Inflation Forecasts: An Empirical Re-examination

Swarna B. Dutt
University of West Georgia

Dipak Ghosh
Emporia State University

Abstract

Inflation forecasts are an important part of modern day activist stabilization policy, and hence their accuracy and credibility is of vital importance. We examine the accuracy of these forecasts by testing for rationality in the expectation formation process using the Survey of Professional Forecasters data surveys of annual inflation forecasts. The non-stationarity of the actual and forecasted series allows for the application of cointegration techniques. We sequentially apply the Johansen Maximum Likelihood and the (recently available) Bieren Non-Parametric Cointegration tests. These techniques are complementary, and thus help strengthen our results. We find evidence indicating the presence of a cointegrating vector between the actual and forecasted series, which is evidence of the presence of rationality in the expectations formation process.

I. Introduction

Inflation forecasts are an important part of macroeconomic models, especially in the context of proactive monetary and fiscal policy. In the corporate world, inflation forecasts are used with the prevailing rate of interest to determine the expected real rate of interest, and these forecasts have a significant impact on firms' future investment on factories, equipment, and inventories. Inflation expectations are also used in household expenditure calculations of consumers, especially for the purchase of consumer durables. Government uses inflation forecasts in calculating the long term solvency of the Social Security system.

Research efforts have been aimed at studying the efficiency, unbiasedness and rationality of forecasts such as the Livingston data, the Decision Makers Poll, Michigan Household Survey, Money Market Survey, etc. These forecasts are of questionable relevance, and their usefulness in macro models is doubtful in the absence of these properties. Thus, a necessary condition for using inflation forecasts in the formulation of policy is the credibility of these forecasts.

Results are far from conclusive. Some surveys are conducted qualitatively, and then subsequently converted to quantitative data, making their use questionable. Other studies, such as Evans and Gulamani (1984), Batchelor (1986) and Kanoh and Li (1990), have examined the credibility of qualitative versus quantitative data and the errors in the conversion from one form to the other. A second line of research examines short horizon and long horizon forecasts and the statistical characteristics of the forecasts as the time frame changes. Studies along these lines are Stenius (1986), Thompson and Ottosen (1993) and Gagnon (1996). But research in this area has revolved around measuring the accuracy and credibility of these forecasts through an empirical examination of economic hypotheses such as unbiasedness, efficiency and rationality. Such studies in this field are

Batchelor and Dua (1987), Hvidding (1987), Engsted (1991) and Marini and Scaramozzino (1993).

Thomas and Grant (2000) compares the relative accuracy of three standard inflation forecasting techniques, namely surveys (Livingston and Michigan Surveys), time series methods (ARIMA processes), and structural econometric models (as developed by the Federal Reserve Bank of San Francisco). They use samples from each forecast and compare these with the actual inflation over the 1980s and 1990s. Findings indicate that the surveys are unbiased over the two decades of forecasts, and clearly superior to backward looking ARIMA, as well as structural models like the Fisher equation. Expert formulation of survey forecasts includes all relevant information in the market regarding monetary and fiscal policy, and therefore indirectly implies the incorporation of bandwagon effects by the forecasters in their predictions. A bandwagon effect would result when market participants use the forecasts of other participants in formulating their forecasts, instead of doing so independently (i.e., participants jump on to the bandwagon.)

This study examines the rationality of inflation forecasts using the Survey of Professional Forecasters data set, and applies a new econometric procedure. The paper is divided into 5 sections. Section 2, discusses the model and the data set used, followed by the non-stationarity tests in section 3. Section 4 sequentially applies the Johansen Multivariate Cointegration test followed by the recently available Bierens Non-Parametric Cointegration technique. These are complementary procedures, and hence enhance the credibility of our results. Section 5 contains some concluding remarks.

II. Model and Data

A standard representation of a rational expectations model is:

$$S_{t+1} = \beta_o + \beta_1\ {}_tS^e_{t+1} + \mu_t \tag{1}$$

where S_{t+1} is the actual inflation rate one period ahead and ${}_tS^e_{t+1}$ is the expected inflation for the next period forecasted (in survey form) in the current period with information set I_t. Thus, the model regresses actual inflation on its forecasted value. Taking expectations of both sides of the equation yields:
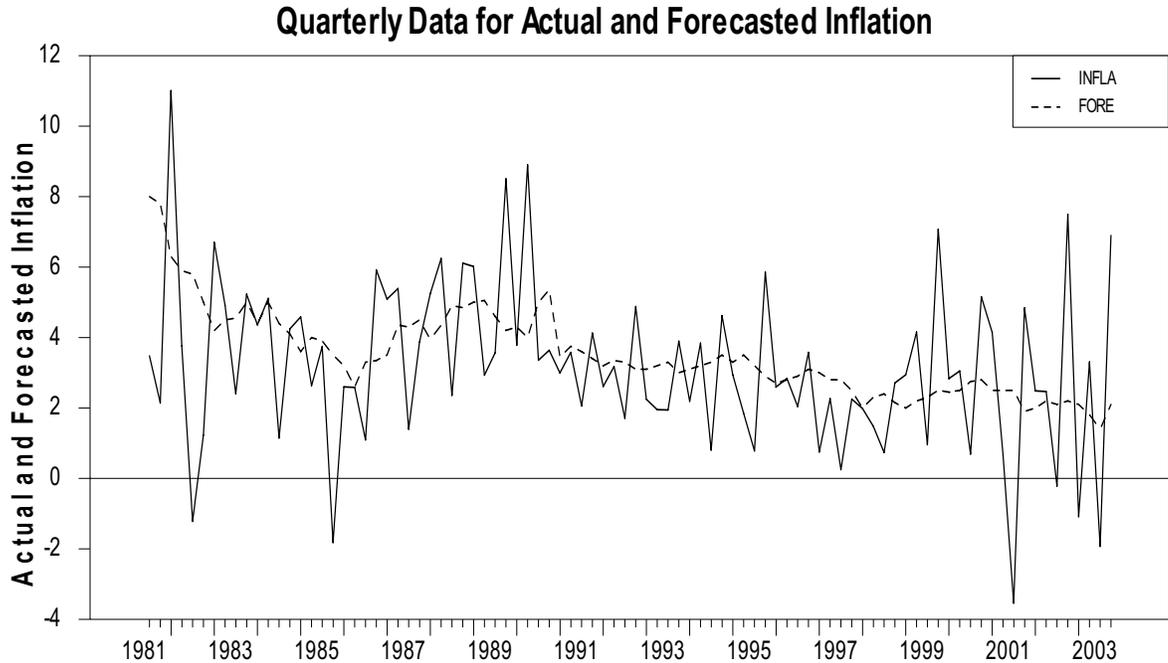
$$E(S_{t+1}) = \beta_o + \beta_1\ {}_tS^e_{t+1} + E(\mu_t) \tag{2}$$

The error term μ_t has an assumed mean of 0, and a test of the Rational Expectations Hypothesis (REH) involves first estimating equation (1), and then testing the error term for stationarity. A stationary error structure would imply that the actual and the forecasted values are cointegrated (move together over time). Cointegration of actual and forecasted values is a necessary condition for the REH since if future inflation and expected inflation rates do not move together over time, we cannot find any relation (including REH) between them.

This paper uses the Survey of Professional Forecasters data set, which is currently maintained by the Philadelphia Federal Reserve Bank (formerly known as the ASA-NBER data set). This survey is an important data source (for policy makers, both public and private) covering a large number of macroeconomic variables over an extended period of time. The survey also includes professional forecasters from business, finance, government and academia, who state their forecasts over multiple time horizons. Data is quantitative, and therefore avoids the controversy associated with any conversion from qualitative data. The model uses quarterly data for actual inflation one quarter ahead (annualized, S_{t+1}) and one quarter ahead (annualized, S^e_{t+1}) forecasts of expected inflation from 1981 (3rd quarter) to 2003 (4th quarter). [(1)] Suitability of the data set for the tests of rationality (since the data is the direct forecast of market participants) along with the

complementary cointegration tests used in this paper (which make the results independent of any particular econometric procedure, and therefore more general) make our results about market rationality particularly relevant to the literature in the field. A graph of the data is given in Figure 1.

Figure 1

Quarterly Data for Actual and Forecasted Inflation

The graph indicates some differences between the two series. Specifically, from approximately 1983-84, 1990-1992, 1997-2001, the two series seem to move in opposite directions, and from 1984-1987 even though both series seem to move together, a large difference in their values is present. An econometric test of cointegration (long run co-movement) is necessary to gather evidence for or against this phenomenon, and the results showing evidence for or against the presence of cointegration between actual and forecasted inflation will have important implications for the use of survey data for research.

III. Tests of Stationarity

Standard tests for stationarity such as the augmented Dickey-Fuller (henceforth ADF, 1979, 1981) and the Phillips-Perron (henceforth PP,1988), are based on the conventional null of no cointegration hypothesis, which is unable to distinguish between unit root and near unit root stationary processes. In addition to these tests, this paper also uses the Kwiatkowski et. al. (henceforth KPSS, 1992) test, where the null is of stationarity and the alternative is the presence of a unit root. The KPSS procedure ensures that the null hypothesis will be rejected only when the results indicate there is strong evidence against it.[2]

Table 1: Unit Root Tests

ADF Test (Null Hypothesis H_0: Unit root without a drift or time trend)		
	t-test Statistic	Critical Value
Actual Inflation	-1.98	-2.897
Forecasted Inflation	-3.88	-2.897
Phillips-Perron Test (Z_t test statistic) (Null Hypothesis H_0: Unit root without a drift or time trend)		
	Test Statistic	Critical Value
Actual Inflation	-10.25	--2.897
Forecasted Inflation	-3.75	-2.897
KPSS Test		
	η_τ	η_μ
Actual Inflation	0.0702	0.4955
Forecasted Inflation	0.1403	3.2575
Critical Value at 5 % level for the KPSS test: $\eta_\tau = 0.146$ and $\eta_\mu = 0.463$.		

Notes:

All three tests were run using RATS 6.02b. The critical values are provided by the software. ADF test indicates that there is a unit root in the actual inflation series, but not in the forecasted series. The Phillips-Perron series indicates that there is no unit root in either series. Due to the drawbacks of these two tests, as described in section 3, we use the results of the KPSS test.

The KPSS test η_τ and η_μ are test statistics that correspond to the null of stationarity with and without a time trend respectively. Since the estimated test statistics for η_τ with time trend) is not significant, and η_μ without time trend) is significant (greater than the critical value), as explained in note 3, the appropriate time series model for our data is without a drift or a trend, η_μ is the appropriate statistic and it indicates that both the current inflation and the one year ahead forecast series have a unit root.

Results for the ADF, PP and KPSS tests are given in Table 1. The ADF test cannot reject the null of a unit root for the actual inflation series; whereas, it does reject the null for the forecasted inflation series. The PP test rejects the null of a unit root for both the actual and forecasted series. The KPSS test statistics η_τ and η_μ are the null of stationarity with and without a time trend respectively. The test statistic for the null hypothesis of stationarity (when the model includes a time trend) is insignificant, whereas, the test statistic without the time trend is significant. As explained in endnote 3, evidence suggests that the data can be accurately described by a model that does not include a time trend. Therefore, the η_μ statistic is appropriate, and indicates rejection of the null hypothesis of stationarity in favor of the existence of a unit root. Based on all three tests, the reasonable observer may therefore conclude that both the actual and forecasted inflation series have a unit root.[3]

V.   System Cointegration Tests

Once the non-stationarity of the relevant series is confirmed, the study runs two systems cointegration tests, namely the Johansen-Juselius (henceforth JJ, 1990) multivariate test, followed by the Bierens non-parametric cointegration (NPC) test. These two tests are complementary procedures and help strengthen our results by making them independent of any specific cointegration test.

In applying the JJ test, first the appropriate lag length was selected using the likelihood-ratio test (Greene, 1993). In case of the trace test (TT), Johansen and Juselius recommend starting with the null hypothesis of the number of cointegrating vectors $r = 0$ and then moving upwards. The idea is that accepting $r \leq n$ implies that one should also accept the $r \leq n+I$, where $I =1,2,3...$, one stops the first time he/she is unable to reject the null. According to JJ (1990) for conclusive evidence regarding the exact number of cointegrating vectors in the system, one has to check the maximum eigenvalue test (MET) results. Here the null hypothesis is specific about the number of cointegrating vectors in the system as $r = I$, where $I = 0, 1, 2, ...$ for the complete model and therefore implies that we are testing a null hypothesis of $r =0$ against an alternate hypothesis that $r =1$, $r =1$ against $r =2$, etc.

Table 2: Johansen-Juselius System Test

NH	TT	Critical Value	NH	MET	Critical Value
$r = 0$	87.56	20.17	$r = 0$	97.19	15.75
$r \leq 1$	9.63	9.09	$r =1$	9.63	9.09

Notes:
NH: Null Hypothesis, TT: Trace Test, MET: Maximum Eigenvalue Test. Since none of the estimated test statistics for either the trace test or the maximum eigenvalue test exceeds the critical values, we are unable to reject the null hypothesis that there are no cointegrating vectors. Critical values are taken from the Johansen-Juselius (1990) paper.

Both the trace (TT) and the maximum eigenvalue test (MET) yield significant statistics at the 5 percent level in Table 2. But the null hypothesis of $r \leq 1$ (trace test) and $r=1$(maximum eigenvalue test) would not be rejected at significance level less than 5% (for example at 1%), indicating that one cointegrating vector exists in the system. Thus, the actual and forecasted series may be cointegrated, that is, they move together over the long run, and this result can be confirmed by estimating the Bierens nonparametric test.

Next, the application of the Bierens nonparametric cointegration (henceforth NPC, 1997) test examines for cointegration between the relevant series.[4] The Bierens procedure is not only a consistent estimator of the number of cointegrating vectors in the system under consideration, it is also a nonparametric test, and hence does not need any specification of the data generating process, making it highly flexible in application and avoiding the estimation of structural (and sometimes nuisance) parameters. Bierens (1997) states that "… our approach is capable of giving the same answers regarding the number of cointegrating vectors and the cointegrating vectors themselves as Johansen's ML method, with less effort." He also includes some Monte Carlo results which indicate that his test performs a little better than Johansen's Lambda-max test. He goes on to conclude that "…. Our approach cannot completely replace Johansen's approach, because the latter provides additional information, in particular regarding possible cointegrating restrictions on the drift parameters, and the presence of linear trends in the cointegrating relations……Thus, rather than being substitutes, the two approaches are complements." Consequently, this study looks at the results of the Johansen and the Bierens' procedures together, and not independent of each other. A brief description of the Bierens procedure is given in Appendix A.

Table 3A: Bierens' Non-parametric Cointegration Test

m	Lambda-min	5 percent critical region	Result
2	0.01059	(0, 0.017)	Reject the H_0

We are unable to reject the null hypothesis of zero cointegrating vectors as the calculated value of lambda-min does not lie in the critical region. The critical region is calculated by the Easyreg international software, which was used to estimate the Bierens procedure.

Here the paper tests the null hypothesis (H_0): r = 0, implying there are no cointegrating vectors in the system against the alternative hypothesis (H_1): r = 1, implying one cointegrating vector. The null hypothesis will be rejected if the calculated value of the lambda-min statistic lies inside the critical region (is less than the critical value). The null hypothesis is rejected in favor of the alternate hypothesis of cointegration, implying that there is one cointegrating vector in the system. The presence of one cointegrating vector is confirmed by estimating the function $\hat{g}_m(r)$ (as outlined in section 4.4 of Bierens (1997), which will converge in probability to infinity if the true number of cointegrating vectors is not equal to r, and will converge in probability to zero if the true number of cointegrating vectors is equal to r. Results are given in Table 3B.

Table 3B

R	$\hat{g}_m(r)$
0	989.63
1	245.40
2	63399.92

Since r=1 has the lowest value of $\hat{g}_m(r)$ this is further proof that there is one cointegrating vector.

Since r = 1 has the lowest value of $\hat{g}_m(r)$ (and this is much lower than the value of $\hat{g}_m(r)$ for the other two values of r), this just confirms the conclusion that the number of cointegrating vectors is one. Thus, the one-quarter ahead (annualized) inflation rate series and its forecasts are cointegrated. Since (as discussed above) Bierens (1997) suggests that the JJ(1990) test and his test are complementary procedures, the JJ results given above and the results from the Bierens tests together indicate that inflation forecasts and actual inflation are cointegrated.

VI.   Conclusion

This study examines the experts' inflation expectations formation process for unbiasedness over a one quarter ahead (annualized) forecast horizon. The empirical results support the presence of cointegration between the actual and the forecasted inflation series, thus indicating rational formation.

Over a time horizon of one year (and longer) it is more than likely that fundamental determinants of inflation will be relevant to forecasting, rather than just historical data. Therefore, the market is possibly unable to accurately forecast inflation over a long time horizon because of a lack of consensus on the fundamental determinants of inflation and how certain economic policies will have an impact on inflation. Englander and Stone (1989) get a similar result, as they conclude that inflation expectations, in spite of having a significant forward looking component, are not

"rational" in the theoretical sense. Engsted (1991), using cointegration (error correction) techniques, also finds evidence against rational expectations in that sense.

Dutt and Ghosh (1999) examined unbiasedness of the CPI expectations process at multiple time horizons (from one to four quarters ahead forecasts). They report the actual and forecasted series to be cointegrated at the short (one quarter horizon) and not cointegrated as the forecast horizon grows longer and longer, which is consistent with the results reported in this paper. Thus cointegration between actual and forecasted inflation supports the bandwagon effect hypothesis in the expectations formation process, at least over the short run (as examined here) i.e., the actual and forecasted series move together in the statistical sense. In the long run, economic fundamentals become more important in forecasting inflation than just the historical values of inflation, leading to a rejection of cointegration between actual and expected inflation, since fundamental determinants of inflation are not included in cointegration studies.

References

Batchelor, R. A. (1986). Qualitative versus quantitative measures of inflation expectations. *Oxford Bulletin of Economics and Statistics*, *48*, 99-120.

Batchelor, R. A., & P. Dua (1987). The accuracy and rationality of UK inflation expectations: Some quantitative evidence. *Applied Economics*, *19*, 819-28.

Bierens, H. J. (1997). Nonparametric cointegration analysis. *Journal of Econometrics*, *77*, 379-404.

Caner, M., & Kilian, L. (2001). Size distortions of tests of the null hypothesis of stationarity: Evidence and implications for the PPP debate. *Journal of International Money and Finance*, *20*(5), 639-657.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of American Statistical Association*, *74*, 27-31.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, 49*, 1057-72.

Dutt, S. D., & Ghosh, D. (1999). Are price forecasts rational? An empirical examination using survey data. *Kentucky Journal of Economics and Business*, *18*, 13-25.

Elder, J., & Kennedy, P.E. (2001). Testing for unit roots: What should students be taught. *Journal of Economic Education*, *32*(2), 137-146.

Elliot, G., Rothenberg, T.J., & Stock, J.H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, *64* (4), 813-836.

Englander, A. S., & Stone, G. (1989). Inflation expectations surveys as predictors of inflation and behavior in financial and labor markets. *Federal Reserve Bank of New York,* Research paper # 8918.

Engsted, T. (1991). A note on the rationality of survey inflation expectations in the United Kingdom. *Applied Economics*, *23*, 1269-76.

Evans, G., & Gulamani, R. (1984). Tests of rationality of the Carlson-Parkin Inflation Expectations Data. *Oxford Bulletin of Economics and Statistics*, *46*, 1-17.

Gagnon, J. E. (1996). Long memory in inflation expectations : Evidence from international financial markets. *International Finance Discussion Papers*, No.538.

Greene, W. H. (1993). *Econometric Analysis*, 2nd edition, New York: Macmillan Publishing Company.

Hvidding, J. M. (1987). Measurements errors and tests for expectational rationality using survey data. *Southern Economic Journal*, 110-18.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration-with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, *52*, 169-210.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54*, 159-78.

Kanoh, S., & Li, Z. D. (1990). A method for exploring the mechanism of inflationary expectations based on qualitative survey data. *Journal of Business and Economic Statistics*, *8*, 395-403.

Marini, G., & Scaramozzino, P. (1993). Inflation expectations, price flexibility and output variability in macro models. *Structural Change and Economics Dynamics*, *4*, 367-83.

Perron, P. (1988). Trend and random walks in macroeconomic time series. *Journal of Economic Dynamics and Control*, *12*(2), 139-162.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrica, 75*, 335-46.

Steinus, M. (1986). Separating short term and long term inflation expectations using observations from financial markets. *Empec, 11*, 57-63.

Thomas, L. B., & Grant, A.P. (2000). Forecasting inflation-surveys versus other forecasts. *Business Economics*, 9-18.

Thompson, D. N., & Ottosen, G.K. (1993). Long term inflation expectations, the personal savings rate and credit standards. *Business Economics*, 45-50.

Appendix A

Bierens' Non-parametric Cointegration Procedure [5]

The data are assumed to be generated by the following expression:

$$z_t = \mu + z_{t-1} + u_t \tag{1}$$

"z_t is a q-variate unit root process with drift, u_t is a zero mean stationary process and μ is a vector of drift parameters. Bierens makes the following assumptions:

Assumption 1: "The process u_t can be written as given above with v_t i.i.d $N_q(0,I_q)$ and $C(L) = C_1(L)^{-1}C_2(L)$, where $C_1(L)$ and $C_2(L)$ are finite-order lag polynomials, with all the roots of $\det(C_1(L))$ lying outside the complex unit circle."

Assumption 2: "Let R_r be the matrix of eigenvectors of $C(1)C(1)^T$ corresponding to the r zero eigenvalues. Then the matrix $R_r^T D(1)D(1)^T R_r$ is nonsingular." It is assumed that the cointegrating relations $R_r^T z_t$ are stationary about a possible intercept but not a time trend.

Assumption 3: $R_r^T \mu = 0$

The Bierens test is based on the matrices

$$\hat{A}_m = \sum_{k=1}^m a_{n,k} a_{n,k}^T, \qquad \hat{B}_m = \sum_{k=1}^m b_{n,k} b_{n,k}^T, \quad m \geq q, \text{ and } a_{n,k}, b_{n,k} \text{ are defined in Bierens}$$
(1997).

Theorem 1 of Bierens (1997) states:

"Let $\hat{\lambda}_{1,m} \geq \ldots \geq \hat{\lambda}_{q,m}$ be the ordered solutions of the generalized eigenvalue problem $\det[\hat{A}_m - \lambda(\hat{B}_m + n^{-2}\hat{A}_m^{-1})] = 0$, and let $\hat{\lambda}_{1,m} \geq \ldots \geq \hat{\lambda}_{q-r,m}$ be the ordered solution of the generalized eigenvalue problem

$$\det\left(\sum_{k=1}^{m} X_k^* X_k^{*T} - \lambda \sum_{k=1}^{m} Y_k^* Y_k^{*T}\right) = 0 \qquad (2)$$

where the $X_i^*{}'s \ and \ Y_j^*{}'s \ are \ i.i.d. \ N_{q-r}(0, I_{q-r})$. If z_t is cointegrated with r linear independent cointegrating vectors, then under assumptions 1-3 ($\hat{\lambda}_{1,m} \geq \ldots \geq \hat{\lambda}_{q,m}$) converges in distribution to ($\hat{\lambda}_{1,m} \geq \ldots \geq \hat{\lambda}_{q-r,m}, 0, \ldots 0$)." The test statistic $\hat{\lambda}_{q-r,m}$ can be used for testing the null hypothesis H_r that there are r cointegating vectors against the alternative H_{r+1} (that there are r+1 cointegrating vectors). This is called the lambda-min test. The critical values for the test are given in an appendix to Bierens (1997). This is a left sided test, i.e., if the test statistic is less than the critical value, then the null hypothesis will be rejected. The choice of the parameter "m" is important to this procedure, as it has an impact on the limiting distribution of the lambda-min statistic and the critical values and the power function. The power of the null hypothesis against the alternative is given by

$$P(\hat{\lambda}_{q-r,m} \leq K_{\alpha,q-r,m}) \approx P(\lambda_{1,m}^* \leq n\sqrt{K_{\alpha,q-r,m}})$$ where $K_{\alpha,q-r,m}$ are the 100% critical

values. We can choose "m" such that the right-hand side of the above expression is maximal, subject to the condition m≥q.

Notes

1. Though most forecasts of macroeconomic variables are available from 1968 onwards, the inflation series was collected from 1981. The forecasts were obtained from the *Survey of Professional Forecasters* data set, available on the web site of the Federal Reserve Bank of Philadelphia. The actual inflation rate was obtained from the web site of the *Bureau of Economic Analysis*.

2. The DF and PP tests are among the first, and most widely used, unit root tests. The KPSS test is an alternate approach to testing for unit roots. We ran all the tests and found similar results. The results of the DF, PP and KPSS results are reported in Table 1. It was pointed out by a referee that a discussion of testing for unit roots with or without a trend should be included. As pointed out by Perron (1988), and Elder and Kennedy (2001), a general to specific strategy (starting with a model with a drift and a trend and working our way down) would be appropriate. Perron (1988) does describe how such a strategy should be implemented. However, Elder and Kennedy (2001) point out that this procedure is needlessly complicated. They suggest that it is straightforward to determine whether a time series model should include a time trend, based on theoretical arguments and graphs, and it is not necessary to have an econometric test for this. They also show that in such cases the ADF t-test is equivalent to testing for the presence of a time trend, and more complicated F-tests, or complicated stepwise procedures like those recommended by Perron (1988) are not necessary. Based on

this strategy, it would seem that the appropriate null hypothesis is a unit root with no time trend and no drift, and we are unable to reject this null hypothesis.

3.   It has been pointed out to us that the KPSS test suffers from severe size distortions, and therefore multiple unit root tests should be included. Therefore, we have included results from ADF and PP tests in addition to the KPSS tests, all of which support the conclusion of unit roots in the data. In addition, Caner and Killian (2001), who discuss this issue of size distortions in some detail, also suggest that the power of the asymptotically efficient DF-GLS test (as discussed in Elliot et. al. (1996)) compares favorably to the ADF test. We ran the DF-GLS test on our data and the results supported the presence of unit roots in both the actual and forecasted inflation series.

4.   The computations for the Bierens procedure were done using the EasyReg International software made available by Herman Bierens on the web-site: http://econ.la.psu.edu/~hbierens/EASYREG.HTM.

5.   We give a brief description about the Bierens' NPC procedure since the details of the procedure can be obtained from that paper.