



ECONOMICS WORKING PAPERS

No. 15

MACROECONOMIC FORECASTING
AND THE PHILLIPS CURVE

Editors:

Peter Griffin
Gareth Olds
J.J. Ooi
Alex Robinson
Karl Storchmann

Benjamin P. Keefer

Fall 2008

WHITMAN COLLEGE



Macroeconomic Forecasting and the Phillips Curve

Benjamin P. Keefer

To copy, abstract, post on servers, or otherwise redistribute is not to be done without the express permission of the author(s).

WORKING PAPER No. 15

MACROECONOMIC FORECASTING AND THE PHILLIPS CURVE

by

Benjamin P. Keefer

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in Economics.

Whitman College

2008

Contents

1	Introduction to Forecasting	2
2	Forecasting Inflation	17
3	Difficulties in Macroeconomic Forecasting	25
4	Empirical Section	30
5	Conclusion	45
A	Deriving the Lucas Rapping Regression	45
B	Instrumental Variables and Data Sources	47

Abstract

In this paper, I first review the historical developments in economic forecasting and then focus on the Phillips curve, one of the most controversial and highly-debated inflation forecasting models.

In the empirical portion of the paper, I extend Robert Lucas and Leonard Rapping's (1969) analysis and then attempt to forecast inflation, much in the spirit of James Stock and Mark Watson's (1999) paper.

1 Introduction to Forecasting

Forecasting is a discipline focused on predicting the future. According to Graham Elliott and Allan Timmerman (2008), forecasting's use lies in its ability to inform present decisions. Forecasting can be used to predict which outcomes are most likely, either through a single point estimate or an interval estimate, which shows a range of values for a chosen confidence level. Policy-makers forecast inflation when trying to determine the appropriate interest rate, financial analysts employ forecasting to select investments, and public officials need forecasting in order to estimate the potential future returns of a capital-intensive infrastructure project in order to assess whether it will have a positive net present value (Graham and Timmerman 2008).

How can we produce forecasts? There are two general types of methods: in-sample and out-of-sample forecasts. Although it may seem that we would only need one forecasting method, in actuality, we need two. The reason is that each method has a specific purpose. To illustrate the differences, consider the simple model:

$$y_{t+1} = \alpha + \beta y_t + \epsilon_t.$$

Suppose our data have one observation per year and cover the years from 1980 to 2000. If our objective were to *explain* the changes in y_t over time, then in-sample forecasting would be appropriate. An in-sample forecast will use all of the available data when estimating parameters. While these parameters

may fit the data, the only way to tell how well these coefficients will match data outside of the time sample will be to use out-of-sample forecasting. However, an in-sample forecast is appropriate for historical analysis. As we will see later on in this paper, Robert Lucas and Leonard Rapping (1969) use in-sample forecasting in order to evaluate the influence of price and wage changes on the unemployment rate.

The key difference between in-sample and out-of-sample forecasting is that out-of-sample forecasting was designed to simulate the real-world forecasting process. For example, we will see that when Stock and Watson (1999) try to forecast future inflation rates, they restrict themselves to data available at least one year prior to the observation they are trying to predict. While they could use in-sample forecasts to model inflation up to that point in time, there is no guarantee that a model excelling at in-sample predictions will be able to accurately predict future inflation rates. In order to determine whether a given model will predict values outside of the time sample accurately, they turn to out-of-sample forecasts.

To determine whether a model is likely to accurately predict future values, we try to judge how well it predicted out-of-sample in the past. First we take the data sample we wish to analyze and partition it into two disjoint pieces; we will choose the first S observations for our regression sample, and the remaining P observations for our prediction sample (Elliott and Timmerman 2008). For example, suppose we select the years 1980 to 1990 as our regression sample and 1991 to 2000 as our prediction sample. We would use the first

interval to estimate the parameters α and β from our original model. Then our prediction for the year 1991 would be

$$\hat{y}_{1991} = \alpha + \beta y_{1990},$$

in which \hat{y} refers to our predicted value and y refers to our actual observation.

We could also estimate

$$\hat{y}_{1992} = \alpha + \beta y_{1991}.$$

We would continue until we had obtained a sequence of ten predictions

$$\{\hat{y}_{1991}, \hat{y}_{1992}, \hat{y}_{1993}, \dots, \hat{y}_{2000}\}$$

and compared each to the sequence of actual values

$$\{y_{1991}, y_{1992}, y_{1993}, \dots, y_{2000}\}.$$

To measure the accuracy of our predictions, we generally use the Root Mean Square Error, or RMSE for short. This measurement is the time-series analogue to the adjusted R^2 value common in cross-sectional data. Thus, the RMSE typically measures the *explanatory* power of our model. To calculate the RMSE, we would compute

$$RMSE = \sqrt{\frac{(\hat{y}_{1991} - y_{1991})^2 + (\hat{y}_{1992} - y_{1992})^2 + \dots + (\hat{y}_{2000} - y_{2000})^2}{10}}. \quad (1)$$

Note that the RMSE can also be used for in-sample forecasts as well. Once we obtain the predicted values for our in-sample forecasts, we can then calculate the RMSE corresponding to our predictions using Equation 1.

Although the process described in the preceding paragraph is likely to produce relatively accurate out-of-sample predictions, it is possible to produce even better forecasts. Consider that when we forecasted y_{1992} we only used the years 1980 to 1990 to estimate α and β *even though the datum was available for the year 1991*. If we estimated α and β using the years from 1980 to 1991 under this new procedure, then we would not only have more data but, more importantly, we would also have more recent data, both of which would help us produce more accurate results. To produce more accurate results, we would want to run a separate regression to estimate α and β using all available data for each observation in the prediction sample. We would continue until we exhausted all observations in the prediction sample. Although this process is more time intensive, the increased accuracy of the forecasts is often worth the expense.

Now suppose that we have a second model, given by

$$y_{t+1} = c + \lambda x_t.$$

Using the out-of-sample forecasting process and the RMSE, we could determine which model is more likely to produce more accurate predictions. While it may seem counterintuitive at first, it is also possible to achieve even

more accurate predictions by combining forecasts. How would we accomplish this? Let $f_1(t)$ and $f_2(t)$ denote our forecasted values for some variable y . We could then combine these forecasts by regressing

$$y_t = \alpha_0 + \alpha_1 f_1(t) + \alpha_2 f_2(t)$$

to estimate α_0 , α_1 , and α_2 .

In fact, combined forecasts are useful because they help minimize a common tradeoff in forecasting between accuracy and precision (Elliott and Timmerman 2008). Due to the prevalence of short-time regression samples for parameter estimation, either because of the instability of our data series or because of data availability issues, the degree of freedom of our regression sample is usually restricted. Constructing a complicated model including all of the necessary variables may lead to imprecise estimates of the coefficients. Conversely, a simple model including only a few of the key variables is likely to produce biased forecasts. However, a combined forecast of multiple models is likely to be relatively precise and unbiased.

When choosing a model, there are several types to consider, three of which are judgmental, structural, and nonstructural. A judgmental model is one that employs human judgment to infer future behavior. For a judgmental model, the term model is deceptive. For example, a judgmental model may have no numbers, may not be replicable, and will likely have estimates that depend on whoever is making the judgment. It is only a model in the

sense that we produce expectations or predictions based on our observations, which could be data, anecdotal evidence, or our intuition. An example of a judgmental model in practice is the Beige Book provided by the Federal Reserve Board, which gives anecdotal evidence, obtained from interviews with businesses, economists, and experts, of the direction of the economy in the near future. This evidence is then interpreted by people to form predictions about the strength of the economy in the near-term.

In practice, judgmental models are somewhat rare in economics, though they may be highly accurate. According to Ben Bernanke (2007), Federal Reserve predictions are often adjusted to incorporate the predicted value of the models and the judgment of the policymakers.

The two most common models in economics, and the ones we will focus on for the remainder of this paper, are structural and nonstructural models. Structural models rely on economic theory and assumptions about the economy to predict the future values of the dependent variables. When researchers estimate a system of equations with consumption and investment determined by independent variables, they are employing structural models to forecast the direction of the economy. Another example of a structural model is the Phillips curve, which we will discuss in the next section.

In contrast, nonstructural models are highly non-theoretical and use past data to predict future movements. For data exhibiting high serial correlation or in which the theoretical underpinnings are not established, nonstructural models are frequently used for prediction. One example would be to predict

inflation using only its past values. For more information about the modeling process, see A. H. Studenmund's *Using Econometrics* (2006).

Francis X. Diebold (1998) traces the growth of forecasting within economics to the development of Keynesian theory in 1936. Using Keynes' work as a springboard, economists sought to model the decision rules of agents (consumers, investors, and governmental) in the model using systems of equations. Each of these equations represented facets of the economy's underlying structure, and this approach became known as structural forecasting. Structural forecasting owed its development to Keynes and to economists' growing acceptance of his theories.

Diebold identifies the breakdown of structural forecasting with growing dissatisfaction over the performance of its models. Some economists objected to the absence of theoretical underpinnings of disequilibrium, while others objected to how expectations were treated.

In fact, modeling expectations has been one of the most controversial topics in the field of macroeconomics. At first, naïve forecasters treated economic actors as having myopic and static expectations. Subsequent forecasters, however, allowed for dynamic expectations. The structural forecasters then modeled expectations as if they were adaptive. On the one hand, this assumption allows for easy estimations of econometric models. On the other hand, adaptive expectations are not very realistic.

Proponents of rational expectations faulted the use of static and adaptive expectations, which assumed that individuals expectations were informed

only by the past or not informed at all. They argued instead that individuals have the capacity to anticipate future events rather than simply adapting their expectations after the fact. Even more damaging to structural forecasting was Lucas' (1976) critique, in which he stated that structural models were inherently unstable because their parameters continually changed. He argued that when public officials tweaked policy, economic agents adjusted their decision making rules, thus necessitating constant estimates of parameters. Since policy is constantly changing, any estimates are likely to be biased.

Further criticisms mounted from the models' poor performance during the 1970s with the rise of stagflation. According to Flint Brayton, Andrew Levin, Ralph Tryon, and John C. William's (1997) article, one of the best-known structural models of this time period was the FRB-MIT-PENN model, developed in 1966 by Franco Modigliani, Frank de Leeuw, and Albert Ando. This model consisted of 60 equations but depended on the IS-LM framework for demand-side analysis and the neoclassical growth models to predict changes in aggregate supply. In spite of its lack of sophistication, Charles R. Nelson (1972) found that a simple nonstructural model consistently outperformed the FRB-MIT-PENN model at forecasting in terms of out-of-sample predictive ability.

Some work has been done to correct the deficiencies of the structural model by implementing rational expectations and better evaluating forecasting performance. According to Diebold (1998), these new models are em-

ployed by the Federal Reserve Board and the IMF, institutions which are concerned with the effects of changes in policy. Yet, significant attention also turned to nonstructural modeling, which provided more robust forecasts though at the expense of the cause-and-effect information provided by the structural models.

The development of nonstructural modeling predated the rise of Keynesian Economics, beginning in the 1920s with the work of Eugene Slutsky (1937), originally printed in 1927 in Russia, and George Yule (1927). These two pioneers found that simple autoregressive processes, in which the previous values of a process were used to forecast future values (e.g., $x_t = f(x_{t-1}, x_{t-2}, \dots)$), provided good predictions for numerous time series. They also studied moving average processes, in which the current value of a system is modeled as a function of previous shocks (or previous error terms). Examples of several types of nonstructural models are given in Table 1.

In 1938, Herman Wold showed that the stochastic component in a time series can be modeled using the methods introduced by Slutsky and Yule. Thirty-two years later, George Box and Gwilym Jenkins' (1970) seminal work on nonstructural time series advocated the use of Autoregressive Moving Average (ARMA) models, which Box and Jenkins showed were equivalent to both moving average and autoregressive models but could be represented with fewer terms. In fact, automatic ARMA modeling can produce forecasts that are competitive to those produced by professional forecasters (Jan G. De Gooijer and Rob J. Hyndman 2006). Box and Jenkins also provided a

Table 1: Example Nonstructural Models

Autoregressive (AR)	$y_t = c_1y_{t-1} + c_2y_{t-2} + \epsilon_t$
Moving Average (MA)	$y_t = c_3\epsilon_{t-1} + c_4\epsilon_{t-2} + \eta_t$
Autoregressive Moving Average (ARMA)	$y_t = c_1y_{t-1} + c_2y_{t-2} + c_3\epsilon_{t-1} + c_4\epsilon_{t-2} + \eta_t$
Autoregressive Integrated Moving Average (ARIMA)	$\Delta y_t = c_2y_{t-2} + c_4\epsilon_{t-2} + \eta_t$
Exponential Smoothing	$y_t = \beta y_t + \beta(1 - \beta)y_{t-1} + \beta(1 - \beta)^2y_{t-2} + \dots$
Vector Autoregressive (VAR) System	$y_t = c_1y_{t-1} + c_2y_{t-2} + c_3u_{t-1} + c_4u_{t-2} + \epsilon_t$ $u_t = c_1y_{t-1} + c_2y_{t-2} + c_3u_{t-1} + c_4u_{t-2} + \epsilon_t$

Here y_t and u_t represent two time-series with ϵ_t , or η_t when the residual ϵ_t is being used as an independent variable, representing the classical error term.

methodology to design AR, MA, ARMA, and ARIMA models based on three steps: identification, estimation, and verification. This approach was called the Box-Jenkins approach while AR, MA, AMRA, and ARIMA models are each generally referred to as a Box-Jenkins method.

Christopher A. Sims (1980) later extended the univariate framework in Box and Jenkins work to create Vector Autoregressive (VAR) models. These model two or more endogenous variables simultaneously using the lagged values of all endogenous variables. An example system is given by

$$y_t = c_1y_{t-1} + c_2y_{t-2} + c_3u_{t-1} + c_4u_{t-2} + \epsilon_t$$

$$u_t = c_1 y_{t-1} + c_2 y_{t-2} + c_3 u_{t-1} + c_4 u_{t-2} + \epsilon_t,$$

in which GDP (y_t) and unemployment (u_t) are modeled using the lagged values of both variables. VARs, however, generally lead to overfitting the data because they estimate too many coefficients (Gooijer and Hyndman 2006). Moreover, in this framework, the causality is inherently muddled. While it may produce accurate predictions, this VAR will likely provide little information as to whether y_t causes changes in u_t or vice versa.

One later development in nonstructural forecasting is cointegration, as discussed by Diebold (1998). A humorous example of the difference between cointegration and a random walk is given by Michael P. Murray (1994). In his example, we first imagine that a drunk woman and her dog are leaving the bar together. If the woman were to forget her dog's leash, then in the absence of conscious thought, we would expect each of their paths to follow a random walk. That is, at any time t , in order to stabilize herself, the drunk will have to put a foot down in the direction in which she is unbalanced. Assuming the such directions are random, we would have a random walk. Similarly for the dog, if we assume the dog follows whatever scent comes his way, and if one and only one new scent develops in each time period, we would expect that the dog's path would follow a random walk. Thus, the change in position between two sequential time periods, denoted as $x_t - x_{t-1}$, may be thought of as random white-noise (e.g. ϵ_t , a normally distributed variable with a mean of zero).

If, however, the woman were trying to find the dog (to avoid being woken up in the middle of the night when the dog would inevitably want to be let in) and if the dog were trying to find the woman to ensure being fed in the morning, then the paths are no longer necessarily random walks. That is, unlike the first scenario, we would expect that the dog and woman would significantly deviate from their starting positions in order to find the other. That is the woman, whose position is denoted x_t , would generally walk in the direction of the dog, as given by

$$x_t - x_{t-1} = \epsilon_{t,y} + c_x(y_{t-1} - x_{t-1})$$

where $(y_{t-1} - x_{t-1})$ denotes direction, either positive or negative. Similarly, the dog, whose position is denoted as y_t , walks according to a similar model:

$$y_t - y_{t-1} = \epsilon_{t,x} + c_y(x_{t-1} - y_{t-1}).$$

In these models, ϵ can be thought of as a normally distributed variable with a mean of zero and c_x and c_y are parameters affected by the drunkenness and length of strides of the two beings in question. In this scenario, the random components are still present in both the dog and woman's paths. However, if $\epsilon_{t,x}$ and $\epsilon_{t,y}$ exhibit identical distributions, then the distance separating the woman and her dog $x_t - y_t$ is largely deterministic. Thus, we can cointegrate the series to minimize the randomness and to assist in predictions.

Yet, even as nonstructural models have grown in popularity, structural

models in the form of dynamic stochastic general equilibrium models, referred to as DSGE by Diebold, have also been developed. In 1982, Kydland and Prescott employed a DSGE to explain business cycle movements as a result of technological shocks, giving rise to the real business cycle theory. Rather than trying to use decision rules of consumers and investors to predict future economic activity, DSGE models instead rely on tastes and technology. Yet even though these models offer significant promise, they are not without disadvantages. Often it is necessary to calibrate the models by estimating parameters (either by previous research or “common sense”), and while these models identify causal links, according to James Stock and Mark Watson (2001), they generally fit the data rather poorly.

Chris Chatfield (1997) discusses some advances in non-linear forecasting that may be promising for future research. One example is a threshold AR model, a piecewise-defined AR model, depending on the previous state of a system. For example, denote real GDP at time t as y_t and let $x_t = 1$ if the economy is in a recessionary state and zero otherwise. Then to forecast real GDP, we would need to first predict x_t and then use these values to estimate the other parameters given by the model:

$$y = \begin{cases} a_0 + a_1 y_{t-1} & \text{if } x_t = 1 \\ b_1 + b_1 y_{t-1} & \text{in } x_t = 0. \end{cases} .$$

Another advance has been the development of neural network, in which algorithms attempt to tease out the relationships between exogenous and en-

dogenous variables. However, neural networks are often unpractical because they require copious amounts of data.

So where does that leave us? Paul Newbold and Clive W. Granger (1974) suggest the following guidelines for univariate forecasts:

- For fewer than thirty observations, exponential smoothing (forecasting using lagged errors with exponentially decreasing weights) is generally good.
- Between 30 and 40 observations, exponential smoothing, in combination with stepwise autoregressive, appears to be best.
- For more than 40 or 50 observations, Box-Jenkins methods will produce good results and will also work well with intractable time series

Newbold and Granger also note that although Box-Jenkins ARMA and ARIMA models generally outperform others, there are also associated time and skill costs. In contrast, exponential smoothing and stepwise autoregressions (using lagged changes) are almost automatic.

For multivariate data, vector autoregressive models are often good. Sims (1980) argues that structural models are flawed because of unknown correct functional specifications, unrealistic assumptions and the faulty treatment of expectations. In comparison, he believes that since vector autoregressive models relax the need for proper identification, they will consequently lead to better forecasting. Stephen Hall (1995) responds that without identifying shocks in the system and without any use of the theory enjoyed by structural

models, VARs yield poor estimates. He argues VARs should only be used to help choose the parameters in structural models. Moreover, he states that with advancements in cointegration, in the treatment of expectations, and greater focus on long-term system behavior, these approaches are slowly converging to a singular methodology. Stock and Watson (2001) similarly argue that the value of VARs lies in their ability to fit the data and provide reasonable estimates for the magnitude of causal effects when constructed in accordance with economic theory. However, they are less useful for policy analysis and for distinguishing correlation and causation, which some authors refer to as the identification problem.

Granger (1996) offers forecasters some practical advice. First, he urges forecasters to provide better information about uncertainty, and rather than providing 95 percent confidence intervals for their forecasts—which are too wide for practical use—he urges the use of 50 percent confidence intervals. In addition, given that specification is a major issue in macroeconomics, he thinks that heteroskedasticity deserves greater attention from forecasters, since specification error may lead to heteroskedasticity and biased estimates of confidence intervals. Furthermore, when trying to forecast further into the future, he advises practitioners to pay attention to the relative strength of the signal compared to the noise. If the signal is relatively weak, long-term forecasts are likely to be inaccurate. Models that only use historical data (or previous data) will only be accurate “to the extent that history repeats itself.”

In systems with structural breaks, evidenced by when forecasts are consistently off target, he advocates the use of leading indicators. Remember that in a typical VAR, we might use lagged values of u_t and y_t to try to predict one time period ahead. In contrast, a leading indicator could be correlated with a value of u_t or y_t that is four or more time periods ahead. Further, in data where errors are primarily associated with outliers, he advises forecasters to expend more effort predicting the unexplained breaks in a time series a priori. One such approach is the Markov-Switching model, also called the regime-switching model. These models are very similar to threshold autoregressives; the only difference is that autoregressives imply that a model only includes lagged values of a variable, whereas regime-switching models may be multivariate. James D. Hamilton (1989) uses the regime-switching model to explain the variation in economic growth by first forecasting the probability of a recession. In his model, a recession naturally coincides with a low-growth state and an expansion with a rate of higher growth.

2 Forecasting Inflation

The Phillips curve, in its original form, summarizes an inverse empirical relationship between wage inflation and the unemployment rate. A. William Phillips' widely cited article of 1958 studied wage and unemployment movements in the United Kingdom from 1861 to 1957. Phillips first conceived of inflation in terms of rising wages. He reasoned that changes in money wage

rates were sensitive to shortages and surpluses in labor. In fact, he states that “when the demand for labour is high and there are very few unemployed we should expect employers to bid up wages quite rapidly.” Thus, when the unemployment rate was low, the low levels of available workers caused wages to increase. These wage increases then caused increases in price level. Phillips suggested that policymakers could use this relationship to choose a desired combination of wage inflation and unemployment.

Phillips was not the first to notice a relationship between inflation and unemployment. In a reprint of his 1926 article, Irving Fisher (1973) finds a 90 percent correlation between inflation and unemployment. However, instead of modeling inflation as a function of unemployment, Fisher models unemployment as a function of inflation. Rather than arguing that low unemployment rates drive wage increases, he reasons that price changes, especially unanticipated price changes, would cause the real wage to fall and the quantity of labor demanded to rise. With a higher level of greater demanded, the unemployment rate would be reduced. At the conclusion of the paper, he notes that the true relationship is likely more complex than the one he hypothesized, and he posits that a stabilization of price levels is consistent with a stabilized level of unemployment.

It was not until Phillips’ (1958) ground-breaking study, however, that the tradeoff between unemployment and inflation was duly noticed by the economics profession. Within ten years after Phillips’ publication, the Federal Reserve incorporated the Phillips curve into the FRB-MIT-PENN model, as

previously mentioned.

Paul Samuelson and Robert Solow (1960) brought the Phillips curve analysis to the United States data, though with some modifications. Like Fisher, they used percentage changes in overall price levels instead of changes in wage levels as Phillips did. Like Phillips, they tried to explain changes in the inflation rate using the unemployment rate. Using these modifications, they found evidence that the Phillips curve shifted outward following the second World War, indicating that after the war, each rate of unemployment was now associated with a higher level of inflation.

This instability would later be documented by other researchers (Robert Lucas and Leonard Rapping 1969). Perhaps the most interesting finding of Samuelson and Solow's study was that price stability was associated with an unemployment rate of approximately 5.5 percent. This level of unemployment would later be referred to as the natural rate of unemployment, although the level would change throughout the years reflecting changes in the economy.

Within ten years, objections to the existence of a long-term and stable Phillips curve mounted. One of the first objections came from Richard Lipsey (1960). He showed that while the variation in wages could be explained by the changes in the unemployment rate, in the rate of change of the unemployment level, and in the rate of change of the price level over the time interval selected by Phillips, the relationship was inherently unstable. Moreover, the relative explanatory power of each variable shifted from 1862 to 1957. As a result,

estimates of parameters in the Phillips curve would be associated with low confidence levels, limiting the usefulness of the Phillips curve to policymakers.

While Lipsey's findings weakened the Phillips curve's empirical basis, Milton Friedman's (1968) conjecture that unemployment was insensitive to anticipated inflation cast doubt on its theoretical underpinnings. He argued that once workers began to anticipate inflation, they would demand pay raises accordingly. At a higher level of pay, the quantity of workers demanded would fall, and unemployment would return to its initial level. The implication is that only unanticipated inflation could reduce unemployment. Thus, in order for policymakers to keep unemployment below its natural rate, they would be forced to increase the inflation rate by a steady amount each year, which would ultimately be unsustainable. Therefore, he argued that people's expectations preclude a long-term tradeoff between inflation and unemployment. He stated that the role of monetary policy makers should therefore be to keep price levels relatively stable.

Building upon Friedman's criticism, Lucas and Rapping (1969) rejected the claim of a stable long-term tradeoff between unemployment and inflation using data from 1904 to 1965. While Friedman provided a theoretical argument against the notion of a stable long-term tradeoff, Lucas and Rapping's work provided an econometric refutation. Using newly-developed microeconomic rationales to explain the Phillips curve, Edmund Phelps (1969) also concluded that any movement in the labor market away from its natural state would lead to rapidly accelerating inflation.

Friedman (1977) conjectured that instead of a long-term tradeoff between unemployment and inflation, the long-run Phillips curve was likely vertical. The implication was that any deviation in unemployment below the natural rate of unemployment would not only bring about inflation but would also cause it to accelerate. He postulated that instead of bargaining for nominal wages, as assumed by previous proponents of the Phillips curve, workers bargained for real wages. Thus, workers would demand wage increases above and beyond the increases in price level whenever unemployment was below its equilibrium, or natural rate.

Since then, the Phillips curve has been controversial. Finn Kydland and Edward Prescott (1982) emphasize the importance of the notion of the classical dichotomy, which claims that nominal variables, such as inflation, and real variables, such as unemployment, are unrelated even in the short-run. Kydland and Prescott proposed the real business cycle model, which assumes that business cycles depend exclusively on real variables, such as technological shocks, with nominal variables having no impact on changes in real output. Greg Mankiw (2001) counters that the classical dichotomy falls apart in the short-term because prices are sticky, and so changes in monetary policy cause inflation and unemployment to move in opposite directions, while Alan Blinder (1997) calls the empirical relationship described by the short-run Phillips curve “the clean little secret of Macroeconomics.” In an attempt to reconcile the notion of the classical dichotomy in the long-run with the apparent existence of the Phillips curve in the short-run, Mankiw (2001)

admits that while the theoretical basis for the short-run Phillips curve has a firm foundation in Keynesian economics, the long-run dynamics are not well understood. He implies that the existence of a short-run Phillips curve depends only on the existence of a short-run tradeoff between inflation and unemployment.

In fact, Stock and Watson (1999) find that the Phillips curve and similarly specified models provide the best out-of-sample forecasts for the years 1959 to 1997. In their article, they conclude that the variables measuring current economic activity, broadly defined, produced the most reliable predictions for future inflation. The three most promising variables across both forecasting samples considered, 1970–1983 and 1984–1997, were unemployment, the rate of capacity utilization, and a composite index of 168 activity measures. Andrew Atkeson and Lee Ohanian (2001), however, find that for the years 1986–2001, this relationship appears to have broken down. In their analysis, the forecasts provided by the Phillips curve do worse than the naïve forecast, measured as the average inflation rate for the last twelve months.

In a later paper, Stock and Watson (2005) reconcile their findings with Atkeson and Ohanian's. While Stock and Watson admit that forecasts from the Phillips curve may perform worse than the naïve forecast, they note that this is likely a result in the decreased volatility of inflation. In other words, if deviations in inflation are small, then it is likely that a larger portion of the variance in inflation is white noise, which, if true, would reduce the difference in predictive powers between the best and worst models. They

argue, however, that claims of a break-down of the Phillips curve are not necessarily true because the relationship between the real activity variables and inflation rate is still present, even if it is weaker.

Others question the appropriateness of using the Phillips curve relationship at all. James Galbraith (1997) argues that the non-accelerating-inflation-rate-of-unemployment (NAIRU), from which the Phillips curve is based, has been subject to unexplained drifts over time. Using historical analysis, Robert Gordon (1997) argues that the NAIRU follows a random walk, which, if true, suggests that the NAIRU cannot be known at any given point in time. Consequently, Galbraith argues that policy-makers should abandon policies based on the Phillips curve. Thomas Laubach (2001) questions whether this relationship is even significant. In his analysis of seven industrialized countries, he wonders whether his estimates for NAIRU for non-U.S. countries are actually valid. Similarly, Robert Gordon (1997) states that the inflation process in Europe and Japan is fundamentally different than the Phillips curve characterizing U.S. inflation.

Even in the U.S., estimates are not easily obtained. Douglas Staiger, James Stock, and Mark Watson find that although their estimates of the NAIRU have moved from 4.9 in 1966 to 7.0 in 1978 and around 5.7 in 1994, the spread of 95% confidence intervals of such estimates vary from about 1.8 to over 5.0, or as George Akerlof (2002) notes, over three times the variation in the unemployment rate for the last fifty years. Staiger, Stock, and Watson suggest that although some could interpret the wide confidence intervals as

an indication that the NAIRU does not exist, they believe that the intervals are so wide because the data is noisy.

If we assume that the NAIRU exists, then the long-run Phillips curve must be vertical. Any other rate of unemployment would either trigger substantial deflation or substantial inflation over prolonged periods of time. That is, for price levels to be stable, the markets must clear and any deviation from the market-clearing point is unsustainable. Akerlof (2002) argues that this analysis is likely overly simplistic. Citing the exceedingly high, and sustained, unemployment rates during the Great Depression and the lack of significant deflation, he believes that the natural rate hypothesis does not hold. Moreover, he reasons that low levels of inflation promote real GDP growth and lower levels of unemployment. It is only when inflation levels are high that wage negotiations cause inflation to spiral out of control. Otherwise, workers ignore inflation and bargain for nominal, and not real, wages. This assumption is justifiable due to rational ignorance. However, levels of inflation that are too low would inhibit employment and growth. Thus, any relationship between unemployment and inflation must incorporate these nonlinearities and is likely to be complex. Akerlof's main contention is with monetary policies in Europe and Canada, which may have been overly restrictive in trying to tame inflation, resulting in significant and prolonged unemployment.

Peter Ireland (1999) argues that past bouts of inflation were caused by a combination of supply shocks causing unemployment to deviate from its natural rate and of policymakers' erroneous belief that there was a long-run

tradeoff between unemployment and inflation during the 1960s and 1970s, they allowed unemployment to fall below the higher non-inflationary level induced by the supply shock, consequently leading to higher inflation in the long term. Ireland believes that periods of higher inflation are unavoidable in the future unless policymakers start to place greater emphasis on combating inflation instead of promoting full-employment and unless they begin to develop the capability to identify changes in the NAIRU. Laurence Ball and Greg Mankiw (2002) subscribe to a different view. They argue that supply-shocks—whether in the form of government policy, changes in demography, and in productivity—shift the tradeoff relationship between unemployment and inflation. According to their analysis, the tradeoff still exists, though is affected by outside variables.

3 Difficulties in Macroeconomic Forecasting

Two of the biggest challenges facing macroeconomic forecasters are unit roots and serial correlation. We have already mentioned unit roots, although we called them random walks (Studenmund 2006). The consequences of both of these phenomena can be severe. For example, if our dependent variable is a random walk, then its changes are random. However, our regressions assume that the changes in the dependent variable are driven by changes in the independent variables, resulting in biased t -statistics and a biased R^2 .

Serial correlation is likewise troubling. In macroeconomic models, the

presence of serial correlation not only biases the estimates of our standard errors but may also bias estimates in the presence of lagged dependent variables (Studenmund 2006). Moreover, Studenmund says that in such models, serial correlation is harder to detect because the DW statistic is biased toward 2.

Given that the DW statistic is no longer an unbiased identifier of serial correlation, we can either employ the Durbin m statistic or the Breusch-Godfrey test, also called the LM test (Peter Kennedy 2003). For our purposes, both will be useful. In the Breusch-Godfrey test, we first run our model, collect p lagged residuals, and then use these lagged residuals as regressors in the model. Note that if our model is a MA, ARMA, or ARIMA model, we would first include the necessary residual terms $\epsilon_t, \epsilon_{t-1}, \dots,$ and ϵ_{t-n} for our model. Then we would look at the $\eta_t, \dots, \eta_{t-p}$, the p lagged residual of our model once the ϵ terms have been included. If an F test suggests that the lagged residuals have a statistically significant effect on the dependent variable, then this is evidence of serial correlation (Kennedy 2003). The Durbin m test is very similar. The only difference is that instead of choosing p lagged values, we choose one lagged value and use the probability value as evidence of serial correlation. Thus, the Durbin m test is useful for identifying first-order serial correlation in the presence of a lagged dependent variable while the Breusch-Godfrey test identifies p^{th} order serial correlation.

Once we have identified that serial correlation exists, we will need to

use the $AR(1)$ method, an extension of the Cochrane-Orcutt method described below (Studenmund 2003). By itself, the Cochrane-Orcutt method cannot identify whether serial correlation exists. Instead we will first need the Breusch-Godfrey or Durbin m test. However, once we have identified that serial correlation exists and if we suspect the serial correlation is of the first order, then we can use the Cochrane-Orcutt method to generate a new equation without first order serial correlation. To make sure that higher order serial correlations are not present, we could then test this new equation using the Breusch-Godfrey or Durbin m tests.

After running $AR(1)$ in the empirical section, we could find no evidence for further serial correlation when replicating Lucas and Rapping's (1969) research. The $AR(1)$ method was sufficient. When trying to replicate Stock and Watson's (1999) out-of-sample forecasting research, we ignored serial correlation and relied exclusively on the Schwarz Criterion and Akaike's Information Criterion. Although the presence of serial correlation may bias our estimates and result in reduced out-of-sample performance, the $AR(1)$ and Cochrane-Orcutt methods are likely to produce a biased estimate of the serial correlation over the small regression samples we will use (Studenmund 2003). Moreover, these methods will also change the interpretation of the dependent and independent variables, which would also be problematic. For these reasons, we will ignore serial correlation when conducting our out-of-sample forecasting.

Suppose we have a simple time series model with first order serial corre-

lation, in which our dependent variable y_t depends only on a constant α_0 and a single independent variable, denoted x_t . Then we could write our model as follows:

$$y_t = \alpha_0 + \alpha_1 x_t + \epsilon_t$$

such that

$$\epsilon_t = \rho\epsilon_{t-1} + \eta_t.$$

This would be an example of first order serial correlation because ϵ_t , the error term in our regression, depends on its lagged value and a classical random error term, denoted η_t . This type of serial correlation is first order because the error term depends only on its first lagged value. In this case, we might notice that

$$\rho y_{t-1} = \rho\alpha_0 + \rho\alpha_1 x_{t-1} + \rho\epsilon_{t-1},$$

and then

$$y_t - \rho y_{t-1} = \alpha_0(1 - \rho) + \alpha_1(x_t - \rho x_{t-1}) + (\epsilon_t - \rho\epsilon_{t-1}).$$

Since $\epsilon_t - \rho\epsilon_{t-1} = \eta_t$, then we have

$$y_t - \rho y_{t-1} = \alpha_0(1 - \rho) + \alpha_1(x_t - \rho x_{t-1}) + \eta_t,$$

where η_t is not serially correlated.

Thus, if we can identify ρ , then we can easily get rid of first order serial

correlations in our models. In the Cochrane Orcutt method, we first regress our first model

$$y_t = \alpha_0 + \alpha_1 x_t + \epsilon_t.$$

Then we collect the residuals and run the regression

$$\epsilon_t = \rho \epsilon_{t-1} + \eta_t.$$

From here, we create new variables x_t^*, y_t^* with

$$x_t^* = x_t - \rho x_{t-1}$$

and

$$y_t^* = y_t - \rho y_{t-1}.$$

We then iterate this process until our value of ρ has converged. This usually occurs after relatively few iterations (Studenmund 2003). The $AR(1)$ approach is almost identical to the Cochrane Orcutt method except that it uses nonlinear techniques to compute superior estimates of the standard errors. According to Studenmund (2003), $AR(1)$ will actually estimate α_0, α_1 , and ρ jointly. In EViews, we implement the $AR(1)$ method by adding $AR(1)$ as an independent variable in our regression.

4 Empirical Section

The fundamental relationship described by the Phillips curve is that any deviation in unemployment from its natural rate is associated with a deviation in inflation from its expected rate. Let u_t denote the unemployment rate and π_t denote inflation at a time period t . Denote the natural rate of unemployment and expected inflation rate at time t as \bar{u}_t and π_t^e , respectively. Then we would write

$$(u_t - \bar{u}_t) = \alpha(\pi_t - \pi_t^e)$$

if we were interested in analyzing unemployment or

$$(\pi_t - \pi_t^e) = \frac{1}{\alpha}(u_t - \bar{u}_t)$$

if we were more interested in forecasting inflation. To see a picture of how these Phillips curves are related, see Figure 1. Note that both sides of Figure 1 illustrate the tradeoff that policymakers face between inflation and unemployment. From the figure, we see that policymakers can achieve a lower unemployment rate in the short-run only at the expense of an increase of inflation above the expected value. In the long-run, however, as people adjust their expectations, deviations from the natural rate of unemployment prove unsustainable, and so inflation will have no effect on the long-run rate of unemployment.

However, unemployment is not the only variable we can use to specify the

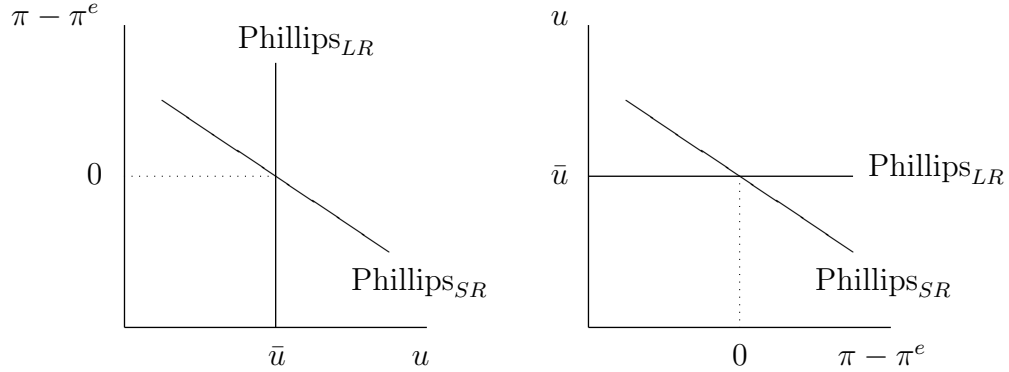


Figure 1: Two Ways of Representing the Phillips Curve

Phillips curve. After all, labor is only one input in the production process. Capital is another. Based on Stock and Watson's (1999) conclusions, we may be equally interested in using capacity utilization instead of labor. Stock and Watson find that forecasts based on capacity utilization provide superior forecasts to those based on unemployment. If we denote the manufacturing capacity gap in our system as $(\bar{x}_t - x_t)$, with \bar{x}_t equal to one and x_t the percentage of the potential capacity currently in use, then the difference measures the excess manufacturing capacity, which is the variable we will use in our regressions later. If we wanted to include these variables in the Phillips curve, then we could rewrite the Phillips curve as

$$(\pi_t - \pi_t^e) = \frac{1}{\alpha}(u_t - \bar{u}_t) + \beta(\bar{x}_t - x_t).$$

If we assumed that expected inflation was equal to the inflation in the last

quarter, then we would get that $\pi_t^e = \pi_{t-1}$ and that

$$\pi_t = \pi_{t-1} + \frac{1}{\alpha}(u_t - \bar{u}_t) + \beta(\bar{x}_t - x_t).$$

If we assume that \bar{u}_t is fixed, and if we denote $c = \frac{-\bar{u}}{\alpha}$, then our new model may be written as

$$\pi_t = c + \pi_{t-1} + \frac{1}{\alpha}u_t + \beta(\bar{x}_t - x_t), \quad (2)$$

which is now in a form suitable for regressions. We should mention that our treatment of expectations in this model is relatively crude. We assume that $\pi_t^e = \pi_{t-1}$. Implicitly we are saying that information about present conditions has no effect on people's expectations. For out-of-sample forecasting purposes, however, this is often the best we can do.

One way to incorporate information about the present is to employ adaptive expectations, although as we hinted at before, these are much easier to implement in in-sample forecasts. In fact, Lucas and Rapping's (1969) paper was significant because they incorporated adaptive expectations into the Phillips curve and found evidence rejecting a stable long term tradeoff between inflation and unemployment. Previously, expectations were treated in a manner that was similar to Equation 2, which led to estimations valid only for short time periods and that were highly unstable. Lucas and Rapping were interested in discovering whether they could find evidence of a long-run and stable Phillips curve once they incorporated expectations into their

model. In order to include expectations, they assumed that people's expectations in a time period t depended on their expectations of the previous time period and their observations in the present. Specifically, Lucas and Rapping assumed that individuals could not perfectly predict future prices and real wages, denoted P_t and w_t . Instead, they assumed that people adapted their expectations in the following manner:

$$\ln(w_t^*) = \lambda \ln(w_t) + (1 - \lambda) \ln(w_{t-1}^*) \quad (3)$$

$$\ln(P_t^*) = \lambda \ln(P_t) + (1 - \lambda) \ln(P_{t-1}^*) \quad (4)$$

where P_t^* and w_t^* denote expected prices and expected real wages at time t , respectively. Here λ denotes how quickly people update their future expectations based on present conditions. If $\lambda = 0$, then we assume that people's expectations, once formed, are fixed. If $\lambda = 1$, then people expect that the future will be no different than the present.

Lucas and Rapping also modeled the equilibrium quantity of labor, and hence the unemployment rate, as a function of the ratio of past expected wages to current wages and the ratio of past expected prices to current prices. Then

$$u_t = \beta_0 + \beta_1 \ln\left(\frac{w_{t-1}^*}{w_t}\right) + \beta_2 \ln\left(\frac{P_{t-1}^*}{P_t}\right).$$

Their model was derived from a labor supply model, which was published

later that year (see Rapping and Lucas 1969). To be precise, w^* does not refer to the expected wage but to the expected *permanent* wage. This distinction will be significant when trying to sign β_1 and β_2 .

What do we mean by *permanent*? In a 1957 book, Friedman first hypothesized the existence of what he called a “permanent income.” The hypothesis was called the permanent income hypothesis, and it would later help him win the Nobel Prize in 1976. To understand the significance of his hypothesis, we will first try to understand the state of economic thinking at that time.

During the 1950s and 1960s, economists were trying to reconcile why the available data conflicted with their understanding of consumer behavior. Keynes had hypothesized that consumption depends only on income and autonomous consumption. Moreover, for every one unit increase in income, he believed that consumption increased by less than one. This implied that as income continued to rise, the ratio of consumption to income would fall. However, Simon Kuznets collected data from 1869 to the 1940s, which showed that the ratio of consumption to income, also called the average propensity to consume (APC), remained constant over long time periods (Mankiw 2003). Kuznets’ finding directly contradicted Keynes’ hypothesis and previous household consumption studies (which were over shorter time intervals), both of which predicted that the APC would fall as consumers became wealthier.

Milton Friedman suggested that perhaps Keynes was looking at the wrong type of income. Friedman suggested that instead of spending based off real-

ized income, consumers spent based off income they could reasonably expect to receive in the future. For example, when wages started to increase due to unanticipated productivity advances—at least when unanticipated by the workers—following World War II, then Friedman’s hypothesis suggested that at first workers might expect such wage gains to be a result of the current state of the business cycle. It would take time, and maybe even a recession, before workers would realize that the wage gains resulting from productivity increases were likely *permanent*. If wages were thought to be transitory, we would expect workers to maintain their current life style. Only once they realize that such wage gains are permanent would we expect them to adjust their consumption.

The difference between observable income and the income that workers consider permanent helps us understand why Kuznets’ data contradicted Keynes’ theory. When economists analyzed time intervals too short for workers to determine how much income was likely to be permanent, they found that consumption failed to keep pace with income. Over time periods long enough for all permanent changes in income to be identified, the average propensity to consume was stable.

Thus, we could interpret the variable w_{t-1}^* as the portion of real wages from the last period that workers considered permanent. If a worker received a bonus, we would expect that if the bonuses were volatile or if this were a one-time event, then the bonus would not be included in w_{t-1}^* but would be included in w_{t-1} . Therefore, if present wage w_t is greater than the estimated

permanent wage from last year w_{t-1}^* , then a worker would likely consider the present a good opportunity to work. He would likely be willing to work in the present, when the reward for working is higher, and would likely defer leisure to a time when current wages are below what workers would expect to receive. Consequently, the unemployment rate should fall.

When $w_t > w_{t-1}^*$, then

$$\ln\left(\frac{w_{t-1}^*}{w_t}\right) < 0.$$

Since we would predict that rising real wages would increase the quantity of labor supplied and reduce the unemployment rate, then β_1 should be greater than zero. Likewise, when expected prices are below actual prices, then a worker suffers a negative income shock, and we would expect him to be willing to work more, lowering the unemployment rate. Thus, β_2 should also be positive.

Now that we have signed the coefficients for our model, we need to convert it to a regressable form. After some algebraic manipulations (see Appendix A), we find that our model is equivalent to:

$$u_t = \lambda\beta_0 - \beta_1 \left[\ln\left(\frac{w_t}{w_{t-1}}\right) \right] - \beta_2 \left[\ln\left(\frac{P_t}{P_{t-1}}\right) \right] + (1 - \lambda)u_{t-1} \quad (5)$$

where λ was the adjustment factor from before. Note that Equation 5 is in a form that we can use for regressions because all of the included variables are observable. The purpose of these algebraic manipulations was to

rid the original model of any explicit use of expectations. Since wages and unemployment are jointly determined in the labor market, we will need an instrument for w_t to use in our regression. For our regression, we will use the reduced-form wage rate equation provided by Lucas and Rapping (1969) as our instrument.

We should mention that Lucas and Rapping assume that prices are determined outside of the labor market. So then the only two cotermined variables in our model are unemployment and wages, though Lucas and Rapping use the natural log of wages as the other dependent variable in their system. If we were interested in estimating $\ln(w_t)$, we would use Lucas and Rapping's reduced-form unemployment rate equation as an instrument for unemployment in order to estimate $\ln(w_t)$. However, since we are only interested in the relationship between the unemployment rate and its sensitivity to changes in real wages and prices, we restrict our attention to Equation 5.

We consider the years from 1947 to 2006. We initially tested for a structural break in our model in 1973 because many other macroeconomic models fundamentally changed following the simultaneous occurrences of the 1973 oil shocks, the decline of the Bretton Woods system, the Nixon price controls, and a U.S. dollar devaluation. However, our Wald test failed to reject a structural break in 1973. Regardless, we will show our regressions for the 1947–1973 and 1974–2006 subsamples in Table 2. Note that in Table 2, significance is measured as two tail significance, though Lucas and Rapping used one tail significance. Furthermore, notice that while the coefficients of the

Table 2: Estimates for Expectations Adjusted Phillips Curve

	1947–2006	1947–1973	1974–2006
c	4.18 (0.98)***	4.75 (1.48)***	4.31 (1.30)***
$\ln\left(\frac{P_t}{P_{t-1}}\right)$	-29.44 (11.1)**	-35.4 (12.9)**	-48.8 (24.3)*
$\ln\left(\frac{P_{t-1}}{P_{t-2}}\right)$	36.1 (10.2)***	18.0 (13.0)	67.8 (25.1)**
$\ln\left(\frac{w_t}{w_{t-1}}\right)$	-58.0 (17.1)***	-50.7 (22.7)**	-57.4 (25.7)**
u_{t-1}	0.76 (0.17)***	0.82 (0.21)***	0.48 (0.31)
u_{t-2}	-0.37 (0.16)**	-0.45 (0.21)**	-0.20 (0.22)
$AR(1)$	-0.07 (0.18)	-0.41 (0.27)	-0.00 (0.24)
Adj. R^2	0.60	0.52	0.58
Breusch-Godfrey Test Prob. with $p = 3$	0.40	0.99	0.17

$\left. \begin{array}{l} * \\ ** \\ *** \end{array} \right\}$ two tail significance at $\left\{ \begin{array}{l} 0.10 \\ 0.05 \\ 0.01 \end{array} \right.$

1947–1973 subsample are generally insignificant, they shared the same sign as the coefficients in the later sample. In fact, our Wald test for a structural break in 1973 suggested that the model appeared to be relatively stable over both time periods.

As we explained before, we would expect that the coefficients β_1 and β_2 to be positive. Recall that in Equation 5 the coefficients on the wage and price variables were $-\beta_1$ and $-\beta_2$, respectively. So we would expect the signs in our regression to be negative. Note that the wage coefficient is negative and significant. Although the summed coefficients on the price variables are positive, we will show that they are not significantly different from zero.

Let H_0 denote the null hypothesis that the sum of the price coefficients is zero. Likewise, let H_1 denote the null hypothesis that the wage coefficient is zero. For example, if we fail to reject H_0 over one time period, then this suggests that there is no long-run tradeoff between price inflation on unemployment over this time period. Hence, the long-term Phillips curve is vertical. If we fail to reject H_1 , then this suggests that policy changes designed to change the real wages, and consequently the demand for workers, have no effect on the unemployment rate. Table 3 provides the F -tests' probabilities for these hypotheses across the different samples considered. Note that when testing H_1 , since there is only one wage variable, the F -test's probability is equal to the p -value of the t -statistic in the original regressions.

Note that the sum of the price variables is statistically different from zero in only one sample but that the wage variable is statistically different from

Table 3: F -tests' Probabilities for Null Hypotheses

Period	H_0	H_1
1947–2006	0.34	0.00
1947–1973	0.18	0.04
1974–2006	0.03	0.03

zero in all three samples. Judging by the relative stability of the coefficients of the Phillips curve in the three regressions from Table 2 and our failure to reject H_0 , our findings suggest that the Phillips curve is relatively stable but vertical for the years 1947 to 2006.

For policy purposes, this suggests that the Phillips curve is vertical over time intervals as short as 25 years. This implies that unemployment and inflation are uncorrelated, which contradicts Stock and Watson's (1999) finding that the Phillips curve is the best predictor of inflation for the last 40 or so years.

Moreover, Stock and Watson's finding was contested by Lee and Ohanian (2001), who found that the forecasting ability of the Phillips curve declined in the years following Stock and Watson's forecasts. We will now assess to what extent changes in the unemployment rate predict changes in the rate of inflation. We will use the three samples described in Table 4. We choose these forecasting intervals in order to account for the 1972 energy crisis and the information technology revolution starting in the early 1990s. For each of these sample periods, we will consider an autoregressive model, a con-

Table 4: Description of Forecasts

	Sample 1	Sample 2	Sample 3
Regression Intervals	1947–1955	1960–1972	1980–1990
Forecasting Intervals	1956–1972	1973–1990	1991–2007

ventional Phillips curve with unemployment as an independent variable, a modified Phillips curve that uses capacity utilization instead of unemployment, and a bivariate model including both unemployment and capacity utilization. In addition, we will also combine the Phillips curves' forecasts into a combined prediction and compare all of these to a naïve prediction of inflation. For the naïve measure, we will assume that inflation one year later is equal to the average inflation rate over the last four quarters.

Our model specifications are given in Tables 5, 6, and 7. We chose the independent variables according to Akaike's Information Criterion and the Schwarz Criterion for the regression intervals. Then we conducted out-of-sample forecasting by predicting inflation using the values of the independent variables in the forecasting interval through EViews' forecast command.

To measure the accuracy of our forecasts, we compared the RMSE values for each of the models. These values are given in Table 8. Notice that in the first sample period, the naïve model outperformed all others. This is because all of our other models either overestimated inflation in the late 1950s and early 1960s or they were unable to predict the inflationary spike preceding 1972. In the other sample periods, the naïve model is one of the weakest

Table 5: Model Specification for Sample 1

AR Model	$\pi_{t+4} = c_0 + c_1\pi_t + c_2\pi_{t-2} + \epsilon_t$
Phillips Curve 1	$\pi_{t+4} = c_0 + c_1\pi_t + c_2u_t + c_3u_{t-1} + c_4u_{t-4} + \epsilon_t$
Phillips Curve 2	$\pi_{t+4} = c_0 + c_1\pi_t + c_2x_t + c_3x_{t-1} + c_4x_{t-4} + \epsilon_t$
Bivariate Model	$\pi_{t+4} = c_0 + c_1\pi_t + c_2x_t + c_3u_t + c_4u_{t-1} + \epsilon_t$
Combined Forecasts	Use predicted values from both Phillips curves
Naïve Model	$\pi_{t+4} = \frac{\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}}{4} + \epsilon_t$

Table 6: Model Specification for Sample 2

AR Model	$\pi_{t+4} = c_0 + \sum_{i=0}^4 (c_{i+1}\pi_{t-i}) + \epsilon_t$
Phillips Curve 1	$\pi_{t+4} = c_0 + \sum_{i=0}^4 (c_{i+1}\pi_{t-i}) + c_6u_t + \epsilon_t$
Phillips Curve 2	$\pi_{t+4} = c_0 + \sum_{i=0}^4 (c_{i+1}\pi_{t-i}) + c_6x_t + \epsilon_t$
Bivariate Model	$\pi_{t+4} = c_0 + \sum_{i=0}^4 (c_{i+1}\pi_{t-i}) + c_6u_t + c_7x_t + \epsilon_t$
Combined Forecasts	Use predicted values from both Phillips curves
Naïve Model	$\pi_{t-4} = \frac{\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}}{4}$

Table 7: Model Specification for Sample 3

AR Model	$\pi_{t+4} = c_0 + c_1\pi_t + \epsilon_t$
Phillips Curve 1	$\pi_{t+4} = c_0 + c_1\pi_t + c_2u_t + \epsilon_t$
Phillips Curve 2	$\pi_{t+4} = c_0 + c_1\pi_t + c_2x_t + \epsilon_t$
Bivariate Model	$\pi_{t+4} = c_0 + c_1\pi_t + c_2x_t + c_3u_t + \epsilon_t$
Combined Forecasts	Use predicted values from both Phillips curves
Naïve Model	$\pi_{t-4} = \frac{\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}}{4} + \epsilon_t$

Table 8: RMSEs of Different Models

	1956–1972	1973–1990	1991–2007
AR Model	2.69	1.40	0.52
Phillips Curve 1	3.59	1.73	0.71
Phillips Curve 2	3.37	1.51	0.57
Bivariate Model	3.45	1.74	0.55
Combined Forecasts	3.55	1.56	1.39
Naïve Model	1.28	2.29	0.84

predictors overall. Of the other models considered, the autoregressive model is the best performing model. This is surprising, considering that all of the other models include the autoregressive terms and some unemployment and/or capacity utilization terms. This suggests that perhaps Lee and Ohanian were correct to argue that the Phillips curve provides poor predictions. Even the capacity utilization specification of the Phillips curve, which Stock and Watson found to be one of the best predictors over their time samples, underperformed.

Of course, it is also possible that our findings simply mean that the SIC and AIC criteria were ill-suited for determining the relevant independent variables. That said, without Stock and Watson's intimate knowledge of inflation models, these two criteria are some of the best tools available to us. One final difference between our work and Stock and Watson's is that we did not stipulate that the inflation coefficients add to one.

In spite of all these differences, it is striking that over all three forecast intervals with different specifications, the autoregressive model was only outperformed in one sample, and that was by our naïve model. From our regressions, there appears to be no evidence that unemployment or capacity utilization terms help to explain the changes in the rate of inflation.

Thus, our research suggests that in the long-run the Phillips curve is vertical, with no tradeoff between unemployment and inflation for policymakers, and that the short-run Phillips curve, if it even exists, is hard to estimate.

5 Conclusion

In this paper, we reviewed the history of forecasting and how it related to the evolution of macroeconomic models. We then turned our attention to inflation forecasts and the Phillips curve, one of the most controversial topics in the field of macroeconomic forecasting, and tried to assess the extent to which we can find evidence supporting or rejecting the existence of the Phillips curve. In our empirical section, we found no evidence for either a tradeoff between inflation and unemployment or a short-run Phillips curve during the last fifty years. This suggests that, by itself, the Phillips curve should have only a limited role when conducting inflation forecasts, although further research is still needed.

A Deriving the Lucas Rapping Regression

Lucas and Rapping (1969) assumed that the quantity of labor, and hence the unemployment rate, depended on the ratio of estimates of past permanent real wages to current real wages and the ratio of past expected prices to current prices. Then:

$$u_t = \beta_0 + \beta_1 \ln \left(\frac{w_{t-1}^*}{w_t} \right) + \beta_2 \ln \left(\frac{P_{t-1}^*}{P_t} \right)$$

and we can solve for u_t by using equations 3 and 4 and through the following algebraic manipulations:

$$\begin{aligned}
u_t &= \beta_0 + \beta_1 \ln\left(\frac{w_{t-1}^*}{w_t}\right) + \beta_2 \ln\left(\frac{P_{t-1}^*}{P_t}\right) \\
&= \beta_0 + \beta_1[\ln(w_{t-1}^*) - \ln(w_t)] + \beta_2[\ln(P_{t-1}^*) - \ln(P_t)] \\
&= \beta_0 + \beta_1[\lambda \ln(w_{t-1}) + (1 - \lambda) \ln(w_{t-2}^*) - \ln(w_t)] + \\
&\quad \beta_2[\lambda \ln(P_{t-1}) + (1 - \lambda) \ln(P_{t-2}^*) - \ln(P_t)] \\
&= \lambda\beta_0 + (1 - \lambda)\beta_0 + \beta_1[\ln(w_{t-1}) + (\lambda - 1) \ln(w_{t-1}) + (1 - \lambda) \ln(w_{t-2}^*) - \ln(w_t)] + \\
&\quad \beta_2[\ln(P_{t-1}) + (\lambda - 1) \ln(P_{t-1}) + (1 - \lambda) \ln(P_{t-2}^*) - \ln(P_t)] \\
&= \lambda\beta_0 + -\beta_1[\ln(w_t) - \ln(w_{t-1})] - \beta_2[\ln(P_t) - \ln(P_{t-1})] + \\
&\quad (1 - \lambda)[\beta_0] + \beta_1[(\lambda - 1) \ln(w_{t-1}) + (1 - \lambda) \ln(w_{t-2}^*)] + \\
&\quad \beta_2[(\lambda - 1) \ln(P_{t-1}) + (1 - \lambda) \ln(P_{t-2}^*)] \\
&= \lambda\beta_0 - \beta_1\left[\ln\left(\frac{w_t}{w_{t-1}}\right)\right] - \beta_2\left[\ln\left(\frac{P_t}{P_{t-1}}\right)\right] + \\
&\quad (1 - \lambda)[\beta_0] + \beta_1[(1 - \lambda) \ln(w_{t-2}^*) - (1 - \lambda) \ln(w_{t-1})] + \\
&\quad \beta_2[(1 - \lambda) \ln(P_{t-2}^*) - (1 - \lambda) \ln(P_{t-1})] \\
&= \lambda\beta_0 - \beta_1\left[\ln\left(\frac{w_t}{w_{t-1}}\right)\right] - \beta_2\left[\ln\left(\frac{P_t}{P_{t-1}}\right)\right] + \\
&\quad (1 - \lambda)[\beta_0 + \beta_1(\ln(w_{t-2}^*) - \ln(w_{t-1}))] + \\
&\quad \beta_2(\ln(P_{t-2}^*) - \ln(P_{t-1})) \\
&= \lambda\beta_0 - \beta_1\left[\ln\left(\frac{w_t}{w_{t-1}}\right)\right] - \beta_2\left[\ln\left(\frac{P_t}{P_{t-1}}\right)\right] + (1 - \lambda)u_{t-1}
\end{aligned}$$

B Instrumental Variables and Data Sources

To extend Lucas and Rapping's research, we needed to first estimate $\ln(w_t)$ in order to predict the rate of unemployment. The regression we conducted to estimate $\ln(w_t)$ is given below:

$$\begin{aligned}\ln(w_t) = & \beta + \beta_1 \ln(w_{t-1}) + \beta_2 \ln\left(\frac{P_t}{P_{t-1}}\right) + \beta_3 \ln\left(\frac{N_{t-1}}{M_{t-1}}\right) + \\ & + \beta_4 \ln\left(\frac{N_{t-1}Q_{t-1}}{y_{t-1}}\right) + \beta_5 \ln\left(\frac{y_t}{y_{t-1}}\right) + \beta_6 \ln(Q_t) + \beta_7 \ln\left(\frac{y_t}{M_t}\right).\end{aligned}$$

Wages were measured as non-farm business compensation and then converted to real terms. The price level was measured by the real GDP deflator. The quantity of labor supplied, N_t , was measured as the total number of hours provided by workers and was calculated from the payroll count and the weekly hours worked per worker. M_t measures the total number of households in the economy according to the Census Bureau. y_t measures the real GDP. Q_t denotes the quality of workers, as measured by the median years of education provided by the Census, which is available at <http://www.census.gov/population/www/socdemo/educ-attn.html>. During the early years for which the education data were missing, we regressed our data on the data provided by Lucas and Rapping. To fill in the later missing observations, we used the information on percentages with each type of education to estimate roughly how many years of education the 50th percentile would have completed assuming a uniform distribution within each category (e.g., high

school degree). The only other variable used in our regression is the unemployment rate, which was obtained from the Federal Reserve Bank of St. Louis. The above variables with unlisted sources were similarly obtained from the Federal Reserve Bank of St. Louis.

The data for the second portion of the empirical research was also primarily gathered from the Federal Reserve Bank of St. Louis. The price level was measured by the GDP deflator. The capacity utilization rate in manufacturing was obtained from the Federal Reserve at <http://www.federalreserve.gov/Releases/G17/caput1.htm>.

References

- [1] G. A. AKERLOF, *Behavioral Macroeconomics and Macroeconomic Behavior*, *The American Economic Review*, 92 (2002), pp. 411–433.
- [2] AKERLOF, GEORGE A., DICKENS, WILLIAM T., PERRY, GEORGE L., BEWLEY, TRUMAN F., AND BLINDER, ALAN S., *Near-Rational Wage and Price Setting and the Long-Run Phillips Curve*, *Brookings Papers on Economic Activity*, 2000 (2000), pp. 1–60.
- [3] ATKESON, ANDREW AND OHANIAN, LEE E., *Are Phillips Curves Useful for Forecasting Inflation?*, *Federal Reserve Bank of Minneapolis Quarterly Review*, 25 (2001), pp. 2–11.
- [4] BALL, LAURENCE AND MANKIW, N. GREGORY, *The NAIRU in Theory and Practice*, *The Journal of Economic Perspectives*, 16 (2002), pp. 115–136.
- [5] B. S. BERNANKE, *Inflation Expectations and Inflation Forecasting*. Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts, July 2007.

- [6] A. S. BLINDER, *Is there a core of practical macroeconomics that we should all believe?*, *The American Economic Review*, 87 (1997), pp. 240–243.
- [7] BRAYTON, FLINT, LEVIN, ANDREW, TRYON, RALPH, AND WILLIAMS, JOHN C., *The Evolution of Macro Models at the Federal Reserve Board*, February 1997.
- [8] C. D. CARROLL, *Macroeconomic Expectations of Households and Professional Forecasters*, *Quarterly Journal of Economics*, 118 (2003), pp. 269–298.
- [9] C. CHATFIELD, *Forecasting in the 1990s*, *The Statistician*, 46 (1997), pp. 461–473.
- [10] F. X. DIEBOLD, *The Past, Present, and Future of Macroeconomic Forecasting*, *The Journal of Economic Perspectives*, 12 (1998), pp. 175–192.
- [11] ———, *Elements of Forecasting*, South-Western, 2001.
- [12] DIEBOLD, FRANCIS X. AND RUDEBUSCH, GLENN D., *Measuring Business Cycles: A Modern Perspective*, *The Review of Economics and Statistics*, 78 (1996), pp. 67–77.
- [13] DOMINGUEZ, KATHRYN M., FAIR, RAY C., AND SHAPIRO, MATTHEW D., *Forecasting the Depression: Harvard Versus Yale*, *The American Economic Review*, 78 (1988), pp. 595–612.
- [14] ELLIOTT, GRAHAM AND TIMMERMAN, ALLAN, *Economic Forecasting*, *Journal of Economic Literature*, 46 (2008), pp. 3–56.
- [15] I. FISHER, *I Discovered the Phillips Curve: “A Statistical Relation between Unemployment and Price Changes”*, *The Journal of Political Economy*, 81 (1973), pp. 496–502.
- [16] M. FRIEDMAN, *The Role of Monetary Policy*, *The American Economic Review*, 58 (1968), pp. 1–17.
- [17] ———, *Nobel Lecture: Inflation and Unemployment*, *The Journal of Political Economy*, 85 (1977), pp. 451–472.

- [18] J. K. GALBRAITH, *Time to Ditch the NAIRU*, *The Journal of Economic Perspectives*, 11 (1997), pp. 93–108.
- [19] GOOIJER, JAN G. DE AND HYNDMAN, ROB J., *25 Years of Time Series Forecasting*, *International Journal of Forecasting*, 22 (2006), pp. 443–473.
- [20] R. J. GORDON, *The Time-Varying NAIRU and its Implications for Economic Policy*, *The Journal of Economic Perspectives*, 11 (1997), pp. 11–32.
- [21] C. W. J. GRANGER, *Can We Improve the Perceived Quality of Economic Forecasts?*, *Journal of Applied Econometrics*, 11 (1996), pp. 455–473.
- [22] S. HALL, *Macroeconomics and a Bit More Reality*, *The Economic Journal*, 105 (1995), pp. 974–988.
- [23] J. D. HAMILTON, *A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle*, *Econometrica*, 57 (1989), pp. 357–384.
- [24] P. N. IRELAND, *Does the Time-Consistency Problem Explain the Behavior of Inflation in the United States?*, *Journal of Monetary Economics*, 44 (1999), pp. 279–291.
- [25] P. KENNEDY, *A Guide to Econometrics*, The MIT Press, 2003.
- [26] KYDLAND, FINN E. AND PRESCOTT, EDWARD C., *Time to Build and Aggregate Fluctuations*, *Econometrica*, 50 (1982), pp. 1345–1370.
- [27] T. LAUBACH, *Measuring the Nairu: Evidence from Seven Economies*, *The Review of Economics and Statistics*, 83 (2001), pp. 218–231.
- [28] R. G. LIPSEY, *The Relationship between Unemployment and Rate of Change of Money Wage Rates in the United Kingdom, 1862–1957: A Further Analysis*, *Economica*, 27 (2001), pp. 1–31.
- [29] R. E. LUCAS, *Econometric Policy Evaluations: a Critique*, *Carnegie-Rochester Conference Series on Public Policy*, 1 (1976), pp. 19–46.

- [30] LUCAS, ROBERT E. AND RAPPING, LEONARD A., *Price Expecations and the Phillips Curve*, The American Economic Review, 59 (1969), pp. 342–350.
- [31] N. G. MANKIW, *The Inexorable and Mysterious Tradeoff between Inflation and Unemployment*, The Economic Journal, 111 (2001), pp. C45–C61.
- [32] ———, *Macroeconomics*, Worth Publishers, 2003.
- [33] M. P. MURRAY, *A Drunk and Her Dog: An Illustration of Cointegration and Error Correlation*, The American Statistician, 48 (1994), pp. 37–39.
- [34] C. R. NELSON, *The Prediction Performance of the FRB-MIT-PENN Model of the U.S. Economy*, The American Economic Review, 62 (1972), pp. 902–917.
- [35] NEWBOLD, P. AND GRANGER, C. W. J., *Experience with Forecasting Univariate Time Series and the Combination of Forecasts*, Journal of the Royal Statistical Society. Series A (General), 137 (1974), pp. 131–165.
- [36] E. S. PHELPS, *The New Microeconomics in Inflation and Employment Theory*, The American Economic Review, 59 (1969), pp. 147–160.
- [37] A. W. PHILLIPS, *The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957*, *Economica*, 25 (1958), pp. 283–299.
- [38] RAPPING, LEONARD A. AND LUCAS, ROBERT E., *Real Wages, Employment, and Inflation*, The Journal of Political Economy, 77 (1969), pp. 721–754.
- [39] SAMUELSON, PAUL A. AND SOLOW, ROBERT M., *Analytical Aspects of Anti-Inflation Policy*, The American Economic Review, 50 (1960), pp. 177–194.
- [40] C. A. SIMS, *Macroeconomics and Reality*, *Econometrica*, 48 (1980), pp. 1–48.
- [41] E. SLUTSKY, *The Summation of Random cause as the Source of Cyclic Processes*, *Econometrica*, 5 (1937), pp. 105–146.

- [42] STAIGER, DOUGLAS, STOCK, JAMES H., AND WATSON, MARK W., *The NAIRU, Unemployment and Monetary Policy*, *The Journal of Economic Perspectives*, 11 (1997), pp. 33–49.
- [43] STOCK, JAMES H. AND WATSON, MARK W., *Forecasting Inflation*, *Journal of Monetary Economics*, 44 (1999), pp. 293–335.
- [44] ———, *Vector Autoregressions*, *The Journal of Economic Perspectives*, 15 (2001), pp. 101–115.
- [45] ———, *Has Inflation Become Harder to Forecast?* Prepared for the conference, Quantitative Evidence on Price Determination, Board of Governors of the Federal Reserve Board, September 29-30, Washington DC, Sept. 2005.
- [46] A. H. STUDENMUND, *Using Econometrics: A Practical Guide*, Pearson Education, Inc., 2006.
- [47] H. WOLD, *A Study in the Analysis of Stationary Time Series*, PhD thesis, University of Stockholm, 1938.
- [48] G. U. YULE, *On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers*, *Philosophical Transactions of the Royal Society of London*, 226A (1927), pp. 267–298.