Forecasting inflation in Mongolia: A dynamic model averaging approach

Doojav, Gan-Ochir and Luvsannyam, Davaajargal

2017

Forecasting inflation in Mongolia: A dynamic model averaging approach

Gan-Ochir Doojav[1] Davaajargal Luvsannyam[2]

Abstract

This paper investigates the use of DMA approach for identifying good inflation predictors and forecasting inflation in Mongolia, one of the most commodity dependent economies, using dynamic model averaging (DMA). The DMA approach allows for both set of predictors for inflation and marginal effects of predictors to change over time. Our empirical work resulted in several novel in findings. First, external variables (i.e., China's growth, China's inflation, change in oil price) play important role in forecasting inflation and change considerably over time and over forecast horizons. Second, among domestic variables, wage inflation and M2 growth are currently the best predictors for short and longer forecast horizons. Third, the use of DMA lead to substantial improvements in forecast performance, and DMA (2,15) with the chosen forgetting factors is the best performer in predicting inflation for Mongolia.

Keywords: Inflation, Dynamic Model Averaging, Time-Varying Parameter, Forecasting.

JEL classifications: C11, C22, C53, E31, E37.

[1] Gan-Ochir Doojav, corresponding author, Chief Economist, Bank of Mongolia, Baga toiruu-3, 15160, Ulaanbaatar-46, Mongolia (Email: doojav_ganochir@mongolbank.mn).

[2] Davaajargal Luvsannyam, Director, Research Division, Research and Statistics Department, Bank of Mongolia, Baga toiruu-3, 15160, Ulaanbaatar-46, Mongolia (Email: davaajargal@mongolbank.mn).

1. Introduction

Forecasting inflation is one of the important, but difficult tasks in conducting monetary policy. Though there are many theoretical approaches to forecast inflation, the most popular are those based on extensions of the Phillips Curve. The influential papers relied on the generalized Phillips curve include Stock and Watson (1999, 2007), Ang et al. (2007), Atkenson and Ohanian (2001). A new front of the literature (i.e., Koop and Korobilis 2012, Ferreira and Palma 2015, Styrin 2019) have focused on dynamic specification of inflation with many predictors where a set of predictors and parameters can potentially change over time. Other front of the literature suggests that commodity-dependent economies are more vulnerable to external shocks such as changes in commodity prices, commodity demands, foreign direct investment (FDI) and foreign demand. The recent studies (i.e., Liu et al. 2014 for UK, Bjørnland and Thorsrud 2016 for Australia and Norway, Allegret and Benkhodja 2015 for Algeria, Hou et al. 2016 for Canada, Bergholt et al. 2017 for Norway, Gan-Ochir and Davaajargal 2019 for Mongolia) have shown the importance of external shocks on business cycles. Therefore, for the economies, understanding interactions between external factors and inflation dynamics is potentially important for improving the forecasting accuracy. Moreover, an issue that whether influences on domestic inflation of external factors such as commodity prices, trading partner's growth and inflation change over time has addressed in the literature.

This paper examines how predictors of inflation change over time and forecasts inflation in Mongolia, one of the most commodity dependent economies. Particularly, we are interested in whether the dynamic specification of inflation leads to forecasting improvements or not. To end this, we estimate the generalized Phillips curve, in which domestic and external factors are chosen as potential predictors, and forecast inflation using a dynamic model averaging (DMA) approach developed by Raftery et al. (2010). The approach is ideally suitable for the problem of inflation forecasting since it allows for the set of predictors (forecasting model) to change over time while, at the same time, allowing for coefficients in each model to evolve over time. Therefore, the approach simultaneously deals with the issues raised by recursive, regression-based methods on generalized Phillips curve models, such as (i) the coefficients on the predictors can change over time, (ii) the number of predictors can be large[3], which can be handled by doing Bayesian model averaging (BMA) or automating the model selection process, and (iii) the model (the set of predictors) relevant for forecasting can potentially change over time[4]. The empirical analysis is applied to Mongolia, facing challenges to strengthen its resilience to cope with negative external shocks and to transform its natural resource wealth into assets that support sustainable growth and prosperity. Hence, evidence from the case of Mongolia would be of high

[3] If the set of models defined by whether each of m potential predictors is included or nor, then there are 2^m models. When facing 2^m models, the computational demands are daunting. For Groen et al. (2010) who used ten predictors, they ended up with 1024 models. This creates the statistical problem of how to select the most efficient model.

[4] If a researcher has 2^m models, and at each point in time, a different forecasting model may apply, then the number of combinations of models that must be estimated in order to to forecast at time τ is $2^{m\tau}$. It can be computationally infeasible to forecast by simply going through all $2^{m\tau}$ combinations.

relevance improving forecasting accuracy of inflation for resource-rich developing countries. This paper contributes the literature in two ways. First, use of the DMA approach to examine the dynamic specification of inflation in a commodity-dependent economy is in its novelty in the literature. Second, since the empirical analysis includes a large set of internal and external predictors, it provides a comprehensive analysis examining effects of external factors on inflation.

Much methodological and empirical works have been done to improve inflation forecasting. By raising potential problems with traditional practices choosing a single model and presenting results based on this model, Koop and Potter (2004) propose methods for implementing BMA with factor models. Using quarterly US data on 162 time series, they show that models which contain factors do out-forecast an AR(p), but only by a relatively small amount and only at short horizon for both GDP and inflation. Using BMA for pseudo out-of-sample prediction of US inflation, Wright (2009) find that it generally gives more accurate forecasts than simple equal-weighted averaging. Maas (2014) presents a time-varying BMA for forecasting US inflation, in which the set of predictors included in the model is automatically selected from a large pool of potential predictors and the set of predictors is allowed to change over time. He finds that it does produce superior density forecasts compared with a range of alternative forecasting models. Koop and Korobilis (2012) employ DMA and dynamic model selection (DSA) where a single (potentially different) model can be used as the forecasting model at each point in time for forecasting US inflation. They find that DMA and DMS leads to substantial forecasting improvements over benchmark regressions and time-varying coefficient models. Ferreira and Palma (2015) reveal that DMA approach deliver good medium-term inflation forecasting for Brazil. Wei and Cao (2017) find that DMA generally outperforms other models including Bayesian model averaging (BMA) in both recursive and rolling forecasting models of housing prices for major Chinese cities. However, Styring (2019) did not find evidence that DMA outperforms simpler benchmark models in the case of Russian inflation forecasting.

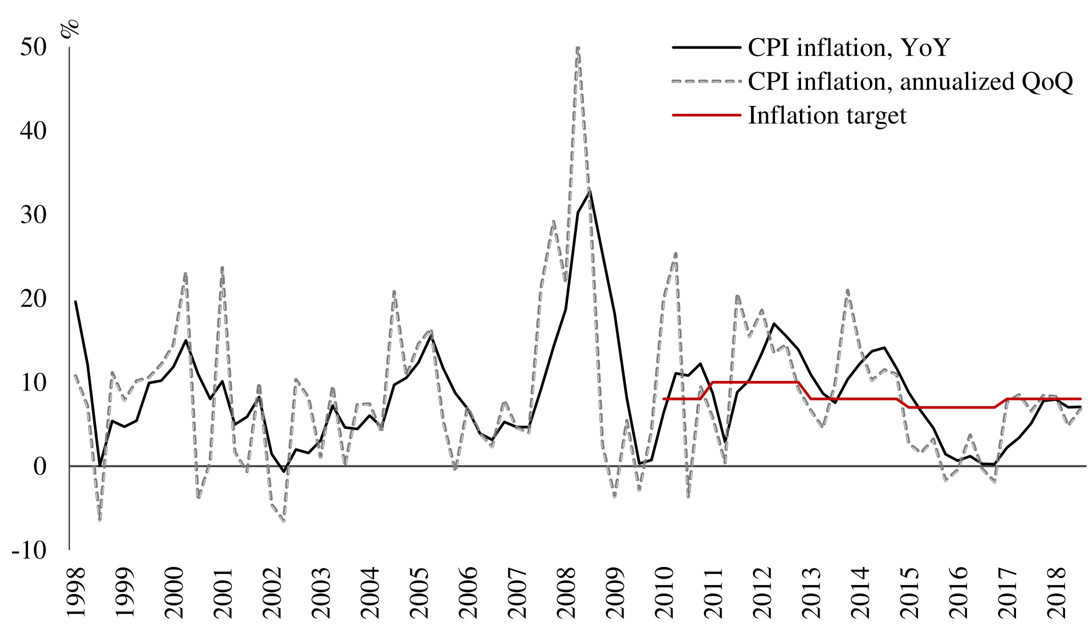
Though several papers have examined inflation dynamics, determinants of inflation and forecasting inflation in Mongolia, no paper has employed BMA or DMA methods to analyze inflation forecasting performance. Based on classic econometric analysis, some papers (i.e., Batrnyam et al. 2008, Gan-Ochir 2011, Gan-Ochir and Dulamzaya 2014) show that (i) inflation is highly seasonal, (ii) inflation fluctuations is dominated by changes in food prices (supply shocks), and (iii) inflation is also driven by demand factors such as government expenditure and exchange rate. Barnett et al. (2012) find that food prices are a key driver of inflation and demand factors (higher fiscal spending through wage hikes and excess demand) are also significant in explaining price movements. Using a Markov-Switching Phillips curve and time-varying VAR models extended with commodity prices, Davaajargal (2015) find that (i) parameters of predictors change over time, and (ii) inflation is mainly driven by government wage growth for the period 2008-2012 and by exchange rate depreciation for the period 2013-2014. Urgamalsuvd et al. (2019) find that (i) the degree of inflation persistence has increased since 2010 and current stands at 0.7, and (ii) the core inflation excluding food and fuel prices has higher persistence compared to overall, food and fuel inflations.

The rest of this paper is structured as follows. Section 2 provides an overview of the inflation dynamics in Mongolia. Section 3 describes the DMA approach employed in the paper. Section 4 presents the data, choice of forgetting factors, and empirical results. Finally, section 5 concludes with implications.

2. Inflation dynamics in Mongolia

Inflation in Mongolia has been volatile and relatively high in last two decades (Figure 1). The economy is subject to large supply and demand shocks. On the supply side, Mongolia experiences harsh winter conditions and is a landlocked and geographically large country, all of which point to high transport costs and the potential for supply bottlenecks (Barnett et al. 2012). Compared to advanced economies, weight of food, gasoline, and administrative price items in consumer price index (CPI) basket are relatively high. In the calculation of headline CPI, administered price items, accounting for roughly 15% of the basket, include items whose prices are adjusted infrequently, such as tuition fees that are adjusted once a year when the new school year starts. Food items accounts for about 35% of the basket, and the 100% of gasoline and diesel fuel items, accounting for 6% of the basket, are imported from overseas. Hence, inflation volatility has been mainly driven by supply factors such as changes in food and gasoline prices. Movements in food prices reflect a large extent supply shock to agriculture rather than changes in demand conditions. Inflation increased in 2000-2001 because of severe winter (Dzud) and drought, reduced supplies of meat, vegetable, and wheat.

Figure 1. Inflation dynamics in Mongolia



Source: Bank of Mongolia

On the demand side, mineral exports are a key driver of the economy, however also volatile due to global commodity price shocks. For instance, commodity exports (mineral and animal products) of the economy account for 95 percent of its total exports, which is about 40 percent of its GDP. Though economic environment was favorable for the period 2002-2007, authorities did not use the advantage to build fiscal and foreign-exchange reserve buffers, strengthening the financial sector, improving the investment environment, and pursuing sound macroeconomic and structural policies. Instead, expansionary monetary and fiscal policies took in place. In the summer of 2008, inflation peaked over 30 percent because of rapid growth in government spending (including child money transfer and government wage growth), 2007-08 world food price crisis and bottleneck at the border. Food price contributed one-third of the headline inflation. As the economy was harshly hit by adverse commodity price shocks, inflation turned negative a year later. As the economy was harshly hit by the collapse of copper prices during the global financial crisis (GFC), inflation turned negative a year later. The economic difficulty led the authorities to request a Stand-by Arrangement from the International Monetary Fund (IMF).

Inflation dipped again in early 2009 to return double digits due to Dzud of 2010-11 and the rapid economic recovery. The quick stabilization and recovery of the economy was mainly driven by positive external shocks such as increases in copper prices, strong coal demand from China, and huge flood of foreign direct investment (FDI), encouraged by development of the Oyu Tolgoi (OT) copper and gold deposit, which is the largest foreign investment project ever in Mongolia and attracted more than $6 billion (50 per cent of GDP) in FDI for its first phase (i.e., the period 2010-2013).

In response to the adverse external shocks (i.e., the super cycle of commodity prices ends in 2012, FDI stopped in 2014[5]), expansionary policies have been implemented for the period 2012-2016. In addition to the state budget, off-budget financing operations have been implemented over the period 2012-2016: (i) central bank's quasi-fiscal operations including the Price Stabilization Program (PSP)[6], (ii) the lending from the Development Bank of Mongolia to publicly-motivated projects (using USD 600 million DBM bond proceeds), (iii) the use of 'Chinggis' bond proceeds (USD 1.5 billion sovereign bonds is issued in the international market at the end of 2012), and (iv) concession agreements. The central bank's quasi-fiscal operations (policy lending programs) were launched in late 2012 when the political demand for higher spending mounted. As the budget revenue growth gradually slowed in the midst of declining FDI and the weakening export revenues, the currency issuance power of the central bank was seen as a reliable financing source that could be tapped to support growing spending demand without revenue constraints. Hence, the government relied on the central bank as an alternative financing source for fiscal operations. Another motivation of the PSP was steamed from an observation that inflation, mostly driven by food inflation, is aggravating poverty, and worsening the income

[5] FDI remained subdued until 2017 because of sharp decline in commodity prices, completion of the first phase of the OT project, and political risk and uncertainties surrounding the mining sector.
[6] The implementation of the PSP is approved by the parliament as it is included in monetary policy guidelines for 2013 and 2014 and the action plan of the government for 2012-2016.

distribution. The main argument is that poorer households feel more pressure when prices of food and fuel increase since the item account for high weight in the CPI basket. The share of food in the consumption basket is about 60% for the lowest expenditure quartile of households.

By explaining the arguments, the central bank has attempted to reduce inflation and spur economic growth using unconventional operations. However, the exceptionally large monetary and quasi-fiscal stimulus led risks ratcheting up inflation, increasing public debt, boosting balance of payment (BOP) pressures, and heightening banking sector vulnerabilities. Eventually, loose monetary and fiscal policies to buffer the economy from the external shocks supported the economic growth for a while (i.e., the period for 2013-2014), but at the cost of economic vulnerabilities. In particular, the policies raised import demands, depleted international reserves, fueling domestic currency depreciation pressure, and pushed up inflation for the period 2012-2015. Consequently, rating agencies were racing to downgrade Mongolian sovereign ratings, and investors' confidence weakened. Therefore, market confidence waned, credit crunch occurred, economic growth dropped, and the economy faced deflation in the second half of 2016. To rebuild the confidence and successfully refinance the short-term public debts, a new government formed based on the 2016 parliamentary election has started to implement the IMF's three-year arrangement under the Extended Fund Facility (EFF) since May 2017. The government's program aims to stabilize the economy, reduce the fiscal deficit and public debt, rebuild foreign exchange reserves, introduce measures to mitigate the boom-bust cycle, and promote sustainable and inclusive growth.

Under the EFF program, Mongolia has made progress in strengthening its economy. Since 2016, the economy has experienced a sharp recovery in real GDP growth, mainly driven by a reform of structural policies, stronger volumes and advantageous prices of coal and copper, high level of FDI for the second phase of the Oyu Tolgoi copper and gold mine, and recovery in confidence. As a result, gross international reserves increased three times, reaching US$3.5 billion. Due to booming tax revenues and relatively contained expenditures, the fiscal balance has improved, and government debt has fallen to 60 percent of GDP. The GDP growth has been above 5% for the last two years, and the inflation rate has been stable at around its target of 8 percent.

3. The methodology of DMA

The generalized Phillips curve with constant coefficient (CC) is given as

$$y_t = \phi + x_{t-1}\,\delta + \varepsilon_t \tag{1}$$

where y_t is inflation, and x_t is a vector of predictors including lagged inflation. This equation is relevant for forecasting at time t given information through time $t-1$. When forecasting $h > 1$ periods ahead, x_{t-1} becomes x_{t-h} in equation (1). Following Stock and Watson (2011), we include p lags of inflation, regardless of the choice of h. For instance, when forecasting 4-quarter ahead inflation ($h = 4$) with $p = 2$, the lags are y_{t-4} and y_{t-5}, and the other predictors are all dated $t-4$. Despite their advantages of providing simple and easy estimation, CC models are also obvious. To be specific, the regressor coefficients are fixed, and are not allowed to change over time. Hence, the time-varying parameter (TVP) model has emerged to overcome the flaws of CC models.

The TVP method allows the parameters of explanatory variables to change over time, incorporating the naturally time-varying relationship between dependent and independent variables. The TVP version of equation (1) can be presented as follows:

$$y_t = z_t\theta_t + \varepsilon_t \tag{2.1}$$

$$\theta_t = \theta_{t-1} + \eta_t \tag{2.2}$$

for $t = 1, \ldots, T$, where $z_t = [1, x_{t-h}]$ is a $1 \times m$ vector of predictors for inflation, $\theta_t$is a $m \times 1$ vector of coefficients (states), $\varepsilon_t \sim i.i.d.N(0, H_t)$, and $\eta_t \sim i.i.d.N(0, Q_t)$. The errors, ε_t and η_t are assumed to be mutually independent at all leads and lags. This kind of TVP model can be estimated using the Kalman filter method. In the TVP model as defined by equations (2.1) and (2.2), it is assumed that the predictors in z_t are fixed throughout all of the time points, which may lead to a substantial loss of forecasting precision and problems of over-parameterization. However, DMA improves the TVP model by allowing the predictor sets (forecasting models) and their coefficients both to change over time. Therefore, following Raftery et al. (2010) and Koop and Korobilis (2012), this paper employs DMA approach to forecast inflation. When allowing a set of predictors to change over time, we will have a set of $K = 2^m$ models that are characterized by having different subsets of z_t as predictors. Denoting these by $z^{(k)}$ for $k = 1, \ldots, K$, the DMA method is illustrated as follows:

$$y_t = \mathrm{z}_t^{(k)}\theta_t^{(k)} + \varepsilon_t^{(k)} \tag{3.1}$$

$$\theta_t^{(k)} = \theta_{t-1}^{(k)} + \eta_t^{(k)} \tag{3.2}$$

where $\mathrm{z}_t^{(k)} \subseteq z_t$, $\varepsilon_t^{(k)} \sim i.i.d.N(0, H_t^{(k)})$ and $\eta_t^{(k)} \sim i.i.d.N(0, Q_t^{(k)})$. Let $L_t \in \{1, 2, \ldots, K\}$ denote which model applied at each time period, $\Theta_t = \left(\theta_t^{(1)'}, \ldots, \theta_t^{(k)'}\right)'$and $y^t = (y_1, \ldots, y_t)'$. As different models hold at each point in time, we will do model averaging ('dynamic model averaging'). To be precise, when forecasting time t variables using information through time $t-1$, DMA involves calculating $Pr(L_t = k|y^{t-1})$ for each $k = 1, \ldots, K$ and averaging forecasts across models using these probabilities (i.e., their historical forecasting performances). In the approach, Dynamic Model Selection (DMS) involves selecting single model with highest value for $Pr(L_t = k|y^{t-1})$ and using this to forecast. Predictions from DMA and DSA can then incorporate the uncertainty factors from these K models' in a dynamic way:

$$\hat{y}_t^{DMA} = E(y_t|y^{t-1}) = \sum_{k=1}^{K} p_{t|t-1,k}\, \mathrm{z}_t^{(k)}\hat{\theta}_{t-1}^{(k)} \tag{4}$$

$$\hat{y}_t^{DMS} = E(y_t|y^{t-1}, k^*) = \mathrm{z}_t^{(k^*)}\hat{\theta}_{t-1}^{(k^*)} \tag{5}$$

where the probability of model k is $p_{t|t-1,k} = Pr(L_t = k|y^{t-1})$ and $\hat{\theta}_{t-1}^{(k)}$ is estimated parameters used in the forecasting, and $k^* = arg\ max_k(p_{t|t-1,k})$ in equation (5) implies the model with the highest probability $p_{t|t-1,k^*}$. Though the specifications such as (3.1) and (3.2) of DMA are potentially great interest in empirical macroeconomics, the problems with such a framework are that many of the models can have a large number of parameters (and hence, risk of being over parameterized) and the computational burden that arises when K is large implies that estimation can take a long time (Koop and Korobilis 2012). In other word, the specification with large K will derail the results of standard recursive algorithms such as Kalman filter.

In order to deal with these restraints, Raftery et al. (2010) propose a simple estimation of DMA. This estimation simplifies calculation without loss of forecasting accuracy by using the Kalman filter method. The initial assumptions of this estimation are that $\theta_{t-1}^{(k)}$ is independent and identically distributed, and that $\theta_{t-1}^{(k)}$ can be determined separately only if $L_{t-1} = k$. In this set of assumptions, λ is recommended as a so-called forgetting factor and fixed to number slightly below one. This forgetting factor plays the most important role in the calculation of $\theta_{t-1}^{(k)}$, where j period gains λ^j weight from the starting period. λ can also simplify the covariance matrix of $\theta_{t-1}^{(k)}$. This process is given below:

$$\theta_{t|t-1}^{(k)} = \theta_{t-1|t-1}^{(k)} \tag{6}$$

$$\Sigma_{t|t-1}^{(k)} = \frac{1}{\lambda}\Sigma_{t-1|t-1}^{(k)} \tag{7}$$

where $\Sigma_{t|t-1}^{(k)}$ denotes the covariance matrix of $\theta_{t-1}^{(k)}$. Then, the estimation of the DMA parameters is completed by the following updating equations:

$$\hat{\theta}_{t|t}^{(k)} = \hat{\theta}_{t-1|t-1}^{(k)} + \Sigma_{t|t-1}^{(k)} \mathrm{z}_t^{(k)\prime} \left(H_t^{(k)} + \mathrm{z}_t^{(k)}\Sigma_{t|t-1}^{(k)}\mathrm{z}_t^{(k)\prime}\right)^{-1} \left(y_t - \mathrm{z}_t^{(k)}\hat{\theta}_{t-1}^{(k)}\right) \tag{8}$$

$$\Sigma_{t|t}^{(k)} = \Sigma_{t|t-1}^{(k)} - \Sigma_{t|t-1}^{(k)} \mathrm{z}_t^{(k)\prime} \left(H_t^{(k)} + \mathrm{z}_t^{(k)}\Sigma_{t|t-1}^{(k)}\mathrm{z}_t^{(k)\prime}\right)^{-1} \mathrm{z}_t^{(k)}\Sigma_{t|t-1}^{(k)} \tag{9}$$

where equation (7) is simplified with $Q_t^{(k)} = (\lambda^{-1} - 1)\Sigma_{t|t-1}^{(k)}$ by using the forgetting factor. If $\lambda = 1$ then $Q_t^{(k)} = 0$, which means that $\theta_t^{(k)}$ equals its value at time $t - 1$. By choosing $0 < \lambda \leq 1$, we introduce time-variation in $\theta_t^{(k)}$. For instance, setting $\lambda = 0.99$ for quarterly data indicates that observations five years ago receive approximately 80% as much weight as last period's observation, which corresponds to gradual time-variation in $\theta_t^{(k)}$. When $\lambda = 0.95$, observations 5 years ago receive only about 35% as much weight as last period's observation, suggesting that a relatively larger shock hits the regression coefficients.

Another forgetting factor is used in the second assumption, in which $0 < \alpha \leq 1$ is applied in equations (4) and (5). If we use a transition matrix of probability, we must consider $K = 2^m$

model comparisons with m predictors at each time point. Once m is larger than 5 ($K > 32$), it is not practicable to operate the Markov switching in the $K \times K$ matrix. Therefore, using forgetting factor α is a practical way to reduce calculation time and error. In this way, the probability in the forecasting model is determined as follows:

$$p_{t|t-1,k} = \frac{p_{t-1|t-1,k}^{\alpha}}{\sum_{l=1}^{K} p_{t-1|t-1,l}^{\alpha}} \tag{10}$$

and the updating equation is

$$p_{t|t,k} = \frac{p_{t|t-1,k} f_k(y_t|y^{t-1})}{\sum_{l=1}^{K} p_{t|t-1,l} f_l(y_t|y^{t-1})} \tag{11}$$

where $f_l(y_t|y^{t-1})$ denotes the predictive normal density of model l (i.e., $N\left(\mathrm{z}_t^{(l)\prime}\hat{\theta}_{t-1}^{(l)}, H_t^{(l)} + \mathrm{z}_t^{(l)} C_{t|t-1}^{(l)} \mathrm{z}_t^{(l)\prime}\right)$) evaluated at y_t. The forgetting factor α refers to the weight applied to model performance. Clearly, the lower the value of α, the lesser weight is given to past performance. For example, if $\alpha = 0.99$, forecast performance 5 years ago receives approximately 80% as much weight as forecast performance last period when using quarterly data. Raftery et al. (2010) and Koop and Korebilis (2012) recommend setting α close to one. On the other hand, Dangl and Halling (2012) fix α at 1. The equations (6)-(11) consists of all the steps of the Kalman filter prediction and the updating process.

Conditional on $H_t^{(k)}$, the estimation and forecasting strategy outlined above only involves evaluating formulae such as in the Kalman filter. All the recursions above are started by choosing a prior for $p_{t|t,k}$ and $\theta_0^{(k)}$ for $k = 1, \dots, K$. Therefore, we need to determine a way to the evolution of $H_t^{(k)}$. Raftery et al. (2010) recommend a simple plug in method where that $H_t^{(k)} = H^{(K)}$ for all t. However, in many macroeconomic applications, allowing for time-variation in the error variance better suits underlying assumptions. Koop and Korobilis (2012) use an Exponentially Weighted Moving Average (EWMA) estimate of $H_t^{(k)}$. In this paper, we follow a generalized approach allowing for time-variation in the error variance employed by Prado and West (2010). The approach adopts a discount factor to induce time-variation in $H_t^{(k)}$. Particularly, it is done by imposing another forgetting factor $0 < \beta \leq 1$, which enters the scale1 and the shape parameters of the inverted-gamma distribution ($H_t^{(k)} \sim IG\left(\frac{1}{2}, \frac{1}{2} S_t^{(k)}\right)$) [7], such that $n_t^{(k)} = \beta n_{t-1}^{(k)} + 1$. This way, $H_t^{(k)}$ is updated according to new data and forgetting past information to reflect changes in volatility. If $\beta < 1$, the time t estimate of $H_t^{(k)}$ is given as:

---

[7] For the inverted-gamma distribution, $S_t^{(k)}$ is given by $S_t^{(k)} = S_{t-1}^{(k)} + \frac{S_{t-1}^{(k)}}{n_t^{(k)}}\left(\frac{e_t^{2(k)}}{W_t^{(k)}} - 1\right)$ where $e_t^{(k)}$ and $W_t^{(k)}$ can be found in Prado and West (2010).

$$S_t^{(k)} = (1-\beta)\sum_{s=0}^{t-1}\beta^s\left(\frac{e_{t-s}^{2(k)}S_{t-s-1}^{(k)}}{W_{t-s}^{(k)}}\right) \tag{12}$$

Equation (12) indicates that $H_t^{(k)}$ has the form of an EWMA and older data are further discounted as time progress. When $\beta = 1$, we obtain $H_t^{(k)} = H^{(k)}$. The generalized evaluation of $H_t^{(k)}$requires setting a value for the forgetting factor β. Catania and Nonejad (2018) highlight that certain combinations of λ and β result in similar estimates if the regression coefficients. For example, in a model with moderate degree of variations in $\theta_t^{(k)}$, the values $\beta = 0.99$ and $\lambda = 0.94$, imply similar dynamics in the regression coefficients as values $\beta = 0.95$ and $\lambda = 0.98$. They conclude that if a practitioner chooses to fix $\beta < 1$, then it is best to fix λ close to 1 (say at 0.96).

However, a higher value of λ implies that the uncertainty introduced by an innovation in $\eta_t^{(k)}$ is small (i.e., $\Sigma_{t|t-1}^{(k)}$ is small). This might be true in tranquil times, but not in in times of turbulence. In order to account for time variation in uncertainty caused by innovations to the vector of coefficients $\theta_t^{(k)}$, Dangl and Halling (2012) allow the forgetting factor λ to take different values. They consider values on a grid $\lambda_j \in \{\lambda_1, \lambda_2, \dots, \lambda_J\}$ where $0 < \lambda_j \leq 1$ and J captures the number of discrete values of λ considered. Conditional on λ_j, $j = 1,2, \dots, J$, a forecast is obtained using equations (4) and (5). Finally, all forecasts $E(y_t|y^{t-1}, \lambda_j)$ are averaged with weights equal to posterior probabilities $p_{t|t-1,\lambda_j} = Pr(\lambda_j|y^{t-1})$, to yield an ultimate forecast:

$$\hat{y}_t^{DMA} = E(y_t|y^{t-1}) = \sum_{j=1}^{J} p_{t|t-1,\lambda_j} E(y_t|y^{t-1}, \lambda_j) \tag{13}$$

$$\hat{y}_t^{DMS} = E(y_t|y^{t-1}, k^*) = \sum_{j=1}^{J} p_{t|t-1,\lambda_j} E(y_t|y^{t-1}, k^*, \lambda_j) \tag{14}$$

For the relation with other alternative models, Raftery et al. (2010) indicates that if $\lambda = 1$, $\alpha = 1$ and $\beta = 1$ then DMA can be treated as BMA without any forgetting. Catania and Nonejad (2018) highlight that if $\lambda = 1$, $\alpha = 1$ and $\beta = 1$, then DMS can be treated as Bayesian model selection (BMS).

4. Data, choice of forgetting factors and empirical analysis

4.1 Data

Data used in the estimation includes quarterly times series of 16 variables for the period 2001Q4-2016Q3. Dependent variable is annual inflation (4-quarter log difference of consumer price index (CPI) (d4lcpi)), and 15 potential predictors include China's GDP growth (gdp_ch), China's inflation (cpi_ch), 4-quarter log difference of copper prices (d4lcopper), 4-quarter log difference of coal prices (d4lcoal), 4-quarter log difference of global oil prices (d4loil), 4-quarter log difference of real GDP (d4lrgdp), seasonally adjusted FDI to GDP ratio (fdi_sa)[8], budget expenditure to GDP ratio (bgdp), 4-quarter log difference of national average wage (d4lwage), 4-quarter log difference of M2 money (d4lm2), 4-quarter log difference of outstanding loan (d4lloant), policy rate per annum (prate), lending rate per annum (lrate), 4-quarter log difference of terms of trade (d4ltot), and 4-quarter log difference of nominal exchange rate (d4lexr).

Foreign variables such as China's growth and China's inflation are directly observed from the Bloomberg database, while copper prices (grade A cathode, LME spot price), coal prices (Australian thermal coal) and crude oil prices (simple average of three spot prices; Dated Brent, West Texas Intermediate, and the Dubai Fateh) are collected from the IMF external statistics. Domestic GDP, CPI, national average wage, and budget expenditure are retrieved from the National Statistical Office. All remaining data are obtained from Statistical Bulletin of the Bank of Mongolia (BOM). All observed variables are in percent.

As a small sample size (60 observations) and many predictors (10 domestic and 5 external predictors) are used, the DMA approach is better equipped to estimate the generalized Phillips curve and forecast inflation. As highlighted in the literature (i.e., Koop and Korobilis 2012, Catania and Nonejad 2018), DMA estimates and forecasts can be sensitive to the setting of three forgetting factors (λ, α and β). To deal with the issue, we set commonly used and suggested values for the forgetting factors.

4.2 Choices of forgetting factors and priors

In the standard setting, we have a total of $2^{17} = 131072$[9] model combinations Furthermore, we set a grid of values for λ as $\lambda_j = \{0.9, 0.91, \ldots, 1\}$ following Dangl and Halling (2012) and Styrin (2019). Hence, total model combinations considered in the analysis are $(2^{17}) \cdot 11 = 1441792$. According to existing papers (i.e., Raftery et al. 2010, Koop and Korobilis 2012, Catania and Nonejad 2018) for forecasting inflation, we set $\alpha = 0.99$ and $\beta = 0.96$, suggested values in the context of working with quarterly data. This choice of forgetting factors allows for reasonable change in coefficients as averaging over λ_j, more weights on 'recent past' forecast performance (slow change in forecasting model over time) and a reasonable time-variation in the error variance. As a standard choice, we impose a uninformative prior over the models (i.e., $p_{0|0,k} = \frac{1}{K}$ for $k = 1, \ldots, K$ so that, initially, all models are equally likely), and a diffuse prior on the initial conditions of the states, such that $\theta_0^{(k)} \sim N(0, 100 I_{n_k})$, where $n_k$is the number of variables in model k, for $k = 1, \ldots, K$.

[8] Seasonal adjustment is made in only FDI data because it has strong seasonal fluctuations.

[9] In the Phillips curve equation, we add two lagged variables of inflation ($p = 2$) on top the 15 exogeneous predictors. Preliminary analysis with lag lengths to up to 4 indicated that two lags of dependent variable leads to the best forecast performance. Since inflation is very persistent, besides these 15 predictors, the inclusion of 2 inflation lags as predictors help capture the persistence.

4.3 Empirical results

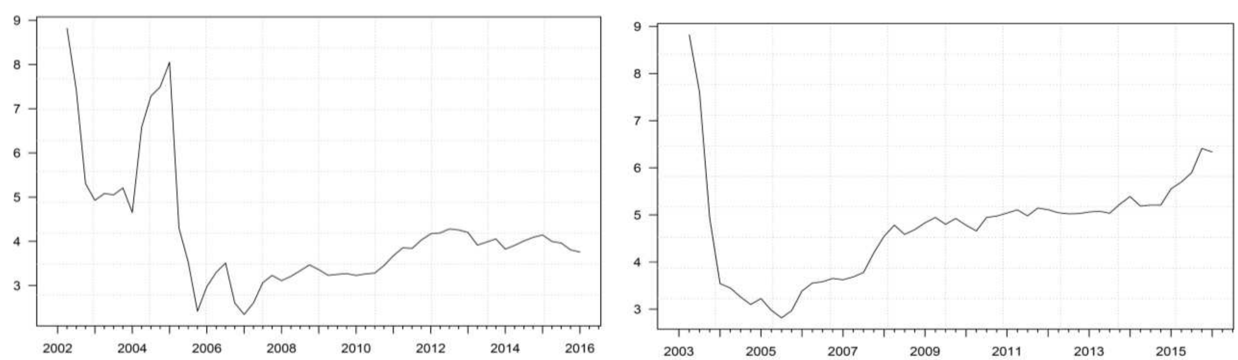
This section focuses on answering two questions: (i) Which variables are good predictors for inflation over time? and (ii) Does DMA improve forecast performance compared to alternative forecasting strategies. In the empirical analysis, all models include an intercept and two lags of the dependent variable.

4.3.1 Time-varying predictors for inflation

In the subsection, we assess how forecasting model of inflation (predictors for inflation) changes over time. Though we have 15 predictors, most probability is attached to very parsimonious models with only a few predictors. If we let $Size_{k,t}$ be the number of predictors in model k at time t, then $E(Size_t) = \sum_{k=1}^{K} p_{t|t-1,k}\, Size_{k,t}$ can be interpreted as the expected or average number of predictors used in DMA at time *t*. Figure 1 plots this measure for our two empirical exercises.

Figure 1. Expected number of predictors, $E(Size_t)$

***(A)* $h = 1$** ***(B)* $h = 5$**



For the shot forecast horizon ($h = 1$), the shrinkage of DMA is particularly strong. For 1 quarter ahead forecast of inflation, DMA (in an expected value sense) between 3 and 4 of the 15 predictors, particularly for the period 2005-2016. At the longer horizons of ($h = 5$), slightly more predictors are included, but almost never more than 6 predictors included. The expected number of predictors is estimated as 5-8 for $h = 1$ and 4-9 for $h = 5$ in the early yeas of the sample (i.e., 2001-2004).

The results show that DMA with chosen forgetting factors favors parsimonious models and provide evidence that predictors changes over time. To know which predictors are important at each pointy in time and how the predictors are changing over time, we show posterior inclusion probability of predictors in Figures 2-3. The posterior inclusion probability implies how the predictor is useful for forecasting at time t, and equivalently, it is the weight used by DMA attached to models which include the predictor. In this regard, DMA can be used as a tool to detect which variables are good predictors at time t.

Figure 2. Posterior probability of inclusion of predictors, $h = 1$

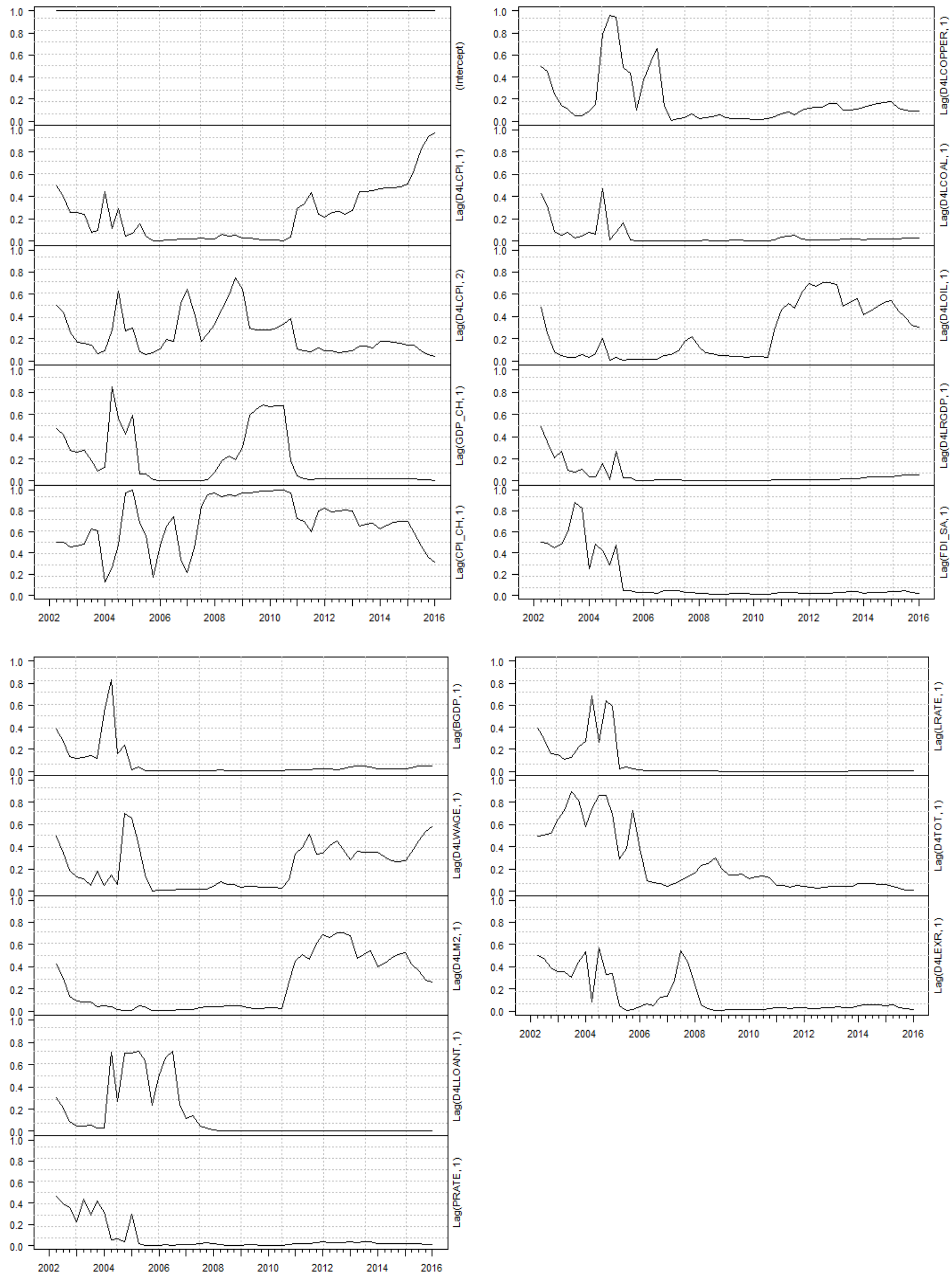
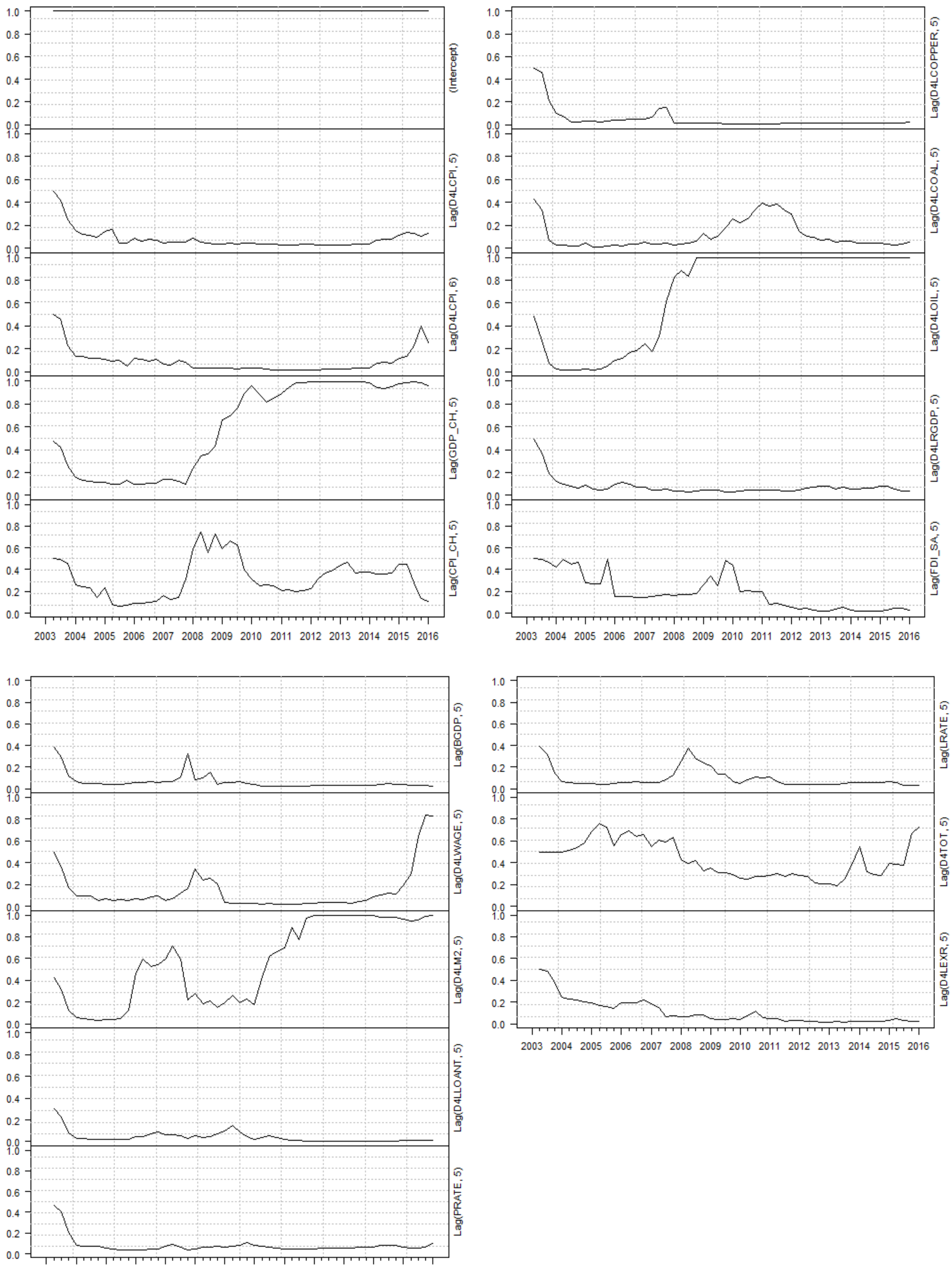


Figure 3. Posterior inclusion probability of predictors, $h = 5$



There is strong evidence of model change for both short and longer forecast horizons. As suggested by time-varying posterior inclusion probabilities, the set of predictors in the forecasting model changes over time. Almost half of our potential predictors come through as being important at some time for both forecast horizons ($h = 1$ and $h = 5$). These predictors include China's GDP growth (gdp_ch), China's inflation (cpi_ch), annual inflation of oil price (d4loil), annual growth of M2 (d4lm2), annual inflation of national average wage (d4lwage), annual change in terms of trade (d4ltot), budget expenditure to GDP ratio (bgdp) and FDI to GDP ratio (fdi_sa), with first four predictors being usually of particular importance. However, there is a large variation over time and over forecast horizons in relation to which one is a good predictor for inflation.

Results for $h = 1$ show that external variables such as China's growth, China's inflation, change in copper price, change in oil price, change in terms of trade and FDI to GDP ratio have been good predictors of annual inflation at some time. The result is completely consistent with the evidence that such external shocks account for almost half of the business cycle fluctuations in Mongolia (Gan-Ochir and Davaajargal 2019). China's growth was a good predictor in the periods 2004-2005 and 2009-2010, however has been become a poor predictor since 2010. China's inflation has been a dominant predictor in almost all the periods, with high posterior inclusion probabilities for the period 2007-2010. These findings suggest that domestic inflation dynamics before and during GFC are highly affected by the biggest trading partner (China)[10]. Change in copper price was a good predictor for the period 2002-2007 when copper export was a major source of export and budget revenues in the economy. As Mongolia still imports 100% of domestic petroleum consumption, change in oil price has been a dominant predictor, particularly since 2010. The change in terms of trade was a useful predictor in the beginning of the sample, however lost its capacity to forecast $t + 1$ inflation after 2005. FDI to GDP ratio had a predictive power in the period 2002-2004 when Mongolia started to receive high amount of FDI in mining sector, but DMA abruptly drops the predictor in 2005 (i.e., the posterior inclusion probability changes from near 0.9 to near zero.

For the domestic variables, budget expenditure to GDP ratio and lending rate were decent predictors in the beginning of the sample but lost their capacity to predict since 2005. Annual growth of loan had a predictive power for the years (2004-2006) when credit growth was high (above 40% per year). Another interesting result is that change in exchange rate plays an important role in short-term forecasting of inflation for the years (i.e., 2004 and 2007) when dollarization increases, pressure on international reserves rises (i.e., expectation of exchange rate depreciation is formed in the FX market). These predictors were historically good predictors, however M2 growth and national average wage growth has become important short-term predictors since 2010. Predictive power of M2 growth gradually increased as M2 growth increases in the economy. In line with the fact that inflation become less volatile and more

[10] Mongolia sells over 90% of total exports to China, and imports buys 40% of total imports from China.

persistent since 2010 (Urgamalsuvd et al. 2019), 1-quarter lagged inflation variable becomes an important predictor after the early-2010s.

Posterior inclusion probabilities for $h = 5$ also indicate that external predictors such as China's growth, China's inflation, change in oil price and change in terms of trade are dominant predictors of annual inflation at some time. In particular, the posterior inclusion probabilities for China's growth and change in oil price have gradually increased during the domestic economic recession driven by GFC (starting from 2008) and become 1 since 2010. Why oil price recently gains a predictive power is due to rapid growth in petroleum consumption driven by mining sector boom, surges in transportation of mineral export and rapid increases in the number of cars in the economy. From the domestic predictors, M2 growth and inflation of national average wage have predictive powers. The importance of M2 growth gradually raised starting after the domestic economic recession and its inclusion probability takes 1 since 2011. The results reconfirm that the domestic inflation (and the economy) is very vulnerable to external shocks and money growth is still a valid domestic predictor of longer horizon forecast of inflation. Wage inflation has recently been gaining its predictive power as it is a main source of income for households with debt burden. Change in terms of trade is one of the few predictors which have been persistently important in predicting inflation for longer horizon.

As a benefit of DMA, the analysis suggests that (i) China's inflation, change in oil price, wage inflation, M2 growth and 1 period lagged inflation must be included in the current short-horizon forecasting model for inflation, and (ii) China's growth, change in oil price, wage inflation, M2 growth, and change in terms of trade are good predictors for longer-horizon inflation forecasts.

4.3.2 Forecast performance

This subsection investigates forecast performance by comparing DMA forecasts to those produced by several alternative methods. An important feature of DMA is out-of-sample forecasting. There are many metrics for evaluating forecast performance and many alternative forecasting methodologies. In this paper, we present two forecast comparison metrics involving point forecasts. These are mean squared forecast error (MSFE) and mean absolute forecast error (MAFE). In terms of forecasting methods, we calculates results for (i) $\mathcal{M}_0$: Plain AR(2) estimated by setting $\lambda = 1$, $\alpha = 1$ and $\beta = 1$; (ii) $\mathcal{M}_1$: Time-varying AR(2) with $\lambda_j = \{0.9, 0.91, \dots, 1\}$ (TVP-AR(2)); (iii) $\mathcal{M}_2$: DMA using only two lags of the dependent variable with $\lambda_j = \{0.9, 0.91, \dots, 1\}$, $\alpha = 0.99$ and $\beta = 0.96$ (DMA (2)); (iv) $\mathcal{M}_3$: DMA using two lags of the dependent variable and 15 exogenous predictors with $\lambda_j = \{0.9, 0.91, \dots, 1\}$, $\alpha = 0.99$ and $\beta = 0.96$ (DMA (2,15)); (v) $\mathcal{M}_4$: DSA with using two lags of the dependent variable and 15 exogenous predictors $\lambda_j = \{0.9, 0.91, \dots, 1\}$, $\alpha = 0.99$ and $\beta = 0.96$; (vi) $\mathcal{M}_5$: BMA as a special case of DMA (2,15) with $\lambda = 1$, $\alpha = 1$ and $\beta = 1$ (BMA (2,15)); and (vii) $\mathcal{M}_6$: BMS as a special case of BMA (2,15) with $\lambda = 1$, $\alpha = 1$ and $\beta = 1$ (BMS (2,15)). As all methods are Bayesian, and log predictive likelihood difference (log(PLD)) (the preferred method of Bayesian forecast comparison) are also presented. Table reports results for our forecasting exercises

(MSFE, MAEF and log(PLD)) of $\mathcal{M}_i$, $i = 1, \dots, 6$, over $\mathcal{M}_0$ (the benchmark) at $h = 1$ and $h = 5$.

Table 1. Comparing different forecasting methods

	$h = 1$			$h = 5$		
	MSFE	MAEF	log(PLD)	MSFE	MAEF	log(PLD)
$\mathcal{M}_1$: TVP-AR(2)	0.902	0.925	4.980	0.961	0.992	0.809
$\mathcal{M}_2$:DMA(2)	1.171	1.054	-0.223	0.972	0.988	0.279
$\mathcal{M}_3$:DMA(2,15)	0.702	0.810	9.883	0.649	0.828	12.096
$\mathcal{M}_4$:DMS(2,15)	0.855	0.875	1.841	0.771	0.916	3.728
$\mathcal{M}_5$:BMA(2,15)	0.808	0.902	5.700	0.750	0.910	4.655
$\mathcal{M}_6$:BMS(2,15)	0.970	0.975	-2.794	0.865	0.935	-7.751

Note: Different methods $\mathcal{M}_i$, $i = 1, \dots, 6$ compared to $\mathcal{M}_0$ (plain AR(2)) for $h = 1$ and $h = 5$ quarters ahead out-of-sample forecasts.

Compared to the benchmark, $\mathcal{M}_1$provides gains in terms of MSFE and MAEF for both short and longer forecast horizons [11]. In terms of log(PLD), $\mathcal{M}_1$ provides gains and losses for $h = 1$ and $h = 5$, respectively. The result of $\mathcal{M}_2$ shows that we do not obtain more gains by using DMA when only lags of inflation are used in the forecasting model. However, DMA using lags of inflation as well as 15 exogenous predictors ($\mathcal{M}_3$) is the top performer, regardless of h. BMA provides the second-best performer in terms of all metrics for both short and longer forecast horizons. Both DMS and BMS produce better forecast performances relative to the benchmark ($\mathcal{M}_0$), TVP-AR(2) and DMA(2). Only exception is the poor performance of BMS at $h = 1$ compared to TVP-AR(2). DMA and DMS (with $\alpha = 0.99, \beta = 0.96$ and average over $\lambda_j = \{0.9, 0.91, \dots, 1\}$) respectively outperform BMA and BMS (setting with $\lambda = 1, \alpha = 1$ and $\beta = 1$) in terms of all metrics for both time horizons. These findings suggest that (i) the exogenous predictors contain enough information to improve forecast accuracy, (ii) averaging of model combinations performs overwhelmingly better than a single best model, and (iii) choice of forgetting factors matters to improve forecasting accuracy. Our exercise certainly recommends to use DMA(2,15) with the chosen forgetting factors in predicting inflation for Mongolia.

5. Conclusion

This paper has investigated the use of DMA approach for identifying good inflation predictors and forecasting inflation in a commodity-exporting economy. The approach is preferably fit for our use as it allows for the set of predictors and their marginal effects to change over time. Using Mongolia as a representative case study, we estimate a DMA that includes several domestic and external variables, capturing characteristics of the economy.

[11] The metric (MSFE or MAEF) is less than one means forecast errors of $\mathcal{M}_i$, , $i = 1, \dots, 6$ is smaller compared to those of the benchmark ($\mathcal{M}_0$).

From our empirical work, three important results stand out. First, external variables (i.e., China's growth, China's inflation, change in copper price, change in oil price, and change in terms of trade) play important role in forecasting inflation. The best external predictors change considerably over time and over forecast horizons. For example, China's inflation is a good predictor for short-horizon forecasting, while China's growth is a dominant predictor for longer-horizon inflation forecasts. However, change in oil price is one of the best predictors for both forecast horizons. The findings support the claim that the domestic inflation is very prone to external shocks in the commodity dependent economies. Second, among domestic variables, wage inflation and M2 growth are currently the best predictors for short and longer forecast horizons. Particularly, we present evidence that M2 growth is a strong predictor of inflation for longer forecast horizon. The importance of wage inflation in predicting inflation has gained since 2015. Third, the use of DMA lead to substantial improvements in forecast performance. Our exercise, comparing seven competing methods, shows that DMA (2,15) with the chosen forgetting factors is the best performer in predicting inflation for Mongolia. There is solid evidence that when using DMA, the inclusion of exogenous predictors, averaging of model combinations and choice of forgetting factors matter to improve forecasting accuracy.

These results provide some implications for modelling inflation dynamics and conducting macroeconomic policies in commodity dependent economies. First, external factors must be explicitly incorporated into a model when analyzing inflation dynamics in a commodity-exporting economy. Second, policy responses to inflation driven by external factors should be different depending on the nature and transmission channels of the factors. Finally, to stabilize inflation, it is important to ensure an optimal policy mix where monetary policy, fiscal policy and FX intervention focus on money growth, revenue expenditure growth and allowing flexibility in exchange rate to absorb external shocks.

References

Allegret, JP and Benkhodja, MT 2015, 'External shocks and monetary policy in an oil exporting economy (Algeria)', *Journal of Policy Modelling,* vol. 37, no. 4, pp. 652-667.

Ang, A, G. Bekaert and M. Wei 2007, 'Do macro variables, asset markets, or surveys forecast inflation better?', *Journal of Monetary Economics,* vol. 54, no. 4, pp. 1163-212.

Atkeson, A. and L. Ohanian 2001, "Are Phillips curves useful for forecasting inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review,* vol. 25, no. 1, pp. 2-11.

Batnyam, D, Gan-Ochir, D and Łyziak, T 2008, 'Small Inflation Model of Mongolia (SIMOM)', Working Paper Series, vol. 5, The Bank of Mongolia.

Bjørnland HC and Thorsrud, LA 2015, 'Bloom or Gloom? Examining the Dutch Disease in two speed economies', *The Economic Journal,* vol. 126, no. 598, pp.2219-2256.

Barnett, S, Bersch, J and Ojima, Y 2012, 'Inflation dynamics in Mongolia: Understanding the roller coaster', IMF Working Paper, WP/12/192, International Monetary Fund.

Catania, L and N. Nonejad 2018, 'Dynamic model averaging for practitioners in economics and finance: The EDMA package, *Journal of Statistical Software,* vol. 84, no. 11, pp. 1-39.

Dangl, T and Halling, M 2012, 'Predictive regression with time-varying coefficients', *Journal of Financial Economics*, vol. 106, no. 1, pp. 157-181.

Davaajargal, L 2015, 'Exchange rate pass-tough in Mongolia', Working Paper Series, vol. 10, The Bank of Mongolia.

Ferreira, D and Palma, A 2015, 'Forecasting inflation with the Phillips curve: A dynamic model averaging approach for Brazil', Revista Brasiliera de Economia, vol. 69, no. 4, pp. 451-564.

Gan-Ochir, D 2011, 'Small model of inflation in Mongolia (SMIM): Effects of cash hand-out on inflation', Working Paper Series, vol. 6, The Bank of Mongolia (In Mongolian).

Gan-Ochir, D and Davaajargal, L 2019, 'External shocks and business cycle fluctuations in Mongolia: Evidence from a Large Bayesian VAR', *International Economic Journal*, vol. 33, no. 1, pp. 42-64.

Gan-Ochir, D and Dulamzaya, B 2014, 'Monetary transmission mechanism: Cost channel', Working Paper Series, vol. 9.2, The Bank of Mongolia (In Mongolian).

Groen, J, Paap, R and Ravazzolo, F 2010, 'Real-time inflation forecasting in a changing world', Staff Report Number 388, Federal Reserve Bank of New York.

Hou K, Mountaiin, DC and Wu, T 2016, ‘Oil price shocks and their transmission mechanism in an oil-exporting economy: A VAR analysis informed by a DSGE model’, *Journal of International Money and Finance,* vol. 68, no. C, pp. 21-49.

Koop, G and Korobilis, D 2012, ‘Forecasting inflation using dynamic model averaging’, *International Economic Review,* vol. 53, no. 3, pp. 867–886.

Koop, G and Potter, S 2004, ‘Forecasting in dynamic factor models using Bayesian Model Averaging’, *The Econometrics Journal,* vol. 7, pp. 550-565.

Liu P, Mumtaz, H and Theophilopoulou, A 2014, ‘The transmission of international shocks to the UK. Estimates based on a time-varying factor augmented VAR’, *Journal of International Money and Finance,* vol. 46, no. C, pp. 1-15.

Raftery, A, Karny, A and Ettler, P 2010, ‘Online prediction under model uncertainty via Dynamic Model Averaging: Application to a cold rolling mill’, *Technometrics,* vol. 52, no. 1, pp. 52–66.

Stock, J and M. Watson 1999, ‘Forecasting inflation’, *Journal of Monetary Economics,* vol. 44, no. 2, pp. 293–335.

Stock, J, and M. Watson 2007, ‘Why has U.S. inflation become harder to forecast?’, *Journal of Monetary Credit and Banking,* vol. 39, no. 1, pp. 3-33.

Styrin, K 2019, ‘Forecasting inflation in Russia using dynamic model averaging’, vol. 78, no. 1, *Russian Journal of Money and Finance*, vol.78, no. 1, pp. 3-18.

Urgamalsuvd, N, Dulamzaya, B and Enkhbayar, J 2019, ‘Inflation persistence in Mongolia’, Chapter 5, : D. Finck and P. Tillmann (Eds), *Price-setting behaviour and inflation dynamics in SEACEN member economies and their implications for inflation,* The SEACEN Centre, pp. 145-157.

Wei, Y and Cao, Y 2017, ‘Forecasting house prices using dynamic model averaging approach: Evidence from China’, *Economic Modelling,* vol. 61, no. 6, pp. 147-155.

Wright, JH 2009, ‘Forecasting US inflation by Bayesian model averaging’, *Journal of Forecasting,* vol. 28, no. 2, pp. 131-144.