Forecasting US Inflation: Phillips Curve, New Keynesian Phillips Curve, Or Something Else?

By

Dalton Kick

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Applied Economics

At

The University of Wisconsin – Whitewater

November 2017

Graduate Studies

The members of the Committee approve the thesis of

Dalton Kick presented on October 30, 2017

Dr. Yamin Ahmad, Chair

Dr. Nick Guo

Dr. Eylem Ersal

Forecasting US Inflation: Phillips Curve, New Keynesian Phillips Curve, Or Something Else?

By

Dalton Kick

The University of Wisconsin – Whitewater, 2017
Under the Supervision of Dr. Ahmad

Abstract

I utilize and compare several common inflation forecasting models, including traditional Phillips curve models and the New Keynesian Phillips curve, as well as several other time series models. I evaluate these models using RMSE over several forecast horizons, using three different measures of inflation: CPI inflation, PPI inflation, and GDP deflator inflation. I find that the theoretical Phillips curve models outperform other time series models, however, the performance is sensitive to the inflation measure used in estimation.

Table of Contents

List of Tables

List of Figures

1. Introduction

Inflation is an increase in prices over time. Existing literature suggests that having some level of inflation encourages economic growth. The reasoning behind this includes, but is not restricted to, the idea that consumers are more willing to buy things now if prices will be increased in the future. Many countries have implemented some form of inflation targeting in order to meet a level of growth-encouraging inflation and it has become important for these countries to have some ability to forecast the inflation rate with confidence so that the central banks can adjust their responses accordingly. Because of this, literature attempting to model inflation dynamics and forecast inflation became prevalent. I intend to compare the forecasting ability of several of these models to see if theoretical models forecast more accurately than time series models.

The New Keynesian Phillips curve (NKPC) seems to be among the most popular and widely accepted methods of analyzing inflation dynamics, and much of the literature in the last fifteen years that utilizes the NKPC borrows extensively from Gali and Gertler (1999). Gwin and Vanhoose (2007) use Gali and Gertler's model exactly, with the only differences being that they use PPI inflation rather than the GDP deflator and that they use a measure for average variable cost as marginal cost. They find that using average variable cost does not result in significantly different results, but that using PPI inflation is more appropriate. Another similar paper is Bratsiotis and Robinson (2014) who also use Gali and Gertler's NKPC but use a different proxy for marginal cost, and find that their new measure is a better fit for empirically estimating inflation dynamics.

Extensive literature comparing the quality of inflation forecast ts across models already exists. Stock and Watson (2009) compare a traditional Phillips curve model to several other time series and theoretical models to forecast US inflation. These other models include autoregressive (AR) models, an unobserved-components stochastic volatility model, a random walk model, as well models with time-varying components. They find that the Phillips curve model's forecasting ability is "episodic" relative to other models (Stock and Watson, 2009). Similarly, Onder (2004) compares a traditional Phillips curve to several other time series models including AR, VAR, and random walk models. She does this with Turkish data and uses output gap as the activity variable rather than the unemployment rate as in Stock and Watson (2009). She finds that the Phillips curve outperforms the other models at all horizons. Rumler and Valderrama (2010) do something similar using Austrian data, however, they also incorporate the New Keynesian Phillips curve (NKPC), as well as a more traditional Phillips curve, a VAR model, an AR model, and a random walk model. They use the NKPC specified by Gali and Gertler (1999), with slight modifications to bring in open economy elements, to forecast inflation and find that, at least at longer horizons, the NKPC is the strongest forecasting model they estimate.

In this paper, I also compare several models' abilities to forecast inflation. However, the main contribution is that I combine several aspects of some of the previously mentioned papers while using a common dataset. It is difficult to compare the results across many of the papers in the existing literature due to the vastly different dataset used to obtain them. For this reason, I borrow models from several different papers and estimate forecasts using only US data in order to try and get a more comprehensive and easily comparable result. The models that I compare include two different Phillips curve specifications, the NKPC, three different vector autoregressive (VAR) models, an autoregressive model, and a simple random walk as a baseline at several horizons. In these exercises, I use three different measures of inflation: CPI inflation, PPI inflation, and GDP deflator inflation. I evaluate the forecasts using the root mean squared error (RMSE) as well as the Diebold-Mariano test to see if forecasts from different models are statistically different.

I find that the best model for forecasting inflation is somewhat sensitive to how inflation is measured. Stock and Watson's (2009) Phillips curve is the best at all horizons for both CPI inflation and PPI inflation, however, it is less accurate at forecasting GDP deflator inflation. The NKPC forecasts reasonably well across all horizons for all three inflation measures, however, it specifically excels at forecasting GDP deflator inflation.

In the next section, I specify the models used for forecasting inflation, as well as give some theoretical background for the Phillips curve, which is followed by a section explaining the data and estimation techniques used to obtain the forecasts for each model in greater detail. I then discuss the results of the forecasts and discuss the evaluation of them before concluding with some explanations for why the results may be what they are as well as suggesting some possible future extensions.

2. Models

2.1 Theoretical Reasoning Behind Phillips Curve

The Phillips curve, named after William Phillips, describes the relationship between inflation and some variable measuring real activity in the economy, generally the unemployment rate or the output gap. Specifically, the Phillips curve describes how the unemployment rate and the inflation rate have an inverse relationship, meaning higher unemployment leads to lower inflation and vice versa. The output gap is the difference between the actual output and potential output of an economy; when the output gap is positive, i.e. when actual output is greater than potential output, inflation increases and unemployment is lower. A negative output gap will have the opposite effect.

2.2 Phillips Curve Models

I use the same interpretation for Phillips curve forecasts that is found in Stock and Watson (2009): any forecast produced using a real activity variable like the unemployment rate, or output gap. I use two different traditional Phillips curve models to forecast inflation; one is a slight modification of Stock and Watson's (2009) modification of Gordon's (1990) triangle model, and the other is from Onder (2004).

The first of these has inflation dependent on past inflation, the unemployment rate u_t, and a vector of supply shock variables, z_t, and is specified by the following equation:

$$\pi_{t+1} = \mu + \alpha(L)\pi_t + \beta(L)u_{t+1} + \gamma(L)z_t + v_{t+1}, \tag{1}$$

where v_{t+1} is the error term. The supply shock variables in z_t include 4 lags of the relative price of imports, 4 lags of food and energy price inflation, and 8 periods of lagged inflation. The unemployment rate consists of the value in period *t+1* and 4 additional lags. As the relative price of imports increases, the domestic price level should also increase. Food and energy inflation is more volatile than overall inflation and should help account for cost-push inflation.

The second of these, from Onder (2004), has inflation dependent on past inflation and the output gap gap_t, and is specified by the following equation:

$$\pi_t = \alpha(L)\pi_t + \beta(L)gap_t + \varepsilon_t\,, \qquad (2)$$

where α is the coefficients on the lags of the inflation rate, β is the coefficients on the lags of the output gap, and ε_t is the error term. This is a standard Phillips curve model using the output gap as the real activity variable, with no additional explanatory variables except for the lags of inflation itself.

2.3 New Keynesian Phillips Curve

The NKPC as described in Gali and Gertler (1999) utilizes marginal cost's relationship with inflation rather than the output gap or unemployment rate to measure real activity. Assume there is a continuum of monopolistically competitive firms, each of which has a time-dependent constraint on their ability to change prices. To account for this, a Calvo (1983) probability, θ, is used to denote that each firm a fixed probability $1-\theta$ that it will adjust its price in a given period. This probability is independent of past time periods.

Assuming the firms face a constant price elasticity of demand, and that the firms are identical except for their differentiated products and pricing history, it is possible to show that the price level is a combination of its lags and the new price that firms that are able to change price set (Gali and Gertler, 1999). Firms able to choose a new price in a given time period will consider future marginal cost, given the likelihood that they may not be able to adjust price again for several time periods. Monopolistically competitive firms mark-up price relative to marginal costs and may be forward-looking so this is intuitive. Because inflation can be thought of as simply an aggregation of prices, it should, therefore, be equal to some measure of expected future marginal costs.

As suggested by the literature, I use the exact NKPC model from Gali and Gertler (1999). The biggest difference between this NKPC and more traditional versions of the Phillips curve is that it is estimated using real marginal cost, rather than unemployment or the output gap. The empirical version of this model that I use is specified as follows:

$$\pi_t = \lambda mc_t + \gamma_f E_t[\pi_{t+1}] + \gamma_b \pi_{t-1} \tag{3}$$

where mc_t is marginal cost, and $E_t[\pi_{t+1}]$, is expected future inflation. The parameters λ, γ_f, and γ_b are described as follows:

$$\lambda \equiv (1-\omega)(1-\theta)(1-\beta\theta)\phi^{-1}$$

$$\gamma_f \equiv \beta\theta\phi^{-1}$$

$$\gamma_b \equiv \omega\phi^{-1}$$

where θ measures the degree of price stickiness, ω measures the degree of 'backwardness' in price setting, and β is a discount factor. In each of the above equations $\phi \equiv \theta + \omega[1-\theta(1-\beta)]$. The structural parameters β, θ, and ω are estimated using a non-linear instrumental variables, Generalized Method of Moments (GMM), estimator. The instrument set used in this estimation include 4 lags of inflation, a measure of marginal cost, the output gap, the long-short interest rate spread, wage inflation, and commodity price inflation.

The instruments should be uncorrelated with the residuals of the equation and should be correlated with the explanatory variables. Both Gali and Gertler (1999) and Bratsiotis and Robinson (2014) find that these instruments satisfy these conditions. Additionally, Bratsiotis and Robinson (2014) claim that this is “the most parsimonious instrument set possible to avoid the estimation bias that arise in small samples with too many over-identifying restrictions.”

2.4 Other Time Series Models

In addition to the standard Phillips Curve models and the New Keynesian Phillips curve model, I also forecast the inflation rate with 3 different VAR models, as well as a standard autoregressive model, and a random walk with drift as a baseline. I determine the number of lags in the autoregressive model using the Ljung-Box methodology to detect when the residuals become statistically indistinguishable from white noise. The VARs I use are borrowed from papers addressing a similar topic as this one: Canova (2007), Onder (2004), and Rumler and Valderrama (2010).

The general form of a VAR can be defined as,

$$Y_t = C + A(L)Y_t + E_t \tag{4}$$

where Y_t is a vector of endogenous variables at time t, C is a vector of constants, $A(L)$ is a polynomial in lag operator L, and E_t is a vector of error terms. The specific variables in each of the three VARs are different except for inflation.

Canova forecasts inflation for G-7 countries, and uses a datamining technique in order to find the best forecasting VAR for each country; he had a large number of explanatory variables, and for each country ran all possible models to minimize the MSE, limiting to 3 variable models in order to avoid "excessive data mining." Because I am only looking at U.S. inflation, I simply use the explanatory variables that he found to give the smallest MSE for the United States. In this case, Y_t in (4) consists of the inflation rate, the output gap, and the civilian labor force. The lags of the output gap and civilian labor force address some of the theoretical Phillips by Onder to forecast inflation specifies Y_t from (4) using money and interest rate, in addition to inflation. She cites Engert and Hendry (1998) and Mirmirani and Li (2001) as using similar variables to forecast inflation. A lower interest rate will result in more borrowed money, and thus more spending and higher inflation will be observed. The opposite is true for a higher interest rate.

Rumler and Valderrama specify their VAR with the assumption that inflation is steered by demand-pull and cost-push, as suggested by common macroeconomic theory. The Y_t vector in (4) includes real GDP, wages, oil prices, and inflation. They include real GDP to account for economic activity, and wages and the oil prices to express supply shocks.

3. Data and Methodology

3.1 Description of Data

I estimate each model using quarterly U.S. data from 1986:1 to 2016:4, with 2000:1 to 2016:4 as the pseudo out of sample forecast period that is estimated recursively in order to forecast the inflation rate. I obtained all of the data from U.S. Bureau of Labor Statistics through the Federal Reserve Bank of St. Louis (FRED), and I chose the start date because of the unavailability of oil price data before then. It is interesting to note that there are two recessions that occur within this sample period (2001, 2007-09), which allows for several instances in seeing which, if any, models can predict recessions, and if so, at what forecast horizons.

I use three different inflation series in these estimations: Consumer Price Index (CPI) inflation, Producer Price Index (PPI) inflation, and GDP deflator inflation[1]. The CPI measures inflation more from the perspective of consumers through purchasing a basket of goods and services in urban areas, the PPI measures inflation from the perspective of domestic producers as the prices they receive for selling their output, and the GDP deflator inflation measures inflation through the overall change in domestic consumption and production. Each of these series is transformed into inflation by taking the log difference. All models are estimated for each inflation series.

For the Phillips curve models, the data needed in addition to inflation is unemployment, food and energy inflation, and the relative price of imports. Unemployment is measured as the total civilian unemployment rate. Food and energy inflation is calculated as Personal Consumption Expenditures inflation (PCE) minus the PCE-core inflation. The relative price of imports is calculated as imports of goods and services index divided by the GDP deflator.

In addition to inflation, the NKPC model also utilizes the output gap, marginal cost, the long-short interest rate spread, wage inflation, and commodity price inflation. The output gap is calculated using the Hodrick-Prescott filter[2]. Marginal cost is

1 Series identifiers for inflation series: CPI – CPIAUCSL; PPI – PPIACO; GDP Deflator - GDPDEF

2 I considered using the CBO estimate; this measure is highly correlated with the HP filter, however, the CBO estimate was non-stationary, whereas the HP filter is stationary. For the sake of being consistent across models, I decided to use the HP-filter.

measured as the log labor income share in the non-farm business sector (Gali, Gertler 1999). The long short interest rate spread is measured as the 10-Year Treasury Constant Maturity minus the 3-Month Treasury Constant Maturity; commodity price inflation is measured using oil prices, and wage inflation is measured using gross domestic income.

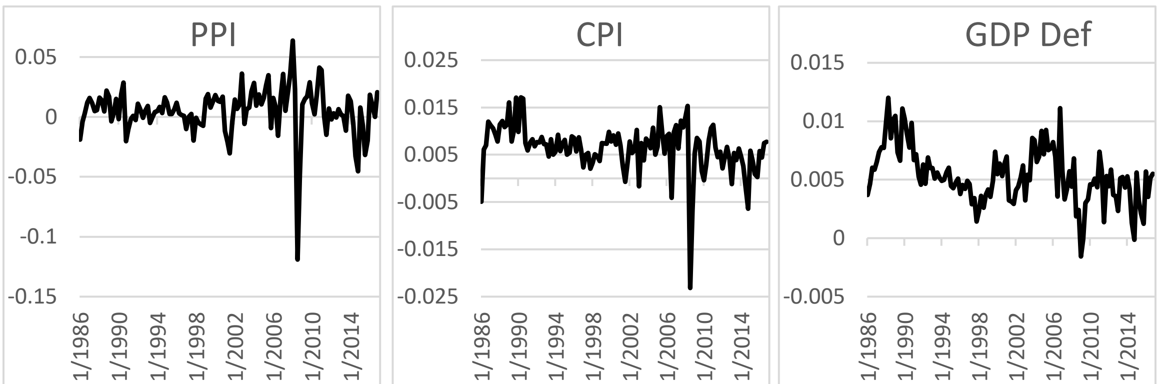
As discussed in the previous section, each of the VARs contain different variables. Canova uses the output gap, again calculated with the HP-filter, and the labor force participation rate. In the VAR from Onder (2004), M1 is used for the money supply, and the 6-month treasury bill is used for the interest rate. The last VAR from Rumler and Valderrama (2010) contains real GDP, wages, and oil prices.

3.2 Methodology

In order to estimate the VARs, each of the variables within them needs to be stationary to avoid dealing with cointegration. The plots of the inflation series can be seen in figure 1. While they appear to be stationary observationally, I run unit root tests as well to verify. These tests include the Augmented Dickey-Fuller (ADF) test, the Phillips-Perron (PP) test, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The ADF and PP tests have a null hypothesis of containing a unit root, or non-stationarity. For the KPSS test, the null hypothesis is that the series is stationary, thus rejection would imply non-stationarity. The results of these tests can be seen in Table 1. For all three inflation measures, the null is rejected in the ADF and PP tests implying stationarity. Similarly, for all three inflation series, the null cannot be rejected in the KPSS test, again implying stationarity.

The other non-inflation series variables within the VARs much be stationary as well. I transform each of these variables into log-difference form, except for the output gap (labeled HP gap). The plots can be seen in Figure 2, and again appear to be stationary. I use the ADF, PP, and KPSS tests for these variables as well in order to verify

Figure 1



their stationarity. Like the inflation series, the results of these tests suggest that each of these variables are stationary and thus can be used in estimating the VARs without worrying about cointegration.

Table 1

Results from stationarity tests for inflation series (p-value)			
	CPI	**PPI**	**GDP Def**
ADF	0.000	0.000	0.005
PP	0.000	0.000	0.000
KPSS	0.815	0.777	0.769
Stationary	Yes	Yes	Yes

In forecasting these models, I estimate 1-quarter, 2-quarter, 4-quarter, and 8-quarter ahead forecast horizons for each model. I do this using recursive forecasts: a forecast for each of these 4 horizons is estimated using the initial sample of 1986:1 to

1999:4 before I increase the sample by 1-quarter and forecast each of the 4 horizons again using the same initial date. This continues, with the sample size in each iteration increasing by one, until the end of the evaluation period. I evaluate these forecasts using the root mean squared error (RMSE) and the Diebold-Mariano test.

Table 2

Results from stationarity tests for non-inflation series in VARs (p-value)							
	LFPR (lnFD)	**HP Output gap**	**ULC (lnFD)**	**Oil Prices (lnFD)**	**M1 (lnFD)**	**Tbill6M (FD)**	**rGDP (lnFD)**
ADF	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PP	0.000	0.001	0.000	0.000	0.000	0.000	0.000
KPSS	0.764	0.860	0.789	0.965	0.871	0.983	0.776
Stationary:	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(FD): Denotes that the series was first differenced; ln: Denotes that this series was log-differenced							

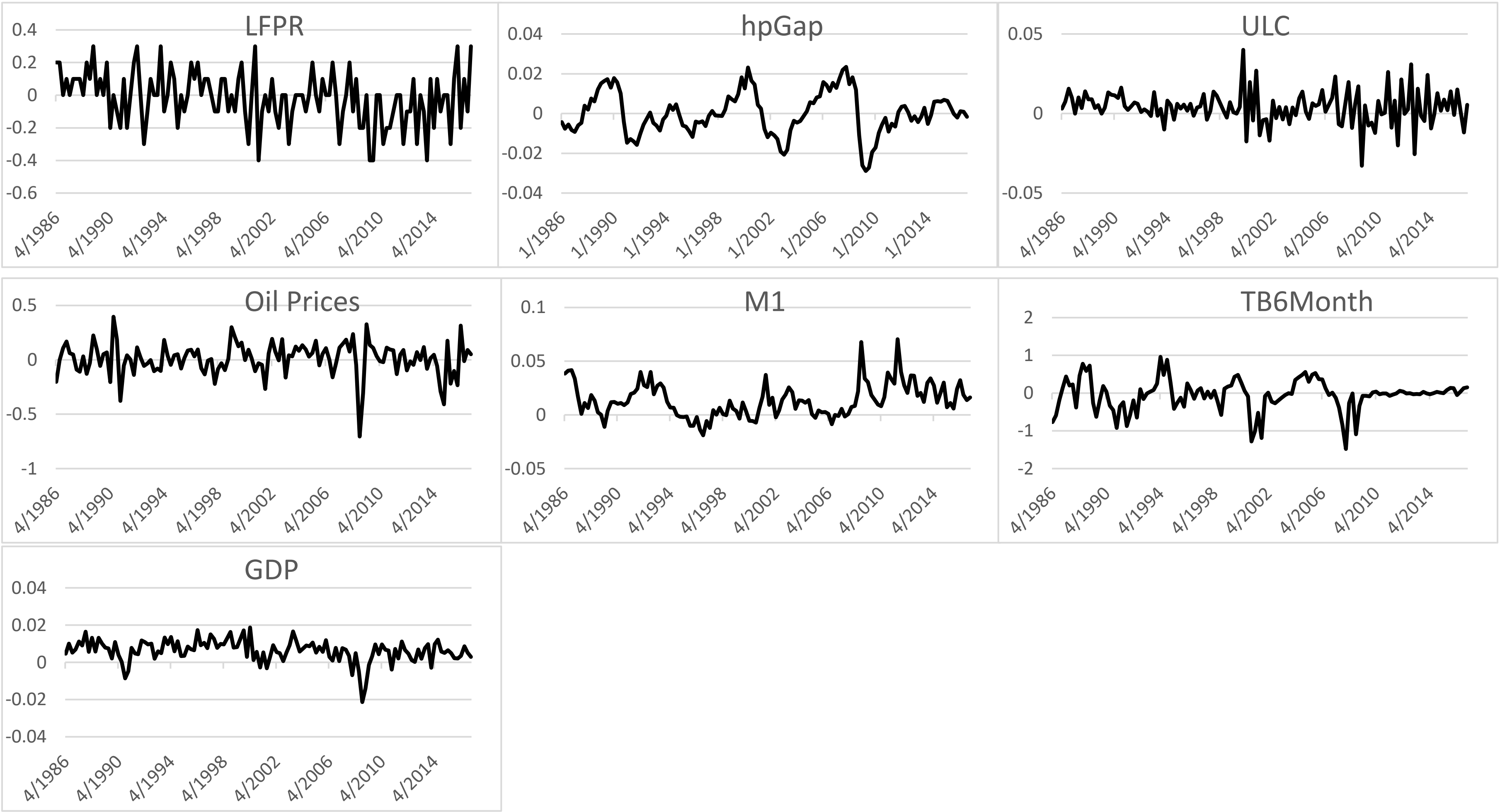
4. Results

4.1 Forecast Evaluation

I evaluate the quality of these forecasts using the RMSE[3]. This number is found by taking the square root of the average squared residual, and thus the closer the number is to zero, the more accurate the model is. For each inflation series, (CPI inflation, PPI inflation, GDP Deflator inflation) the point-forecasting ability of each model varies. See the results in Table 3.

3 I considered using mean average error (MAE) rather than RMSE. I chose RMSE over MAE due to being a more common measure in economic forecasting literature, and because of the results in Chai and Draxler (2014) which show that MAE is not superior to RMSE.

Figure 2



This shows the graphs of the data series used in the VARs in their transformed state

For each of the inflation series, the Phillips curve specified in Stock and Watson (2009) has the lowest RMSE for the 1-quarter ahead forecasts. This model also has the lowest RMSE for the 2-quarter, 4-quarter, and 8-quarter forecasts for both the CPI and PPI inflation series. It's interesting to note that not only does model have the lowest RMSE at each horizon, but for CPI inflation the RMSE is less than half of the NKPC at 8-quarters ahead, which has the second lowest RMSE, and for PPI inflation the RMSE is less than 60% of the NKPC at 8-quarters ahead. This suggests that for these inflation series, this Phillips curve model is clearly superior at point forecasts.

The NKPC does a better job of forecasting compared to the VARs, AR, and random walk models. At 1-quarter ahead, the NKPC is either the second (CPI inflation) or the third (PPI inflation and GDP deflator inflation) best in terms of RMSE. For 2-quarter, 4-quarter, and 8-quarter ahead horizons, the NKPC is either the second best (CPI inflation and PPI inflation) or the best (GDP deflator inflation) in terms of RMSE. This is largely consistent with the literature that suggests that the NKPC is the best current model for capturing inflation dynamics, and one of the better models for forecasting inflation.

The Phillips curve detailed in Onder (2004) also does a relatively good job of forecasting compared to the other time series models. It is the only model besides Stock and Watson's Phillips curve that ever beats the NKPC in terms of RMSE at any horizon and does so in 1-quarter ahead forecasts for PPI inflation and GDP deflator inflation. Recall that Onder used the output gap in her Phillips curve model rather than the unemployment rate due to limitations in the available data for Turkey. The results of

Table 3

RMSE (Annualized)				
Model	**1 qtr**	**2 qtr**	**4 qtr**	**8 qtr**
CPI Inflation				
Random Walk	0.023516	0.026212	0.028067	0.028056
AR(1)	0.019444	0.019332	0.020424	0.022896
PC Stock and Watson	**0.006288**	**0.0067**	**0.00746**	**0.00826**
PC Onder	0.022144	0.022952	0.02352	0.023584
VAR Rumler and Valderrama	0.021396	0.022604	0.021748	0.023008
VAR Onder	0.021068	0.02202	0.02154	0.023644
VAR Canova	0.022392	0.024164	0.023772	0.0254
NKPC	0.018588	0.018672	0.019476	0.022256
PPI Inflation				
Random Walk	0.0894	0.11964	0.12724	0.1182
AR(2)	0.07904	0.09	0.09	0.09488
PC Stock and Watson	**0.05084**	**0.05388**	**0.05192**	**0.05296**
PC Onder	0.076	0.08484	0.08736	0.09328
VAR Rumler and Valderrama	0.08228	0.09236	0.0972	0.0978
VAR Onder	0.08132	0.0928	0.08892	0.09572
VAR Canova	0.08156	0.0888	0.9104	0.09524
NKPC	0.07688	0.08224	0.08296	0.0894
GDP Deflator Inflation				
Random Walk	0.008572	0.008656	0.00992	0.012416
AR(2)	0.007692	0.007568	0.008416	0.009976
PC Stock and Watson	**0.007208**	0.007368	0.009168	0.011936
PC Onder	0.00728	0.007172	0.00778	0.010144
VAR Rumler and Valderrama	0.007944	0.008344	0.00956	0.012316
VAR Onder	0.007896	0.007444	0.007884	0.009604
VAR Canova	0.0075	0.007112	0.008188	0.010048
NKPC	0.007316	**0.007016**	**0.007368**	**0.007332**

these forecasts suggest that any Phillips curve specification will provide better forecasts than VARs or atheoretical models.

Both the VAR specifications and the AR model has similar forecasting ability based on RMSE, and that ability seems mediocre. None of the VAR specifications or AR model are ever superior to either of the traditional Phillips curve models or the NKPC, nor are they ever inferior to the random walk baseline model. They have varying degrees of success, but always somewhere in the middle of the best and worst forecasts. This is largely consistent with what is found in Rumler and Valderrama (2008) and Onder (2004). Among only these four models, the successfulness varies across inflation series and forecast horizons. In terms of CPI inflation, the pure AR model is the best of the four across each horizon, but that is the only inflation series with any sort of clarity in terms of which model is best. Because the inflation series are stationary (detailed in section 3), autoregressive processes should outperform the random walk. Unsurprisingly, the random walk model is the worst in terms of RMSE for all three inflation series, and across all 4 forecast horizons relative to the VAR and AR models.

I also check to see if the forecasting ability across these models is statistically significant using the Diebold-Mariano test (1995) with a quadratic loss function. The null hypothesis for this test is that two models have equal predictive accuracy. I compare each model to every other model over each forecast horizon for all three inflation series, which gives the most possible information. The results of this test can be seen in table 4. In this table, the first listed model is used as the baseline; a positive DM test statistic suggests that the model being tested against the baseline is more accurate in forecasting inflation. The opposite is true if this number is negative.

For CPI inflation, the forecasts from Stock and Watson's (2009) Phillips curve is significantly better than every other model at every horizon. This is reasonable given that the RMSE is much lower from this model than any other model. The other strong forecasting models indicated by RMSE perform inconsistently here. The NKPC is significantly better than every model except the previously mentioned Phillips curve and the AR model at the 4-quarter ahead forecast horizon. Curiously, at shorter horizons and at the 8-quarter ahead horizon, the NKPC forecasts are not significantly better than the other models. Onder's (2004) Phillips curve is not significantly better than any model at any horizon. These results reinforce that Stock and Watson's (2009) Phillips curve is the best forecaster for CPI inflation.

PPI inflation has similar results as CPI inflation. Stock and Watson's (2009) Phillips curve again performs strongly, performing significantly better than all other models at all horizons. The NKPC performs somewhat better for this inflation series. It is significantly better than every other model at most horizons with some exceptions. Obviously, the previously mentioned Phillips curve again is better than the NKPC, but also the NKPC is not significantly different than the AR model at the 1-quarter and 8-quarter ahead horizons, Onder's (2004) Phillips curve at the 1-quarter horizon, and both Canova's (2007) and Rumler and Valderrama's (2008) VARs at 1-ahead. Despite these exceptions, these results for the NKPC suggest that it is a stronger predictor than other models for PPI inflation relative to CPI inflation. Onder's (2004) Phillips curve performs better here relative to how it performs for CPI inflation, as it is significantly better than Rumler and Valderrama's (2008) VAR at 4-quarter and 8-quarter horizons, Canova's (2007) VAR at 1-quarter ahead horizons, and the random walk model at 2-quarter and 4-quarter ahead horizons. Once again, the Stock and Watson (2009) Phillips curve is suggested to be the best forecasting model which is consistent with what was found from the RMSEs for PPI inflation.

As is the case with the RMSEs, the results for the GDP deflator inflation series is somewhat dissimilar to those found for CPI and PPI inflation. Unlike the other inflation series, here all of the models are significantly better than the random walk baseline model for one or more horizons. The only model that is not significantly better than the random walk at the 8-quarter horizon is the Stock and Watson Phillips curve, which is already in contrast with what was found in the other inflation series where this was the best model all the way through. Besides the random walk series, there are very few examples of the models being significantly different to one another. The NKPC is significantly better than every other model at the 8-quarter ahead horizon except the AR model. This is consistent with what was found in the RMSE analysis where the NKPC had a much lower RMSE than every other model at this horizon.

For all three inflation series, there is very little difference between each of the VARs and the AR model. Again, this is similar to what was seen in the RMSE analysis. There are a few sporadic cases of the one of the VARs being superior to another, or the AR being superior to the VARs, but for the most part, there are no significant differences between them.

4.2 Additional Results

These evaluations suggest that the traditional Phillips curve specified in Stock and Watson (2009) is the overall best forecasting model at all horizons when using CPI inflation or PPI inflation as the inflation measure, and that the NKPC is the best overall forecasting model at all except 1-quarter horizons when using GDP deflator inflation as the inflation measure. This suggests that the relative success of these models is at least

Table 4

Percentage of Times that Model is Superior in Head-to-Head DM Test

	CPI Inflation				**PPI Inflation**				**GDP Deflator inflation**			
Model	1 qtr	2 qtr	4 qtr	8 qtr	1 qtr	2 qtr	4 qtr	8 qtr	1 qtr	2 qtr	4 qtr	8 qtr
Random Walk	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
AR	28.6%	28.6%	42.9%	14.3%	28.6%	28.6%	14.3%	0.0%	14.3%	14.3%	14.3%	28.6%
PC Stock and Watson	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	57.1%	14.3%	0.0%	0.0%
PC Onder	0.0%	0.0%	14.3%	0.0%	28.6%	14.3%	14.3%	28.6%	14.3%	14.3%	28.6%	14.3%
VAR Rumler and Valderrama	0.0%	0.0%	0.0%	0.0%	0.0%	14.3%	0.0%	14.3%	0.0%	0.0%	14.3%	28.6%
VAR Onder	0.0%	0.0%	14.3%	0.0%	0.0%	14.3%	0.0%	0.0%	0.0%	14.3%	14.3%	14.3%
VAR Canova	0.0%	0.0%	0.0%	0.0%	0.0%	14.3%	0.0%	0.0%	14.3%	14.3%	14.3%	14.3%
NKPC	0.0%	57.1%	57.1%	0.0%	28.6%	71.4%	57.1%	57.1%	14.3%	14.3%	28.6%	**85.7%**

somewhat sensitive to what measure of inflation is being used.

Gali and Gertler (1999) utilized GDP deflator inflation in their NKPC estimations and found that it modeled inflation dynamics reasonably well. Their results seem to be consistent with what I find in this paper, specifically that the NKPC models GDP deflator inflation rather well, and thus forecasts well. Curiously, Gwin and VanHoose (2007) utilize the NKPC from Gali and Gertler (1999) but they argue that PPI inflation is a more appropriate measure of inflation than GDP deflator inflation. My results indicate that the NKPC does reasonably well forecasting PPI inflation, but does not do so as accurately as Stock and Watson's (2009) traditional Phillips curve.

Table 5

Forecast Error Variance Decomposition			
Food and Energy Inflation			
Horizon	2	4	8
CPI Inflation	0.038	0.091	0.121
PPI Inflation	0.026	0.027	0.027
GDP Deflator Inflation	0.004	0.006	0.007
Unemployment Rate			
Horizon	2	4	8
CPI Inflation	0.006	0.016	0.02
PPI Inflation	0.007	0.017	0.021
GDP Deflator Inflation	0.027	0.097	0.148

There are two main variables in Stock and Watson's (2009) Phillips curve that do not appear in the NKPC (or in any other models): food and energy inflation and the unemployment rate. A variance decomposition for both of these variables with each inflation measure showed that the unemployment rate contributes to CPI and PPI inflation less than GDP deflator inflation, whereas food and energy inflation contributes much more to CPI inflation and PPI inflation than it does to GDP deflator inflation. This means that food and energy inflation has more impact on CPI and PPI inflation, which may explain to some extent why the Phillips curve model that accounts for food and energy inflation forecasts CPI and PPI inflation more accurately than any other model. The results of this variance decomposition can be seen in Table 5.

5. Conclusion

This paper intends to see how various commonly utilized theoretical and atheoretical time series models perform in terms of forecasting quarterly inflation in the United States. These models include two variations of the traditional Phillips curve, the NKPC as detailed in Gali and Gertler (1999), VARs as they were specified in similar papers, and simple AR and random walk. I evaluated these over 1-quarter, 2-quarter, 4-quarter, and 8-quarter horizons using the RMSE and DM-statistic.

I find that Stock and Watson's (2009) Phillips curve model is a more accurate predictor of CPI and PPI inflation at most horizons and that the NKPC is a more accurate predictor of GDP deflator inflation at most horizons. This suggests that inflation forecasting is sensitive to the inflation measure being used in estimation. Specifically, Stock and Watson's (2009) Phillips curve model accounts for food and energy inflation which appears to be captured more by CPI and PPI inflation than GDP deflator inflation.

This could be explored further by potentially altering the NKPC to account for food and energy inflation which could improve its forecasting ability for other measures of inflation. Additionally, running forecast estimations using core inflation measures could help to see whether the NKPC would be superior to Stock and Watson's traditional

Phillips curve when estimated with an inflation measure without food and energy. This would help verify that food and energy inflation is the key piece as to why Stock and Watson's (2009) Phillips curve outperforms the NKPC.

Other potential extension would be finding the density forecast (Rossi, Sehkposyan 2013) to evaluate more than just point forecasts for these models, and using different measures of marginal cost in estimating the NKPC such as in Rumler and Valderrama (2008), Gwin and VanHoose (2007), and Bratisiotis and Robinson (2014). The latter may result in changes to the forecasting performance of the NKPC.

References

Bratsiotis, George J. & Robinson, Wayne A. (2014). “Unit total costs: An alternative marginal cost proxy for inflation dynamics.” *Centre for Growth and Business Cycle Research, Economic Studies, University of Manchester*, vol. 192.

Brissimis, Sophacles N. & Magginas, Nicholas S. (2008). “Inflation forecasts and the new Keynesian Phillips curve.” *International Journal of Central Banking*, vol. 4, pages 1-22.

Canova, Fabio. (2007). “G-7 inflation forecasts: Random walk, Phillips curve or what else?” *Macroeconomic Dynamics,* vol. 11, pages 1-30.

Chai, T. & Draxler, R. R. (2014). “Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature.” *Geoscientific Model Development*, vol. 7, pages 1247-1250.

Engert, Walter & Hendry, Scott. (1998). “Forecasting inflation with the M1-VECM: Part two.” *Staff Working Papers, Bank of Canada*.

Gali, Jordi & Gertler, Mark. (1999). “Inflation dynamics: A structural econometric analysis.” *Journal of Monetary Economics*, vol. 44, pages 195-222.

Gwin, Carl R. & VanHoose, David D. (2008). “Alternative measures of marginal cost and inflation in estimations of new Keynesian inflation dynamics.” *Journal of Macroeconomics*, vol.30, pages 928-940.

Gordon, Robert J. (1990). "U.S. Inflation, Labor's Share, and the Natural Rate of Unemployment." In *Economics of Wage Determination* (Heinz Konig, ed.). Berlin: Springer-Verlag.

Larrain, Borja & Yogo, Motohiro. (2008). “Does frim value move too much to be justified by subsequent changes in cash flow?” *Journal of Financial Economics,* vol. 87, pages 200-226.

Matsuura, Kenji & Willmott Cort J. (2005). “Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance.” *Climate Research,* vol. 30, pages 79-82.

Mirmirani, S. & Li, H., (2001). “The United States' Inflation Forecasting with Neural Networks.” *International Review of Economics and Business*.

Önder, A. Özlem. (2004). “Forecasting Inflation in Emerging Markets by Using the Phillips Curve and Alternative Time Series Models.” *Emerging Markets Finance & Trade,* vol. 40, no. 2, pages 71–82.

Rossi, Barbara & Sehkposyan, Tatevik. (2013). “Conditional predictive density evaluation in the presence of Instabilities.” *Journal of Econometrics,* vol. 177, pages 199-212.

Rumler, Fabio & Valderrama, Maria Teresa. (2010). "Comparing the New Keynesian Phillips Curve with time series models to forecast inflation," *The North American Journal of Economics and Finance*, Elsevier, vol. 21(2), pages 126-144.

Stock J, Watson MW. (2009). "Phillips Curve Inflation Forecasts." In *Understanding Inflation and the Implications for Monetary Policy (Fuhrer J, Kodrzycki Y, Little J, Olivei G). Cambridge: MIT Press*, pages 99-202.

Vega, Marco & Winkelried, Diego. (2005). "Inflation targeting and inflation behavior: A successful story?" *International Journal of Central Banking*, vol. 1.