IOWA STATE UNIVERSITY

Empirical Phillips Curves in OECD Countries: Has There Been A Common Breakdown?

Matthew Doyle

September 2006

Working Paper # 06032

Department of Economics Working Papers Series

Ames, Iowa 50011

Empirical Phillips Curves in OECD Countries: Has there Been a Common Breakdown?

Matthew Doyle*

September 25, 2006

Abstract

Recent work on U.S. data calls into question the ability of simple Phillips curve models to forecast inflation. This paper asks whether there is similar evidence of a breakdown in the forecasting ability of Phillips curve models in other OECD countries. The results suggests that the ability of a Phillips curve to out-forecast simpler models has deteriorated in many OECD countries. The evidence is less clear as to whether this breakdown can be attributed to structural breaks in the parameters of the Phillips curve.

KEYWORDS: Phillips curve, structural break, forecast breakdown.
JEL CLASSIFICATION: E31.

*Department of Economics, Iowa State University, Ames, IA, 50011, USA. Phone: (515) 294-0039. Email: msdoyle@iastate.edu

1 Introduction

Are empirical Phillips curves useful tools for forecasting inflation? Many empirical and policy oriented macroeconomists view the Phillips curve, broadly defined as relationships between inflation and some measure of real activity, as a stable and reliable relationship (prominent examples include see Fuhrer (1995), and Blinder (1997)). Despite concerns about it's theoretical foundations, the Phillips curve concept remains at the heart of monetary policy makers' views of the economy.

One of the most common uses of the Phillips curve is as a tool for forecasting of inflation. The usefulness of the Phillips curve as an inflation forecaster is predicated on the view that forecasts from Phillips curve models are generally more accurate than forecasts from simpler models. Moreover, it is not enough for Phillips curves to outperform simpler models in only some periods. To be of value to economic decision makers the superiority of Phillips curve forecasts ought to hold across different periods. Recent work on U.S. Phillips curves, however, suggests that Phillips curves do not outperform more naive forecasts of inflation in recent years.

This paper extends the existing literature by documenting a breakdown in the forecasting performance of the Phillips curve in recent years in a number of other OECD countries. In particular, the paper uses both in and out of sample tests of predictive ability to show that, across OECD countries, a Phillips curve model outperforms a simple autoregressive model in earlier periods, but not in more recent years. The paper goes on to investigate whether shifts in the parameters of the Phillips curve can explain the forecasting breakdown, but finds that the evidence in favor of this hypothesis is mixed.[1]

The paper begins by establishing the existence of a Phillips curve relationship between inflation and the output gap of 11 OECD countries over the past 4 decades. The results suggest that a positive relationship between inflation and measures of the cyclical component of real activity is a robust feature of the data for all of the countries in the sample. The Phillips curve is robust to alternate measures of inflation and real activity, as well as to various specifications of the inflation process. In most countries, the coefficients suggest that an episode where output is 1% above trend is accompanied by increase in annualized inflation in the range of 0.3-1%. The U.S. estimate falls in the middle to low end of this range. These results suggest that a standard Phillips curve is no less reasonable a specification for other OECD countries than it is for the U.S..

The paper then turns to address the question of whether Phillips curve models provide

[1]The empirical models studied in this paper are all versions of backwards looking Phillips curve, as opposed to forward looking and hybrid New Keynesian style Phillips curves. While many authors prefer the latter type models, due to their micro founded theoretical roots, backwards looking models are still widely used in the empirical literature. Furthermore, it is often claimed that backwards looking models are better fitting and more stable than their forward looking counterparts (Fuhrer & Moore (1995), Eller & Gordon (2003), Estrella & Fuhrer (2003), Rudebusch (2005)).

superior forecasts of inflation than a simple autoregressive model of inflation. The question of whether U.S. Phillips curves outperform more naive forecasts of inflation has been the subject of recent debate. Most notable amongst those who argue that the Phillips curve is a good inflation forecaster, is the study by Stock and Watson (1999), which examines the one-year-ahead forecasts of U.S. inflation from both Phillips curve and autoregressive models, over the sample 1959-1997. They conclude that Phillips curve specifications outperform the more naive models.

This conclusion has been challenged by a number of authors. Atkeson & Ohanian (2001) argue that the conclusion reached by Stock & Watson does not apply over the most recent period (from 1985 onwards). Fischer, Liu,& Zhou (2002) find that Phillips curve models do outperform more naive models in earlier year (1977-84) but not in more recent subsamples (they consider both 1985-1992, and 1993-2000). Brave & Fischer (2004) confirm this result in a comparison of the performance of several alternate inflation forecasting models across different subsamples.[2]

This paper compares the forecasting ability of a Phillips curve versus a simple AR specification in 11 OECD countries. Following the literature, I test the ability of the models to forecast inflation at the one-quarter-ahead and one-year ahead horizons. Since it is unclear whether the best metric is in or out of sample forecasting ability,[3] the tests are performed on both in and out of sample forecasts. To capture the possibility of a forecast breakdown, the tests are run on early and late samples. I report the results for two possible break dates (1987:1, and 1992:1), corresponding to 1/2 and 1/3 of the sample, though the results are not sensitive to alternate break dates.[4]

The in sample tests are simple F-tests of the null hypothesis that the coefficient on the real activity term is zero. If the null hypothesis is correct, then the use of a Phillips curve, as opposed to a simple AR, doesn't reduce the in sample forecast errors. For out of sample forecasts, I report the results of forecast encompassing tests of the type proposed by Clark & McCracken (2001, 2005a).[5]

Both the in sample and out of sample test suggest that while a Phillips curve specification out performs the AR model in the earlier part of the sample, the Phillips curve forecasts no better than the AR over the past decade. Although no one test is conclusive, the overall pattern is of a breakdown of the performance of the Phillips curve relative to the autoregressive model across the sample of countries.

If there has been a common the deterioration in performance of the Phillips curve as a forecasting tool in recent years, then it is worth considering a common explanation for

[2]Also see Bachmeier & Swanson (2005).

[3]Inoue & Killian (2004), Clark & McCracken (2005b).

[4]These dates are also similar to the sub samples chosen by Fischer,Liu, & Zhou (2002) and Black & Fischer (2004), to correspond to changes in U.S. monetary policy regimes.

[5]Tests of equal forecast accuracy based on equal MSE (for example, Diebold & Mariano (1995), West (1996), and McCracken (2004)) provide broadly similar results.

this breakdown. The literature suggests a number of plausible causes of a breakdown in U.S. Phillips curves: changes in the conduct of monetary policy[6], changes in productivity growth[7], and mis-measurement of real activity.[8] Most of these hypotheses suggest that there has been a structural shift in the parameters of the Phillips curve. In the U.S., there is evidence of breaks in the parameters on lagged inflation terms[9] as well as a possible shift in the slope of the Phillips curve.[10]

This paper takes a first step in asking whether any of these hypotheses might serve as an explanation of the breakdown of the Phillips curve in OECD countries by asking whether the decline in the forecasting ability of the Phillips curve is due to structural breaks.[11] I apply several well known tests for detecting a break at an unknown date, to investigate whether theories positing a structural shift in the Phillips curve are likely to be able to explain the observed forecast breakdowns. I focus on structural breaks in the slope of the Phillips curve, rather than the in the parameters of lagged inflation, as shifts in the slope are most likely to affect the forecasting performance of the Phillips curve relative to a univariate autoregressive model of inflation.

The evidence is mixed. When the Phillips curve is estimated in differences in inflation (a specification that imposes a unit root on inflation), the null hypothesis that the slope did not change can generally not be rejected. There is more evidence of structural change in the coefficients of inflation, though it is unclear that structural breaks in inflation would affect Phillips curve forecasts more severely than forecasts from a univariate AR model.[12] Furthermore, although formal tests of structural change do not reject the hypothesis that the slope of the Phillips curve is constant throughout the sample, estimating of the Phillips curve on data from the latter part of the sample shows that the slope of the Phillips curve is not statistically significant from zero in most OECD countries.

In addition to the literature on the U.S. Phillips curve cited above, this paper is also related to a smaller literature that investigates instability in international Phillips curves. Haldane & Quah (1999) examine changes in the Phillips curve in the U.K., Nishizaki & Watanabe (2000) document a decline in the slope of the Phillips curve in Japan, while Beaudry & Doyle (2000) present evidence of a similar decline in Canadian data. These studies, however, do not focus on the role of the Phillips curve as a tool for forecasting inflation.

The paper proceeds as follows: Section 2 discusses the data, the empirical model, and

[6] Estrella & Fuhrer (2003), Clark & McCracken (2005b), and Giacomini & Rossi (2005).

[7] Ball & Moffit (2001), Mankiw (2001), Staiger, Stock & Watson (2001)

[8] Anderson, Barrow, & Butcher (2005)

[9] Levin & Piger (2003)

[10] Lown & Rich (1997), Brayton, Robert, & Williams (1999), and Roberts (2004).

[11] A forecast breakdown could also be caused by factors, such as instability in the distribution of the regressors.

[12] If the Phillips curve is specified in levels of inflation, rather than changes, stability is generally rejected regardless of whether one looks for breaks in the inflation coefficients or just in the Phillips curve parameters.

results of baseline Phillips curve estimations in 11 OECD countries. Section 3 presents evidence concerning breakdowns in the forecasting performance of OECD Phillips curves, using both formal and informal metrics. Section 4 presents evidence of structural break in international Phillips curves using both known and unknown breakpoint tests. Section 5 offers concluding remarks.

2 Data and Empirical Framework

Broadly defined, the Phillips curve is a relationship between inflation, its lags and some measure of real activity. In the absence of compelling theoretical guidance, it is not immediately apparent which of the many possible measures of prices and real activity ought to be employed. A common recourse in the empirical Phillips curve literature is to run the estimation with a variety of different measures and check for robustness of result. That is the approach employed here.

The baseline measure of inflation is the annualized percentage change in the all items Consumer Price index as a measure of inflation. To provide a check for robustness I also use the annualized percentage change in the all items Consumer Price index as an alternate series the implicit GDP deflator.

The issue of how to measure real activity is more problematic. In general the Phillips curve literature takes the approach that the relationship to be estimated is between inflation and some measure of excess demand. Here I de-trend real output and estimate a Phillips curve in inflation-output gap space, where the deviation of output from trend is taken to measure excess demand. The literature employs a variety of techniques (HP filters, structural VARs, structural macroeconomic models, simple time trends) to arrive at trend series for output. I use the HP-filtered output gap as the baseline measure of real activity. I have replicated the results using an output gap arrived at by de-trending output using both the bandpass filter of Christiano and Fitzgerald (1999) and a cubic time trend as checks of robustness.

The sample consists of quarterly data for the period 1964:1-2003:3, drawn from the IMF's International Financial Statistics data base and supplemented with data from the OECD's Main Economic Indicators and the Eurostat website where necessary.[13]

2.1 Phillips Curve Models

The baseline Phillips curve specification used in the paper relates the change in inflation to its own lags and a measure of real activity. This approach follows Stock and Watson

[13] The countries in the sample are Australia, Austria, Canada, Finland, France, Germany, Italy, Japan, Sweden, the United Kingdom and the United States. Due to data limitations, the sample period is not identical for all countries. The reported results are based on the longest available sample for each country. Estimates on a consistent sample, from 1970:2-1999:4, are qualitatively similar to the results presented in the paper.

(1999, 2003), Clark and McCracken (2005), among others, who treat inflation as being close enough to an I(1) process to warrant imposing a unit root.

The baseline Phillips curve model, in differences, is:

$$\pi_{t+\tau}^{(\tau)} - \pi_t = \alpha_0 + \beta y_t + \sum_{n=0}^{N} \Delta\pi_{t-n} + \epsilon_t, \tag{2.1}$$

where $\pi_{t+\tau}^{(\tau)} \equiv (400/\tau) * \ln(p_t/p_{t-\tau})$, τ is the forecast horizon, y_t is the measure of real activity, specified here as an output gap in percentage terms, and ϵ_t is an error term. In the forecasting literature, it is common to use the Phillips curve to forecast inflation at different horizons (a one year time horizon is commonly used). For a one quarter ahead forecast, $\tau = 1$, for a one year ahead forecast from a model based on quarterly data, $\tau = 4$.

An common alternative is to estimate the model in levels:

$$\pi_{t+1} = \alpha_0 + \beta y_t + \sum_{n=0}^{N} \pi_{t-n} + \epsilon_t, \tag{2.2}$$

where inflation is $\pi_t \equiv 400 * \ln(p_t/p_{t-1})$, y_t is the measure of real activity, specified here as an output gap in percentage terms, and ϵ_t is an error term. In many cases, the performance of a model in levels of inflation is qualitatively to the model in differences, though the results of structural break tests are sensitive to the choice of specification.

The number of lags of the inflation terms used is the value that minimized the Bayesian Information Criterion in the full sample estimates. The number of lags is determined individually for each country and model combination using the estimate of the relevant model over the full sample. For simplicity, the same lag length is used across applications, so, for example, the optimal lag length is not re-estimated for models allowing for the possibility of a structural break. For comparability with much of the forecasting literature, only one lag of the output gap is used throughout.

2.2 Baseline Phillips Curve Estimates

This section documents the existence of a Phillips curve in the 11 countries in the sample. Table 1 presents results for the baseline specification using the annualized percentage change in the quarterly CPI as the measure of inflation. The output gap is de-trended quarterly GDP, where the trend is determined via the HP-filter. The model is specified in differences in inflation, as in Equation 2.1 with $\tau = 1$, which imposes the constraint that inflation possesses a unit root. The number of lags of the change in inflation used as right hand side variables is the value that minimizes the BIC (the number of lags included in each regression is indicated next to each country's name).[14]

[14]Coefficient estimates on up to four lags are reported. Coefficient estimates for lags of the change in inflation after the 4th lag are not reported.

The results suggest the existence of a Phillips curve with positive slope in all countries. The slope coefficient is positive in all 11 countries, and statistically significant in 9 of these countries. Lagged values of the change in inflation are generally negative and at least two lag are statistically significant in every case. The model provides a reasonably good fit, as measured by the $\bar{R}^2$ coefficient, which is between .3 and .6 for most countries.

In addition to being statistically significant, the magnitude of the slope coefficients are of economic significance in many cases. The slope coefficient, for example, lies between 0.2 and 1.25 in 10 of the 11 cases. This suggests that a plausibly sized recessionary gap, in which output lay 2% below trend, would be associated with a decline in annualized inflation of anywhere between 0.4 and 1.5%.

Importantly, aside from the inclusion of relatively few lags of the change in inflation, the estimated U.S. Phillips curve is not notably different from the estimates in other countries. The Phillips curve fits neither especially well or especially poorly in the U.S., and the estimated slope is neither exceptionally large or exceptionally small relative to the other countries. This suggests that a simple Phillips curve model is as reasonable a specification for the 10 other countries in the sample, as it is for the U.S.

The results presented in Table 1 are not altered substantially by employing the alternate measures of inflation or real activity or by varying the number of lags of inflation. Specifying the Phillips curve in levels, as in Equation 2.2, does not alter the main conclusion that there is evidence a Phillips curve relationship in most OECD countries, and that the U.S. Phillips curve is not an outlier.

3 Forecasting

This section investigates the forecasting performance of the Phillips curve in OECD countries. In particular, the focus is on whether there is evidence of a breakdown in the forecasting performance of Phillips curves relative to more naive models across countries. Since it is unclear whether the best metric is in or out of sample forecasting ability,[15] the tests are performed on both in and out of sample forecasts.

3.1 In Sample Tests

Table 2 presents the results of tests of the value of adding measures of real activity to an AR in the change in inflation over different sub samples. Essentially, the test is an F-test of the null hypothesis that the coefficient on the real activity term is zero. If the null hypothesis is correct, then the use of a Phillips curve, as opposed to a simple AR, doesn't reduce the in sample forecast errors. To capture the possibility of a forecast breakdown, the tests are run on early and late samples.

[15] See the discussion in Inoue & Killian (2004), and Clark & McCracken (2005b).

Columns 1 and 2 report the results of the test that measures of real activity do not improve the 4 quarter ahead forecast of the change in inflation, relative to a simple AR specification. The first column reports the results for the test on data up to 1987:1.[16] The null hypothesis is rejected in 9 of 11 countries (Australia and Finland being the exceptions) for the early period. Column 2 reports results of the same test for the more recent period. Here, the null of no difference is only rejected in 3 of 11 countries. In 10 of 11 countries, the p-value is higher in the latter period, implying that we are more likely to reject the null of no difference in the earlier period than in the latter period.

Columns 3 and 4 of Table 2 present the results when the break is chosen to split the sample into the first 2/3rds and last 3rd. Table 3 presents similar test statistics for 1 quarter ahead forecasting models. The results are qualitatively similar to the baseline case, although the evidence of a forecast breakdown is somewhat more mixed when the later break date is employed.

3.2 Out of Sample Tests

This section employs tests of forecast encompassing to evaluate whether the out of sample forecasting performance of the Phillips curve has deteriorated relative to a simple AR specification in recent years. Both t and F type tests of forecast encompassing (due to Harvey, Leybourne, and Newbold (1998) in the former, and Clark and McCracken (2001) in the latter case) are employed.[17]

In order to test the out of sample forecasting ability of the model, it is first necessary to divide the data into a in sample period, which is used to estimate the model's parameters, and a out of sample period, used for assessing the model's forecasting performance. Here I focus on a rolling scheme. Letting T be the total number of observations, τ the forecast horizon, and m the in sample window of data, at each data t the model is estimated on the most recent m observations, and these parameter estimates are used to generate the forecast for $t + \tau$. m is set to 60 quarters, or 15 years worth of data.

Table 4 provides the results of tests of forecast encompassing for the 4 quarter ahead model when forecasting the change in inflation in the pre and post 1992 samples.[18] The null hypothesis in all cases is that the addition of the real activity terms adds nothing to the forecasting ability of the model.

Columns 1 reports the statistic of the Harvey, Leybourne, and Newbold (1998) test of

[16] 1987:1 is about 1/2 of the way through the sample for countries starting in 1970. The results are generally robust to alternate choices of break dates.

[17] Tests of equal forecast accuracy based on equal MSE (for example, Diebold & Mariano (1995), West (1996), and McCracken (2004)) provide similar results, though the null of no difference tends to be rejected less frequently in both periods when these tests are used. Clark & McCracken (2005a) show that the forecast encompassing tests have greater power than the corresponding MSE based tests.

[18] 1992:1 is about halfway through the remaining sample, after the first 15 years have been used to generate the first set of estimates of the model parameters.

forecast encompassing (hereafter referred to as ENC-T, following Clark & McCracken (2001) applied to the pre-1992 data. Columns 3 reports an alternate test of the same hypothesis, applying the Clark & McCracken test of forecast encompassing (hereafter referred to as ENC-F) to the same, pre-1992 data. We can see that the, regardless of the test, the null hypothesis is rejected in the majority of countries.

Columns 2 and 4 report the results of the same tests applied to the post-1992 data. The ENC-T test, reported in column 2, rejects the null hypothesis only in the case of Sweden using the post 1992 data, and even there, the null is rejected at a weaker significance level than in the earlier period. Using the ENC-F test, the null is rejected more often in the recent period, though still only in about 1/3rd of the countries in the sample. Overall, the results support the idea that there has been a deterioration in the ability of a Phillips curve to out forecast a simpler AR model.

Table 5 reports results from the same tests, applied to the same data, using a 1 quarter, rather than one year, forecast horizon. Columns 1 and 3 report the results for the ENC-T and ENC-F tests for the early period. The null hypothesis that the AR forecasts as well as the Phillips curve is rejected in 7 and 9 of 11 cases for the ENC-T and ENC-F tests respectively. In the more recent period, the null is rejected in only 2 of 11 cases, regardless of the test. Overall, the results presented in this section suggest that, the Phillips curve out performs an AR in the early part of the sample, but fails to do so in recent years.

4 Structural Break Tests

This section investigates the stability of the parameters of the simple Phillips curve via structural break tests. Since the literature on the U.S. experience has largely focussed on the possibility of slope and intercept shifts, the paper begins by investigating the stability of these two parameters. A second stability test considers the possibility of instability in the Phillips curve as a whole, which incorporates the possibility that the autoregressive component of inflation may also have changed over time.

I assess parameter instability using the well known tests proposed by Andrews (1993) and Andrews & Ploberger (1994), for structural change at an unknown date. Elliot & Mueller (2003), however, show that these traditional structural break tests have most power against alternatives in which the parameters exhibit “large” breaks, but may have low power against alternatives where the parameters in question exhibit small, gradual drift. To encompass the possibility that Phillips curve parameters may have multiple, gradual breaks, the paper also employs the test put forth by Elliot & Mueller (2003), which has optimal power against this kind of parameter instability.

Columns 1 to 3 of Table 6 report the results of tests of the null hypothesis of no structural break against the alternative hypothesis that the intercept and slope parameters of the Phillips curve changed at unknown date(s). Column 1 reports the Andrews (1993) sup-LM test statistics, for the differenced version of the model. The null hypothesis that there is no structural break is rejected in only 3 of the 11 countries, though at the 1% level in each of these. The Andrews & Ploberger (1994) avg.-exp-LM test statistic, reported in column 2, suggests that the null hypothesis can be rejected in an additional 4 countries. The results of Elliot and Mueller's (2005) test for unknown structural change, which has power against the alternative of gradual change, are reported in column 3. This test uncovers the weakest evidence of structural break, rejecting the null in only 3 countries, and in no case at the 1% significance level.

Columns 4 to 6 report the results of tests of the null hypothesis of no structural break, where the alternative is that all parameters of the model may have changed.[19] Again, column 4 reports the Andrews (1993) sup-LM statistic, column 5 reports the Andrews & Ploberger (1994) avg.-exp-LM statistic, and column 6 the Elliot and Mueller (2005) statistic. I this case, the evidence of structural change is stronger. The Andrews & Ploberger (1994) test rejects the null for all but one country, and the Andrews (1993) rejects the null of stability in about 2/3rds of the countries in the sample (7 of 11). Only the Elliot and Mueller test fails to reject the null in a majority of countries (rejecting in 5 of 11) cases.

The results depend, to some extent, on the fact that the model is specified in differences (equation 2.1) as opposed to levels (equation 2.2). Table 7 reports the results of the same set of structural break tests on the levels specification of the model. In this case, both the Andrews (1993), and Andrews & Ploberger (1994) tests reject the null of no break, with few exceptions, regardless of whether all of the parameters or just the slope and intercept are allowed to vary. As before, the Elliot & Mueller (2003) test rejects less frequently.

It is worth noting, that point estimates of any change in slope or intercept, generally imply that any such change is negative. That is, for breaks in the mid portion of the sample, estimated coefficients on the slope and intercept terms are lower in the latter portion of the sample than in the former. Table 8 presents the results of estimating the baseline Phillips curve model on data from the most recent decade. [20] In stark contrast to the results presented in Table 1, the slope coefficients, when estimated using only the data the sample, are generally not statistically different from zero in most countries.

5 Conclusion

This paper has documented that the forecasting ability of the Phillips curve, relative to a simple AR specification, has deteriorated in recent years, not only in the U.S., but across OECD countries. Simple F-tests, for in sample forecasts, as well as forecast encompass-

[19] In this case the critical values depend on the number of lags of the change of inflation included in the regression for a given country, which is indicated in parentheses beside the country names.

[20] 1992:1 is chosen as the start of the sample, to correspond with the forecast breakdown tests of the previous section.

ing tests for out of sample predictive ability support the view that while a Phillips curve specification out performs the AR model through the 1980s, the Phillips curve forecasts no better than the AR over the past decade for most countries.

The evidence concerning whether this decline in forecasting ability is due to structural change in Phillips curve parameters is mixed. When the Phillips curve is estimated in differences in inflation (a specification that imposes a unit root on inflation), the null hypothesis of stability can not be rejected very compellingly. If the Phillips curve is specified in levels of inflation, however, the null is more frequently rejected.

Researchers inclined to be skeptical of the use of Phillips curves for forecasting can point to the results as additional evidence that the Phillips curve is an unreliable tool. On the other hand, if the deterioration in performance of the Phillips curve as a forecasting tool in recent years is common across OECD countries, then the cross-country evidence may serve as a useful tool in distinguishing between the competing theories that have been offered as explanations for the break down of the Phillips curve in the U.S..

REFERENCES

ANDERSON, K., BARROW, L., and BUTCHER, L. (2005), "Implications of Changes in Men's and Women's Labor Force Participation for Real Compensation Growth and Inflation," *Berkeley Electronic Press, Topics in Economic Analysis and Policy*, **5**, issue 1, article 7.

ANDREWS, D. W. K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point", *Econometrica*, **61**, 821–856.

ANDREWS, D. W. K., and PLOBERGER, W. (1994), "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative", *Econometrica*, **62**, 1382–1414.

ATKESON, A., and OHANIAN, L. (2001), "Are Phillips Curves Useful for Forecasting Inflation?", *Federal Reserve Bank of Minneapolis Quarterly Review*, **25**, 2–11.

BACHMEIER, L., and SWANSON, N. (2005), "Predicting Inflation: Does the Quantity Theory Help?", *Economic Inquiry*, **43**, 570–585.

BALL, L., and MOFFITT, L. (2001), "Productivity Growth and the Phillips Curve", NBER Working Paper 8421.

BEAUDRY, P., and DOYLE, M. (2000), "What Happened to the Phillips Curve in the 1990s in Canada?", in *Proceedings of the Bank of Canada Conference: Price Stability and the Long-Run target for Monetary Policy*, Bank of Canada, Ottawa, pp. 51-82 .

BLINDER, A. (1997), "Is There A Core of Practical Macroeconomics That We Should All Believe?", *American Economic Review*, **87**, 240–243.

BRAYTON, F., ROBERTS, J., and WILLIAMS, J. (1999), "What's Happened to the Phillips Curve?", Federal Reserve Board, Finance and Economics Discussion Series: 1999-49.

BRAVE, S., and FISHER, J. (2004), "In Search of a Robust Inflation Forecast", *Federal Reserve Bank of Chicago Economic Perspectives*, 12–31.

CLARK, T., and MCCRACKEN, M. (2001), "Tests of Equal Forecast Accuracy", *Journal of Econometrics*, **105**, 85-110.

CLARK, T., and MCCRACKEN, M. (2005a), "Evaluating Direct Multistep Forecasts", *Econometric Reviews*, **24** 369–404.

CLARK, T., and MCCRACKEN, M. (2005b), "The Predictive Content of the Output Gap for Inflation: Resolving the In-Sample and Out-of Sample Evidence", ...

ELLER, J., and GORDON, R. (2003), "Nesting the New Keynesian Phillips Curve within the Mainstream Model of U.S. Inflation Dynamics", **need a full citation**

ELLIOT, G., and MUELLER, U. (2003), "Optimally Testing General Breaking Processes in Linear Time Series Models", UC San Diego, Department of Economics, Working Paper 2003-07.

ESTRELLA, A., and FUHRER, J. (2003), "Monetary Policy Shifts and the Stability of Monetary Policy Models", *Review of Economics and Statistics*, **85**, 94–104.

FAUVEL, Y., GUAY, A., and PAQUET., A. (2004), "What Has the U.S. Phillips Curve Been Up To?", mimeo.

FISHER, J., LIU, C., and ZHOU, R. (2002), "When Can We Forecast Inflation?", *Federal Reserve Bank of Chicago Economic Perspectives*, 30–42.

FUHRER, J. (1995), "The Phillips Curve Is Alive and Well", *New England Economic Review*, 41–56.

GIACOMINI, R., and ROSSI, B. (2005), "Detecting and Predicting Forecast Breakdowns", *Duke University Working Paper.*

HALDANE, A., and QUAH, D. (1999), "UK Phillips Curves and Monetary Policy", *Journal of Monetary Economics*, **44**, 259–178.

HOOKER, M. (2002), "Are Oil Shocks Inflationary? Asymmetric and Nonlinear Specifications versus Changes in Regime", *Journal of Money, Credit, and Banking*, **34**, 540–561.

INOUE, A., and KILIAN, L. (2004), "In-Sample or Out-of-Sample Tests of Predictability? Which One Should We Use?" *Econometric Reviews*, **23**, 371-402.

LEVIN, A., and PIGER, J. (2003), "Is Inflation Persistence Intrinsic in Industrial Economies?", Federal Reserve Bank of St. Louis, Working Paper 2002-023E.

LOWN, C., and RICH, R. (1997), "Is There an Inflation Puzzle?", *Federal Reserve Bank of New York Economic Policy Review*, **3**, 51–69.

NISHIZAKI, K., and WATANABE, T. (2000), "Output-Inflation Trade-Off at Near-Zero Inflation Rates", *Journal of the Japanese and International Economies*, **14**, 304–326.

ROBERTS, J. (2004), "Monetary Policy and Inflation Dynamics", Federal Reserve Board, Finance and Economics Discussion Series: 2004-62.

RUDEBUSCH, G. (2005), "Assessing the Lucas Critique in Monetary Policy Models," *Journal of Money, Credit, and Banking*, **37**, 245–72.

STOCK, J. and WATSON, M. (1999), "Forecasting Inflation", *Journal of Monetary Economics*, **44**, 293–335.

Table 1. A Baseline Phillips Curve Model

	α	β	δ_1	δ_2	δ_3	δ_4	$\bar{R}^2$
Australia (3)	-0.10	0.39*	-0.76***	-0.50***	-0.27**		0.37
	(0.30)	(0.22)	(0.12)	(0.13)	(0.12)		
Austria (8)	-0.21	0.54***	-0.91***	-0.77***	-0.74***	-0.11	0.77
	(0.17)	(0.18)	(0.11)	(0.13)	(0.13)	(0.13)	
Canada (3)	-0.06	0.48***	-0.56***	-0.48***	-0.28***		0.29
	(0.18)	(0.12)	(0.10)	(0.11)	(0.08)		
Finland (8)	-0.19	0.11	-0.64***	-0.40***	-0.21	0.18	0.43
	(0.22)	(0.11)	(0.09)	(0.11)	(0.13)	(0.11)	
France (3)	-0.09	0.42***	-0.42***	-0.33***	-0.21		0.15
	(0.17)	(0.15)	(0.10)	(0.12)	(0.13)		
Germany (3)	-0.10	0.22**	-0.76***	-0.75***	-0.60***		0.56
	(0.15)	(0.10)	(0.07)	(0.08)	(0.07)		
Italy (8)	-0.11	0.95***	-0.50***	-0.52***	-0.37***	-0.25**	0.32
	(0.25)	(0.33)	(0.17)	(0.17)	(0.14)	(0.12)	
Japan (8)	-0.23	1.06***	-0.79***	-0.47***	-0.22***	0.08	0.61
	(0.31)	(0.29)	(0.08)	(0.09)	(0.15)	(0.18)	
Sweden (8)	-0.15	0.22	-0.76***	-0.66***	-0.33***	-0.15	0.44
	(0.31)	(0.22)	(0.10)	(0.12)	(0.12)	(0.11)	
U.K. (8)	-0.17	1.02***	-0.54***	-0.42***	-0.58***	-0.04	0.57
	(0.36)	(0.32)	(0.14)	(0.14)	(0.11)	(0.20)	
U.S. (2)	-0.02	0.37***	-0.31***	-0.59***			0.34
	(0.14)	(0.11)	(0.08)	(0.10)			

Table 2. 4 Q-ahead, P-Curve vs. AR: In Sample Test (H_0 : no difference)

	Pre-1987	Post-1987	Pre-1992	Post-1992
Australia (3)	10.56***	0.06	6.69***	3.36*
Austria (8)	6.02**	2.74*	5.56**	8.87***
Canada (3)	6.29**	8.10***	12.49***	1.70
Finland (8)	4.68**	2.77*	0.07	5.43**
France (3)	10.39***	0.53	8.87***	0.14
Germany (3)	6.43**	0.11	3.07*	1.44
Italy (8)	8.43***	0.08	10.30***	0.80
Japan (8)	33.19***	0.02	31.90***	3.53*
Sweden (8)	6.78***	4.12**	12.84***	2.28
U.K. (8)	27.37***	10.94***	30.11***	3.56*
U.S. (2)	13.51***	0.53	16.12***	0.15

Table 3. 1 Q-ahead, P-Curve vs. AR: In Sample Test (H_0 : no difference)

	Pre-1987	Post-1987	Pre-1992	Post-1992
Australia (3)	2.80	1.17	2.13	4.08**
Austria (8)	5.41**	4.26**	4.26**	15.10***
Canada (3)	7.35***	3.13*	8.50***	1.46
Finland (8)	2.70	0.18	0.49	3.67*
France (3)	4.97**	0.08	5.00**	0.03
Germany (3)	12.12***	0.83	12.74***	0.01
Italy (8)	9.00***	1.14	11.06***	0.38
Japan (8)	18.83***	2.16	17.86***	4.83**
Sweden (8)	0.25	1.02	2.72*	1.05
U.K. (8)	10.67***	12.41***	12.74***	3.17*
U.S. (2)	12.09***	1.55	13.98***	0.61

Table 4. 4 Q-ahead, P-Curve vs. AR: Out of Sample Test (H_0 : no difference)-15year rolling

	ENC-T test		ENC-F test	
	Pre-1992	Post-1992	Pre-1992	Post-1992
Australia (3)	0.97	-1.73	7.35***	-1.15
Austria (8)	1.28*	-1.37	2.89*	-2.41
Canada (3)	2.24***	-0.85	13.38***	-1.56
Finland (8)	1.04*	-1.87	4.95*	-0.90
France (3)	0.95	0.20	2.38***	0.31***
Germany (3)	1.67**	-1.70	9.10***	-4.72
Italy (8)	-1.22	-1.12	-1.98	3.60*
Japan (8)	2.66***	-1.51	10.42***	-1.29
Sweden (8)	3.27***	1.42**	4.43***	7.92***
U.K. (8)	2.19***	0.78	44.87***	1.13***
U.S. (2)	1.78**	-0.33	12.01***	-0.57

Table 5. 1 Q-ahead, P-Curve vs. AR: Out of Sample Test (H_0 : no difference)–15 year rolling

	ENC-T test		ENC-F test	
	Pre-1992	Post-1992	Pre-1992	Post-1992
Australia (3)	1.30*	0.39	2.04**	0.21
Austria (8)	1.09*	2.44***	1.77**	4.37***
Canada (3)	2.20***	0.91	7.08***	0.83
Finland (8)	-0.03	-0.02	-0.11	-0.01
France (3)	0.56	-0.53	1.24*	-0.30
Germany (3)	1.45**	-0.32	3.94***	-0.29
Italy (8)	1.94**	-0.02	11.08***	-0.04
Japan (8)	0.91	0.42	3.13**	0.33
Sweden (8)	1.69**	-0.83	0.64	-1.38
U.K. (8)	2.36***	1.73**	12.03***	12.36***
U.S. (2)	0.89	-0.03	3.48***	-0.04

Table 6. Tests of Structural Change: Differences

	α and β			All Parameters		
	sup-LM	exp-W-LM	E-M	sup-LM	exp-W-LM	E-M
Australia (3)	3.75	1.91	-7.44	12.03	7.44**	-14.16
Austria (8)	8.65	5.52***	-12.56	40.15***	36.70***	-87.83***
Canada (3)	5.67	3.31**	-6.86	7.37	4.67	-13.91
Finland (8)	8.01	4.88***	-12.38	25.92*	21.59***	-71.06***
France (3)	15.53**	10.88***	-13.64*	16.86*	12.21***	-22.30
Germany (3)	4.90	2.55	-10.75	21.07**	16.38***	-31.61**
Italy (8)	25.14***	20.37***	-12.87*	46.94***	42.41***	-44.70
Japan (8)	8.68	6.38***	-16.56**	31.16**	27.59***	-54.45*
Sweden (8)	6.45	2.48	-7.23	20.58	16.22***	-62.58***
U.K. (8)	11.26*	6.99***	-9.17	31.01**	27.57***	-42.36
U.S. (2)	9.88	5.96***	-10.76	10.25	6.91**	-16.77

Table 7. Tests of Structural Change: Levels

	α and β			All Parameters		
	sup-LM	exp-W-LM	E-M	sup-LM	exp-W-LM	E-M
Australia (4)	5.35	3.20*	-10.06	13.34	9.35***	-21.19
Austria (8)	11.38*	8.33***	-15.01**	41.55***	37.94***	-93.25***
Canada (4)	7.35	5.38***	-10.51	23.27**	18.82***	-20.38
Finland (8)	22.85***	18.90***	-16.04**	40.91***	36.70***	-73.34***
France (3)	18.08***	14.63***	-11.95	19.07**	15.52***	-25.48
Germany (4)	10.28*	7.06***	-12.92*	26.70***	22.77***	-37.91**
Italy (8)	30.85***	26.54***	-15.68**	49.18***	44.48***	-50.87
Japan (8)	13.86**	10.65***	-15.07**	31.89**	28.66***	-46.96
Sweden (8)	15.55**	11.45***	-8.23	24.03	19.83***	-62.88***
U.K. (8)	12.80**	8.74***	-10.52	45.71***	40.96***	-42.57
U.S. (3)	15.05**	11.33***	-14.06*	30.26***	25.49***	-23.42

Table 8. The Phillips Curve After 1992:1

	α	β	δ_1	δ_2	δ_3	δ_4	$\bar{R}^2$
Australia (3)	0.19	0.89	-0.76***	-0.36*	-0.15		0.32
	(0.46)	(0.78)	(0.23)	(0.20)	(0.17)		
Austria (8)	-0.22	0.97***	-1.18***	-1.11***	-0.88***	-0.27	0.84
	(0.17)	(0.20)	(0.14)	(0.16)	(0.16)	(0.20)	
Canada (3)	0.09	0.38	-0.57***	-0.62***	-0.20***		0.30
	(0.32)	(0.24)	(0.18)	(0.16)	(0.13)		
Finland (8)	-0.19	0.11	-0.64***	-0.40***	-0.21	0.18	0.39
	(0.22)	(0.11)	(0.09)	(0.11)	(0.13)	(0.11)	
France (3)	-0.12	0.17	-0.87***	-0.81***	-0.58***		0.55
	(0.24)	(0.14)	(0.14)	(0.17)	(0.16)		
Germany (3)	-0.26	0.19	-0.93***	-0.77***	-0.64***		0.62
	(0.26)	(0.52)	(0.11)	(0.14)	(0.12)		
Italy (8)	-0.11	0.24	-0.36*	-0.42**	-0.28	-0.05	0.14
	(0.17)	(0.27)	(0.18)	(0.18)	(0.17)	(0.17)	
Japan (8)	-0.24	0.86***	-1.23***	-1.26***	-1.34***	-1.08**	0.79
	(0.30)	(0.38)	(0.17)	(0.31)	(0.40)	(0.38)	
Sweden (8)	-0.17	-0.14	-0.89***	-0.63***	-0.60***	-0.37	0.60
	(0.30)	(0.23)	(0.16)	(0.20)	(0.21)	(0.22)	
U.K. (8)	-0.03	0.27	-0.72***	-0.60***	-0.57***	-0.17	0.82
	(0.21)	(0.42)	(0.12)	(0.16)	(0.13)	(0.15)	
U.S. (2)	-0.01	0.18	-0.35**	-0.67***			0.44
	(0.15)	(0.17)	(0.16)	(0.10)			