4. IMPROVING PHILLIPS CURVE'S INFLATION FORECASTS UNDER MISSPECIFICATION

Mamdouh Abdelmoula M. ABDELSALAM[1]

Abstract

The Philips Curve (PC) is empirically criticized as falling short on many occasions in its predictability power of inflation due to an inherent deficiency in its specification features. This study is an attempt to improve the accuracy of Philips Curve forecasts. It considers various econometric specifications and estimation methods and different measures of the business cycle. In addition to the traditional New Keynesian open economy PC, we analyze some augmented versions with other information which incorporates the monetary variables such as the price gap. Additionally, we propose two different identifications for PC with time varying coefficients: the Time-Varying Coefficients with Random Walk (TVCR) coefficients and the Time Varying Coefficient (TVC). TVC allows us to confront directly specification biases and spurious relationships; this is usually the case for PC under the traditional estimation approaches. Moreover, we employ some static and dynamic forecast combination techniques. We find that PC with TVC provides the most accurate forecasts.

Keywords: forecasting inflation; Phillips Curve; misspecification; time-varying coefficients; model averaging, business cycles

JEL Classification: E32, E37, E58, E17, E31

1. Introduction

The Philips Curve was a well-known workhorse in economic literature over many decades. It relates the real activity side to the nominal side of the economy. Initially, Phillips (1958) and Samuelson and Solow (1960) analyzed the inference of the statistical relationship between unemployment and wages. However, this idea has been developed to numerous identifications based on outstanding assumptions to interpret the dynamics of inflation. For instance, in the New Keynesian (NK) PC framework, inflation depends on real marginal cost and expected inflation. Previous studies employed many proxies to express the marginal cost. These proxies include output gap, unemployment gap, and real unit labour cost or an index of the leading real variables (Stock and Watson, 2009; Marta, 2013).

Several specifications of PCs were applied to predict inflation, and the results were heterogeneous. While Stock and Watson (1999) showed that PC gives credible forecasts for

[1] *Department of Economics, University of Leicester, UK and Menoufia University, Egypt. E-mail: maa64@le.ac.uk, Mamdouh.economist711@gmail.com;mamdmohamed@yahoo.com.*

US inflation, Atkeson and Ohanian (2001) reported that naive models forecasts outperform the PC's forecasts. Additionally, Fisher *et al.* (2002) found that PC could be an effective tool for capturing the direction of changes in inflation, especially because its prediction were superior to the naive models in case of high volatile inflation. However, in the periods that witness changes in monetary regimes, PC can lose all or most of its predictive ability.

A serious matter, which arises in modeling the PC, is the existence of high levels of uncertainty either in the basic relation or the incorporated independent variables. There is significant evidence of existence of possible instability in the PC relationship (Dolado, Maria-Dolores, and Naveira, 2005; Canova, 2007; Musso *et al.*, 2009; Furhrer *et al.*, 2009). Consequently, some empirical work analyzed the variations in the dynamics of inflation and their potential effects on the PC's forecasting ability of inflation (Stock and Watson, 2007; Olivier *et al.*, 2015). These studies found that the prediction power of PC depends massively on its specification, the business cycle phase, and the analyzed sample (Stock and Watson, 2009; 2010). Furthermore, since the business cycle is an unobserved variable, there are many measures of the real activity inside the PC which are associated with large degree of uncertainty (Orphanides and Van Norden, 2005; Hilde & Brubakk & Jore, 2008; Marta and Harun, 2013). Moreover, recent studies confirm that inflation uncertainty has a considerable effect on other economic factors. For instance, Wright (2007) argues that uncertainty of inflation is a major factor in determining the level of term premium on nominal bonds and, consequently, the slope of the term structure of the economy.

This study provides a comprehensive analysis of the issues above for the South African inflation by covering a rich set of econometric identifications and explanatory variables. Additionally, most of the previous studies of PC forecasting performance were conducted mainly on the developed economies, and little efforts have been devoted to the less developed countries[2]. However, the latter countries are characterized by higher levels of uncertainty and instability in both policy regimes and economic structure relationships, respectively, which implies that the results of the developed countries cannot be generalized to these countries. Thus, current research analyzes the predictability precision of four different forms of PC using a univariate model which acts as a benchmark for all competitive models. Initially, the hybrid PC in a small open NK economy is assessed; this is common for inflation forecasting and policy analysis in both academia and central banks. Then, some augmented versions are analyzed with other information, such as the augmented PC with price gap to check the impact of this variable on improving the PC's forecast accuracy. Moreover, this study proposes two different specifications for PC with time-varying coefficients. Firstly, the Time-Varying Coefficients with Random Walk PC (TVCRPC) is introduced. Secondly, we introduce Time Varying Coefficient Philips Curve (TVCPC); this represents the most developed technique which overcomes the instability problem and approximates the actual specification for the real inflation function. Also, it allows for directly confronting the specification biases and spurious relationships, which is usually the case for PC under the traditional estimation approaches. For instance, Hall *et al.* (2009) proved that the traditional estimations approaches, such as Generalized Method of Moments (GMM), produce inconsistent estimations for coefficients inside NKPC.

The TVCPC is based on the approach or Swamy *et al.* (2010), who proposed the Time-Varying Coefficient (TVC) estimation method for estimating consistent parameters although we are uncertain about the true functional form. Additionally, the estimated coefficients are

[2] *Examples, by US data (e.g. Stock and Watson, 2009; Faust and Wright, 2012) and the euro area (e.g. Marta and Harun, 2013).*

consistent in cases of omitting some important variables, or if there are measurement errors in the included variables. Hall *et al.* (2014) suggested a more formal approach to select the split of the variables that can investigate its movements at least partially, and which can account for both the nonlinearity and omitted variables features in both the actual data and the econometric relationship, respectively. The latter approach will be called the Extended Time Varying Coefficient (ETVC). For many reasons, ETVC can be a very useful way for modeling PC. Firstly, in times of regime change, it allows us to imitate the unknown future function form for the dynamics of inflation. Secondly, the ETVC approach adds more information inside the PC equation; this depends on some transformations from the regressors in the state variables and not only the linear form. Also, we can augment different measures for the variable that we are uncertain about (for instance, we can add various measures for the unobserved business cycle indicator or different inflation indexes inside the state equation). Thirdly, the ETVC provides more flexibility in the identification process in comparison to traditional approaches; we can try different forms of considered variables to obtain the best identification of the inflation dynamics.

Given the possible instability in both the PC specification and many suggested model forms, many studies applied the forecast combination approach to improve the prediction accuracy (see for example, Stock and Watson, 2010; Clark and McCracken, 2010). In this research, we follow this literature by employing the simple Equal Weighting (EQ) method and the Dynamic Model Averaging (DMA) approach inspired by Raftery, Karny, and Ettler (2010). Furthermore, to overcome some drawbacks of the DMA, such as depending on many assumptions and constraints, we propose simple dynamic forecast combination approach which does not require any previous assumptions and can work more efficiently.

Therefore, this study contributes to the literature in several ways. The main contribution is to obtain a robust forecasting for inflation through combining different forecasts in the existence of many sources of uncertainty. Another contribution is to examine the usefulness of augmenting some monetary variables inside PC. In addition, we propose a non-traditional time-varying coefficient approach for PC which can be advantageous for forecasting inflation in circumstances of existence of instability, misspecification, or uncertainty problems. Moreover, in addition to model uncertainty, we take account of the effect of the uncertainty on the unobserved business cycle measure, which is one of the main determinants inside PC. Indeed, we consider a number of indicators which cover both statistical and structural methods for estimating the output gap or marginal cost indicator. This procedure is done to explore whether the specific business cycle measure can give robust forecasts for inflation in comparison to forecasts with different methods.

The main outcomes of current research can be summarized as follows: in terms of in-sample forecasts, the PC with Time-Varying Coefficients under Marginal Cost (TVC_MC) gap provides the most accurate historical forecasts. On the other hand, for the out-sample forecasts, the TVC with HP output gap (TVC_HP) produces the best forecasting accuracy over both the short term and medium term horizons. Additionally, the Bayesian Autoregressive Vector model for the Philips curve augmented with the price gap under MC measure (BVAR_MC) dominates all other models' forecasts over the long term. Moreover, the traditional NKPC[3] under different output gap gives the worst and least accurate forecasts. These results imply that depending on misspecified PC might lead to a significant loss in terms of forecasting power. Furthermore, there is no significant difference in the forecast performance of the TVC under different output gap measures. That is it to say that all of their

[3] *Without augmented variables or time varying coefficients.*

forecasts are close over all different horizons; they are better than the naive benchmark model. The latter result indicates that the underlying uncertainties do not have large effects in the case of TVC despite its significant impacts on the forecasting performance of the other approaches. This could be attributed to the advantages of this approach by including the other business cycle measures as driving sets in each case to overcome the noticeable misspecification problem. Thus, the paper is structured as follows. Section 2 presents the different employed models and Section 3 describes the combination methods. Section 4 is devoted to the discussion of the data, analysis of the main results, and forecasting evaluation criteria, while the conclusion is presented in Section 5.

2. Models

This section presents the different models that are used to predict future inflation. These models include the Random Walk (RW) model of Atkeson and Ohanian (2001). This model is considered the benchmark model as it is found that its forecasts are superior to other forecasts when the data is characterized by a high degree of persistence. Since it is a simple and well-known model, we do not present it in detail. Then, two models that are time-invariant are presented in details. These two models are the Hybrid Small Open NK Economy and the Augmented Philips Curve (APC). Finally, two models that assume time-varying coefficients are introduced, namely, the Time-Varying Coefficient Approach Philips Curve (TVC) and Time-Varying Philips Curve with Random Walk Updating (TVCR).

2.1. Hybrid Small Open NK Economy

The Small Open NK Economy Model analyses the PC relationship in depth. The NK hybrid PC is commonly estimated in both academia and central banks. The fundamental model is similar to the one developed by Gali and Gertlet (2002) and extended in 2008 to the open economy version of the IMF research team. This extension is done mainly for policy analysis and forecasting purposes and contains four main equations with other identities[4].

The output gap y_t depends on both backward and forward looking, which reflects the formation of the household expenditure habits, the lagged foreign output gap $y_{F,t-1}$[5], the lag of the real effective exchange rate gap z_{t-1}[6] and the lagged real interest rate gap r_{t-1}.

$$y_t = B_1^j y_{t-1} + B_2^j y_{t+1} - B_3^j r_{t-1} + B_4^j z_{t-1} + B_5^j y_{F,t-1} + \epsilon_t^y \qquad (1)$$

Concerning the PC, inflation is modelled as a function of current forecasting of future inflation π_{t+1}, lagged inflation π_{t-1}, the lag of output gap y_{t-1}, the changes in real exchange rate (ΔZ_t) and the disturbance term ϵ_t^{π}. Note that *j* refers to the identifications with the different business cycle measures. This study uses this notation which accounts for the uncertainty of the unobserved business cycle measure by using different measures and not as usual by employing one indicator.

$$\pi_t = \lambda_1^j \pi_{t+1} + (1\text{-}\lambda_1^j)\, \pi_{t-1} + \lambda_2^j y_{t-1} + \lambda_3^j \Delta Z_t + \epsilon_t^{j\pi} \qquad (2)$$

---

[4] *The full model with all identities is Arbatli & Moryoima (2011).*

[5] *The foreign output gap is calculated as the weighted average of the output gap of main trade partners of south Africa :* $Y_{F,t-1} = \sum_j W_j Y_{j,t-1}$, *where* W_j *are the weights of main trade partners for South Africa according to OCED indicators in the period of study.*

[6] *Real effective exchange rate is weighed by main trade partners.*

Real exchange rate takes the form of real interest rate gap differentials where its rise means the depreciation of the currency. Additionally, the expected exchange rate has forward and backward looking components, and the error part is interpreted as the financial assets risk premium. Further, the expected exchange rate, Z_t^e, is a function of the weighted average of both backward real exchange rate and expected consistent model level of it.

$$4 * (Z_t^e - Z_t) = \left(R_t - R_{f,t}\right) - (\bar{R}_t \text{-} \overline{R}_{f,t}) + \epsilon_t^{Z-Z^e} \quad (3)$$

$$Z_t^e = \phi Z_{t+1} + (1-\phi) Z_{t-1}$$

The monetary policy is represented in the form of the common Taylor Rule:

$$I_t = \gamma_1 I_{t-1} + (1-\gamma_1)[\bar{R}_t + \pi_{t+4} + \gamma_t\left(\pi_{t+4} - \pi_{t+4}^{targ}\right) + \gamma_3 y_t] + \epsilon_t^I \quad (4)$$

where: $\bar{R}_t$ is the real equilibrium interest rate and π_{t+4}^{targ} refers to the central bank inflation target policy.

2.2. The Augmented Philips Curve (APC)

As a result of the criticisms of PC in the literature, there were many trials to include other variables which might help to improve the identification and forecasting accuracy. In this paper, we analyze the advantages of adding the price gap inside the traditional PC model.

Our price gap is derived from the famous traditional quantity theory of money equation as:

$$M \times V = P \times Y \qquad (5)$$

where: M reflects the money stock, V is the income velocity of money, P refers to the price level, and Y is the level of real output. Equation (5) can be rewritten in logarithmic form (small letters) and by giving time indexation (t) as:

$$p_t = v_t + m_t - y_t \qquad (6)$$

Equation (6) is subject to the theoretical assumptions that p_t, v_t and y_t take an equilibrium path. Hence, the long run price quantity can be expressed in the form:

$$p_t^* = m_t + v_t^* - y_t^* \quad (7)$$

In equation (7) the equilibrium elements are marked by stars. Then, as shown in equation (8), the price gap is expressed as the deviation between the equilibrium price and current price:

$$(p_t^* - p_t) = (v_t^* - v_t) + (y_t - y_t^*) \quad (8)$$

Thus, the price gap incorporates the liquidity gap $(v_t^* - v_t)$ and the output gap $(y_t - y_t^*)$.

Given that the relationship between v_t and y_t can be written as: $v_t = y_t + p_t - m_t$, any increase (or decrease) in the current income y_t is associated by movement in v_t with the same amount and in the same direction (see, e.g. Robert Czudaj, 2011). As a result, inflation is a monetary phenomenon and it is committed completely to the changes in the stock of money. In addition, equation (8) can be reformulated as in equation (9):

$$-v_t = m_t - p_t - y_t \qquad (9)$$

This can be augmented with equation (8) to yield:

$$(p_t^* - p_t) = (m_t + v_t^* - p_t) - y_t^* \quad (10)$$

The last equation states that the price gap is determined by the difference between the real money stock $(m_t - p_t)$ modified by the equilibrium amount of velocity v_t^* and the equilibrium real output y_t^*.

To analyse the last augmented PC case, we follow Orphanides and van Norden (2005) and Garratt *et al.*, (2009) by considering PC forecasting models in the Vector Autoregressive

(VAR) form. The included variables are inflation inf_t, different output gaps y_t^j, real exchange rate Z_t and $pgap_t$ denotes the price gap which we derived above. Also, we include a supply shock which is the changes in oil price op_t.

2.3. Time Varying Coefficient Philips Curve (TVCPC)

As indicated in the introduction, TVC formulation depends on Swamy *et al.* (2010). This approach provides consistent parameters even if: the actual functional form of the considered relationship is unrecognized, the interested variables are measured with errors, or the model is identified with missing some important variables. The approach is based on Swamy and Mehta (1975) theorem as verified by Granger (2008). This theorem declares that any nonlinear relationship can be formalized in the form of a linear variables model but with time-varying coefficients. The significance of this theorem is that, even if the specification of the actual function of the considered relationship is unknown, we can still depend on the TVC method to model this relationship. Thus, the TVC approach might help to estimate the structural parameters inside the PC without the need to specify the true functional form for inflation.

Then, given that the true functional form of this relationship is incorrectly specified; as some important explanatory variables are omitted, and both the independent variable and the explanatory variables are measured with some errors we need to model this relationship by specific way to overcome all of these problems. More specifically, our main aim is to find consistent estimators for the coefficient inside the inflation dynamics equation. The application of the TVC techniques initially depends on some arbitrary assumptions regarding the choice of the set of drivers in the TVC model to explain at least part of coefficient movements (coefficient drivers) and to distinguish between these drivers' sub-sets. Hall *et al.* (2014) propose a more formal approach for choosing this split which can take account of the nonlinearity features in the actual data and the omitted variables inside the considered relationship. In more detail, they state that, when the considered true function is believed to be non-linear, some of the drivers should attain this nonlinearity. On the other hand, if the true function is thought to be linear, the sets of drivers should incorporate only the constant components. Furthermore, they note, also, that drivers should fulfill two conditions. Namely, they should be significant, and they should interpret a considerable proportion of the time-varying coefficient movements. They call the latter condition "predictive power and relevance".

We refer to the last extension as the Extended Time-Varying Coefficient (ETVC). We apply ETVC approach on the hybrid NKPC similar to that adopted by Christiano, Eichenbaum and Evan (2005)[7]:

$$\pi_t = \beta_{1t}^j E_t(\pi_{t+1}) + \beta_{2t}^j \pi_{t-1} + \beta_{3t}^j y_t^j \qquad (11)$$

where: β_{1t}^j, β_{2t}^j and β_{3t}^j are time-varying coefficients for the independent variables and j indicates, as usual, the different business cycle measures and $E_t(\pi_{t+1})$ is the expected inflation in next period. For the expected inflation, it is difficult to obtain an accurate expectation series for the whole required period in the case of the emerging countries. As suggested by Ball (2000) and Rumler and Maria (2010), such a prediction can be produced from a supplementary time series model. Following the above-mentioned literature, in order to generate their prediction, we apply an adaptive expectation hypothesis where agents'

[7] *Christiano, Eichenbaum and Evan (2005) expressed PC in the form of:* $\pi_t = wfE_t(\pi_{t+1}) + wb\pi_{t-1} + \lambda y_t^j$. *However for simplicity, in time varying models we will express coefficents as* β_{it}.

information set depends only on the realization of past inflations. Under these assumptions the explicit process of inflation expectation can be written in the form:

$\pi_t = \rho_1\pi_{t-1} + \rho_2\pi_{t-2} + \cdots + \rho_p\pi_{t-p} + \varepsilon_t$.

Then, $E_t(\pi_{t+1}) = \pi_t + \hat{\rho}_1\pi_t + \hat{\rho}_1\pi_{t-1} + \hat{\rho}_2\pi_{t-2} + \hat{\rho}_p\pi_{t-p+1}$. The appropriate lag is selected by employing the ARMA model.

Notations and Assumptions: as indicated before, our main aim is to overcome the misspecification of PC to improve the generated forecasts. In order to get that, we try to capture the consistent estimations for the coefficients of the explanatory variables $x^*_{1t}, \dots, x^*_{k-1,t}$ in (11) equation; or, in other words, to find the best formulation for the partial derivative of π_t with respect to x^*_{it} as the other variables are constant. In this case, k-1 may not be the complete set of π_t. Then, the estimations are exposed to some problems. Firstly, any particular functional form may imitate incorrectly the true function (specification biases) which may cause specification errors. Secondly, the available data on π_t and $x_{1t}, \dots, x_{k-1,t}$ may contain some errors. Assuming that there are *T* measurements on π_t , x_{1t} , ..., $x_{k-1,t}$, the true "unobserved" variables may be measured as: $\pi_t = \pi^*_t + v_{\pi t}$, $x_{it} = x^*_{it} + v_{xit}$, where: π^*_t is the available data for inflation, *i= 1, 2,...,k-1* and $v_{\pi t}$ is the measurement error of the dependent variable and $v_{xi,t}$ is the measurement error in the different explanatory variables. Thirdly, this relationship is subject to the biases of omitted variables as a result of not including the complete set of explanatory variables that can interpret the dynamic of inflation. For instance, if m_t refers to the total number of the explanatory variables for π^*_t in-context of the true functional form; it is usually unknown and it varies by time. In addition, the total number of the explanatory variables in the true function is larger than the explanatory variables included in the estimated relationship ($m_t > K_{t-1}$).

The latter assumption indicates that some explanatory variables are not included inside (11); the econometric specification of the true function. Suppose, $x^*_{g,t}$, g=k,…, m_t refers to the omitted variables, β^*_{it} is the coefficients of the included variables inside (11), β^*_{gt} is the coefficients of the omitted variables. However, the underlying equation may depend on linear variables it properly represents nonlinear relationship as the coefficients are time varying form. Furthermore, for $g = k, \dots, m_t$, let $x^*_{gt} = \lambda^*_{0gt} + \lambda^*_{1gt}x^*_{1t} + \cdots + \lambda^*_{K-1,gt}x^*_{K-1,t}$. As λ^*_{0gt} can be interpreted as the remaining effect of the omitted variables after removing the included explanatory variables (x^*_{1t},.., $x^*_{K-1,t}$) effect on x^*_{gt}. Thus, the time paths of $\lambda^*_{i,t}$s are determined according to the unknown true generation function.

To select driver sets that are used to get the unbiased coefficients, we have to follow some assumptions. Some of these assumptions are similar to those adopted in Hall *et al*. (2014). However, other assumptions than those in Hall *et al*. (2014) are needed to generate out-sample forecasts; the aim of the Hall *et al.* (2014) was to evaluate the impact of rating agencies decisions on the spread of the sovereign bonds but without discussing how to generate forecasts from TVC. The whole assumptions can be summarized as follows:

Assumption 1: Any particular econometric representation for the underlying relationship includes some errors. Therefore, correction is required to overcome these biases. TVC does the required correction by allowing coefficients to vary over time as a function of the selected drivers sets to encompass the true value of the coefficients in the main equation.

Assumption 2: Each coefficient is relevant to particular driver sets in linear relationship, as well as a random component error.

$$\beta_{it} = \gamma_{j0} + \sum_{d=0}^{P-1} \gamma_{id}\, z_{dt} + \epsilon_{it} \qquad (i = 1, \ldots., k-1) \qquad (12)$$

where: γ_{i0}, and γ_{id} are "fixed parameters", the z_{dt} are "coefficient drivers". In addition, given the incorporated coefficient drivers, each regressor and the coefficient of (12) is conditionally independent of each other. When the coefficient drivers are bunch of variables, these should explain a significant portion of movements in β_{it}. The coefficients of (12) satisfy the following form: $\beta_{it} = \alpha^*_{it} + \sum_{g=k}^{m_t} \alpha^*_{gt}\lambda^*_{igt} - \left(\alpha^*_{it} + \sum_{g=k}^{m_t} \alpha^*_{gt}\lambda^*_{igt} \right)\left(\frac{v_{it}}{X_{it}}\right)$ $(i = 1, \ldots, K-1)$, where: X_{it} is defined as a continuous variable, λ^*_{igt} is the time varying coefficient of the omitted variables X_{gt} on the involved variables X_{it}, α^*_{it} reflects the partial derivatives of the unknown true function in equation (11) and v_{it} reflects the measurement error might exist in the included variables X_{it} (More details in Dimitrios and Hall, 2016, PP.427:428).

Assumption 3: The set of constant component and drivers in (12) can be divided into three components: A_{1J}, A_{2J}, A_{3J}. The first component is related to the fluctuations which may be induced by the potential nonlinearity in the true functional form. The second component is correlated with specification bias due to excluding any relevant variables from the underlying relationship, where the last component is related to bias resulting from data measurement error. The assumption allows us to identify separately the bias-free, omitted-variables and measurement-error bias components of the coefficients which certainly have an effect on the prediction accuracy. Consequently, accounting for them might improve the forecasts.

Assumption 4: The vector of errors in (12) $\epsilon_t = \epsilon_{1t}, \ldots\ldots.., \epsilon_{K-1,t})'$ is based on the following stochastic form:

$$\epsilon_t = \varpi\, \epsilon_{t-1} + u_t \quad . \qquad (13)$$

where: ϖ is a (K-1) x (K-1) matrix whose all eigenvalues are less than unit. Also, u_t follows normal distribution with mean $E(u_t) = 0$ and variance $E(u_t u_t') = \{ \sigma_u^2 \Sigma_u \; if\; t = t';\, and\; 0\; if\; t \neq t')$.

In Matrix notations, the equation (11) under the previous assumptions 1-4 can be written as follows:

$$\pi_t = x_t'\, \beta_t \qquad (14)$$

where: $x_t = (x_{1t}, \ldots..., x_{K-1,t})'$ and $\beta_t = (\beta_{1t}, \ldots..., \beta_{K-1,t})'$.

Also, the vector formulation for the equation (13) is:

$$\beta_t = \Phi\, z_t + \epsilon_t \qquad (15)$$

where: $\Pi = [\beta_{id}]_{0\le i\le K-1, 0\le d\le p-1}$ is a $(K-1)\times P$ matrix as β_{id} is a matrix with a dimension of $(K-1)\times P$, whose elements are β_{id} as (i,d+1)–th element and $z_t = (z_{0t}, z_{1t}, \ldots\ldots, z_{P-1,t})'$. Then, after substituting (15) in (14) this yields:

$$\pi_t = (z_t' \otimes x_t')\phi + x_t'\epsilon_t \qquad (16)$$

where: $\otimes$ is a Kronecker product, and ϕ is a $(K-1)\times P$ vector which represents column stack of Φ.

Moreover, we can represent the observation in the equation (16) in the form: $\pi_t = X_z'\phi + D_x\epsilon$.

Where: $\pi_t = (\pi_1, \ldots\ldots, \pi_T)'$ is a T-dimensional vector, $X_z = (z_1 \otimes x_1, \ldots\ldots.., z_T \otimes x_T)'$ is a $T\times(K-1)P$ matrix, D_x is the diagonal of x_t' as it has a dimension of $(T\times(K-1)T)$, and $\epsilon_t = \epsilon_1', \ldots\ldots.., \epsilon_T')'$ is a $T\times(K-1)$.

Under assumptions 1-4, $E(\pi_t|X_z) = X_z\phi$ has a variance which is equal to $D_x\sigma_u^2 \Sigma_\epsilon D_x'$, where variance-covariance matrix is $\sigma_u^2 \Sigma_\epsilon$. (Proof in Swamy *et al*., p.11). Consequently, the vector ϕ is identified under the condition that X_z is a full rank matrix. This requires that $T > (K - 1)P$. In addition, the matrix $D_x\epsilon$ is given by $(\pi_t - X_z\phi)$, and conditional on ϕ is identified , the D_x matrix has a full rank. Then, "the Best Linear Unbiased Predictor (BLUP)" of $D_x\epsilon$ may be utilized to get consistent estimated values for $\Sigma_u, \Phi \text{ and } \sigma_u^2$.

Moreover, under assumptions from 1 to 4, the utilized explanatory variables are conditionally independent of their coefficients and the best linear unbiased predictor of $D_x\epsilon$ exists (the details of the proof in Swamy, Yaghi, Mehta and Chang 2007, p. 3387).

Assumption 5: In order to obtain forecasts for inflation from the equation number (11), some more assumption other than those of Hall *et al*. (2014) should be adopted; the aim of the last paper assumptions is to explore the relationship between the sovereign bond spread and rating agencies decisions. To generate forecasts we take one lead form from the equation number (11) and we assume the one step ahead forecast of any exogenous variable in the system X is $\hat{X}_{t+1} = AX_t$. Hence, the next period inflation forecast can be expressed only depending on the current variables. Then, this forecast for inflation in period t+1 can be used to generate the forecast for the next period t+2 and so on iteratively for the following period t+3,t+4,...,t+h. The final outcome of this forecasting process yields a h step ahead forecasts for future inflation upon the present value of the Time Varying New Keynesian Phillips Curve.

This model can be specified in state space form as the first equation for the observed interested variable and a state equation for each time varying coefficient (*More details about state space representation for time varying model are given in the Appendix*).

The estimation is based on five coefficient drivers which include: nominal interest rate on three month deposits (as an indicator for policy changes), real exchange rate (as an indicator of external sector fluctuations), wage inflation, and the other two business cycle measures for each estimation (as it was mentioned before, we have three measures of business cycle and each time we estimate the main equation by any one of them, we include the two others as coefficient drivers; this is in order to get the marginal information in all measures together).

2.4. Time Varying Coefficients with Random Walk Updating Phillips Curve (TVCRPC)

In this section, we present another proposed form of the time-varying PC as written in equation (17):

$$\pi_t = \beta_{1t}^j E_t(\pi_{t+1}) + \beta_{2t}^j \pi_{t-1} + \beta_{3t}^j y_t^j + \epsilon_t^j \qquad (17)$$

$$\beta_{it}^j = \beta_{it-1}^j + u_t$$

where: β_{1t}^j, β_{2t}^j and β_{3t}^j are time varying coefficients for the independent variables and j indicates the different business cycle measures and $E_t(\pi_{t+1})$ is the next period expected inflation, which we obtain by using the adaptive expectation approach as explained in the previous section. The β_{it} coefficients are modeled as a random walk process, and both ε_t *and* u_t are independent and follow normal distribution with zero mean and constant variances (for more details, please see Song *et al*. (2011)). Similar to the previous approach, the model is formulated in a state space form.

3. Methodologies of Forecast Combination

Given the numerous available model specifications, therefore, the paper applies forecast combination techniques to form a combined forecast, based on the assumption that we have k available forecasts $\hat{y}_{T+h,1}, \hat{y}_{T+h,2}, \ldots, \hat{y}_{T+h,k}$ that result from k different models. These models are utilized to generate a forecast of y_{T+h}, on the assumption that the combined forecast is a function of those individual k forecasts in addition to the vector of weights associated with each alternative forecast, $w_{i,T+h}$. Then, this combined relationship can be represented in the form: $y_{T+h} = g(\hat{y}_{T+h,1}, \hat{y}_{T+h,2}, \ldots, \hat{y}_{T+h,k}, w_{i,T+h})$. Additionally, if we assume that the forecast error is equal to $e_{T+h} = y_{T+h} - g(\hat{y}_{T+h,1}, \ldots, \hat{y}_{T+1,k}, \hat{w}_{T+h})$, then the optimum values of the weights can be computed by minimising the following loss function:

$$\min_{w_{T+h}} E[L(e_{T+h}(w_{T+h})[\hat{y}_{T+h,1}, \ldots, \hat{y}_{T+h,k}])] \qquad (18)$$

Thus, the vector $\hat{w}_{T+h}$ incorporates the optimal weights set which satisfy equation (18). In the following subsection, we present different methods for calculating the combination weights w_i for individual models.

3.1. Simple Static Forecasts Combination Schemes

According to Timmermann (2006), in simple forecast combination schemes there is no need to estimate any parameters or to estimate the variance-covariance matrix.

Equal Weight (EQ) which is regarded as the simplest method for calculating the combination weights as it is the mathematical average of all available individual forecasts. Despite its simplicity, many studies find that it works better than many complicated techniques for calculating combination weights.

$$w_i = 1/k \qquad (19)$$

Inverse Root Mean Square Forecast Error (IRMSFE): This combination scheme depends on the assumption that models, characterized by less forecasting errors, should attain higher weights. Hence, as shown in equation (20), calculating the combination weights depends on the inverse forecasting error for the available forecasting models.

$$w_i = \frac{(1/IRMSFE_i)}{\sum_{i=1}^{k}(1/IRMSFE_i)} \qquad (20)$$

3.2. Dynamic Model Averaging

Dynamic Model Averaging (DMA), was recently developed by Raftery *et al*., (2010). It combines forecasts from different models based on the predictive likelihood of each model as approximate to the past forecasting performance. The DMA allows for the weights associated with the different models to vary over time. Raftery *et al*., (2010) calculate the time-varying weight for each available model $w_{i,t}$ as $A_i \in \{1, \ldots, k\}$ based on the predictive likelihood associated with model A_i :

$$w_{i,t|t-1,i} = \frac{pr^{\alpha}_{t|t-1,A_i}}{\sum_{i=1}^{k} pr^{\alpha}_{t|t-1,A_i}} \qquad (21)$$

As a consequence, the different model weights at each current period t are conditional on this model performance in the recent past periods. In this respect, length is the "recent past" and is determined by the α value which is referred to as the “forgetting factor”. We depend that in case of quarterly data the last 5 years performance receives around 80% in the weighting criteria. Then, once the next realization y_t becomes available $w_{t|s,i}$ can be updated by recalculating the predictive density depending on the conditional predictive density and the new available information ψ_s :

$$w_{i,t|s,i} = \frac{pr_{t|t-1,A_i}^{\alpha}\ f_{A_i}(y_t|\psi_s)}{\sum_{i=1}^{k} pr_{t|t-1,A_i}^{\alpha}\ f_{A_i}(y_t|\psi_s)} \qquad (22)$$

Hence, the weight is defined in terms of the probability that the model Ai engages at time t conditional on the information set up to point s. Equations (21) and (22) are re-estimated successively for each t, with initial assuming equal weighting scheme for the individual models at the first period t=1. Although the DMA approach provides a flexible weighting average scheme which may be very useful in many cases with time series data, it suffers from one main drawback since it depends on many assumptions and constraints to obtain the weights. In addition, its results are very sensitive to the previous assumptions which the researcher should determine in advance. These may cause very poor results in comparison to the individual models or other combination techniques.

3.3. Dynamic Inverse Mean Square Forecast Error (DIRMSFE)

To
overcome the DMA drawbacks, we propose calculating the dynamic form of the simple RootMean Square Forecasting Error (RMSFE) by considering this value for each period not for the whole sample. Therefore, the optimal weight can be computed as follows:

$$w_{it} = \frac{(1/RMSFE_{t,i})}{\sum_{i=1}^{k}(1/RMSFE_{t,i})} \qquad (23)$$

4. Empirical Results

4.1. The Data

The paper employed the South Africa quarterly data covering the period from 1975Q1 to 2014Q2, depending on the sources described in detail in Table 1. Inflation data is computed as the quarterly change in the logarithm of the CPI. Inside all models, we use data from 1975Q1 to 2005Q4 to fit these models and to obtain the in-sample forecasts, while we keep the observations of the period 2006Q1 to 2014Q2 to evaluate the out-sample forecasts.

Table 1

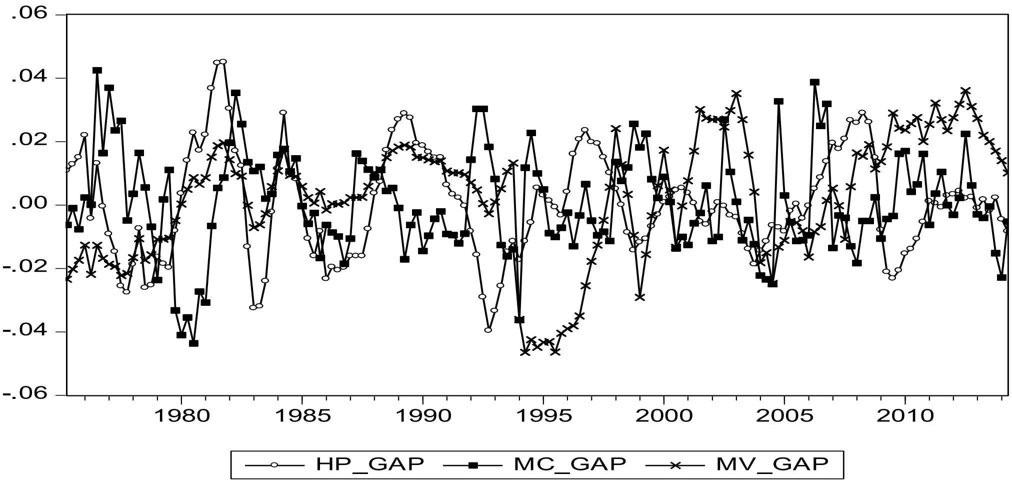
Data Sources

Variable	Source
CPI (LCPI)	IFS
Real GDP (LGDP)	SARB-South Africa
Nominal Interest Rate(I)	IFS
Unemployment Rate(UNR)	SARB
Real exchange index(LZ)	Bloomberg
Nominal exchange (S)	Bloomberg
Employee wages	SARB
Employee number	SARB
Current GDP (CGDP)	SARB
Foreign CPI (LCPI)for the 6 main trade partners.	IFS
Foreign Real GDP(LGDP)by US Dollar or the 6 main trade partners.	OCED
Foreign Nominal interest rate (FI) for the 6 main trade partners.	IFS

The estimation of output gap is used in different areas, such as non-inflationary steady state growth. It is also used to evaluate macroeconomic policies and to decide on what conditions the aggregate demand policies should be either contractionary or expansionary. Given that business cycle is an unobserved measure, it has to be estimated. Not surprisingly, different assumptions about potential output and different estimation methodologies yield different estimates of the output gap. Hence, it is associated with a large degree of uncertainty (Orphanides and Norden, 2005; Camba-M´endez and Rodriguez-Palenzuela, 2003). In contrast to the majority of empirical studies, which depend only on one business cycle indicator, we use different indicators of the business cycle. The common methods to estimate either the output gap or business cycle indicator can be classified into statistical and structural methods. Other studies used some approximations like the labor marginal cost. However, this paper applies different types, since we estimate univariate statistical (Hedrick–Prescott) and unobserved multivariate as a structural approach and we estimate, also, the real marginal labor cost. Figure 1 shows the different estimated business cycle measures, while the calculations of these measures are detailed in Appendix.

Figure 1

Different Business Cycle Measures (1975Q2 to 2014Q1)



4.2. Estimation and Forecasting Methodology

In this subsection, we describe each of the estimation and forecasting methodologies which we apply to the employed models. Firstly, we illustrate the case of constant coefficient models and, then, we move on to the time varying ones.

Constant coefficient models: We use the Bayesian approach for both the structural model and the autoregressive approach. The scheme of the structural NK model equations is expressed in the form of a state space model. This model is solved based on Blanchard and Kahn technique and the Metropolis-Hastings algorithm is utilized to derive the posterior distribution via simulation. We compute 100,000 draws from the Metropolis-Hastings algorithm after 30,000 burn-in iterations to calibrate the scale such that the outcome is an acceptance ratio of 0.3[(8)]. Then, we calculate the mean of the considered parameters from the posterior distribution of the included parameters. Then, to calculate the required forecasts, we calculate density forecasts and, then, the mean forecasts are computed as the average of the driven density forecasts. The posterior distribution is used to generate a large number of drawn values; 25,000, and then the forecast density is given by the ordered set of forecast draws. Finally, the point forecast is given as the mean of the derived forecast density.

Time-varying parameter estimation and forecasting: The two Time Varying Philips curve models can be presented in the form of the state space form. They can be solved by using predictive Kalman Filter approach as explained by Harvey (1987). To generate forecasts, we take one lead form from the equation (11), and we assume the one step ahead forecast of any exogenous variable in the system X is $\hat{X}_{t+1} = AX_t$. Hence, the next period inflation forecast can be expressed only depending on the current variables. Then, this forecast for inflation in period t+1 can be used to generate the forecast for the next period t+2 and so on iteratively for the following period t+3,t+4,...,t+h. Figure 2 shows the estimation results of PC with TVC and TVCR for inflation with the different business cycle measures.

Figure 2

Time-varying Coefficients Values by One Step Ahead Kalman Filter Prediction

Notes: Each row represents one of our models where each column for particular coefficient under all models.

Results are estimated by utilizing one step ahead prediction as explained in the previous sub-section for the data from 1978Q1 to 2005Q4. Consequently, we leave three years to initialize the models.

---

[8] *Estimations are implemented inside the Dynare software package (more details in Juillard, 1996).*

4.3. Forecasting Evaluation

Table 2 displays the different measures used to assess the predictive power of the employed models and different combination approaches. The forecast error statistics RMSE and MSE depend on the scale of the dependent variable. Therefore, it is a relative measure to compare forecasts across different models. According to this criterion, the smaller the error, the better is the forecasting ability of the related model. With respect to Theil inequality coefficient, it must lie between zero and one, where zero is a sign of a perfect fit.

Table 2

Different Criterions of Predictive Power

Root Mean square error	$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2}$
Mean square error	$\mathrm{MSE} = \frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2$
Mean Absolute Error	$\mathrm{MAE} = \frac{1}{N}\sum_{t=T+1}^{T+N}\lvert\hat{\pi}_t - \pi_t\rvert$
Theil inequality coefficient	$\mathrm{TIC} = \frac{\sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\pi}_t - \pi_t)^2}}{\sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}\hat{\pi}_t^2} + \sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}\pi_t^2}}$

Notes: π_t , $\hat{\pi}_t$ is the actual inflation values and the forecasted ones, respectively.

4.3.1. In-sample Forecasting Performance of the Estimated Models

However, the in-sample forecasts are not the main aim for any forecaster. On the other hand, they provide an enormous interest in showing the history predictive power of each particular model with changes in either breaks or historical events. Consequently, to obtain an in-depth analysis of how the alternative models' forecasting accuracy varies with historical changes, we generate in-sample forecasts over the period 1980Q1 to 2005Q4. Table 3 summarizes the main evaluation criterions for the in-sample forecasts for the different models. As shown in Table 3, PC with time varying coefficients under marginal cost gap provides the most accurate historical in-sample forecasts.

4.3.2. Out-sample Forecasting Performance

To assess the forecasting performance of different models over different horizons, we create predictions for one-step, four-step and eight-step ahead forecasts to represent short term, medium term and long term, namely over the period from 2006Q1 to 2014Q2. Table 4 summarizes the results of the four different criterions which we use to assess the predictive power of the competitive models. The results show that over both the short term and the medium term, the time-varying coefficient with HP output gap (TVC_HP) provides the best forecasting accuracy while, in the long term, the augmented PC (BVAR_MC) dominates all other models' forecasts. Moreover, the PC without augmented variables or time-varying coefficients (the misspecified PC) under the different measures of output gap provides the worst forecasting over all horizons according to all the criterions. This implies that depending on the misspecified PC there may be a significant loss in terms of forecasting power. Also, we may see that there is no significant difference in the forecast performance of the different time-varying models. This indicates that the underlined uncertainties do not have a huge effect in the case of time-varying coefficients compared to models with constant coefficients whose forecasting performance varies significantly with changes in the business cycle indicator. The last result is expected because other output gap measures are used as driving sets in each of the cases above.

Table 3

In-sample Forecasting Performance for One Step Ahead of the Models

	RMSE	MSE	MAE	TIC
SC_HP	0.01105[10]	0.000122[10]	0.008161[10]	0.212[11]
SC _MV	0.01158[13]	0.000134[13]	0.008447[11]	0.222[13]
SC _MC	0.01138[12]	0.000129[12]	0.008817[12]	0.218[12]
TVC_HP	0.00907[5]	0.0000823[5]	0.00661[3]	0.162[3]
TVC_MV	0.00850[2]	0.0000722[2]	0.006453[2]	0.152[2]
TVC_MC	0.00837[1]	0.0000701[1]	0.00615[1]	0.150[1]
TVCR_HP	0.0099[8]	0.0000986[8]	0.00784[8]	0.180[9]
TVCR_MV	0.01008[9]	0.000102[9]	0.00795[9]	0.179[8]
TVCR_MC	0.01126[11]	0.000127[11]	0.00910[13]	0.203[10]
BVAR_HP	0.00979[6]	0.0000958[6]	0.00783[7]	0.177[7]
BVAR_MV	0.0090[3]	0.00008118[3]	0.007121[5]	0.166 [4]
BVAR_MC	0.009055[4]	0.00008199[4]	0.007185[6]	0.167[5]
RW	0.009798[7]	0.000096[7]	0.0071[4]	0.172[6]

Notes: The numbers in the square brackets indicate rankings of the models where [1] indicates the best models according to the corresponding measure. SC indicates the structural model, TVC is sophisticated time varying coefficients method, TVRC is time varying with random walk updating, and BVAR represents the augmented PC. The model, shown in bold, is the best model.

Table 4

Evaluation of Out-of-Sample Forecasts Horizons One, Four and Eight Step Ahead

	One Step ahead H=1				Four Step ahead H=4			
	RMSFE	MSFE	MAFE	TIC	RMSFE	MSFE	MAFE	TIC
SC_HP	0.00815 [1]	0.0000665 [11]	0.00649 [11]	0.230 [9]	0.05394 [13]	0.0029 [13]	0.0422 [13]	0.74 [12]
SC _MV	0.00872 [13]	0.000076 [13]	0.00668 [12]	0.252 [12]	0.0511 [12]	0.00261 [12]	0.0394 [12]	0.774 [13]
SC _MC	0.00839 [12]	0.0000705 [12]	0.00636 [10]	0.248 [11]	0.0344 [11]	0.00118 [11]	0.0290 [11]	0.642 [11]
TVC_HP	0.00490 [1]	0.000024 [1]	0.00371 [2]	0.138 [1]	0.00581 [1]	0.0000338 [1]	0.00420 [1]	0.154 [1]
TVC_MV	0.00490 [2]	0.0000241 [2]	0.00366 [1]	0.139 [2]	0.00612 [3]	0.0000375 [3]	0.00478 [3]	0.182 [3]
TVC_MC	0.00519 [3]	0.000027 [3]	0.00409 [3]	0.152 [3]	0.00601 [2]	0.000036 [2]	0.00477 [2]	0.175 [2]
TVR_HP	0.00612 [6]	0.0000376 [6]	0.00453 [6]	0.189 [6]	0.00873 [8]	0.000076 [8]	0.00676 [8]	0.268 [8]
TVCR_MV	0.00812 [10]	0.000066 [10]	0.00615 [9]	0.254 [13]	0.00985 [9]	0.0000969 [9]	0.0081 [10]	0.314 [10]
TVCR_MC	0.00765 [8]	0.000058 [8]	0.00556 [8]	0.234 [10]	0.01037 [10]	0.000107 [10]	0.00797 [9]	0.294 [9]

	One Step ahead H=1				Four Step ahead H=4			
	RMSFE	MSFE	MAFE	TIC	RMSFE	MSFE	MAFE	TIC
BVAR_HP	0.005747 [5]	0.000033 [5]	0.00444 [5]	0.168 [5]	0.00643 [4]	0.0000414 [4]	0.0052 [4]	0.185 [4]
BVAR_MV	0.008007 [9]	0.000064 [9]	0.00699 [13]	0.216 [8]	0.00758 [6]	0.000057 [6]	0.0063 [6]	0.207 [6]
BVAR_MC	0.005588 [4]	0.0000312 [4]	0.00424 [4]	0.164 [4]	0.00689 [5]	0.0000475 [5]	0.0053 [5]	0.199 [5]
RW	0.006489 [7]	0.0000421 [7]	0.00490 [7]	0.191 [7]	0.00824 [7]	0.0000679 [7]	0.0064 [7]	0.239 [7]
	Eight Step ahead H=8							
	RMSE	MSE	MAE	TIC				
SC_HP	0.190 [13]	0.0363 [13]	0.150 [13]	0.916 [13]				
SC _MV	0.171 [12]	0.0295 [12]	0.1366 [12]	0.912 [12]				
SC _MC	0.115 [11]	0.01323 [11]	0.0898 [11]	0.857 [11]				
TVC_HP	0.01281 [8]	0.000164 [8]	0.00947 [7]	0.307 [5]				
TVC_MV	0.03035 [9]	0.000921 [9]	0.02768 [9]	0.543 [9]				
TVC_MC	0.00789 [3]	0.000062 [3]	0.00681 [4]	0.217 [4]				
TVCR_HP	0.01078 [5]	0.000116 [5]	0.00793 [5]	0.326 [7]				
TVCR_MV	0.01210 [7]	0.000146 [7]	0.00975 [8]	0.454 [8]				
TVCR_MC	0.0598 [10]	0.003584 [10]	0.0373 [10]	0.696 [10]				
BVAR_HP	0.00745 [2]	0.000055 [2]	0.00616 [2]	0.212 [2]				
BVAR_MV	0.00795 [4]	0.000063 [4]	0.00680 [3]	0.216 [3]				
BVAR_MC	0.00695 [1]	0.000048 [1]	0.0056 [1]	0.200 [1]				
RW	0.01128 [6]	0.000127 [6]	0.00875 [6]	0.324 [6]				

Notes: The numbers in the square brackets indicate rankings of the models where [1] indicates the best models according to the corresponding measure. SC indicates the structural model, TVC is sophisticated time varying coefficients method, TVRC is time varying with random walk updating, and BVAR represents the augmented PC. The model, shown in bold, is the best model.

4.4. Combination Results

The main aim of any forecast combination process is to improve the accuracy of the individual forecasts. In our analysis, we compare the forecasting performance of the different forecasting combination schemes and the best model. Table 5 shows the forecasting power comparisons for the different combination methods and the best model in terms of forecasting performance. Table 6 shows individual weights under the static IRMSE combination method over the different horizons, where the individual weights under the two dynamic combination methods for the first forecasting horizon are depicted in figures 3 and 4. Table 5 shows that the dynamic approaches, DMA and DIRMSE combination, dominate the best model forecasting and both the EQ and MSFE combination approaches. The two simple static combination schemes EQ and MSFE fail to beat the best model accuracy over the different horizons. Moreover, DIRMSE outperforms the DMA combination approach over all horizons.

Figure 3. Individual Models Weights According to DIRMSFE for H=1

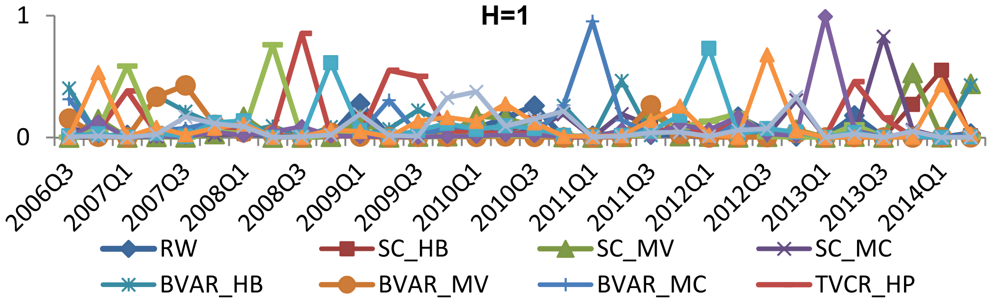


Figure 4. Posterior Probabilities of Individual Models in DMA for H=1

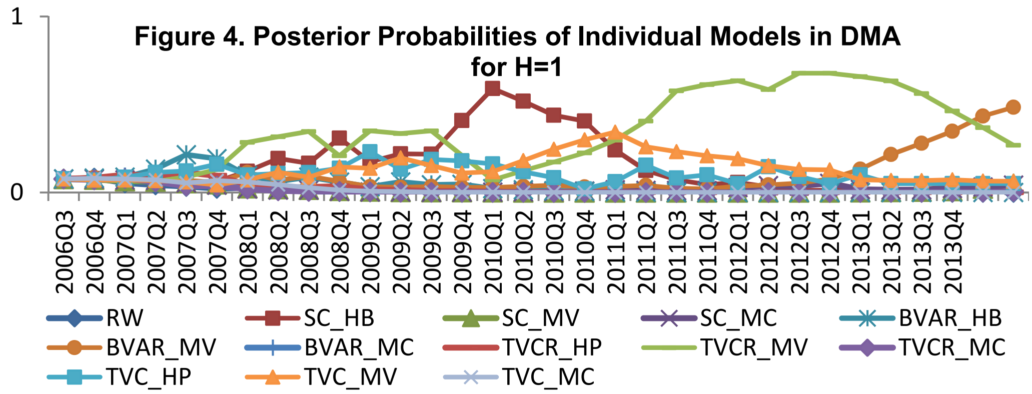


Table 5

Forecasts Power of Combinations over Different Horizons

	H=1		H=4		H=8	
	RMSFE	MSFE	RMSFE	MSFE	RMSFE	MSFE
BM	0.00490[3]	0.0000240[3]	0.00581[3]	0.0000338[3]	0.00695[2]	0.000048[2]
EQ	0.00562[5]	0.0000317[5]	0.0123[5]	0.000151[5]	0.0350[5]	0.00122[5]
IRMSE	0.00539[4]	0.00000292[4]	0.005811[4]	0.0000338[4]	0.0089[4]	0.000078[4]
DMA	0.00477[2]	0.0000228[2]	0.00577[2]	0.0000338[2]	0.0073[3]	0.000053[3]
DIRMSE	0.00349[1]	0.0000122 [1]	0.00430 [1]	0.0000186 [1]	0.0042[1]	0.0042[1]

Table 6

Individual Weights under IRMSE Combination Method over Different Horizons

	H=1	H=4	H=8
SC_HP	0.061151	0.012917	0.005566
SC _MV	0.057206	0.013614	0.006184
SC _MC	0.059379	0.020221	0.00923
TVC_HP	0.101634	0.119934	0.082853
TVC_MV	0.101593	0.113821	0.034983
TVC_MC	0.095961	0.115769	0.13457
TVCR_HP	0.081362	0.079838	0.098501
TVCR_MV	0.061374	0.070782	0.087704
TVCR_MC	0.065186	0.067195	0.017731
VAR_HP	0.08677	0.108302	0.142411
VAR_MV	0.062285	0.091944	0.133508
VAR_MC	0.089246	0.10109	0.152659
RW	0.076854	0.084573	0.0941

5. Conclusion and Future Research

This paper aims to improve the prediction accuracy of inflation from PC to identify the robust one. We consider various econometric specifications, estimation methods, and different measures of the business cycle to forecast inflation. In addition to the traditional PC, we analyze some augmented versions with other information, such as the augmented version with monetary variables. Then we propose two time-varying approaches, the first one is Time-Varying Coefficients PC with random walk updating (TVCR), while the second one is a more sophisticated approach which can help to reduce the level of uncertainty associated with PC forecasts. The last proposed approach, which we call Time Varying Coefficient PC (TVCPC), represents the most developed techniques that overcome the instability problem and also approximates the actual specification for the real inflation function. Given the large number of possible model specifications, we also evaluate the forecast performance of different forecast combination techniques.

The results indicate that, in case of in-sample forecasts, the PC with Time-Varying Coefficients under Marginal Cost gap (TVC_MC) gives the most accurate historical forecasts. However, in the case of the out sample forecasts, over both the short and the medium terms, the Time-Varying Coefficient with HP output gap (TVC_HP) produces the best forecasting accuracy. Additionally, the BVAR model for the PC augmented with the price gap under MC output gap (BVAR_MC) dominates all other models' forecasts over the long term. Moreover, according to all evaluation criteria, the NK model without either augmented variables or time-varying coefficients, which is regarded as the misspecified PC, provides the worst forecasts under the different output gap indicators. That is to say that, in both in-sample and out-sample, its performance is less accurate over different horizons even when compared to the naive model as a benchmark model. This implies that using misspecified PC in forecasting leads to misleading results. In addition, there is no significant difference in the forecast performance of the sophisticated Time-Varying Coefficients model under the different output gap measures. This is because all of their forecasts are close and, in the case of both in-sample and out-sample forecasts over all different horizons, they are better than the naive benchmark model. This indicates that the underlying uncertainties or misspecification does not have a great impact in the case of time-varying despite the other approaches whose forecasting performance varies significantly with changes in the business cycle indicator. The latter result could be investigated by the using the other output gap measures as driving sets to overcome the misspecification in each case. Regarding the forecast combinations, the two dynamic combinations approaches are superior to the best model forecasting.

Future studies may compare our approach with different time-varying approaches of PC like Time-varying with stochastic volatility approach as proposed by Primiceri (2005). Moreover, the ETVC approach may be augmented inside the whole structural models like the NK model by including state variables for all behaviour equations in the model as a tool to overcome a chronic misspecification problem inside this kind of models. Lastly, the density forecasts are also another avenue for future studies.

References

Arbatli E.C. and Moriyama K., 2011. Estimating a Small Open-Economy for Egypt Spillovers, Inflation Dynamics, and Implications for Monetary Policy. *IMF working paper 11/108*.

Atkeson, A. and Ohanian L.E., 2001. Are Philips Curves Useful for Forecasting Inflation?. *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1), pp.2-11.

Ball, L., 2000. Near-rationality and Inflation in Two Monetary Regimes. *NBER Working Paper* 7988.

Canova, F., 2007. G-7 Inflation Forecasts: Random Walk, Phillips Curve, or What Else? *Macroeconomic Dynamics*, 11(1), pp.1-30.

Christiano, L. Eichenbaum, M. Evans, C., 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, 113, pp.1- 45.

Dimitrios Asteriou & Hall, Stephen G., 2016. Applied Econometrics. Palgrave Macmillan; 3rd edition.

Dolado, J.J. Maria-Dolores, R. and Naveira M., 2005. Are Monetary-policy Reaction Functions Asymmetric? the Role of Nonlinearity in the Phillips Curve. *European Economic Review*, 49 (2), pp.485–503.

Faust, J., and Wright, J., 2013. Forecasting Inflation. In G. Elliott, and A. Timmermann, Eds. Handbook of Economic Forecasting, vol. 2. Amsterdam, Netherlands: Elsevier.

Fisher, J. D. M. Liu, C.T. and Zhou R., 2002. When Can We Forecast Inflation? Economic Perspectives. *Federal Reserve Bank of Chicago (Q I)*, pp.30-42.

Gali J. and M. Gertler, 1999. Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics,* 44(2), pp.195-222.

Gali, J. Gertler, M. and opez-Salido J.D.L., 2002. European Inflation Dynamics. *European Economic Review*, 45(7), pp. 1237-1270.

Garratt A. Mitchell, J. and Vahey S., 2010. Measuring Output Gap Uncertainty. *CEPR Discussion Papers* no. 7742.

Hall, S.G. Swamy, P.A. and George S.T., 2014. Time Varying Coefficient Models: a Proposal for Selecting the Coefficient Driver Sets. *Leicester University, Working Paper* No. 14/18.

Harvey, A.C., 1987. Applications of the Kalman Filter in Econometrics. In T.F. Bewley, ed., Advances in Econometrics, Fifth World Congress, Volume I., Cambridge, England, Cambridge University Press.

Hilde, B. Brubakk, L. and Jore, A., 2008. Forecasting Inflation with an Uncertain Output Gap. Empirical Economics, Springer, 35(3), pp.413-436.

Hodrick, R. Prescott, E.C., 1997. Post-war U.S. Business Cycles: an Empirical Investigation. *Journal of Money, Credit, and Banking,* 2**9** (1), pp.1–16.

Koop, G. and Korobilis, D., 2012. Forecasting Inflation Using Dynamic Model Averaging. *International Economic Review*, 53(3), pp. 867-886.

Marta, B. Harun M., 2013. Forecasting Euro Area Inflation with the Phillips Curve. European Central Bank. [online] Available at: <https://a793f7de-a-62cb3a1a-s-sites.googlegroups.com/site/hmirzahome/research/Forecasting-Euro-Area-Inflation-with-the-Phillips-Curve.pdf>.

Musso, A. Stracca, L. and van Dijk, D., 2009. Instability and Nonlinearity in the Euro Area Phillips Curve*. International Journal of Central Banking*, 5(2), pp. 181-212.

Olivier, C. and Gorodnichenko, Y., 2015. Is the Phillips Curve Alive and Well After All? Inflation Expectations and the Missing Disinflation. A*merican Economic Journal: Macroeconomics*, 7(1), pp.197-232.

Orphanides, A. and van Norden, S., 2005. The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real time. *Journal of Money, Credit and Banking*, 37(3), pp. 583-601.

Phillips, A.W., 1958. The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957. *Economica*, 25(100), pp. 283-299.

Primiceri, G., 2005. Time Varying Structural Vector Auto-regressions and Monetary Policy. *Review of Economic Studies,* 72, pp. 821-852.

Raftery, A.E. Karny, M. and Ettler, P., 2010. Online Prediction under Model Uncertainty via Dynamic Model Averaging*:* Application to a Cold Rolling Mill. *Technometrics*: 52(1), pp. 52-66.

Robert Czudaj, 2011. Improving Phillips Curve-Based Inflation Forecasts: a Monetary Approach for the Euro area. *Banks and Bank Systems*, 6(2), pp. 5-14.

Rumler, F. and Valderrama, M.T., 2010. Comparing the New Keynesian Phillips Curve with Time Series Models to Forecast Inflation. *The North American Journal of Economics and Finance*, 21(2), pp. 126-144.

Runstler G., 2002. The Information Content of Real-time Output gap Estimates: an Application to the Euro Area. *European Central Bank Working Paper Series* no. 182.

Samuelson, P.A. and Solow R. M., 1960. Analytical Aspects of Anti-inflation Policy. *American Economic Review*, 50(2), pp. 177-194.

Song, H. *et al*., 2011. Forecasting Tourist Arrivals Using Time-Varying Parameter Structural Time Series Models. *International Journal of Forecasting*, 27(3), pp.855-869.

Hall, S.G. *et al*., 2009. The New Keynesian Phillips Curve and Lagged Inflation: A Case of Spurious Correlation?. *Southern Economic Journal*, 76(2), pp. 467-481.

Stock, J.H. and Watson, M.M., 2007. Why has U.S. Inflation Become Harder to Forecast?.. *Journal of Money, Credit, and Banking*, 39 (7), pp. 3-34.

Stock, J.H. and Watson, M.M., 2009. Phillips Curve Inflation Forecasts. Federal Reserve Bank of Boston, vol. 53.

Stock, J.H. and Watson, M.W., 1999. Forecasting Inflation. *Journal of Monetary Economics*, 44(2), pp. 293-335.

Swamy, P.A. Tavlas, G.S. Hall, S.G. and Hondroyiannis, G., 2010. Estimation of parameters in the Presence of Model Misspecification and Measurement Error. *Studies in Nonlinear Dynamics and Econometrics*, 14(1), pp. 1-33.

Swamy, P.A.V.B. Yaghi, W. Mehta, J.S. and Chang, I-L., 2007. Empirical Best Linear Unbiased Prediction. In Misspecified and Improved Panel Data models with an Application to Gasoline Demand. *Comput Statist Data Anal*, 51, pp.3381-3392.

Timmermann, A., 2006. Forecast Combinations. In: Elliott, C. Granger, and Timmermann, A., Eds. Handbook of economic forecasting, Volume 1, Chapter 4, pp. 135-196, Amsterdam: North-Holland.

Todd E.C. and Michael M.W., 2010. Averaging Forecasts from VARs with Uncertain Instabilities. *Journal of Applied Econometrics*, 25(1), pp. 5-29.

Wright, J., 2007. Forecasting US Inflation by Bayesian Model Averaging. *Journal of Forecasting*, 28(2), pp. 131-144.

Appendix

State Space Representation for Time Varying Parameters Models

The two Time Varying Phillips curve models can be represented in the form of the state space form. In order to illustrate the utilized estimation process, we follow Harvey (1987) who depends on the standard state-space equation with the appropriate Kalman Filter equations for the univariate case. This supposes that:

$$y_t = \delta' s_t + \epsilon_t \qquad (24)$$

where: the equation (24) is called the measurement equation, y_t is a measurement variable, s_t is the state vector of unobserved variables inside the system, δ is a vector of considered parameters and $\epsilon_t \sim NID(0, \Gamma_t)$. The state equation can be represented in the form:

$$s_t = \Psi s_{t-1} + \psi_t \qquad (25)$$

where: Ψ refer to parameters, $\psi_t \sim NID(0, \mathrm{Q}_t)$ and Q_t is called hyperparmeters.

Then, the equation of appropriate Kalman filter prediction can be derived by defining s_t^* as the best estimate of s_t based on information up to t-1, and P_t is the covariance matrix of s_t^*, and is expressed, also, as:

$$s_{t|t-1}^* = \Psi\, s_{t-1|t-1}^* \qquad (26)$$

And

$$P_{t|t-1} = \Psi\, P_{t-1} \Psi' + Q_t \qquad (27)$$

Once the new information of y_t is attained, then estimations can be updated depending on the following equations:

$$s_{t|t} = s_t^* + P_{t|t-1}\delta(y_t - \delta' s_t^*)/(\delta' P_{t|t-1}\delta + \Gamma_t) \quad (28)$$

And

$$P_t = P_{t|t-1} - P_{t|t-1}\delta\delta' P_{t|t-1}/(\delta' P_{t|t-1}\delta + \Gamma_t) \qquad (29)$$

Equations from (24) to (29) represent the predictive Kalman filter process.

Then, one can define the one step ahead prediction errors by:

$$\epsilon_t = y_t - \delta' z_{t-1}^* \qquad (30)$$

Then, the log-Likelihood equation can be represented in the form (For more details we refer to Hall and Nixon, 1997):

$$log(\mathcal{L}) = \sum log(\mathcal{F}_t) + N\, log(\sum \epsilon_t^2 / N\, \mathcal{F}_t) \qquad (31)$$

Since $\mathcal{F}_t = \delta' P_{t|t-1}\delta$, $N = T - k$ and k is the number of required periods to estimate the state vector. Hence, the description of the log likelihood function depends on one step ahead prediction errors.

Given these equations, we can estimate time varying parameter PC models such as in (11) or (17) based on the maximum likelihood estimation for the state space model. Using our process, we define firstly δ in the form of a vector of known variables and s_t is a vector of time varying parameters. In addition, Ψ is defined as a constant identity matrix and Q_t is a diagonal matrix. The previous elements can be estimated by utilizing the maximum likelihood based on the likelihood function in equation (31).

Calculation of the Measures of Output Gap and Real Wage Share

A. Univariate Output Gap (UN)

Univariate methods of the output gap estimation depend only on the GDP data. One such detrending procedure is that suggested by Hodrick and Prescott (1997) which contains the linear trend as a special case. The Hodrick-Prescott (HP) filter sets the potential component of output to minimize the loss function, L:

$$L = \sum_{t=1}^{S}(y_t - y_t^T)^2 + \lambda\sum_{t=2}^{S-1}(\Delta y_{t+1}^T - \Delta y_t^T)^2 \quad (32)$$

Where: λ is the smoothing weight on potential output growth and S is the sample size.

B. Unobserved Multivariate (MU)

Multivariate methods explore the relationship between the GDP and other observed variables. The relationship between the unemployment rate and the output gap given by the Okun's law are explored, see Okun (1962). The potential output is determined its own lag and the quarterly growth rate and disturbance term:

$$Y_t = \overline{Y}_t + ygap_t \quad (33) \qquad \overline{Y}_t = \overline{Y}_{t-1} + \frac{G_t^{\overline{Y}}}{4} + \epsilon_t^{\overline{Y}} \quad (34)$$

Potential GDP growth is a function in its steady state, its lagged value and disturbance term:

$$G_t^{\overline{Y}} = \tau G_i^{\overline{Y\,ss}} + (1-\tau)G_{t-1}^{\overline{Y}} + \epsilon_t^{\overline{Y_G}} \quad (35)$$

The output gap follows AR (2) process:

$$ygap_t = beta1 * ygap_t(-1) + beta2 * ygap_t(-2) + RES\_ygap_t \quad (36)$$

Similarly, dynamics of the unemployment rate is governed by the dynamic version of Okun's law.

$$u_t = \alpha_1 u_{t-1} + \alpha_2 ygap_t + \epsilon_t^u \quad (37)$$

This in general can be expressed in the form: $u_t = U_t - \overline{U}_t$ (38)

E quilibrium unemployment rate depends on its lag, growth term and disturbance term

$$\overline{U}_t = \overline{U}_{t-1} + G_t^u + \epsilon_t^{\overline{U}} \quad (39)$$

When the growth element is defined as a function in its own lag and disturbance term

$$G_t^u = (1-\alpha_3)G_{t-1}^u + \epsilon_t^{G^u} \quad (40)$$

C. Real Wage Share

Gali and Gertler (1999) real wage share rather than output gap as a representative for a business cycle measure. They state that this strength the New Keynesian's Philips Curve validation. Similarly, we utilize the deviation from Real Unit Labor Cost which is defined as a portion of compensation per employee and his productivity in terms of the current prices:

$$RULC_t = \frac{Compensation\ for\ employees}{employess} \Big/ \frac{GDP\ in\ current\ price}{employment} \quad (41)$$