Graduate Institute of International and Development Studies

International Economics Department

Working Paper Series

Working Paper No. HEIDWP05-2017

Forecasting Inflation in Vietnam with Univariate and Vector Autoregressive Models

Tran Thanh Hoa
The State Bank of Vietnam

Chemin Eugène-Rigot 2
P.O. Box 136
CH - 1211 Geneva 21
Switzerland

FORECASTING INFLATION IN VIETNAM WITH UNIVARIATE AND VECTOR AUTOREGRESSIVE MODELS[1]

Tran Thanh Hoa[2]
The State Bank of Vietnam

Abstract

In this paper, I apply univariate and vector autoregressive (VAR) models to forecast inflation in Vietnam. To investigate the forecasting performance of the models, two naïve benchmark models (one is a variant of a random walk and the other is an autoregressive model) are first built based on Atkeson-Ohanian (2001), Gosselin-Tkacz (2001) and the specific properties of inflation in Vietnam. Then, I compute the pseudo out-of-sample root mean square error (RMSE) as a measure of forecast accuracy for the candidate models and benchmarks, using rolling window and expanding window forecasting evaluation strategies. The process is applied to both monthly and quarterly data from Vietnam for the period from 2000 through the first half of 2015. I also apply the forecast-encompassing Diebold-Mariano test to support choosing statistically better forecasting models from among the different candidates. I find that VAR_m2 is the best monthly model to forecast inflation in Vietnam, whereas AR(6) is the best of the quarterly forecasting models, although it provides a statistically insignificantly better forecast than the benchmark BM2_q.

Keywords: Inflation, Forecast, Univariate Models, Vector Autoregressive Models, Forecast Accuracy
JEL Classifications: C22, C32, C51, C53, E31, E37

[1]This working paper was written at the Graduate Institute of International and Development Studies (IHEID) in Geneva, Switzerland, under the Visiting Program Framework of Bilateral Assistance and Capacity Building for Central Banks (BCC), which is a part of the cooperating agreement signed between the Swiss State Secretariat for Economic Affairs (SECO) and the State Bank of Vietnam (SBV).

[2]The author sincerely thanks Professor Ugo Panizza (IHEID) for all of his valuable guidance and comments as a supervisor. The author also gratefully thanks the BCC visiting program (IHEID and SECO) and The State Bank of Vietnam for all of the facilitation and support. The views expressed in this paper are those of the author and do not necessarily represent those of The State Bank of Vietnam.

Please send further comments to one of the following email addresses: hoa.tranthanh@sbv.gov.vn or hoattsbv@gmail.com; they are sincerely appreciated by the author.

1. INTRODUCTION

There is no official announcement specifying the time when an inflation-targeting framework will be conducted; however, in recent years the State Bank of Vietnam (hereafter SBV) has been successfully controlling inflation via monetary policy, which focuses not only on promoting economic growth but also on stabilizing the macroeconomy.

This achievement has been supported significantly by forecasting inflation. The SBV has incentives to perform inflation forecasting because one of its main tasks, as specified by the State Bank Law of 2010, is to stabilize price levels to stabilize the domestic currency (VND). The Forecasting and Statistics Department (hereafter called FSD) was established in December 2008 to implement the task of performing statistical and macroeconomic forecasts to support monetary policy, in which forecasting inflation is a core function. The FSD has been submitting its monthly inflation reports to the SBV Management Board since mid-2012. Additionally, forecasting inflation is indispensable work to prepare for an inflation-targeting framework, which will hopefully be conducted by the SBV in the near future. Forecasting inflation is also in accordance with the developing trend of all central banks in the world.

In every monthly inflation report, nowcasting and forecasting for one time unit in the future are implemented using Microsoft Excel calculations, which are based on the categories of eleven first-level groups or eighty-six third-level groups of commodities and services' price indexes. These groups are included in the collection used to compute the headline consumer price index (CPI). Updated information from markets on any commodity or service's price in the collection within the month is gathered in an Excel spreadsheet to compute the aggregate CPI for the current month and combined with prior market information to compute the inflation forecast for the following month. The point forecasts and inflation forecast intervals are submitted in the monthly inflation report. For the final month of each quarter, inflation forecasts for the following quarter are also released in the report. For semi-annual or annual reports, inflation forecasts for the second half of the year or for the following year are also reported, respectively. To produce future forecast horizons (next quarter, next half or next year), SBV now uses reduced-form vector autoregressive models (reduced-form VARs) and vector error correction models (VECMs). According to the SBV's plans, other types of models that are more structurally oriented, such as Bayesian VARs (BVARs) and dynamic stochastic general equilibrium (DSGE) models, will be built and developed in the near future to produce inflation forecasts.

This paper attempts to apply univariate and VAR models to forecast inflation in Vietnam. The work is expected to enrich the SBV's toolkit for forecasting inflation. Forecasting inflation is an important task, especially in the context of the SBV's preparation for its explicit inflation-targeting framework. To investigate the forecasting performance of the models, two naïve benchmark models (one is a variant of a random walk and the other is an autoregressive model) are first built based on Atkeson-Ohanian (2001), Gosselin-Tkacz (2001) and the specific properties of inflation in Vietnam. Then, I compute the pseudo out-of-sample root mean square error (RMSE) as a measure of forecast accuracy for candidate models and benchmarks, using rolling window and expanding window forecasting evaluation strategies. The process is applied for both monthly and quarterly data from Vietnam for the period from 2000 through the first half of 2015. I also apply the forecast-encompassing Diebold-Mariano test to support choosing statistically superior forecast models from among the different candidates. I find that VAR_m2 is the best monthly model to forecast inflation in Vietnam, whereas AR(6) is the best of the quarterly forecasting models, although it provides a statistically insignificantly better forecast than the benchmark BM2_q.

This paper includes five sections in addition to the introduction. Section 2 presents an overview of Vietnam's inflation and its developments. Section 3 contains the literature review, which includes working papers related to inflation forecasting. Section 4 explains the data and methodology used and introduces the benchmark and candidate models. Section 5 presents the forecasting performance results. The last section summarizes the conclusions derived from the empirical results. Additionally, related tables, figures and boxes are included in the Appendix.

2. OVERVIEW OF VIETNAM'S INFLATION AND ITS DEVELOPMENTS

Vietnam had a very high inflation rate in the second half of the 1980s, with a peak of nearly 800% year-on-year (yoy) in December of 1986 (Figure 1). This reflected a period of instability and a dismal situation in macroeconomic development, which was characterized by a vicious cycle of price-wage-currency caused by maintaining the central-planning mechanism in the economy for too long. However, the "Doi Moi" (Renovation) was initiated in 1986, with an important transformation in the economy from centrally planned to market-oriented, which helped Vietnam to successfully restructure its economy. Therefore, inflation has not been as high because the vicious cycle was resolved as one of Doi Moi's achievements.

Figure 1. Inflation Developments in Vietnam 1986-2014 (% yoy, annual data)

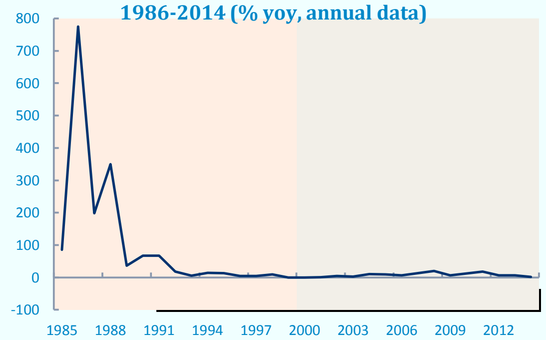
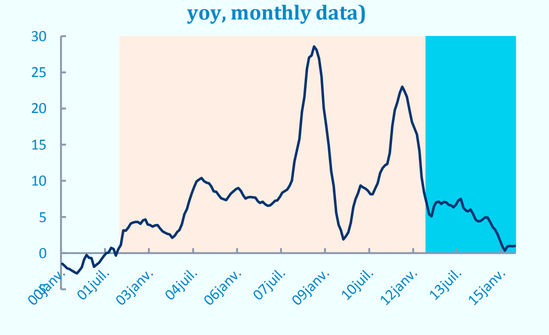
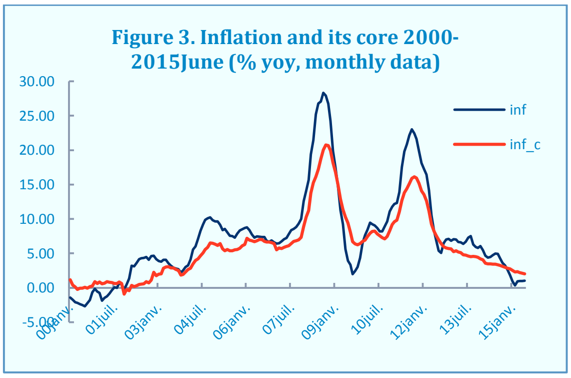


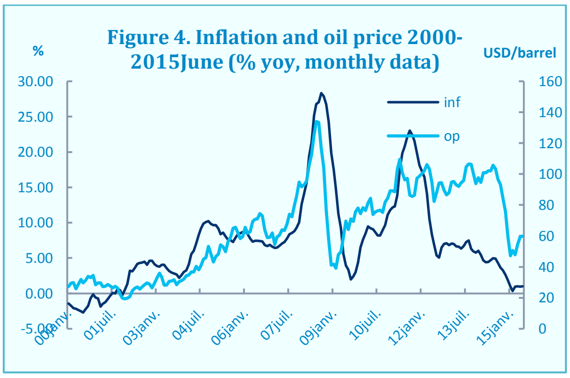
Figure 2. Inflation from 2000-2015 June (% yoy, monthly data)



Since 2000, inflation has been much more stable and significantly lower than in the previous period. Nevertheless, high inflation is still sometimes experienced. The highest peak of 28% yoy was recorded in September 2008, and the second peak of 23% yoy was recorded in August 2011 (Figure 2). Vietnam joined the World Trade Organization (WTO) in 2007 and was then shocked by sudden capital inflows at the beginning of 2008. This is the primary explanation for the first peak of inflation in this period. An economic stimulus package in 2009 provided a rapid and large credit expansion to the economy in the following years to address the slowed economic circumstance caused by the global crisis in 2007. This might explain the second peak of inflation. Some observers even believed in a pattern of a "higher rate for 2 years and then a lower rate in the following year" of inflation since 2004. However, since May 2012, inflation has maintained a single-digit level and keeps decreasing, therefore negating the belief that inflation would come back at a higher rate again.

The factors that affect inflation in Vietnam can be divided into 3 groups: (1) monetary policy, (2) price administration, and (3) price fluctuations in the markets.





Monetary policy makes an important contribution to controlling inflation, which can be foreseen through the core inflation. Specifically, the downward trend of inflation in the sub-period since May 2012 has been led by the decrease in core inflation followed by the rapid decline of oil prices which occurred in the second half of 2014 (Figures 3 and 4). Price administration appears to be a product of a transition economy, in which the government adjusts prices of some special goods and services that were subsidized for such a long time that their prices do not reflect their full costs (for example water, electricity, health and education services), to make their prices equal to market prices. The adjustments are conducted according to specific schedules over a period of several years. In addition, the government must also control other goods (for example gasoline price or other types of energy) to harmonize the domestic price level, using stabilized funds. Therefore, the administrative prices group is still an important factor that has a significant effect on Vietnamese inflation, not only historically but also presently and in the future. However, over the long-term, while these goods and services' prices will be asymptotic or equal to the market price, the proportion of the group with this factor will be reduced remarkably. For the third factor group, in the domestic market the price of food (raw foods, processed foods and cereal) is critical because accounts for the largest weighted share of the CPI. In the international market, rice and oil prices are more important than others because the two prices can account for the fluctuations in the domestic rice price and the gasoline price, respectively. However, the government has instruments to use as buffers to mitigate exogenous oil price shocks to domestic gasoline and energy prices.

Working with inflation forecasts based on calculations in Excel often requires detailed and disaggregated information about each minuscule good and service. This process is sometimes efficient, but sometimes it is complex and confusing. Therefore, there is also an incentive to develop other types of models to create inflation forecasts, especially for the long-term.

3. LITERATURE REVIEW

The purpose of this paper is to forecast inflation by applying univariate and vector autoregressive models. Therefore, the literature review is focused on influential papers concerning inflation forecasts in general, and specifically cases in Vietnam with similar approaches, which are the basis of my paper.

Atkeson and Ohanian (2001) build random walk models as naïve benchmarks to compare the forecast accuracy of the three Phillips curves to these benchmarks based on US data: (i) for the Phillips curves based on the textbook NAIRU[3] models that used quarterly data from the first quarter of 1984 to the third quarter of 1999, the authors build a random walk model of a previous fourth quarter as a naïve benchmark ($\pi_t = \pi_{t-4} + u_t$, where π_t is the inflation rate calculated by the percentage change of GDP deflator between quarter t-4 and t); (ii) for the Phillips curves from Stock and Watson (1999b), which used US monthly data from January 1959 to September 1997, the authors build a random walk model of the previous twelfth month as a benchmark model ($\pi_t = \pi_{t-12} + u_t$, where π_t is the inflation rate calculated by the percentage change of the consumer price index (CPI) between month t-12 and t); and (iii) for the Phillips curve model based on the Federal Reserve's Green book, the authors build a random walk model of historical data as a benchmark ($\pi_{t-1} = \pi_{t-5} + u_t$, where π_t is the inflation rate calculated by the percentage change of GDP or GNI deflator between quarter t-4 and t). The authors' findings demonstrate that such random walk models for the annual rate of inflation forecast (benchmarks) provide better results than do multivariate models that use measures of economic activities as predictors (Phillips curves models).

Gosselin and Tkacz (2001) also build a random walk model based on Atkeson and Ohanian's work as a naïve benchmark in their paper applied to the Bank of Canada. In addition, an autoregressive model of the previous fourth quarter is used as another benchmark model ($\pi_t = c + \beta\pi_{t-4} + u_t$), which is much more general than a random walk AO model because it is not necessary to set the restrictions of c=0 and $\beta = 1$ as in a random walk. The authors also build a vector error correction model (VECM) as the third benchmark and primarily focus on factor models (combining several useful explanatory variables for the rate of inflation into one or a few representative factors) as candidates, using quarterly data from 1969Q1 to 2000Q1. After performing the forecast exercise, the authors conclude that in terms of forecast accuracy, the factor models are statistically equal to the benchmarks. Additionally, the authors highly recommend the use of factor models to forecast inflation because they can provide information that is useful for "at least predicting changes in the direction of inflation", even if a model’s RMSE is not significantly less than the benchmarks', and "this is an important feature for monetary policy decisions".

Stock and Watson (2008) use US quarterly data from the first quarter of 1953 to the first quarter of 2008 to build three univariate models (in which the random walk model of the previous fourth quarter recommended by Atkeson and Ohanian (2001) is used again by Stock and Watson and is called the "AO model" in this paper), two backward-looking Phillips curves and an autoregressive distributed lag model. Stock and Watson provide evidence of improved performance with the

[3] NAIRU is an acronym for "non-accelerating inflation rate of unemployment", or the baseline unemployment rate.

Phillips curves relative to the AO model during some sub-periods, but they do not affirm the entire sample.

The three papers demonstrate the rational reasons for continuing to use univariate models to forecast inflation, and this is one methodology that I apply in this paper. Simultaneously, the papers suggested building structurally oriented models (Phillips curves, multivariate models with economic variables, or factor models) to forecast inflation. Although they cannot provide better results than the univariate (or other type of benchmark) models in terms of forecast accuracy, they are still useful in predicting the change of the direction of inflation; therefore, they make an important contribution to policy makers in making monetary policy decisions. Based on this perspective, VARs are applied concurrently with univariate models to forecast inflation in this paper.

The term VAR (vector autoregressive model) was first used and made popular by Sims (1980). In a VAR model, every explanatory variable is a lag of the endogenous variables in the system. Moreover, Sims (1992) introduces a structural VAR with recursive identification. Using monthly data for France, Germany, Japan, the UK, and the US (from 1958M4 to 1991M2 with the US), Sims provides empirical evidence for the "price puzzle" in the impulse response functions. This means that inflation statistically increases following a tightening of monetary policy or an increase in the policy interest rate. The result is robust because the author uses a six-variable or four-variable VAR. Sims also gives two explanations for the price puzzle. First, inflation would increase more significantly without an increase in the policy rate. Second, it might take time for the monetary policy to have any real impact on inflation (the policy lag).

In another paper, Waggoner and Zha (2010) summarize theories and references concerning identification of a non-recursive structural VARs and also provide examples of using a structural VAR with non-recursive identification.

Recursive or non-recursive identification of a structural VAR, however, is more popularly used to analyze the monetary or other policy than to forecast, not only in the final two papers I referenced above but also in the majority of the literature, because the orthogonal identification of SVAR does not affect the forecasting result compared to the reduced-form VAR. Consequently, this paper uses reduced-form VARs in addition to univariate models to focus on forecasting inflation.

In the case of Vietnam, specific papers have been written using a VAR approach, but the majority of them focus on investigating the monetary transmission mechanism rather than on forecasting inflation.

Camen (2006) builds VAR models using monthly data from February 1996 to April 2005 with the following variables: US aggregate money supply (M3US), petrol price, rice price, USD/VND exchange rate, domestic aggregate money supply (M2) and consumer price index (CPI), while M2 is replaced by credit to the economy (CTE) or lending rate (LR) or both in alternative models. The author's finding is that the lending rate does not contribute to the explanation of inflation, whereas the exchange rate contributes as an important factor. The paper also analyses the variance decomposition but does not focus on forecasting inflation.

Nguyen and Fujita (2007) use monthly data from January 1992 to April 2005 for VAR models (two alternative VAR models with five endogenous variables – the index of industrial production (IIP) as a proxy for output, consumer price index (CPI), exchange rate, money supply, and trade balance – and an exogenous variable, the Fed rate). The findings indicate that the primary source of variance in output and CPI comes from "own shocks" and the exchange rate has a larger impact on output than on inflation. No focus on forecasting inflation is provided.

Le and Pfau (2008) use quarterly data from 1996Q2 to 2005Q4 for a VAR approach with one basic model and three channel models. The basic model is a VAR of three endogenous variables (output, CPI and M2) and three exogenous variables (oil price, rice price and Fed fund rate). In interest rate, exchange rate, and credit rate channel models, the interest rate, real effective exchange rate and credit to the economy, respectively, are added as one more endogenous variable relative to the basic model. The finding is that there is no obvious relationship between the money supply and inflation, whereas there exists a statistical relationship between the money supply and the real GDP. The authors do not focus on forecasting inflation.

Bui and Tran (2015) use quarterly data from 2000Q1 to 2011Q4 to build a six-variable VAR model of money demand, GDP, inflation, lending rate (LR), exchange rate (EXR) and stock exchange index (VNI). Their research indicates that for a shorter horizon, the variation of prices is mostly explained by internal shocks and monetary demand, whereas for a longer horizon, much of the movement is caused by monetary demand, interest rates and the stock exchange.

As stated above, all of the papers concerning Vietnam focus on investigating the monetary transmission mechanism in Vietnam rather than on forecasting inflation. Therefore, my paper is expected to make a contribution to the literature regarding forecasting inflation in Vietnam.

4. DATA AND METHODOLOGY

4.1. Data

In this paper, I use quarterly data from 2000Q1 to 2015Q2 (62 observations) and monthly data from January 2000 to June 2015 (186 observations). Whereas univariate models only use the inflation rate series to forecast itself, multivariate models (specifically VARs) require other variables to perform inflation forecasts. The data include the following series based on the real economy (real GDP or its proxy) and policy variables, as reported in Table 1 in the Appendix.

The inflation rate (inf) is defined as the year-on-year growth rate of the CPI, as computed using the following equations:

$$inf_t = \left(\frac{CPI_t - CPI_{t-12}}{CPI_{t-12}}\right) \times 100 \quad (1)$$

$$inf_t = \left(\frac{CPI_t - CPI_{t-4}}{CPI_{t-4}}\right) \times 100 \quad (2)$$

Equation (1) is used to compute the monthly inflation rate from the monthly CPI, whereas equation (2) is used to compute the quarterly inflation rate from the quarterly CPI, which is the simple average of the three months' CPI within a quarter.

The real GDP growth rate (ry_g) is also computed using equation (2) by replacing the quarterly CPI series with the quarterly series of real GDP in billion VND. Specifically, the real GDP growth rate is the change in percentage of the real GDP between this quarter (t) and the previous four quarters (t-4).

The exchange rate (exr) is computed as an average value for the period (it is the average rate in a month for monthly data and the average rate in a quarter for the quarterly data). The exchange rate series is computed from the daily exchange rate of a commercial bank (the representative is Vietcombank (VCB), the bank that has the largest share of foreign trade activities). It is used as a proxy for the official exchange rate or policy exchange rate. The SBV has its own official exchange rate, which is announced by SBV as a reference for interbank trading among commercial banks. However, the official rate is kept constant for a long period, with infrequent adjustments, so there are many large changes in its series[4]. The exchange rate that is used for trades between VCB and its clients must lie within the bounds of the official exchange rate; however, it reflects more closely the markets and real activities.

The lending rate that commercial banks apply to their clients is also used as a proxy for the policy interest rate (ir). This lending rate is collected from the International Financial Statistics of International Monetary Funds (IFS-IMF). The refinancing rate or discounting rate is similar to the official exchange rate, which cannot easily be changed. The other rates, such as the official interest rate announced by SBV for the interbank market, the executing rate in the interbank market, or even the open market operation (OMO) rate, will be other good proxies for policy rate; however, they have only been observed in the past several years, so the series are shorter and therefore not the same length as the other variables.

In general, the monthly and quarterly data sets include the same variables but with a different frequency. There are some points that distinguish the monthly and quarterly data series. First, quarterly data includes the real GDP growth representing the real economy, whereas the monthly data includes the real retail sales (rrs) being used as a proxy for the real GDP. Second, the quarterly data includes the series "credit to the economy" (cred), which is collected from Reuters, whereas the monthly data does not include this series.

In the case of the inflation series, it is interesting that when the frequency changes, the stationarity also changes, as indicated in Tables 2 and 3 in the Appendix. The null hypothesis of the unit root is rejected at the 5% level with the monthly inflation rate; however, it fails to reject the null hypothesis of the unit root with the quarterly inflation rate.

The series "real retail sales" is computed using the nominal retail sales (from the General Statistics Office) divided by the series of the CPI-base of 2009 (computed by the author). The "WTI oil price" (op) and "effective Fed funds rate" (ifed) are used as exogenous variables in the VAR models.

Whereas the monthly inflation rate and effective Fed funds rate are stationary variables, the other variables are I(1) and must be transformed into a stationary form before being used to estimate

[4] This occurred in Vietnam until the beginning of 2016. A new exchange rate regime has been in place since January 1, 2016, such that the "central rate USD/VND" is announced by the SBV daily, based on the weighted average rate of the foreign currency interbank market, the foreign currency exchange rate versus the USD of main trading partners in the international market, and macroeconomic and monetary balances. This means that the official exchange rate of the USD/VND fluctuates in both directions following external markets, which is more flexible than the prior rate.

models and forecast. The Augmented Dickey-Fuller (ADF) unit root test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are used to investigate the stationarity of variables. The KPSS test is only considered seriously when the ADF test fails to reject the null hypothesis of the unit root (this sometimes occurs when the ρ in the unit root test is very close to, but still smaller than, 1). Tables 2 and 3 in the Appendix present the results of these unit root tests (ADF test) and stationarity tests (KPSS test). Figures 1 and 2 in the Appendix present graphs of the variables included in both datasets.

4.2. Methodology

4.2.1. Strategy for evaluating forecast accuracy

The purpose of all forecasters is to build models that make accurate forecasts or, at a minimum with, models with the smallest possible forecast errors. Therefore, evaluating forecast accuracy is the main function of forecasting work. There are many measures for evaluating forecast accuracy (Bias, MSE, RMSE, SE, MAE, MAPE, Theil IC, etc.). Nevertheless, to be concise, in this paper, only the root mean square error (RMSE) is used to assess the models' forecast accuracy. The smaller RMSE a model has, the more accurate the forecast or better performance in forecasting it is able to achieve compared with other models.

In practice in SBV, to evaluate the forecasting properties of a model, forecasting accuracy measures are computed for in-sample forecasts. This method is simple to apply; however, one model that might be good at in-sample forecasting is not necessarily good at out-of-sample forecasting. The alternative solution for computing forecast accuracy measures for out-of-sample forecasting, without the real out-of-sample outcomes, is by simulating the computation of pseudo out-of-sample forecast errors, as suggested by many econometricians (see also Stock and Watson, 2008).

This paper makes a contribution in computing the pseudo out-of-sample RMSE for Vietnamese inflation, which has not yet been computed in the SBV or presented in any working paper applied to Vietnam. There are two approaches to implementing the forecasting strategy: (i) applying a recursive window (a so-called expanding window) and/or (ii) applying a rolling window. In this paper, I apply both types of windows to compute pseudo out-of-sample RMSEs among models. More details regarding the in-sample and pseudo out-of-sample RMSEs, expanding window and rolling window can be found in Box 1 of the Appendix.

4.2.2. Benchmark setting

Based on Atkeson-Ohanian (2001), Gosselin-Tkacz (2001), and the specific properties of Vietnamese inflation, I establish the naïve or benchmark models as follows:

For monthly data, two benchmarks are built. The first is a variant of a random walk that includes on the right-hand side the lag one of inflation, its lag twelve and a constant (c), with the restriction of $\rho_1 + \rho_{12} = 1$(BM1_m). The other is an AR process of inflation, with the lag one and lag twelve, which is more general than the first benchmark model because there is no need to set $\rho_1 + \rho_{12} = 1$:

$$inf_t = c + \rho_1 inf_{t-1} + (1 - \rho_1) inf_{t-12} + u_t \qquad \text{(BM1\_m)}$$

$$inf_t = c + \rho_1 inf_{t-1} + \rho_{12} inf_{t-12} + u_t \qquad \text{(BM2\_m)}$$

The quarterly benchmark models are established with the regressors of lag one and lag four of inflation as below. In BM2_q, it is not necessary to set the restriction $\rho_1 + \rho_4 = 1$, as in BM1_q.

$$inf_t = c + \rho_1 inf_{t-1} + (1 - \rho_1) inf_{t-4} + u_t \qquad \text{(BM1\_q)}$$

$$inf_t = c + \rho_1 inf_{t-1} + \rho_4 inf_{t-4} + u_t \qquad \text{(BM2\_q)}$$

Then the forecasting exercise is to execute all of the benchmarks and the others (so-called candidate models) for both rolling and expanding windows. The model that has the smaller RMSEs is chosen as more preferable for inflation forecasting.

The above benchmark models can capture two important features. First, they consider the seasonality issue by the presence of lag twelve (for monthly data) and lag four (for quarterly data) of inflation on the right-hand side as regressors. This feature is in accordance with the literature review of Atkeson-Ohanian (2001) and Gosselin-Tkacz (2001). Second, the high level of Vietnamese inflation's inertia or momentum is addressed by the presence of lag one of inflation as another regressor in the above benchmark models. This feature also reflects the strong anchor of the expected inflation rate and that of the previous period in Vietnam.

4.2.3. Univariate Models

Using univariate models to perform forecasting is to use only one variable's historical data to predict its future. The first advantage of this approach is that a forecaster only needs to use one variable that must be forecast. It does not require a large dataset with many different variables, which is very difficult to obtain and is normally not available in developing or emerging economies. Second, this approach does not require much economics, but rather econometrics; thus, forecasters who are inexperienced in macro- or micro-economics can feel comfortable and confident to follow the approach. Third, some evidence in the literature indicates that a univariate model approach is an effective methodology because it may have better forecasting accuracy than methodologies that are more structurally oriented (Atkeson and Ohanian, 2001).

Forecasting by using univariate models follows the Box-Jenkins approach. A stationary series of inflation (inf with monthly data, d_inf with quarterly data) is considered if it fits one or more. The equations are written as follows:

(i) an AR(p) process (with p lag): $inf_t = c + \emptyset_1 inf_{t-1} + \emptyset_2 inf_{t-2} + \cdots + \emptyset_p inf_{t-p} + u_t$

(ii) an MA(q) process (with q lag): $inf_t = c + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q} + u_t$

or the combination between the two as an ARMA(p,q) process or an ARIMA(p,d,q), where d is the integrated degree. The Box-Jenkins approach allows forecasters to choose better candidate models or the best from many different options using information criteria (e.g., the AIC, SBIC, or HQIC).

4.2.4. VAR Models

Different from univariate models, VAR models are vector autoregressive models, which include several variables in one system. To perform inflation forecasts, other variables explaining inflation fluctuations are also included in the models.

In most of the literature, VAR is primarily used for analyzing policy through investigating the impulse response function and variance decomposition rather than being used for forecasting purposes. In the forecasting aspect, a reduced-form VAR is preferable. The term "VAR" was first used by Sims (1980). A VAR is a system that is a vector of an AR(p) process. This means that one endogenous variable is explained by the lag of the other variables in the system and the lag of itself.

In this paper, the specifications of reduced-form VAR models are generally written as

$$y_t = M(L)y_{t-1} + Bx_t + v_t \tag{3}$$

where

y_t is a vector of endogenous variables at time t;

$$y_t = \begin{bmatrix} inf_t \\ dl\_rrs_t \\ dl\_exr_t \\ d\_ir_t \end{bmatrix} \text{ in monthly VAR (hereafter called VAR\_m);}$$

$$y_t = \begin{bmatrix} d\_inf_t \\ dl\_cred_t \\ dry\_g_t \\ dl\_exr_t \\ d\_ir_t \end{bmatrix} \text{ in quarterly VAR (hereafter called VAR\_q);}$$

x_t is a vector of exogenous variables; $x_t = \begin{bmatrix} dl\_op_{t-1} \\ ifed_{t-1} \end{bmatrix}$ in VAR_m; $x_t = \begin{bmatrix} dl\_op_t \\ ifed_t \end{bmatrix}$ in VAR_q;

v_t is vector of reduced-form residuals;

$M(L) = 1 - M_1L - M_2L^2 - \cdots - M_pL^p$ is a matrix size Nxp, where N is the number of endogenous variables; p is the lagged number of endogenous variables that describes the coefficients on lagged endogenous variables; and

B is a vector of coefficients on exogenous variables.

In the two VAR models, oil price (dl_op) and Fed rate (ifed) are set as exogenous variables with their first lags in VAR_m and their current time (lag zero) in VAR_q. This is in accordance with the fact that the oil price in the international market cannot have a contemporaneous effect on the domestic inflation within a month, but it may have effects on the domestic inflation within a quarter, because buffers created by administrative price management provide lags in the effects. The same explanation is also suitable for the Fed rate. The crawling peg exchange rate regime that SBV had implemented until the end of 2015 only considers the changes of the Fed rate, also with lag.

Because all of the variables in VAR models are a stationary series, the residuals vector v_t is an unobservable zero mean white noise vector process (serially uncorrelated or independent) with a time-invariant covariance matrix. Put differently, the systems satisfy the assumptions of ordinary least squares (OLS) such that each equation in these systems can be estimated using an OLS approach independently.

Whereas univariate models do not seem to use any economic theory because they only use one variable to estimate and predict itself, it is necessary to consider basic economic theories with VAR models. Based on the inflation determinants, in a small VAR model, the real sector and monetary factors are normally considered as explanatory variables for inflation.

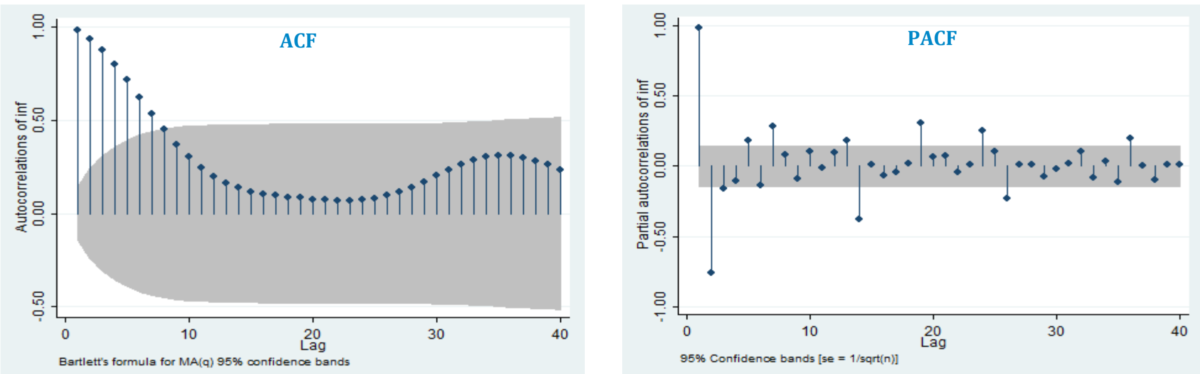
5. FORECASTING PERFORMANCE RESULTS

5.1. Monthly Models

5.1.1. Univariate Models

Following the Box-Jenkins approach, first, visual diagnosis is performed by reviewing the graphs of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the inflation rate (the "correlogram"), as shown in Figure 5.

Figure 5. The correlograms of monthly inflation rate



Whereas the ACF declines exponentially, which suggests an AR(p) process, the PACF appears to go to zero after a lag of thirty-six. The inflation rate therefore seems to be an AR model with a high degree of lags or an ARMA model, which is a combination of an autoregressive process and a moving average process.

In this paper, eleven experimental univariate models are built, and from them, three candidates are chosen for forecasting evaluation. They are the following: (i) AR(6): the current inflation is explained by its lags from lag one to lag six; (ii) ARMA(4,4): the current inflation is explained by its lags from lag one to lag four and the four lags of the error term; and (iii) ARMA(6,6): the current inflation is explained by its lags from one to six and also the lags of the error term from one to six.

Table 1 presents the estimated results and the reasons that one candidate model is chosen over the others. First, a good univariate should have white noise[5] residuals. These results are reasonably consistent with the visual diagnosis through the correlograms, whereas a separate MA(q) process is not feasible. Second, the majority of the chosen models also have statistically significant estimated coefficients.

In the following step, information criteria are used to choose the best univariate model for monthly data. Table 2 presents the model selection suggested by Akaike's information criterion (AIC) and the Bayesian information criterion (BIC). Whereas the AIC suggests selecting ARMA(6,6) with the smallest AIC of 376.9, BIC suggests selecting ARMA(4,4) with the smallest BIC of 417.1. The AR(1) model, which is not chosen because the residual series is not a white noise process, also has the worst AIC and BIC, compared with the other three.

Table 1. Univariate models and appropriate candidates for monthly data

Explanatory Variables		Monthly Univariate Models									
	Dependent variable: inf									Dependent variable: D.inf	
	AR(1)	AR(2)	AR(4)	AR(6)	MA(2)	MA(6)	ARMA(2,2)	ARMA(4,4)	ARMA(6,6)	ARIMA(4,1,4)	ARIMA(6,1,6)
Constant	4.10	6.83***	7.14***	7.08***	7.35***	7.21***	7.08***	7.12***	6.41**	0.025	0.031
AR											
L1	0.99***	1.74***	1.596***	1.64***			1.82***	0.58***	1.01***	1.86***	1.18***
L2		-0.76***	-0.53***	-0.59***			-0.85***	0.54***	0.68*	-0.39	-0.55***
L3			0.003	0.11				0.57***	-0.03	-0.98***	0.53***
L4			-0.102	-0.46***				-0.76***	-1.27***	0.49	-0.76***
L5				0.40***					0.22		0.94***
L6				-0.14**					0.36*		-0.47***
MA											
L1					1.11	1.90***	-0.23***	1.22***	0.71***	-1.20**	-0.43***
L2					1.00	2.68***	0.01	0.84***	-0.25***	-0.44*	0.31**
L3						3.23***		-0.17	-0.997***	1.04**	-0.35***
L4						2.99***		-0.27***	0.01	-0.37*	0.54***
L5						2.14***			0.83***		-0.47***
L6						0.71***			0.53***		-0.45***
WN	X	X	X	√	X	X	X	√	√	X	X
Conclusion	X	X	X	√	X	X	X	√	√	X	X

[5] A white noise process has a constant mean (sometimes assumed to be a zero mean), constant variance and zero autocovariance with all lags non-zero. A white noise test uses the Ljung-Box Q* statistics with the null hypothesis "the autocorrelation function has no significant elements at all lags different from zero". The test statistic follows a chi-squared distribution, where the number of degrees of freedom is the number of lags.

***Note**: AR - Autoregressive process; MA - Moving Average process; WN - The residual series is a white noise process; ***, **, and * represent the significance level at 1%, 5% and 10%, respectively.*

Based on the information criteria, ARMA(6,6) and ARMA(4,4) appear to be better candidates than AR(6). Because a model that is not the best for making estimates may be best for forecasting, all three candidate models are selected for forecasting and evaluating their forecasting performance.

Table 2. AIC and BIC for monthly models

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
ar1	186	.	-285.9255	3	577.8509	587.5282
ar6	186	.	-197.3269	8	410.6537	436.4597
arma44	186	.	-182.4271	10	384.8543	417.1117
arma66	186	.	-174.4535	14	376.9071	422.0675

5.1.2. VAR models

In section 4.2.4, the monthly reduced-form VAR model is specified by the four stationary endogenous variables ($inf_t, dl\_rrs_t, dl\_exr_t, d\_ir_t$) and the lag one of two stationary exogenous variables ($dl\_op_{t-1}, ifed_{t-1}$), and named VAR_m. In its specification, the inflation rate is explained by other endogenous variables' lags, the lags of itself, and the lag one of exogenous variables. To make the discussion more interesting, one can argue that exogenous variables may not affect domestic inflation. VAR_m1 is built with the exogenous variables, whereas VAR_m2 is considered without the exogenous variables.

For the next step, the lag length and stability condition of the models must be checked. The critical values of the lag length criteria are different between VAR_m1 and VAR_m2; however, the conclusions converge to the same results based on the same criterion for both models. Whereas the SC and HQ statistics suggest that VAR_m1 and VAR_m2 should be included in two lags, the FPE and AIC statistics suggest a lag length of three for both models, and the LR statistic suggests a lag length of five (see Table 4 in the Appendix for the lag length criteria suggestion). However, all of the criteria do not indicate the same lag length for one specific model. To have parsimonious models in the context of a small sample size, two lags are applied for both VAR_m1 and VAR_m2.

The two models also satisfy the stability condition because all of the inverse roots of the AR characteristic polynomial points are inside the unit circle (see Figure 3 in the Appendix for the graphs of the inverse root of the AR characteristic polynomial. The two graphs on the left-hand side are depicted for monthly VARs.)

Although some of the estimated coefficients of the explanatory variables are not statistically significant, the estimated results appear to be good because the fitness of the two models (the R_squared), with and without exogenous variables (VAR_m1 and VAR_m2, respectively), are very high. Table 6 in the Appendix presents that the presence of exogenous variables does not increase the R_squared significantly compared to the model in which they are absent. However, models in general without exogenous variables can make the forecasts directly compared to the models with exogenous variables, where the exogenous variables will be predicted by other satellite models or other organizations.

5.1.3. Forecasting performance of monthly models

To evaluate the forecasting performance of the models, this paper applies forecasting evaluation strategies of rolling window and expanding window to compute the pseudo out-of-sample RMSEs of the models and then compare them within one type of window to choose the best. Because only in-sample forecasting measures have been applied in SBV to evaluate the forecasting performance of all of the models, computing pseudo out-of-sample RMSEs by simulation is one of the main contributions of this paper. Descriptions and details of the rolling and expanding window forecasting strategies are given in Box 2 of the Appendix.

There is a reason for selecting a rolling window length of three years. Recall from part 2 of inflation developments, the pattern "higher rate in 2 years then lower rate in following year" of the inflation rate in the period between 2004 and 2012 suggests that a structural break would occur every three years. Therefore, three years may be a suitable length of a business cycle that many observers who believe in the three-year cycle of the inflation rate would understand.

Table 3 presents the forecasting accuracy in terms of pseudo out-of-sample RMSEs of monthly candidate models and benchmarks. There are some remarkable results:

Table 3. Comparison of the pseudo out-of-sample RMSEs among monthly models

Models	**Horizon**											
	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9	h=10	h=11	h=12
Panel 1: Rolling window (window=36, simulations=50, last observations of the first sample=2010m04, horizon=12)												
BM1_m	1.20	2.26	3.17	3.98	4.71	5.35	5.91	6.40	6.80	7.13	7.40	7.58
BM2_m	1.28	2.31	3.11	3.71	4.15	4.48	4.72	4.95	5.16	5.36	5.57	5.83
AR(6)	0.83	1.73	2.44	3.03	3.56	4.01	4.45	4.89	5.29	5.65	6.01	6.37
ARMA(4,4)	0.91	2.00	3.00	3.94	4.82	5.58	6.33	6.99	7.51	7.90	8.23	8.50
ARMA(6,6)	0.98	2.23	3.50	4.69	5.84	6.81	7.58	8.17	8.63	8.95	9.07	9.04
VAR_m1	1.15	2.46	3.87	5.20	6.43	7.33	7.93	8.31	8.43	8.27	8.00	7.82
VAR_m2	1.08	2.16	2.98	3.53	3.99	4.27	4.49	4.74	4.94	5.04	5.09	5.20
Panel 2: Expanding window (simulations=50, last observations of the first sample=2010m04, horizon=12)												
BM1_m	1.19	2.26	3.23	4.13	4.99	5.80	6.55	7.25	7.88	8.43	8.90	9.28
BM2_m	1.15	2.16	3.06	3.88	4.63	5.32	5.94	6.50	6.97	7.34	7.63	7.80
AR(6)	0.87	1.79	2.54	3.25	3.94	4.55	5.13	5.66	6.09	6.46	6.82	7.09
ARMA(4,4)	0.82	1.72	2.48	3.16	3.84	4.42	4.94	5.44	5.89	6.28	6.66	7.01
ARMA(6,6)	0.91	1.88	2.65	3.41	4.09	4.62	5.07	5.51	5.92	6.29	6.67	6.98
VAR_m1	0.84	1.82	2.68	3.39	4.04	4.62	5.15	5.64	6.04	6.33	6.56	6.73
VAR_m2	0.89	1.90	2.73	3.39	3.98	4.49	4.95	5.39	5.77	6.07	6.33	6.53

(i) In the strategy of the rolling window (Panel 1- Table 3): (1) The candidate model AR(6) appears to provide better results than the benchmark model BM1_m, other univariate candidates and VAR_m1 with smallest RMSEs at all twelve months ahead of the horizon. AR(6) also produces better forecasts than BM2_m and VAR_m2 at some first forecasting horizons (h is from 1 to 8 in the case of comparison to BM2_m, from 1 to 7 in the case of comparison to VAR_m2); however, it produces worse forecasts than BM2_m and VAR_m2 at later forecasting horizons; (2) VAR_m2 is likely to forecast better than BM2_m at every forecasting horizon in terms of the smallest RMSEs and becomes the best choice to produce inflation forecasts.

(ii) In the strategy of the expanding window (Panel 2- Table 3): Every candidate model appears to produce a better forecast than the two benchmark models, in which VAR_m2 provides the best forecast in the last 6 months horizon, whereas ARMA(4,4) seemingly achieves the best forecast in the first 6 months horizon. The RMSEs are not significantly different among the candidate models; instead, the results are based only on visual assessment.

The results also prove that at a few points on the horizon the difference in RMSEs among models is so small that they are not significantly different from each other. Therefore, I cannot conclude which model forecasts better than the other, only basing my results on visual assessment. To examine whether the models are significantly different from the others in terms of forecasting accuracy, I apply a forecast-encompassing test called the Diebold & Mariano test (dmariano test). The null hypothesis of the "dmariano test" states that the "forecast accuracy is equal". Where the probability value (p-value) is smaller than the critical value (normally using a critical value of a 5% level), the null hypothesis is rejected by the test.

Applying the "dmariano test" to Panel 1 of Table 3, the two preferable candidate models of VAR_m2 and AR(6) are compared to the better benchmark model BM2_m. Whereas VAR_m2 is statistically a better forecast than BM2_m at the critical value of 5% (p-value=0.0032), it is certain that AR(6) is statistically insignificantly better than BM2_m because the null hypothesis of equal forecast accuracy cannot be rejected at the 5% level (p-value=0.6310). For Panel 2, all of the candidate models are compared to BM2_m. As a result, they all statistically provide a better forecast than BM2_m as recorded in Table 4. To choose the most preferable of the candidates, the "dmariano test" is applied for AR(6), ARMA(4,4), ARMA(6,6), and VAR_m1 compared to VAR_m2. Non-surprisingly, VAR_m2 is statistically a better forecast than VAR_m1 at the 5% level, better than AR(6) and ARMA(6,6) at the 10% level; however, it is not statistically significantly better than ARMA(4,4).

Table 4. Diebold-Mariano forecast comparison test for monthly models

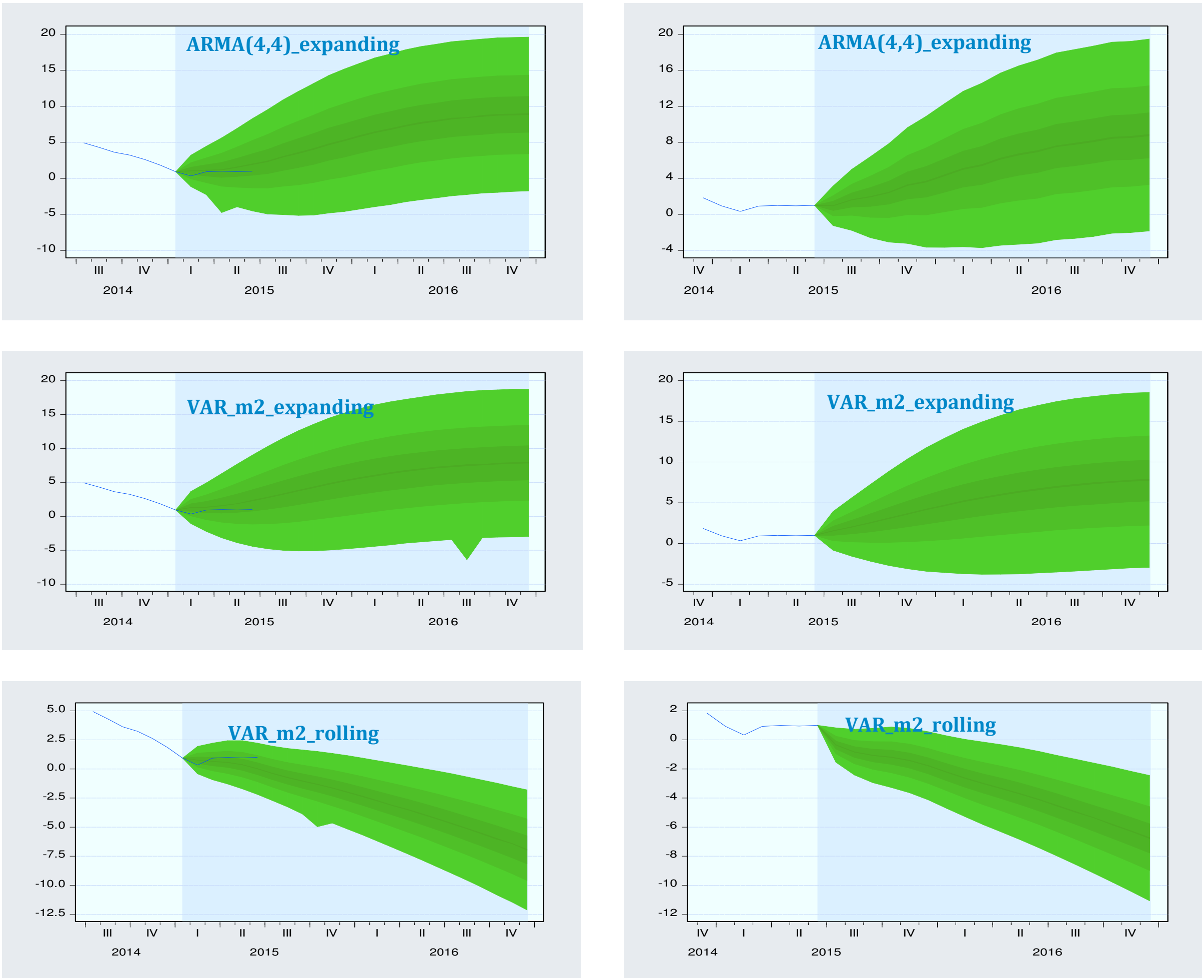
(using the loss function of pseudo out-of-sample RMSEs)

The alternative hypothesis H_1	**The null hypothesis H_0: Forecast accuracy is equal**			
	MSE difference	***Computed S(1)***	***p-value***	***Conclusion***
Panel 1: For the rolling window forecast strategy (The alternative model provides a better forecast compared to BM2_m)				
AR(6)	0.5804	0.4803	0.6310	cannot reject H_0
VAR_m2	2.346	2.945	0.0032	reject H_0
Panel 2: For expanding window forecast strategy				
The alternative model provides a better forecast compared to BM2_m				
AR(6)	7.337	4.123	0.0000	reject H_0
ARMA(4,4)	8.627	4.17	0.0000	reject H_0
ARMA(6,6)	7.852	3.56	0.0004	reject H_0
VAR_m1	7.969	3.233	0.0012	reject H_0
VAR_m2	9.454	3.103	0.0019	reject H_0
VAR_m2 provides a better forecast compared to the alternative model				
AR(6)	-2.117	-1.713	0.0868	reject H_0 at 10%
ARMA(4,4)	-0.8269	-0.9084	0.3637	cannot reject H_0
ARMA(6,6)	-1.602	-1.943	0.0520	reject H_0 at 10%
VAR_m1	-1.485	-2.584	0.0098	reject H_0

Figure 6 shows fan charts of the out-of-sample monthly inflation forecasts of VAR_m2 (rolling), ARMA(4,4) and VAR_m2 (expanding) as suggested from the above results.

Figure 6. Fan charts of the out-of-sample inflation forecasts for monthly models

Comparison to actual values in the first 6 months | **No comparison to actual values**



***Note**: "*_rolling" represents models that have an estimated sample of 36 months from 2012M07 to 2015M06; "*_expanding" represents models that have the largest estimated sample from 2000M01 to 2015M06.*

There is currently no method to evaluate the real out-of-sample forecasting accuracy owing to the absence of actual values. However, a strategy of making fan charts may provide visual diagnostics such that one can make judgments on the forecasting accuracy of the real out-of-sample forecasts. This paper draws two fan charts for each chosen model, in which (i) the fan charts on the left-hand side panel apply the following strategy: because the actual inflation rate series ends at 2015M06, the first point of the fan charts begins at 2015M01 so that the first 6 months of the fan charts (which are forecasted by models) can be compared to the actual inflation rates (called the "in-sample forecast"). (ii) The fan charts on the right-hand side panel do not apply the strategy; thus, the first point of the fan charts begins at 2015M06.

In the logical inference, the right-hand side fan charts will have a more concise interquartile range compared to the left-hand side charts of the chosen models.

The rolling window evaluation strategy tends to produce a negative inflation rate at the end of 2016 because the dominant trend of the last rolling window sample of 36 months is downward-sloping. However, this will not be the case for the outlook for inflation at the end of 2016 in Vietnam. The oil price in 2016 obviously will not decline sharply from the low base of 2015, whereas the administrative price (health service and education service) will be adjusted proportionally to increase. This context will lead to a higher inflation rate in 2016 than in 2015. In contrast, the expanding window tends to make the inflation forecasts for 2016 much more suitable with the real context than those of the rolling window. The left-hand side panel of Figure 6 of VAR_m2_expanding also shows that the actual inflation rates in the first half of 2015 are very close to the median point forecasts (the dark line in the middle of the fan charts) and are included in the narrowest 30% of the interquartile range with the expanding window. The rolling window with only the last three years of estimated sample appears to have a bias or under-predicted inflation forecasts compared to the actual values.

Even the expanding window tends to over-predict inflation forecasts compared to the actual values. The trend of higher inflation rates in 2016 compared to those in 2015 is more or less confirmed by the models with the expanding window. According to the ARMA(4,4), the inflation yoy will be approximately 9% at the end of 2016 while it will be approximately 7.5% as forecast by VAR_m2 (at median point forecasts). Nevertheless, the forecasted values do not converge to an identical number because the median point forecasts are different among models. Therefore, expert adjustment is always important to determine which number will be more reliable. Given the over-forecast recorded at 2015M06 of VAR_m2-expanding (the second graph from the top on the left-hand side of Figure 6), the lower line of the narrowest 30% of the interquartile range should make the inflation forecast closer to inflation's actual values. According to this judgment, the inflation point forecast at the end of 2016 is approximately 5%.

5.2. Quarterly Models

5.2.1. Univariate Models

Repeating all of the steps executed with the monthly data for the quarterly data, some candidate models are found from the pilot models. The correlograms of ACF and PACF also suggest an AR(p) or ARMA(p,q) or even an ARIMA(p,d,q) process over an MA(q) process, and the AR process appears to be more dominant than the MA process with the combination of ARMA or ARIMA (Figure 7).

Table 5 presents the estimated results of the univariate quarterly models. From the twelve experimental models, through the first round, seven candidate models that have white noise residuals are selected: AR(6), MA(6), ARMA(2,2), ARMA(4,4), ARMA(6,1), ARIMA(4,1,1) and ARIMA(6,1,1). However, through the second round, three of the seven models are removed because all of the MA coefficients of these three models are statistically insignificant. The three removed models are: MA(6), ARMA(4,4) and ARIMA(6,1,1). Though the last filter using information criteria, AR(6) is the best model in terms of the smallest BIC at 259.7, whereas ARMA(6,1) is the best model in terms of the smallest AIC, 242.1 (Table 6).

Figure 7. The correlograms of the quarterly inflation rate

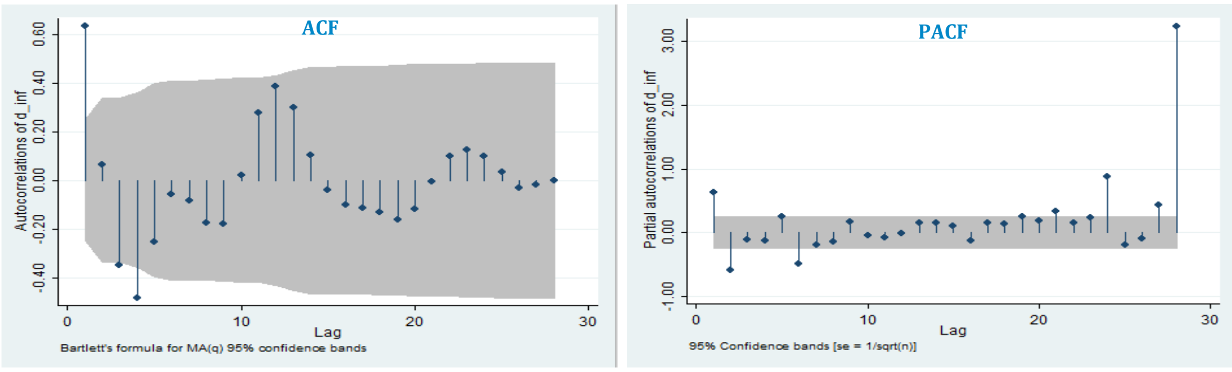


Table 5. Univariate models and appropriate candidates for quarterly data

Explanatory Variables	Quarterly Univariate Models											
	Dependent variable: d_inf								Dependent variable: D.d_inf			
	AR(1)	AR(2)	AR(6)	MA(6)	ARMA(2, 2)	ARMA(4, 4)	ARMA(4,1)	ARMA(6,1)	ARIMA(2,1, 2)	ARIMA(2,1, 1)	ARIMA(4, 1,1)	ARIMA(6,1, 1)
Constant	0.03	0.07	0.09	0.11	0.10	0.12	0.08	0.09	-0.02	-0.02	0.00	-0.02
AR												
L1	0.63***	0.99***	1.07***		1.16***	1.45***	0.18	1.36***	1.18***	0.99***	-0.20	1.05***
L2		-0.53***	-0.70***		-0.47***	-1.15***	0.09	-0.97***	-0.68***	-0.56***	-0.24	-0.71***
L3			0.23			0.40	-0.20	0.37			-0.17	0.23
L4			-0.64***			0.08	-0.32**	-0.69***			-0.57***	-0.65***
L5			0.72***					0.83***				0.71***
L6			-0.47***					-0.55***				-0.48***
MA												
L1				0.82	-0.10	-0.51	0.85***	-0.39*	-1.3	-1	0.54***	-1
L2				0.11	-0.65***	0.16			0.3			
L3				-0.36		0.24						
L4				-1.28		-0.89						
L5				-1.02								
L6				-0.21								
WN	X	X	√	√	√	√	X	√	X	X	√	√
Conclusion	X	X	√	X	√	X	X	√	X	X	√	X

***Note**: AR - Autoregressive process; MA - Moving Average process; WN - The residual series is a white noise process; ***, **, and * represent the significance levels at 1%, 5% and 10%, respectively.*

Table 6. AIC and BIC for quarterly models

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
ar6	61	.	-113.4347	8	242.8694	259.7564
arma22	61	.	-121.4034	6	254.8068	267.472
arma61	61	.	-112.073	9	242.146	261.1439
arima411	60	.	-122.0613	7	258.1227	272.7831

However, the experience obtained from working with the monthly data suggested that I retain all four candidate models listed in Table 6 to compute the forecast accuracy, using pseudo out-of-sample RMSEs (Table 7).

5.2.2. VAR models

As mentioned in section 4.2.4, the quarterly reduced-form VAR model is specified by the five stationary endogenous variables ($d\_inf_t, dl\_cred_t, dry\_g_t, dl\_exr_t, d\_ir_t$) and the lag zero (current time) of two stationary exogenous variables ($dl\_op_t, ifed_t$); this model is called VAR_q. In its specification, the inflation rate is explained by other endogenous variables' lags, the lags of itself, and the lag zero of exogenous variables. Similar to the monthly VARs, VAR_q1 is built with the exogenous variables, while VAR_q2 is considered without the exogenous variables.

Determining the lag length and checking the stability condition of the models are required to draw their exact specifications. The lag length criteria suggested for VAR_q1 and VAR_q2 diverge to different results for each criterion (see Table 5 in the Appendix for lag length criteria suggestions for quarterly VARs). To have parsimonious models in the context of a small sample size, two lags are applied for VAR_q1, and one lag is applied for VAR_q2. The two models also meet the stability condition because all of the inverse roots of the AR characteristic polynomial points are inside the unit circle (see Figure 3 in the Appendix for the graphs of the inverse root of the AR characteristic polynomial, the right-hand side panel for the quarterly VARs).

Table 7 in the Appendix presents the estimated results of the quarterly VARs. With a very small sample size (only 59 observations for VAR_q1 and 60 observations for VAR_q2 after adjustment), some coefficients of the explanatory variables are not statistically significant. In addition, the goodness-of-fit values of the two quarterly VAR models (the R_squared) are much smaller than those of the monthly VARs. In the case of the quarterly VARs, the R_squared of the model with exogenous variables (VAR_q1) is much higher than the model without exogenous variables (VAR_q2). However, the gap between the R_squared value of the two models originates not only from the presence or absence of exogenous variables but also from the lag length of the endogenous variables because VAR_q1 includes two lags, whereas VAR_q2 includes only one lag. To have parsimonious models, choosing the short lag length can reduce the degrees of freedom; however, it leads to a trade-off of more concise R_squared values and sometimes the insignificance of estimated coefficients, in terms of statistics and/or economic theory.

5.2.3. Forecasting performance of quarterly models

This section attempts to repeat the forecasting evaluation exercise, which is applied for monthly models as computing pseudo out-of-sample RMSEs by simulation using expanding windows.

Table 7. Comparison of pseudo out-of-sample RMSEs among quarterly models

Models	**Expanding window** (ir=20, last obs of 1st Sample=2009Q2, h=4)			
Horizon	*h=1*	*h=2*	*h=3*	*h=4*
BM1_q	3.20	5.86	7.85	9.17
BM2_q	3.33	6.23	8.15	8.87
AR(6)	2.32	4.81	6.95	9.03
ARMA(2,2)	2.79	5.83	8.49	10.70
ARMA(6,1)	2.82	5.88	8.55	11.20
ARIMA(4,1,1)	2.86	6.12	8.78	10.86
VAR_q1	3.23	6.52	8.94	11.02
VAR_q2	3.01	6.13	8.88	11.12

Table 7 presents the forecasting accuracy in terms of pseudo out-of-sample RMSEs of the quarterly candidate models and benchmarks. There are some visually apparent results: (1) for the benchmarks, BM2_q appears to provide a better forecast than BM1_q at horizon four, whereas it provides a worse forecast than BM1_q at other horizons; (2) AR(6) appears to provide a better forecast than the benchmarks, whereas ARMA(6,1) is no longer better, especially at horizon three and horizon four; and (3) other candidates cannot provide better results than the benchmarks in forecasting, even at some horizons, whereas some can provide better results than the benchmarks (for example, all of the univariate candidate models and VAR_q2 provide a better forecast at horizon one than the benchmarks).

The Diebold-Mariano tests are performed for quarterly models to examine whether one model provides a statistically better forecast than the others (Table 8). As a result, equal forecast accuracy between BM1_q and BM2_q cannot be rejected, and AR(6) provides a statistically better forecast than BM2_q, whereas there is no evidence that other candidates are better than BM2_q.

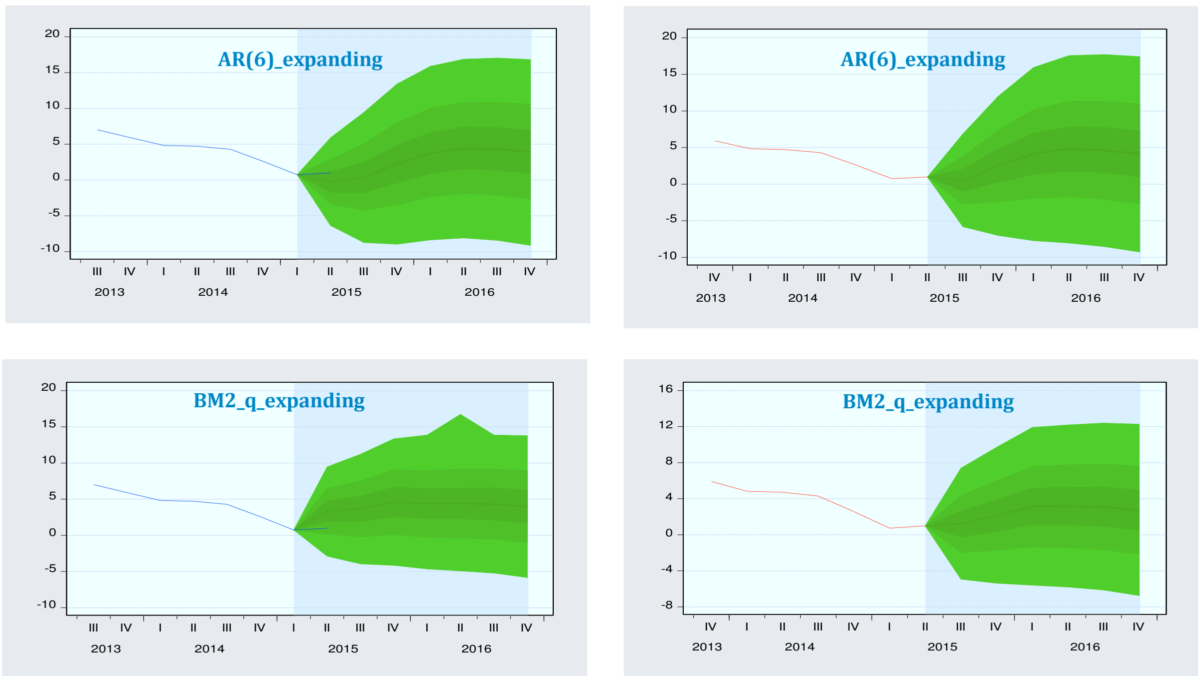
Table 8. Diebold-Mariano forecast comparison test for quarterly models

(using the loss function of pseudo out-of-sample RMSEs)

The alternative hypothesis H_1	The null hypothesis H_0: Forecast accuracy is equal			
	MSE difference	*Computed S(1)*	*p-value*	*Conclusion*
For expanding window forecast strategy				
BM1_q is better forecast than BM2_q	-1.178	-0.7313	0.4646	cannot reject H_0
AR(6) is better forecast than BM2_q	-9.16	-3.147	0.0017	reject H_0
BM2_q is better forecast than ARMA(2,2)	8.335	9.89e+07	0.0000	reject H_0
BM2_q is better forecast than ARMA(6,1)	11.52	1.433	0.1518	cannot reject H_0
BM2_q is better forecast than ARIMA(4,1,1)	11.42	1.35e+08	0.0000	reject H_0
BM2_q is better forecast than VAR_q1	14.83	2.016	0.0438	reject H_0
BM2_q is better forecast than VAR_q2	13.54	1.731	0.0835	reject H_0 at 10%

Figure 7.Fan charts of out-of-sample inflation forecasts for quarterly models

LHS panel_Comparison to actual values in the first 2 quarters **RHS panel_No comparison to actual values**



Note: *"*_expanding" represents models which have the largest estimated sample from 2000Q01 to 2015Q02.*

Figure 7 presents fan charts of out-of-sample inflation forecasts as suggested by the pseudo out-of-sample RMSEs from the above results. The left-hand side panel presents fan charts with the first point of 2015Q1 such that the visual comparison of the median point forecasts to actual inflation values can be depicted. On the right-hand side panel, the first point of the fan charts is 2015Q2, without the comparison to the actual values. The goal is to draw the right-hand side panel to depict the smaller interquartile ranges. The ranges, however, are not tightened as much as they are in the case of quarterly models because only one quarter is dropped back from the left-hand side fan charts.

While AR(6) with expanding window tends to produce an under-forecasted value compared with actual inflation values, the actual value is still inside the narrowest 30% of the interquartile range. However, BM2_q tends to provide an over-forecasted value in very short horizon. However, at the end of 2016, the forecasted inflation rate will be approximately 4% yoy. AR(6) reports a rate a little greater than 4%, whereas BM2_q estimates a rate a little less than 4%.

6. CONCLUSIONS

This paper has evaluated the forecasting performance of different inflation-forecasting models, which included both univariate and vector autoregressive models. Their forecast accuracy was evaluated using frequency datasets, both monthly and quarterly. For the results, some additional accurate candidate models are chosen in terms of more concise pseudo out-of-sample RMSEs and passing the Diebold-Mariano test (to be a statistically better forecast) to produce the real out-of-sample inflation forecasts by the end of 2016.

Focusing on the year-on-year growth rate of headline CPI, two sets of univariate models are constructed: (i) the monthly set initially includes eleven univariate models, then three of them are chosen for evaluating forecasting accuracy. They are AR(6), ARMA(4,4) and ARMA(6,6). They all provide a statistically better forecast than benchmark BM2_m; only ARMA(4,4) has a forecast accuracy equal to the multivariate candidate model VAR_m2, whereas AR(6) and ARMA(6,6) provide a worse forecast than VAR_m2. ARMA(4,4) with expanding window therefore is chosen to provide the out-of-sample forecast. By the end of 2016, the inflation rate is forecast to be approximately 6% yoy, based on the lower line of the narrowest 30% of the interquartile range depicted from ARMA(4,4) with expanding window. (ii) The quarterly set initially includes twelve univariate models, then four of them are chosen for the forecasting evaluation process. AR(6) is finally chosen to produce the out-of-sample forecasts with the benchmark model BM2_q, applying expanding window. The forecast numbers converge to approximately 4% yoy by the end of 2016, based on median point forecasts of AR(6) and BM2_q.

Two sets of multivariate models are constructed simultaneously in this paper for forecasting: (i) the monthly set includes one VAR model with exogenous variables (VAR_m1) and one VAR model without exogenous variables (VAR_m2). They both have the same four endogenous variables with two lags of them. In the case of monthly models, VAR_m2 provides a statistically better forecast than the benchmark models and almost all other candidates, except ARMA(4,4), which has a statistically equal forecast accuracy to VAR_m2. By the end of 2016, VAR_m2 with expanding window produces the forecast inflation rate of 5% yoy, based on the lower line of the narrowest 30% of the interquartile range. (ii) The quarterly set includes two VAR models of five endogenous variables with exogenous (VAR_q1) and without exogenous (VAR_q2). However, some estimated coefficients of the two quarterly VAR models are statistically and/or economically insignificant. In addition, the pseudo out-of-sample RMSEs of the two models are not as small as univariate candidate AR(6) and benchmark BM2_q. Therefore, they are not chosen for producing out-of-sample inflation forecasts.

In summary, this paper can contribute to the forecasting work/practice of SBV in the following aspects: (i) computing pseudo out-of-sample RMSEs by simulations is the first and most important contribution in evaluating forecasting performance of models, which has not previously been provided in SBV. Applying the forecast-encompassing Diebold-Mariano test is another technique contributing to determining whether one model provides a statistically better forecast than others.

(ii) Quarterly models are preferable for long-term forecasts because the forecast numbers tend to converge, whereas monthly models over-forecast and are more useful for a shorter forecast horizon. (iii) The simple approach does not always produce worse forecasts. In the case of quarterly models, multivariate VAR models provide worse results than the univariate AR(6) and the benchmark model BM2_q. Therefore, some good candidate models must be operated simultaneously to produce inflation forecasts and evaluated frequently to obtain the best forecasting results.

Several topics can be pursued in future research. Because different models produce different forecasts, and only few of them have converging results, forecasting combinations will be one of the next directions for inflation forecasts; combinations have been implemented at several central banks and international organizations. Whereas the quarterly VAR with macro-economic variables does not exhibit good estimated results and good forecasting performance, an alternative VAR with the endogenous variables of inflation's components (such as core inflation rate, administrative price, energy price, and raw food price) should be a good choice for forecasting purposes. Additionally, factor models, which are constructed from a set of dozens or hundreds of macroeconomic and financial variables by many developed central banks, should become another type of model that SBV must consider in its new plan for developing and enriching the forecasting toolkit. Even structural VARs do not contribute more efficient forecasting results than reduced-form VARs because both types of models have identical forecasting outcomes; however, they are very useful to determine the important factors that drive inflation in Vietnam. To conclude, regularly performing ex-post forecasting evaluation is indispensable to enhancing the efficiency of making inflation forecasts in the SBV.

REFERENCES

Atkeson, A. and Ohanian, L.E. (2001), “Are Phillips curves useful for Forecasting Inflation?”, Quarterly Review, Winter 2001, Federal Reserve Bank of Minneapolis, Vol. 25, No. 1, 2-11.

Bui Van Hai and Tran Thi Minh Trang (2015), "The Transmission Mechanism of Monetary Policy in Vietnam: a VAR approach", Graduate Institute of International and Development Studies' website, Working Paper No.IHEIDWP15-2015.

Camen, U. (2006), “Monetary Policy in Vietnam: The case of a transition country” in “Monetary Policy in Asia: approaches and implementation”, BIS paper, No. 31, 232-252.

Enders, W. (2013), "Applied Econometric Time Series", Textbook, Third Edition.

Gosselin, M.A. and Tkacz, G. (2001), "Evaluating Factor Models: An Application to Forecasting Inflation in Canada", Working paper No. 18, Bank of Canada.

Le Viet Hung and W. Pfau (2008), “VAR Analysis of the Monetary Transmission Mechanism in Vietnam”, Applied Econometrics and International Development, Vol.9-1.

Nguyen ThiThuyVinh and S. Fujita (2007), "The impact of Real Exchange Rate on Output and Inflation in Vietnam: a VAR approach", Kobe University's website, Discussion paper, No. 0625.

Panizza, U. (2011), Lecture on "Univariate Time Series", Graduate courseatGraduate Institute of International and Development Studies.

Panizza, U. (2013), Lecture on "Vector Autoregresive Models", Graduate course at Graduate Institute of International and Development Studies.

Rubio-Ramirez, J.F., Waggoner, D.F. and Zha, T. (2010), "Structural vector autoregression: Theory of Identification and Algorithms for Inference", Review of Economic Studies 77, 665-696.

Sims, C.A. (1980), "Macroeconomics and Reality", Econometrica 48, 1-48.

Sims, C.A. (1992), "Interpreting the macroeconomic time series facts: The Effects of Monetary Policy", European Economic Review, Vol. 36, Issue 5, 975-1000.

Stock, J.H. and Watson, M.W. (1999b), “Forecasting Inflation”, Journal of Monetary Economics 44: 293-335.

Stock, J.H. and Watson, M.W. (2008), “Phillips curve inflation forecasts”, Working paper No. 14322, National Bureau Economic Research.

APPENDIX

Table 1. Monthly and Quarterly Data

Variables	Legends and Unit	Frequency	Source
Level series			
inf	Inflation rate, % yoy	Monthly, Quarterly	General Statistics Office of Vietnam
rrs	Real retail sales, in billion VND	Monthly	GSO and author's calculation
ry_g	Real GDP growth rate, % yoy	Quarterly	General Statistics Office of Vietnam
cred	Credit to (or Claims on) the economy, in billion VND	Quarterly	Thomson Reuters
exr	Nominal exchange rate (USD/VND)	Monthly, Quarterly	Vietcombank
ir	Interest rate, annual percentage (%/year)	Monthly, Quarterly	International Financial Statistics - IMF
op	West Texas Intermediate (WTI) oil price, USD/barrel	Monthly, Quarterly	Federal Reserve Bank of St. Louis
ifed	Effective Fed funds rate, annual percentage (%/year)	Monthly, Quarterly	Federal Reserve Bank of St. Louis
Transforming into stationary series			
inf	Monthly inflation rate series is stationary one at level		
d_inf	Take the first difference of the quarterly inflation rate series		
dl_rrs	Take the log and then the first difference of the real retail sales series		
dry_g	Take the first difference of the real GDP growth rate		
dl_cred	Take the log and then the first difference of the credit to the economy series		
dl_exr	Take the log and then the first difference of the exchange rate series		
d_ir	Take the first difference of the interest rate series		
dl_op	Take the log and then the first difference of the WTI oil price series		
ifed	Stationary series at level		

Table 2. Unit Root and Stationarity Tests on the Monthly Data Set

	p-values					
	H_0: Series has a unit root			H_0: Series is stationary		
Variables	ADF Tests			KPSS Test		**Conclusion**
	None	Intercept	Intercept and trend	Intercept	Intercept and trend	
inf						
Full sample (2000M1 - 2015M6)	0.1930	0.0440	0.3080	0.01<p<0.05	p<0.01	stationary
Post_WTO&cris (2008M1 - 2012M4)	0.2620	0.2370	0.4640	p>0.10	0.01<p<0.05	stationary
rrs	0.8883	0.7099	0.0021	p<0.01	0.01<p<0.05	non-stationary
dl_rrs	0.0000	0.0000	0.0000	p>0.10		stationary
exr	0.8049	0.3407	0.0000	p<0.01	p<0.01	non-stationary
dl_exr	0.0000	0.0000	0.0000			stationary
ir	0.4281	0.1329	0.3839	0.01<p<0.05	0.01<p<0.05	non-stationary
d_ir	0.0000	0.0000	0.0000	p>0.10	p>0.10	stationary
op	0.4232	0.1912	0.1118	p<0.01	0.01<p<0.05	non-stationary
dl_op	0.0000	0.0000	0.0000	p>0.10	p>0.10	stationary
ifed	0.0208	0.2246	0.5335			stationary

Table 3. Unit Root and Stationarity Tests on the Quarterly Data Set

Variables	*p-values*					Conclusion
	H_0: Series has a unit root			H_0: Series is stationary		
	ADF Tests			KPSS Test		
	None	Intercept	Intercept and trend	Intercept	Intercept and trend	
inf	0.4242	0.1127	0.4252	0.05<p<0.10	0.01<p<0.05	non-stationary
d_inf	0.0000	0.0003	0.0008	p>0.10	p>0.10	stationary
ry	0.9826	1.0000	0.9444	p<0.01	0.01<p<0.05	non-stationary
dry_g	0.0000	0.0030	0.0050			stationary
cred	0.9434	0.9976	0.8726			non-stationary
dl_cred	0.3809	0.2365	0.1941	p=0.05	0.01<p<0.05	stationary
exr	0.9829	0.8973	0.2856	p<0.01	0.01<p<0.05	non-stationary
dl_exr	0.0000	0.0000	0.0000	p>0.10		stationary
ir	0.5036	0.3594	0.7834	0.05<p<0.10	0.01<p<0.05	non-stationary
d_ir	0.0000	0.0000	0.0000	p>0.10	p>0.10	stationary
ifed	0.0050	0.0170	0.0190			stationary
op	0.609	0.4434	0.0197	p<0.01	0.01<p<0.05	non-stationary
dl_op	0.0000	0.0000	0.0000	p>0.10	p>0.10	stationary

Table 4. Suggested Lag Length Criteria for Monthly VARs (VAR_m1 and VAR_m2 top-down, respectively)

VAR Lag Order Selection Criteria
Endogenous variables: INF, DL_RRS, DL_EXR, D_IR
Exogenous variables: C, DL_OP(-1), IFED(-1)
Sample: 2000M01 to 2015M06
Included observations: 177

Lag	LogL	LR	FPE	AIC	SC	HQ
0	29.18597	NA	9.68e-06	-0.194192	0.021140	-0.106861
1	425.5450	761.3676	1.32e-07	-4.492034	-3.989592	-4.288263
2	493.4928	127.4501	7.32e-08	-5.079014	-4.289463*	-4.758803*
3	510.9136	31.88896	7.21e-08*	-5.095069*	-4.018408	-4.658417
4	520.8555	17.74936	7.74e-08	-5.026616	-3.662845	-4.473523
5	536.8188	27.77802*	7.76e-08	-5.026202	-3.375321	-4.356669
6	545.5313	14.76690	8.46e-08	-4.943857	-3.005867	-4.157883
7	558.1871	20.87844	8.83e-08	-4.906069	-2.680969	-4.003655
8	566.7674	13.76735	9.66e-08	-4.822231	-2.310022	-3.803377

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at the 5% level)
FPE: Final prediction error
AIC: Akaike's information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

VAR Lag Order Selection Criteria
Endogenous variables: INF, DL_RRS, DL_EXR, D_IR
Exogenous variables: C
Sample: 2000M01 to 2015M06
Included observations: 177

Lag	LogL	LR	FPE	AIC	SC	HQ
0	9.309102	NA	1.11e-05	-0.059990	0.011788	-0.030880
1	415.8578	790.1285	1.34e-07	-4.472969	-4.114082	-4.327419
2	483.0189	127.4922	7.53e-08	-5.051061	-4.405064*	-4.789069*
3	499.5493	30.63277	7.49e-08*	-5.057055*	-4.123949	-4.678623
4	509.8606	18.64176	7.99e-08	-4.992775	-3.772559	-4.497903
5	527.8131	31.64515*	7.84e-08	-5.014837	-3.507512	-4.403525
6	536.6377	15.15634	8.53e-08	-4.933759	-3.139324	-4.206006
7	549.4546	21.43388	8.88e-08	-4.897792	-2.816247	-4.053598
8	557.6471	13.33024	9.75e-08	-4.809572	-2.440918	-3.848938

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at the 5% level)
FPE: Final prediction error
AIC: Akaike's information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Table 5. Suggested Lag Length Criteria for Quarterly VARs (VAR_q1 and VAR_q2 top-down, respectively)

VAR Lag Order Selection Criteria
Endogenous variables: D_INF, DL_CRED, DRY_G, DL_EXR, D_IR
Exogenous variables: C, DL_OP, IFED
Sample: 2000Q1 to 2015Q2
Included observations: 57

Lag	LogL	LR	FPE	AIC	SC	HQ
0	29.09099	NA	4.20e-07	-0.494421	0.043225	-0.285473
1	84.45994	95.19575	1.46e-07	-1.559998	-0.126278*	-1.002805
2	132.3814	73.98408	6.74e-08	-2.364261	-0.034466	-1.458823*
3	162.5409	41.27084*	6.04e-08	-2.545295	0.680575	-1.291612
4	191.5604	34.61980	5.98e-08*	-2.686331*	1.435614	-1.084403

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at the 5% level)
FPE: Final prediction error
AIC: Akaike's information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

VAR Lag Order Selection Criteria
Endogenous variables: D_INF, DL_CRED, DRY_G, DL_EXR, D_IR
Exogenous variables: C
Sample: 2000Q1 to 2015Q2
Included observations: 57

Lag	LogL	LR	FPE	AIC	SC	HQ
0	6.438780	NA	6.54e-07	-0.050483	0.128732	0.019166
1	67.48790	109.2458	1.85e-07	-1.315365	-0.240075*	-0.897470
2	116.9896	79.89747	8.01e-08	-2.175073	-0.203708	-1.408934
3	152.2638	50.74535	5.89e-08	-2.535572	0.331869	-1.421187*
4	183.9444	40.01757*	5.17e-08*	-2.769978*	0.993538	-1.307348

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at the 5% level)
FPE: Final prediction error
AIC: Akaike's information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Table 6. Estimated Results of Monthly VARs (VAR_m1 and VAR_m2) for Inflation

Explanatory variables		Dependent variable: inf	
		VAR_m1	VAR_m2
Lags of endogenous variables	inf(-1)	1.69 [27.31]	1.68 [26.58]
	inf(-2)	-0.71 [-11.54]	-0.71 [-11.28]
	dl_rrs(-1)	-0.43 [-0.89]	-0.53 [-1.07]
	dl_rrs(-2)	-0.22 [-0.46]	-0.20 [-0.40]
	dl_exr(-1)	10.69 [1.24]	7.78 [0.89]
	dl_exr(-2)	1.79 [0.21]	5.57 [0.64]
	d_ir(-1)	-0.02 [-0.20]	0.08 [0.68]
	d_ir(-2)	0.07 [0.71]	0.08 [0.79]
Constant	c	0.15 [1.35]	0.20 [2.39]
Exogenous variables	dl_op(-1)	2.21 [3.42]	
	ifed(-1)	0.01 [0.25]	
Other statistics	R_squared	0.989	0.988
	F_statistic	1541.50	1823.70
	Log likelihood	-191.67	-197.69
	Akaike's AIC	2.21	2.26
	Schwarz SC	2.41	2.42

***Note**: t_statistic in [ ].*

Table 7. Estimated Results of Quarterly VARs (VAR_q1 and VAR_q2) for Inflation

Explanatory variables		Dependent variable: d_inf	
		VAR_q1	VAR_q2
Lags of endogenous variables	d_inf(-1)	0.77 [5.26]	0.46 [3.32]
	d_inf(-2)	-0.49 [-3.38]	
	dl_cred(-1)	13.98 [1.33]	30.23 [3.53]
	dl_cred(-2)	15.47 [1.40]	
	dry_g(-1)	0.004 [0.01]	-0.31 [-1.01]
	dry_g(-2)	0.03 [0.12]	
	dl_exr(-1)	-5.27 [-0.19]	-47.34 [-1.67]
	dl_exr(-2)	-25.23 [-0.89]	
	d_ir(-1)	0.11 [0.37]	0.51 [1.61]
	d_ir(-2)	0.39 [1.24]	
Constant	c	-1.30 [-2.06]	-1.38 [-2.28]
Exogenous variables	dl_op	3.62 [1.98]	
	ifed	-0.11 [-0.76]	
Other statistics	R_squared	0.732	0.549
	F_statistic	10.48	13.17
	Log likelihood	-108.62	-125.57
	Akaike's AIC	4.12	4.39
	Schwarz SC	4.58	4.59

***Note**: t_statistic in [ ].*

Figure 1. Graphs of Monthly Variables

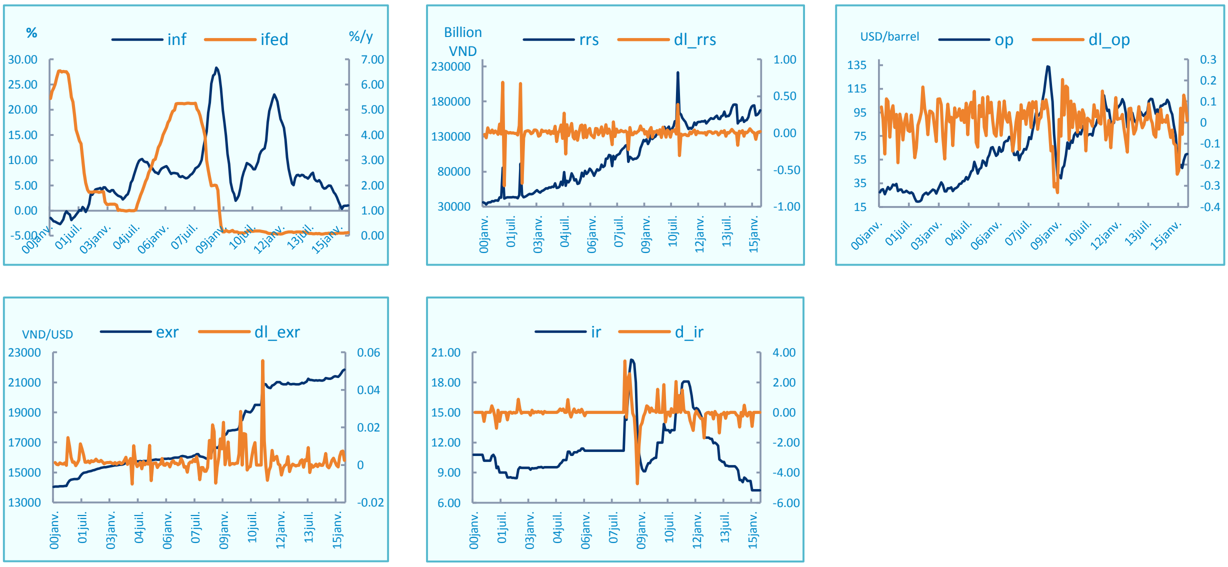


Figure 2. Graphs of Quarterly Variables

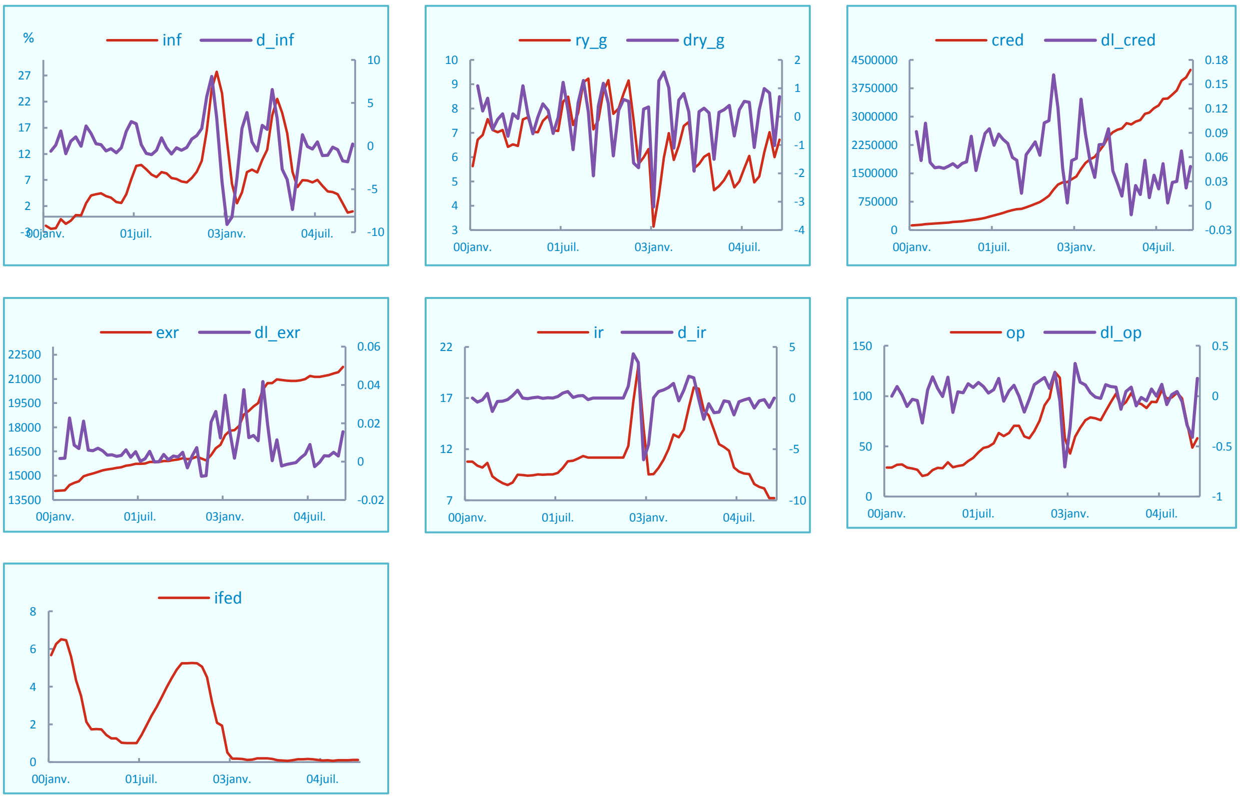
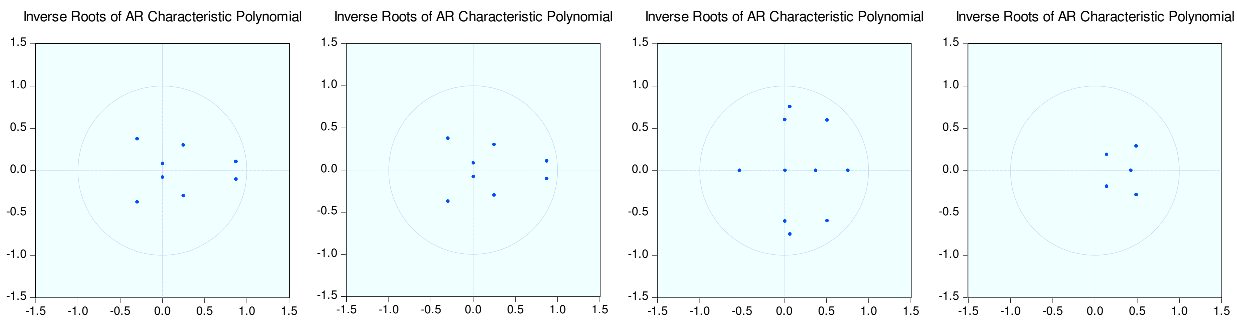


Figure 3. Graphs of the Inverse Root of the AR Characteristic Polynomial

VAR_m1 | VAR_m2 | VAR_q1 | VAR_q2



***Note**: With the same lag length, the graphs of the inverse root of the AR characteristic polynomial of monthly VARs are identical in both cases, regardless of whether exogenous variables are included.*

Box 1. Computing the RMSEs of in-sample and pseudo out-of-sample forecasts

Forecast errors (FE) are computed using the following formula: $FE_t = \hat{y}_t - y_t$ (1)

where, $\hat{y}_t$ is the predicted or forecasted value at time t given by the model and y_t is actual value of variable y at time t.

There are many measures for evaluating forecast accuracy (Bias, MSE, RMSE, SE, MAE, MAPE, etc.). This paper attempts to use the RMSE to assess the models' forecast accuracy. The smaller the model's RMSE, the more accurate its forecast compared with other models. The basic formula used to compute RMSE is

$$RMSE = \sqrt{\frac{1}{h}\sum_{t=1}^{h} FE_t^2} \quad (2)$$

where h is the forecast horizon or h periods before the end of the estimated sample.

Computing the RMSE (or whatever type of forecast accuracy measures) for an in-sample forecast is a type of forecasting exercise that uses T-h observations (or periods of time) to estimate different types of models and then uses the results to forecast the H observations that are not included in the estimated sample. The dataset has a sample size of T observations. To make an in-sample forecast, one just uses T-h first observations to estimate models (from t=1 to t=T-h) and then produces the forecast for the last h observations from T-h+1 to T. Formula (2) is used to compute the RMSE for each model and then compare those RMSEs.

To evaluate an out-of-sample forecast (the actual values are not observable at the current time), one can use the solution suggested by many econometricians, which is to use simulations to compute a pseudo out-of-sample forecast error. The strategy is that instead of accepting the last H observations for the in-sample forecast, a forecaster should allow (h+n) last observations to repeat the forecast (n+1) times. The pseudo out-of-sample RMSE of h-period-ahead forecasts for the inflation rate (inf) made over the period t_1 to $t_2$therefore is computed using the following formula[6]:

$$RMSE_{t_1,t_2} = \sqrt{\frac{1}{t_2-t_1+1}\sum_{t=t_1}^{t_2}(inf_{t+h}^{h} - inf_{t+h|t}^{h})^2} \quad (3)$$

where $inf_{t+h|t}^{h}$ is the pseudo out-of-sample forecast of inf_{t+h}^{h} made using data through date t and $t_2 - t_1 + 1 = n + 1$ is the number of times to repeat the forecast or the number of iterations that a forecaster estimates through time t and provides h-period future forecasts.

There are two methods to implement the strategy:

(i) Recursive window or expanding window: lengthening the estimated window by adding one observation (or period) at a time. Specifically, the first time a forecaster can use the sample from the first period to the $(t_1 - 1)^{th}$ period to estimate and then use the result to forecast the inflation rate for h-periods ahead from period t_1. The second time, the forecaster uses the sample from the first period to the t_1^{th} period and then forecasts the inflation rate for h-periods ahead from the period (t_1+1), repeating these steps until the estimated sample is as large as possible (from the first period to t_2).

(ii) Rolling window: maintaining the same estimated window length. The first time will be the same as in the recursive window. However, the second time, a forecaster uses the sample from the second period to the t_1^{th} period to estimate; the third time, the sample will be from the third period to the $(t_1+1)^{th}$ period, etc., repeating the work until the end period of the estimated sample is t_2. For any time of (t_2-t_1+1) iterations, the estimated sample size is always kept equal to the same length of (t_1-1) observations (periods).

Whereas the recursive or expanding window can assist the estimated model in becoming more informative by adding one period at a time, which is suitable and useful for developed economies or more stable systems, the rolling window may be helpful for developing or emerging systems because it is suitable for evaluating a system that has more frequent structural changes.

[6] This formula is based on Stock and Watson (2008).

Box 2. Description of the rolling window and expanding window forecasting evaluation strategies

The rolling window and expanding window forecasting strategies applied to monthly data are described in detail below:

(i) For both types of estimated window, whether rolling or expanding, these properties are applied uniformly: (1) the last observation of the first sample is April 2010, (2) the forecast horizon is twelve months ahead, and (3) the number of simulations is 50.

(ii) For the rolling window: a window of thirty-six months, equivalent to three years, is applied. This means that the sample size of the estimated window is a constant of thirty-six months for each iteration. Therefore, the first rolling window sample is from May 2007 to April 2010, the second rolling window sample is from June 2007 to May 2010, etc.

(iii) For the expanding window: the estimated window is expanded by one in each iteration. Therefore, the first expanding window is from the first observation of the initial sample (January 2000) to April 2010, the second expanding window is from January 2000 to May 2010, etc.

The following describes the expanding window forecasting strategy applied to quarterly data:

(i) For the expanding window, these properties are applied: (1) the last observation of the first sample is the second quarter of 2009, (2) the forecast horizon is four quarters ahead, and (3) the number of simulations is 20.

(ii) The estimated window is expanded by one quarter at a time. Therefore, the first expanding window is from the first observation of the initial sample (2000Q1) to 2009Q2, the second expanding window is from 2000Q1 to 2009Q3, etc.