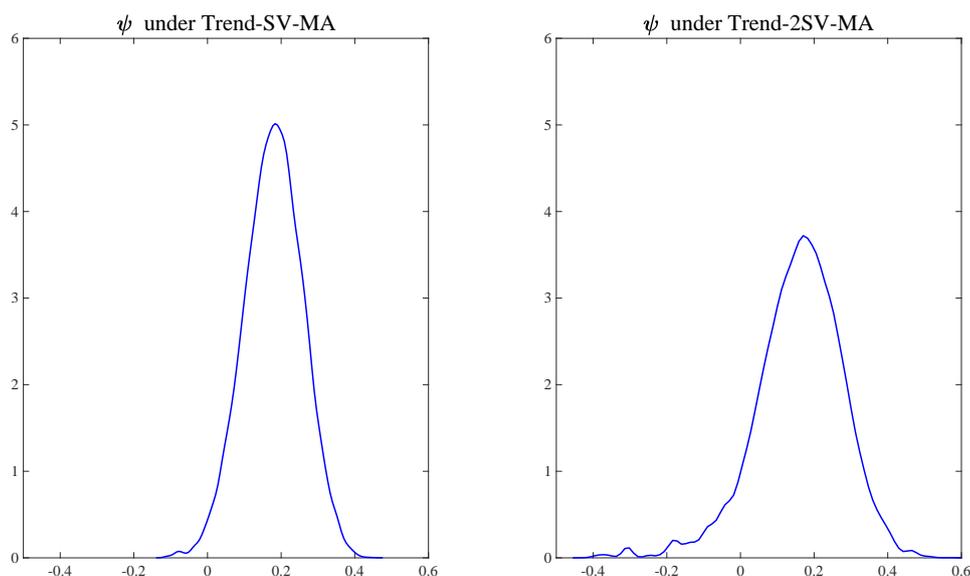


Figure 5: Posterior distribution of MA coefficients under Trend-SV-MA and Trend-2SV-MA.

4 Forecast Results

In this section, we first discuss the forecast method used in the main analysis, including forecast horizons, the evaluation period, and accuracy metrics. We then present the results from each of the individual models, followed by a combination forecast and a robustness check with a rolling window method in place of the expanding window.

4.1 Forecast Method

To conduct the forecasts, we use a pseudo-out-of-sample forecast method and conduct the forecasts from 1993Q3-2019Q4. Let $T_0 = 2004Q4$, then the forecasts are first calculated at time T_0+1 and compared with the actual data at time T_0+1 . We then step to T_0+2 , T_0+3 , and so on, until we reach the end of the sample. The forecast horizons used are one-quarter ahead, one-year ahead, two-year ahead, and three-year ahead forecasts, denoted $k = 1, 4, 8$ and 12 . Set in this manner we are able to compare short, medium, and long-run forecasts from each of the competing models. Since the PC group does not have an iterative formula for the unemployment rate, we use a direct forecasting methodology for all models.

The forecast performance is evaluated in terms of two metrics. Point forecasts are evaluated using the relative mean square forecast error (MSFE) and density forecasts are

evaluated using the relative average log predictive likelihood (ALPL).

When calculating the MSFE of a model, the forecast $\hat{y}_{T_0+t+k-1}$ is evaluated by averaging all the posterior means $\mathbb{E}(y_{T_0+t+k-1} | \mathbf{y}_{1:T_0+t})$ at $T_0 + t$. Then, the forecasting error is calculated by $\mathbf{e}_{T_0+t+k-1}^2 = \mathbf{y}_{T_0+t+k-1}^0 - \mathbb{E}(y_{T_0+t+k-1} | \mathbf{y}_{1:T_0+t})$. The MSFE is calculated as:

$$\text{MSFE} = \frac{1}{T - T_0 - k + 1} \sum_{t=1}^{T-T_0-k+1} \mathbf{e}_{T_0+t+k-1}^2.$$

The relative MSFE reports the ratio of the MSFE between a candidate model and the benchmark. Set in this manner, a value that is smaller than one indicates that the related candidate model has better forecasting performance than the benchmark, while a value that is larger than one means that the competing model has worse forecasting performance than the benchmark.

The predictive likelihood $p(\hat{y}_{T_0+t+k-1} = \mathbf{y}_{T_0+t+k-1} | \mathbf{y}_{1:T_0+t})$ is used to evaluate the density forecast performance and is given by:

$$\text{ALPL} = \frac{1}{T - T_0 - k + 1} \sum_{t=1}^{T-T_0-k+1} \log p(\hat{y}_{T_0+t+k-1} = y_{T_0+t+k-1} | \mathbf{y}_{1:T_0+t}).$$

When the observed data with falls into a higher probability density region of the posterior predictive distribution, the estimated parameters, which are conditional on the observed data $\mathbf{y}_{T_0+t+k-1}$, will produce a larger predictive likelihood value. The relative ALPL reports the difference of the ALPL between a candidate model and the benchmark. This means that a positive relative ALPL value suggests that the candidate model has better performance than the benchmark, while a negative ALPL value indicates that the benchmark forecasts better.

For forecast accuracy comparison, we use a one-sided sign test of equal predictive accuracy of Diebold and Mariano (1995). When the competing models are all nested, test statistics introduced by Clark and McCracken (2001) can be used and interpreted in a Bayesian manner.⁵ Since models are allocated into four groups and are not all nested, we do not report test results of nested models. The rejection of equal forecast accuracy relative to the benchmark at credible level 0.05 and 0.10 are denoted by one and two asterisks, respectively, and reported in the result tables below.

Considering that the inflation targeting framework was implemented by the RBA in early 1990s (Macfarlane, 1999), Stevens (1999) suggests that mid-1993 is the time at which the medium-term inflation target was explicitly articulated by the RBA. This time point is widely accepted as the beginning of conducting a new monetary policy, so we report the forecasting results from sample 1993Q3-2019Q4. The forecasts are estimated recursively by expanding the window for parameter estimation. The first 6 points are separated for lags and the following 40-time points are used as the initial data for the parameter estimation. Thus, the forecasting period is 2005Q1 to 2019Q4.

⁵In the Bayesian interpretation of probability, the word *credibility* is used in place of the classical notion of *statistical significance*. For a textbook reference, see Koop (2003).

4.2 Recursive Forecast Results

We present the results across three sub-sections. First, we present the average results for the full evaluation period. Second, we look at how the point forecast performance of each model has evolved. Finally, we look at how the density forecast performance of each model has evolved over time.

4.2.1 Average Forecast Results

The forecast results for the time period 1993Q3-2019Q4, thereby reflecting the period in which Australia had an inflation target, are reported in Table 2. In the table, the relative MSFE results are reported in the left column and the relative ALPL results are in the right column. Both metrics use the AR model as a benchmark. In terms of point forecasts, we observe that the trend model group generally outperforms the AR and PC groups at medium to long-run horizons. The density forecast results are less definitive, however, the Trend-SV, Trend-2SV, and Trend-2SV-MA also have superior forecast performance than the benchmark on short to medium runs.

For the trend group, Trend-SV and Trend-SV-MA with SV in the measurement equation outperform other trend models, suggesting the underlying trend inflation captured by Trend-SV and Trend-SV-MA accurately depicts the persistence of inflation than other trend models. For AR, PC, and PCd groups, models with SV or SV-MA also forecast better than models with Gaussian distribution or just MA specification in most cases. Taken together, these forecast results show that models with SV or SV-MA provide better point and density forecasts than their counterparts for inflation in the inflation targeting period.

When comparing models in the AR and PC groups to the benchmark, we find that only the AR-SV-MA can outperform the benchmark on a one-quarter ahead point forecast. Moreover, None of the relative MSFE and the relative ALPL of competing models in these two groups is better than those of the benchmark on three-year ahead forecasts.

Table 2: Recursive forecast results for all groups, 1993Q3–2019Q4

	Relative MSFE				Relative ALPL			
	$k=1$	$k=4$	$k=8$	$k=12$	$k=1$	$k=4$	$k=8$	$k=12$
AR	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
AR-SV	1.00	1.00*	1.00	1.00	0.05	0.03	-0.01	-0.04
AR-MA	1.01	1.03**	1.05	1.04	-0.03	-0.05**	-0.06**	-0.06*
AR-SV-MA	0.99	1.00**	1.00	1.00	0.05	0.04	0.00	-0.03
Trend	1.11	0.95	0.92	0.97	-0.15**	-0.14**	-0.18**	-0.23**
Trend-SV	1.11	0.94	0.92	0.98	0.03	0.07	0.03	-0.03
Trend-2SV	1.17	0.99	0.92	0.99	0.02*	0.05	-0.02	-0.11
Trend-SV-MA	1.11	0.94	0.93	0.97	0.01	0.04	0.02	-0.02
Trend-2SV-MA	1.10	0.98	0.91	0.97	-0.06**	-0.07**	-0.12**	-0.18**
PC	1.43**	1.16	1.53**	1.28	-0.12**	-0.04	-0.11**	-0.08
PC-SV	1.22*	1.10	1.38	1.08	0.01*	0.08	-0.03	-0.01
PC-MA	1.42**	1.17	1.47*	1.25	-0.13**	-0.09**	-0.15**	-0.14**
PC-SV-MA	1.22*	1.11	1.39	1.07	0.01*	0.07	-0.03	-0.01
PCd	1.42**	1.11	1.36	1.10	-0.12**	-0.04	-0.08	-0.04
PCd-SV	1.31**	1.22	1.41	1.08	-0.03**	0.04	-0.07	-0.02
PCd-MA	1.42**	1.10	1.32	1.06	-0.13**	-0.09*	-0.12**	-0.11**
PCd-SV-MA	1.30**	1.19	1.37	1.05	-0.02**	0.05	0.06	-0.01

Notes: Bold entries are the smallest relative MSFE or the largest relative ALPL for the corresponding horizons. ** and * indicates rejection of equal forecast accuracy relative to AR model at significance level 0.05 and 0.10, respectively, when using an asymptotic test in Diebold and Mariano (1995).

4.2.2 Point Forecast Results Over Time

To give a closer examination of the point performance of all the competing models, in Figure 6 we present recursively computed MSFEs of the one-year ahead point forecasts across for the remaining models from 2006Q1-2019Q4. We see that the MSFEs of all the models increased from 2006 to 2009 due to the Global Financial crisis and then gradually decreased to the end of the sample. This indicates that all of the models capture the flexible specification of the inflation target in Australia during the first few years, but fail to predict that the inflation decreased substantially with the onset of the Global Financial Crisis on long-run forecasts. Once the Australian economy recovered gradually, the MSFEs decline smoothly. A possible mechanism is that when the demand pressures from the real economy were eased, inflation declined, and the MSFEs raised to a platform with a sustained delayed response from model-based forecasts.

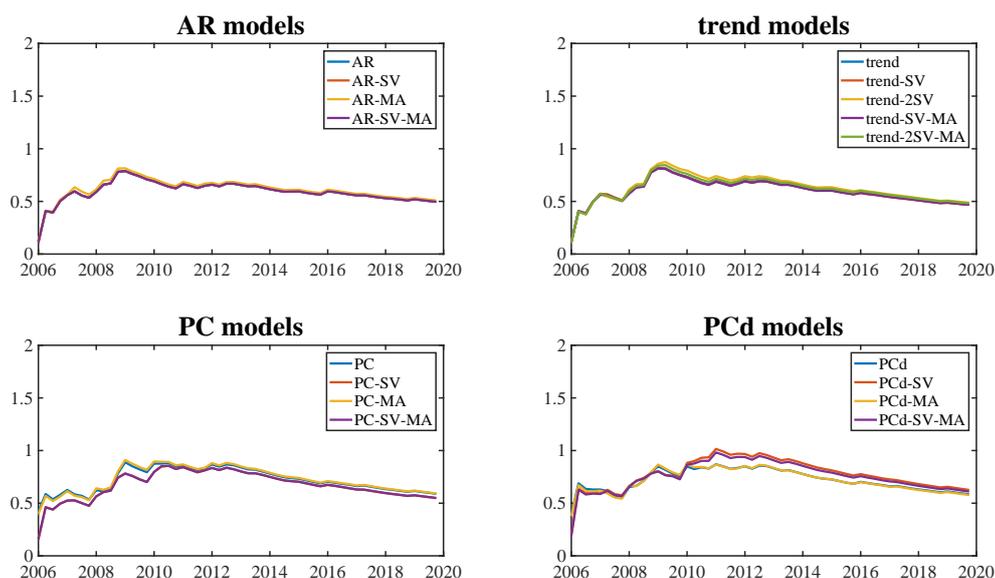


Figure 6: MSFEs of one-year ahead forecast under all the competing models on a recursive basis.

While trend models struggle to outperform AR models at the one-step-ahead horizon, they generally provide superior forecast performance at medium to long runs horizons. To investigate why this is the case, we plot the actual inflation time series and one-year ahead point forecasts in Figure 7. The competing models are allocated by groups to facilitate a model comparison. The main insight is that the forecasts from models in the trend group adjust rapidly to abrupt jumps in the actual inflation series. It is also important to note that forecasts from models with SV and SV-MA specifications (especially Trend-SV and Trend-SV-MA) are relatively close to the actual value. This ability to adapt to abrupt changes potentially explains why Trend-SV and Trend-SV-MA can outperform other models in the medium to long runs forecast exercise.

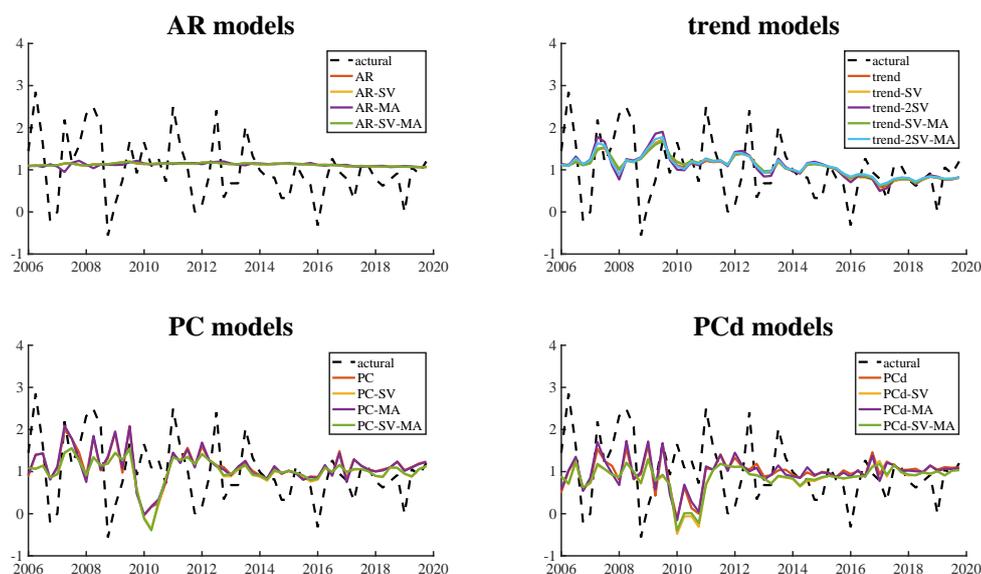


Figure 7: Actual inflation time series and one-year ahead point forecast under all the competing models.

Specifically, the Reserve Bank of Australia has undershot the inflation target and the inflation rate is under its lower bound of 2 per cent since 2014. The results in Figure 7 show that models with an underlying trend specification capture this feature and have smoother forecasts for inflation. The MSFEs of all the models have similar performance and all the values fall back below 1 by the end of the sample.

4.2.3 Density Forecast Results Over Time

To give a closer examination of the density performance of all the competing models, in Figure 8 we present recursively computed ALPLs of the one-year ahead density forecasts across the remaining models from 2006Q1-2019Q4. We find that when actual inflation either drops or rises dramatically, e.g. Global Financial Crisis in 2008, none of the competing models can have the consistent forecasting performance as before, coming with large absolute log-likelihood values. However, it is worth noting that trend models with flexible error specifications have less volatility and are generally smoother.

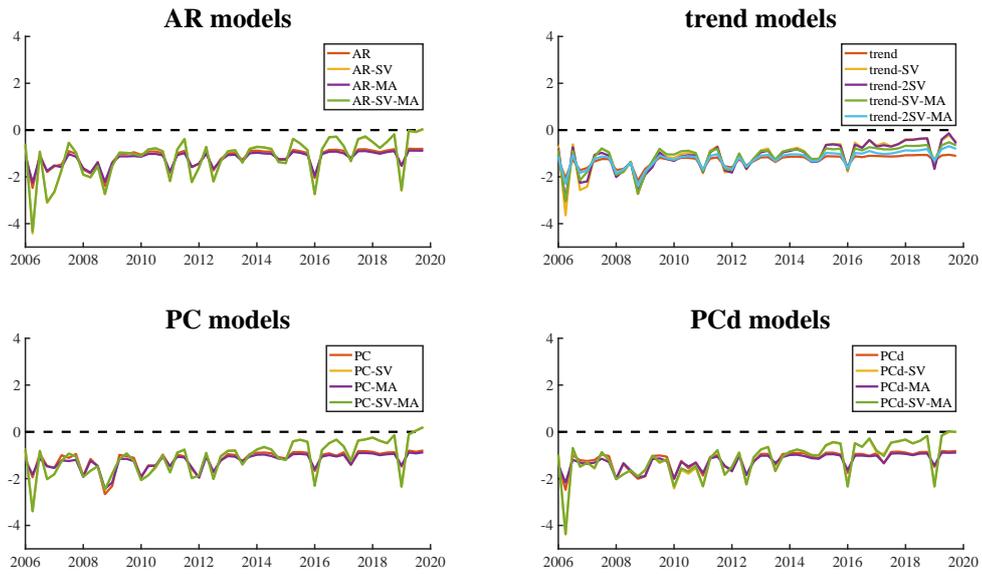


Figure 8: Log likelihood of one-year ahead forecasts under all the competing models on a recursive basis.

The primary reason for considering density forecasts in addition to point forecasts is that they can provide a summary of the forecast uncertainty in the posterior predictive density. They can therefore be used to examine tail risks in macroeconomic outcomes (Carriero et al., 2020). To have a closer look at the conditional density forecasts from each model at a certain time point, in Figure 9 we plot the conditional predictive distributions for each model in 2016Q2. As mentioned before, inflation fell steadily and maintained a low level since 2014, and this quarter is typical of this period. The figure highlights the fact that trend models tend to have heavier tails than the AR and PC models. This allows them to place a much higher probability weight on inflation values that fall within the tails of the distribution. The fact that each model has the same error structure highlights the benefits of specifying a flexible trend component when forecasting Australian inflation.

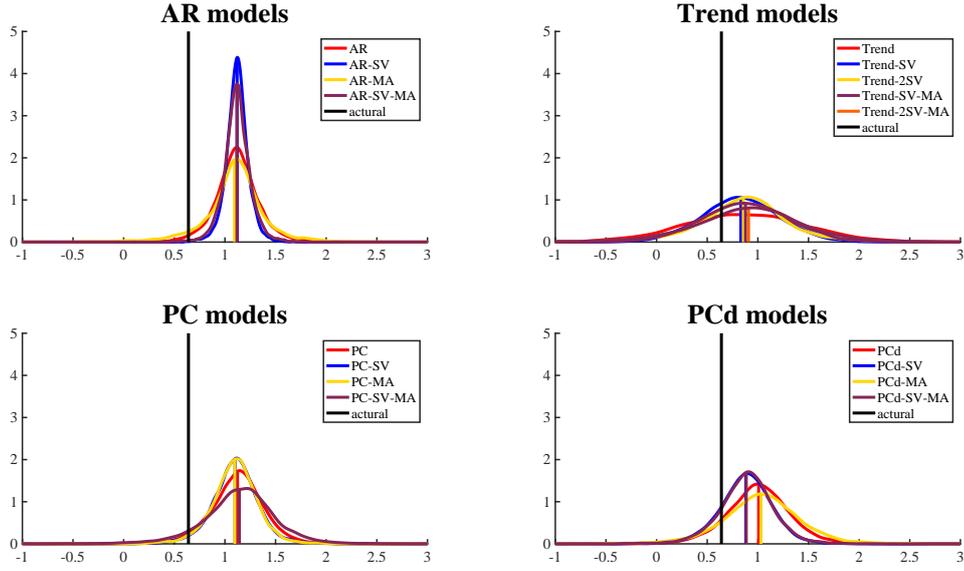


Figure 9: Conditional density forecasts in 2016Q2 on one-year ahead forecast horizon under all the competing models.

4.3 Combination Forecast

In this section, we conduct a forecasting analysis by combining the forecasting results in each group. Specifically, the forecasting results are combined by both equal weight (EW) and time-varying weight (TVW) methods for the entire sample period. Following (Zhang, 2019), the window width is set to be twenty quarters in the time-varying weight combination forecast. The results for both equal weight and time-varying weight forecasting results for AR group, trend group, PC, and PCd groups, respectively, are shown in Table 3.

Overall, the best point combination performance is given by the AR group at short-run forecasts and the trend group on medium to long runs. The AR group also has superior density forecasts at one quarter ahead, one year ahead, and three-year ahead forecast horizons. It is also notable that the forecast performance of PC and PCd groups greatly improves on the individual models, with both their point and density forecast results being closer to those of the benchmark. Finally, it is worth noting that the relatively close performance between the equal weight and time-varying weight forecasts is also consistent with the similar forecast results on US inflation data in Zhang (2019).

Table 3: Combination forecast results for all groups

	Relative MSFE				Relative ALPL			
	$k=1$	$k=4$	$k=8$	$k=12$	$k=1$	$k=4$	$k=8$	$k=12$
AR	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
AR-EW	1.00	1.00	1.18	0.98*	0.03	0.04	-0.01	0.02
AR-TVW	1.00	1.00	1.18	0.98*	0.03	0.04	0.00	0.03
Trend-EW	1.08	0.91	0.97	0.85	-0.01	-0.01	-0.07	-0.10
Trend-TVW	1.08	0.91	0.97	0.85	-0.01	0.01	-0.07	-0.09
PC-EW	1.19	1.28	1.55	1.07	-0.02	-0.01	-0.05	-0.01
PC-TVW	1.20	1.29	1.55	1.06	-0.01	0.00	-0.04	0.00
PCd-EW	1.29**	1.36	1.33	0.97	-0.05**	-0.03	-0.02	0.00
PCd-TVW	1.29**	1.36	1.32	0.96	-0.04**	-0.02	-0.02	0.00

Notes: See the notes to Table 2.

4.4 Rolling Window Forecast

As a sensitivity analysis, we also conducted forecasts using a rolling window approach to examine whether or not the trend model group still maintains better forecasting performance on medium to long runs. Unlike the expanding window approach used in our main analysis, this method fixes the parameter estimation period to a certain number of time period, in our case 40-quarters (i.e. 10 years). The relative MSFE and ALPL metrics are reported in Tables 4 for the forecasting sample period 1993Q3-2019Q4.

The results confirm our main finding that trend models generally provide superior point forecasts at medium and long-run horizons. Interestingly, in contrast to our main results, the PC and PCd groups now outperform the AR models and trend models on the one-year ahead point forecast. However, the benchmark is hard to beat in terms of density forecast. In summary, the results of the rolling window forecast method indicate that the trend model group is less influenced by the change of information available for estimation. Generally, trend models can provide better point forecast performance than other groups. Models with SV and SV-MA specifications have better forecasting performance than those without them, which are the same as the findings in the recursive forecast.

Table 4: Rolling window forecast results for all groups, 1993Q3–2019Q4

	Relative MSFE				Relative ALPL			
	$k=1$	$k=4$	$k=8$	$k=12$	$k=1$	$k=4$	$k=8$	$k=12$
AR	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
AR-SV	1.00	1.00	1.00**	1.00	0.00	-0.04	-0.07	-0.10
AR-MA	1.01	0.99	1.02	1.00	-0.03	-0.04	-0.05*	-0.06
AR-SV-MA	1.05*	1.00	1.00*	1.00	-0.01	-0.04	-0.08	-0.11
Trend	1.05	0.95	0.88**	0.91*	-0.24**	-0.25**	-0.26**	-0.30**
Trend-SV	1.06	0.94	0.87**	0.92	-0.03	0.00	-0.01	-0.07
Trend-2SV	1.13	0.99	0.92	0.89	-0.07*	-0.09	-0.15**	-0.24**
Trend-SV-MA	1.13	0.96	0.90*	0.92	-0.09*	-0.06	-0.04*	-0.08**
Trend-2SV-MA	1.16	0.98	0.90	0.90	-0.18**	-0.22**	-0.25**	-0.32**
PC	1.25	0.95	1.19	1.28	-0.06	0.00	-0.08**	-0.13
PC-SV	1.10	0.94	0.94*	0.98	-0.02	0.01	-0.01	-0.08
PC-MA	1.25	0.95	1.13	1.25	-0.11	-0.10	-0.17**	-0.22**
PC-SV-MA	1.11	0.93	0.95	0.98	-0.03	0.01	-0.01	-0.07
PCd	1.33	0.94	1.23*	1.22**	-0.11*	-0.02	-0.10*	-0.11**
PCd-SV	1.17*	0.93	0.99	1.08	-0.08*	0.00	-0.08	-0.11
PCd-MA	1.31*	0.90	1.19	1.22*	-0.13*	-0.09	-0.17**	-0.19**
PCd-SV-MA	1.15*	0.92	0.99	1.06	-0.07*	0.00	-0.07	-0.10

Notes: See the notes to Table 2.

5 Concluding Remarks

In this paper, we have compared the forecast accuracy of trend models against commonly used autoregressive and Phillips curve models when forecasting the CPI inflation rate in Australia. Overall, the results showed that while autoregressive models are tough to beat at the one-step-ahead forecast horizon, trend models generally provide superior point and density forecasts at medium to long-run forecast horizons. The best medium-term point forecasts come from a trend model with stochastic volatility in the transitory component and that with a moving average component, while long-run point forecasts are better made by trend models with stochastic volatilities and a moving average component. These improvements were found to be robust to both expanding and rolling window forecast methodologies.

In an in-sample analysis, we also found that trend models can capture various dynamics in periods of significance which the AR and PC models cannot. This includes the dramatic reduction in inflation when the RBA adopted inflation targeting, a one-off 10 per cent Goods and Services Tax inflationary episode in 2000, and then gradually decline in inflation since 2014. Taken together, our results suggest that policymakers would benefit from adopting trend models when forecasting inflation in Australia.

References

- M. Beechey and P. Österholm. Forecasting inflation in an inflation-targeting regime: A role for informative steady-state priors. *International Journal of Forecasting*, 26(2): 248–264, 2010.
- F. Brayton, J. M. Roberts, and J. C. Williams. What happened to the Phillips curve? *Finance and Economics Discussion Series*, 1999-49, 1999.
- A. Carriero, T.E. Clark, and M. Marcellino. Capturing macroeconomic tail risks with Bayesian vector autoregressions. *Federal Reserve Bank of Cleveland working paper*, 20-02, 2020.
- J.C.C. Chan. Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2):162–172, 2013.
- J.C.C. Chan and I. Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1/2):101–120, 2009.
- T.E. Clark and M.W. McCracken. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105(1):85–110, 2001.
- J. Cross and A. Poon. Forecasting structural change and fat-tailed events in Australian macroeconomic variables. *Economic Modelling*, 58:34–51, 2016.
- J.L. Cross. On the reduced macroeconomic volatility of the Australian economy: Good policy or good luck? *Economic Modelling*, 77:174–186, 2019.
- F.X. Diebold and R.S. Mariano. Comparing predictive accuracy. *Journal of Business and Econometric Statistics*, 13(3):134–144, 1995.
- R. Dixon and G.C. Lim. Underlying inflation in Australia: Are the existing measures satisfactory. *Economic Record*, 80(251):373–386, 2004.
- C. Garnier, E. Mertens, and E. Nelson. Trend inflation in advanced economies. *International Journal of Central Banking*, 11(4):5–136, 2015.
- A. Garratt, J. Mitchell, S. Vahey, and E. Wakerly. Real-time inflation forecast densities from ensemble Phillips curves. *The North American Journal of Economics and Finance*, 22(1):77–87, 2011.
- D. Gruen, T. Robinson, and A. Stone. Output gaps in real time: How reliable are they? *Economic Record*, 81(252):6–18, 2005.
- G.M. Koop. *Bayesian econometrics*. John Wiley & Sons Inc., 2003.
- I.J. Macfarlane. Australian monetary policy in the last quarter of the twentieth century. *Economic Record*, 75(3):213–224, 1999.

- T. Robinson, A. Stone, and M. van Zyl. Results—RDP 2003-12: The real-time forecasting performance of Phillips curves. *Reserve Bank of Australia Research Discussion Papers*, (December), 2003.
- P. Sheehan and R.G. Gregory. The resources boom and economic policy in the long run. *Australian Economic Review*, 46(2):121–139, 2013.
- D. Staiger, J. Stock, and M. Watson. The NAIRU, unemployment and monetary policy. *Journal of Economic Perspectives*, 11(1):33–49, 1997.
- G. Stevens. Six years of inflation targeting. *Reserve Bank of Australia Bulletin*, 5:46–61, 1999.
- J.H. Stock and M.W. Watson. Forecasting inflation. *Journal of Monetary Economics*, 44(2):293–335, 1999.
- J.H. Stock and M.W. Watson. Why has U. S. inflation become harder to forecast? *Journal of Money Credit and Banking*, 39(1):3–33, 2007.
- B. Zhang. Real-time inflation forecast combination for time-varying coefficient models. *Journal of Forecasting*, 38(3):175–191, 2019.
- B. Zhang and B. Nguyen. Real-time forecasting of the Australian macroeconomy using flexible Bayesian VARs. *CAMA working paper*, 91/2020, 2020.
- B. Zhang, J.C.C. Chan, and J.L. Cross. Stochastic volatility models with ARMA innovations: An application to G7 inflation forecasts. *International Journal of Forecasting*, 36(4):1318–1328, 2020.