Are Hodrick-Prescott 'Forecasts' Rational?

by

J.C.K. Ash,[a] J.Z. Easaw,[b] S.M. Heravi[c] and D.J. Smyth[b]

Abstract

We examine the proposal that inflationary expectations might be proxied using the Hodrick-Prescott (HP) filter - specifically that the HP filter might stand as an *ex post* proxy for corresponding *ex ante* rational expectations. We apply a battery of tests for rationality to long time series of US data . Our conclusion is that while the HP series are not fully rational in the sense of Muth (1981), they do meet the criterion of 'weak rationality' recently proposed by Grant and Thomas (1999). They are also rational proxy predictors of direction for, following Merton (1981), agents would not change their prior in the opposite direction to these 'forecasts'.

JEL classifications: C53; E37.

Keywords: Hodrick-Prescott filter; Rational expectations; Proxy forecasts.

1. Introduction

Consider the following problem. An economic model-builder requires a time-series of expected inflation for a relatively distant historical period. (Perhaps the model builder's intention is to estimate an expectations augmented Phillips Curve or Lucas surprise supply function.) Further, these expectations are required to be rational - that is, they are unbiased and take into account all available information.

Suppose that no contemporaneous *ex-ante* inflation forecasts are extant. This is quite likely. For example, quantitative forecasts for the UK began to be published by H.M. Treasury in 1947, becoming more confident and elaborate until 1951. Then, for a decade, official quantitative macroeconomic forecasts were not published at all (Dow,

[a] Department of Economics, University of Reading.
[b] Economics Group, Middlesex University Business School.
[c] Cardiff Business School, University of Wales.

1965). It was not until 1970 that a formal computerised econometric version of the Treasury's short-term aggregate model was fully operational, providing the basis of the forecasts published at the time of the annual Budget (Ash and Smyth, 1973). Though formal model-based forecasts for the USA were published somewhat earlier than in the UK, Zarnowitz (1967), in his pioneering appraisal of short-term forecasts made by business, government and academic institutions, reports that few reached further back than the early 1950's, and none before 1947.

In any case, as a recent survey article reveals, there is no guarantee that even if these forecasts existed they would be rational (Stekler and Fildes, 1999). Laster *et al* (1999) go so far as to propose a theory of 'rational bias' in macroeconomic forecasting, in which forecasters compromise the accuracy of their projections to gain publicity; their paper provides evidence that professional forecasting has an important strategic component. The same caveat applies to the Livingston and Michigan series of surveys of expected inflation, which date back to 1946 and 1948 respectively: while some studies support their rationality, for example Bryan and Gavin (1986), Mullineaux (1978) and Rich (1989), at least as many reject it - see Carlson (1977), Gramlich (1983), Figlewski and Wachtel (1981), and Smyth (1992).

Rational forecasts are model-consistent: they are the forecasts that would be generated by the 'true' model of the economy, were such a model to exist. Conceivably therefore, our researcher might attempt to estimate the economic structure generating the inflationary process with a view to obtaining *ex-post* rational predictions. However, the prospect of successfully doing so must be remote, certainly if the sample period pre-dates the Second World War, not least because of data limitations. Fortunately, a much simpler, more mechanical approach has recently been suggested. In recent OECD models (Orr *et al*. 1995, and Martins and Scarpetta, 1999), inflation expectations series are proxied using the Hodrick-Prescott (HP) filter, formulated in Hodrick and Prescott (1997). Smant (1998) goes further, arguing that the HP filter generates a price series which is consistent with price expectations formed rationally.

The HP filter is a two-sided linear filter that computes a smoothed series x_t of y_t, for $t = 1,\ldots T$. This is done by minimizing the variance of y_t around x_t, while subjecting it to a penalty that constrains the second difference of x_t. The HP filter chooses x_t to minimize as follows

$$\sum_{t=1}^{T}(y_t - x_t)^2 + l\sum_{t}^{T-i}((x_{t+1} - x_t) - (x_t - x_{t-1}))^2$$

The penalty parameter, λ, controls the smoothness of the x_t series. The larger is the parameter, the smoother the series. This is not the place to discuss the statistical properties of the HP filter; for such information the reader is referred to Cogley and Nason (1995), Harvey and Jaeger (1993), and King and Rebelo (1993). Rather, we focus on its economic applications. Introduced as a detrending process, the HP filter has been widely used to estimate potential output and hence the output gap (Laxton *et al*, 1995, and St.-Amant and van Norden, 1997). Because both the HP filter and rational forecasts both incorporate a substantial element of 'perfect foresight,' Smant (1998) suggests that the former might stand as an *ex-post* proxy for the latter. This is the hypothesis tested in the remainder of this paper.

2. The Data.

We evaluate the usefulness of HP-filtered time series as a proxy for rational expectations using long runs of annual data for the US economy. (Following Hodrick and Prescott's (1997) recommendation for annual data, we set the smoothness parameter at $\lambda = 100$.) Our focus is on expected price inflation, both in the GDP deflator and the consumer price index. We analyse percentage changes in these deflators. In addition, we take the second difference in these series, i.e. changes in percentage changes, in order to assess whether the HP filter retrospectively 'predicts' accelerations and decelerations as well as it forecasts the inflation rate. For comparison, we also evaluate the rationality of a genuine *ex ante* inflation forecast, the Livingston survey of consumer price expectations. Finally, because of the attention the HP filter has received as a means of estimating potential output, we investigate whether it would also serve as a useful proxy for a rational expectation of GDP.

The sample period differs somewhat between variables: all are tested over the post-World War II period, while HP filters of GDP and its deflator are also tested over a span of 119 years, from 1878 to 1996. Details of the variables, our notation, and the sample periods are as follows:

P = Gross Domestic Product implicit price deflator, percentage change, 1878-1996 and 1946 - 1996. Source *Gordon (1997)*.

DP = change in P.

PC = Consumer Price Index, annualised percentage change, 1947-1997. Source *Citibase*.

DPC = change in PC.

LIV = Livingston 12-month-ahead survey of expected Consumer Price Index, annualised percentage change, 1947-1998. Source *The Philadelphia Inquirer* and *Federal Reserve Bank of Philadelphia*.

DLIV = change in LIV.

Y = Real Gross Domestic Product, percentage change, 1878 - 1996 and 1946 - 1996. Source *Gordon 1997).*

3. Techniques of Analysis.

A number of rationality tests are applied to the HP-filtered time series. The essence of the rational expectations hypothesis is that “agents make use of all available information by deriving their expectations of the future values of variables from the underlying true economic model that (theoretically, at least) generates the variable to be forecast” (Holden et al., 1985, p. 25). Rationality implies that forecasts are unbiased, efficient and consistent. Following, for example, Holden et al. (1985), we define unbiasedness, efficiency and consistency as follows. Unbiased forecasts have the same mean as the actual outcomes. Efficiency requires that the forecasts utilize all relevant information available at the time that they are made; if they did not, accuracy could be improved by incorporating the extra unused information. Forecasts which span overlapping time periods should use known information consistently, such that one cannot predict the way in which these forecasts are revised. Consistent forecasts are therefore revised only in the light of new information. (Here, as we do not attempt to use the HP filter to simulate forecast revisions, no tests for consistency are relevant.) If forecasts do not possess these properties of rationality, their accuracy could be improved by using publicly available information apparently ignored by the forecasters.

It can be argued that the presence of serial correlation in forecast errors is not wholly inconsistent with rationality. First, due to overlapping forecasts intervals, forecasters might be unaware of their recent errors at the time new forecasts are made. Expecting agents to learn quickly from their mistakes when they have incomplete information concerning the nature of those mistakes is too stringent. Second, Cukierman and Meltzer (1982) maintain that survey based expectations series may exhibit serial correlation if agents are unaware of the stochastic nature of the inflation-generating process. If shocks to inflation are assumed to be temporary when in fact they are permanent, agents may consistently under- or overpredict inflation for several consecutive periods. Grant and Thomas (1999) define expectations series which are unbiased but have serially correlated forecast errors as ‘weakly rational’.

Let A_t, t = 1,2,…,T, denote a time series of outcomes, and F_t, t = 1, 2, …, T, the corresponding HP-filtered series. We hypothesize that F_t is an *ex post* proxy for a rational forecast of A_t. To test for unbiasedness we first estimate the regression of outcomes, A_t, on the corresponding proxy prediction, F_t, obtained from the HP filter:

$$A_t = \alpha_0 + \alpha_1 . F_t + u_t \tag{1}$$

Then, following Batchelor and Dua (1989), Holden and Peel (1985) and Rich (1989), we test the joint hypothesis that $\alpha_0 = 0$ and $\alpha_1 = 1$ using a $\chi^2(2)$ static denoted by *XB* in the results tables. Also, Holden and Peel (1990) demonstrate that testing the joint hypothesis that $\alpha_0 = 0$ and $\alpha_1 = 1$ is a sufficient but not necessary condition for unbiasedness. They propose a further test performed directly on the forecast errors,

$$A_t - F_t = E_t = m + u_t \tag{2}$$

TM is a *t*-test with (*T* - 1) degrees of freedom for $m = 0$, where m is the mean error and *N* is the number of observations in the sample.

As highlighted earlier in this section, an expectation series exhibiting serial correlation does not necessarily violate the rational expectations hypothesis, and an expectation series that is tested to be only unbiased can be deemed weakly rational. Nevertheless, as pointed out by Grant and Thomas (1999), even when establishing weak rationality the test for unbiasedness as outlined by equation (1) needs to account for actual and expected series that maybe nonstationary. They argue that while the test for unbiasedness is a necessary condition for establishing weak rationality, it is not a sufficient one. In the presence of nonstationary series, establishing whether a linear combination the actual and expected series is stationary is an essential prelude to establishing weak rationality.

If the time paths of the actual and expected series evolve in different ways, the error will be nonstationary and the estimation of equation (1) may overstate the influence of one variable on the other. Furthermore, deviations from the equality of predictions and outcomes will not be eliminated over time. On the other hand if deviations from this equality are temporary, the actual and expected series are said to be cointegrated, and the coefficient vector [1, - α_1] is termed the cointegrating vector. Thus suppose that $A_{t-1} > F_{t-1}$: equality between prediction and outcome can be restored in period *t* by a decline in the actual series, an increase in the expected series, or a combination of the two, as indicated in the following system of regression equations:

$$\begin{aligned} \Delta A_1 &= d_A\,(A_{t-1} - a_1 F_{t-1}) + e_{1A} \qquad & d_A &< 0 \\ \Delta F_1 &= d_F\,(A_{t-1} - a_1 F_{t-1}) + e_{1F} \qquad & d_F &> 0 \end{aligned} \tag{3}$$

Grant and Thomas (1999) strongly argue that equation system (3) highlights a number of factors relevant to assessing expected rationality and the behaviour of actual and expected series in general. The error correction representation depends critically on the structure of the expectation errors. Serial correlation in equation (1) reveals information about the adjustment of the expected series to past mistakes and provides useful information in estimating the behavioural equations in equation system (3).

Furthermore, the speed of adjustment coefficients are insightful when assessing the rationality of the expected series. If forecasters do not respond to previous prediction errors, δ_F will not be significantly different from zero. In such a scenario agents systematically ignore useful information, and we can reject full rationality. Finally, the system above highlights the two-way feedback between actual and expected series. If both δ_F and δ_A are significantly different from zero, then not only do forecasters respond to the behaviour of the actual series, but the actual series responds to the behaviour of the forecasters. This is one of the fundamental propositions of the rational expectations paradigm.

Grant and Thomas (1999) advocate that the system of error-correction regressions can be extended and specified within a VAR framework. Hence the method outlined in Johansen (1988) would be an appropriate method for determining cointegrating relationships between nonstationary actual and expected series. Of the various series investigated here, only the actual inflation series based on the Consumer Price Index (PC) and the Livingston survey of expected inflation series (LIV) are nonstationary: both Paquet (1992) and Grant and Thomas (1999) also found similar trends in the actual inflation series and expected inflation series based on the Livingston survey. We now examine whether the linear combination of the expected series based on the Livingston survey and actual inflation, and the expected series based on the HP filter and actual inflation are respectively stationary. Table 1 gives the results using the Johansen cointegration method: they show that for each expected series there exists a single cointegrating relationship. This enables us to test for unbiasedness in order to establish whether the expected series are at least weakly rational.

All relevant, available information is incorporated by efficient forecasts. As a test of efficiency, Mullineaux (1978) proposes that the forecast error should be uncorrelated with any element in the set of information available at the time the forecasts are prepared. Important elements in this information set are past outcomes, forecasts and errors. Forecasts for *s* steps ahead ought to have error autocorrelation of order *s* or greater equal to zero, for otherwise the error E_t will be correlated with information available to the forecasters which could have been used to improve the forecast. *BLE* and *BLR* are Box-Ljung tests of, respectively, the forecast errors, E_t, and the residuals from equation (1). The Box-Ljung statistic, *BL*, is defined as

$$BL = \frac{(T+2)\sum_{j=s}^{n}(T-j)^{-1}\hat{r}_j^2}{\left(1+2\sum_{j=1}^{s-1}\hat{r}_j^2\right)\Big/ T}$$

where the $\hat{r}_j$ are estimated autocorrelation coefficients, *T* is the number of observations, and n = 4 are the number of lags used in these tests. The number of forecast steps ahead is denoted by *s*; as we use the HP filter as if it were a one step ahead annual predictor, s =1 here. Box-Ljung tests follow the c^2 distribution, here therefore with four degrees of freedom. Also, following McNees (1978) and Figlewski and Wachtel (1981), we focus on the most recent known forecast error as a test for the orthogonality of errors to available information. Thus *XO* is a $c^2(2)$ test of the joint hypothesis $b_0 = b_1 = 0$ in the regression.

$$A_t - F_t = b_0 + b_1(A_{t-1} - F_{t-1}) + u_t \qquad (4)$$

Throughout these tests, reported in Table 2 below, an asterisk denotes departure from the rationality criterion at the 5 percent significance level.

HP filtered data may be used *ex post* as a proxy forecast irrespective of rationality criteria. So we also report, in Table 4 below, a measure of the accuracy with which the HP filter predicts the outcome series, along with various diagnostic checks on these *ex-post* ‘forecasts’. Our accuracy measure is the widely used Theil (1966) inequality coefficient, *U*:

$$U = RMSE / \sqrt{(\sum A_t^2 / T}$$

where RMSE is the root-mean-square error, and summation $\sum$ is for t = 1 to T. The inequality coefficient is zero only in the case of perfectly accurate forecasts, rises with inaccuracy, and has no upper bound. As all the series we examine here are changes, not levels, U = 1 for any HP series as inaccurate as a naïve repetitive no-change prediction.

The square of the numerator of the inequality coefficient is the mean squared error which, as Theil (1961) demonstrates, can be decomposed in two alternative ways. The decomposition preferred by Granger and Newbold (1973) results in the following inequality proportions:

$$UM = (FM - AM)^2 \;/\; MSE$$

$$UR = (FSD - R.ASD)^2 \ / \ MSE$$

$$UD = (1 - R^2).ASD^2 \ / \ MSE.$$

FM and *AM* denote the means of the forecast and actual series respectively, the corresponding standard deviations are *FSD* and *ASD*, and *R* is the sample correlation coefficient between predictions and outcomes.

The inequality proportions *UM*, *UR* and *UD* are best explained in the context of equation (1), the regression of outcomes on predictions. *UM* is the proportion of forecasting error (as measured by *MSE*) due to bias, in the sense of over- or underpredicting the mean outcome. The regression coefficient a_1 takes the form *R.ASD/FSD*. For optimal forecasts, this coefficient equals unity. Therefore Theil (1966) calls *UR* the ‘regression proportion’, because it deals with the deviation of the regression slope from unity. It is the proportion of *MSE* due to misforecasting the systematic component of the variance of outcomes. *UD* is the disturbance or residual proportion, because it deals with the variance of the regression disturbances or residuals.

For a series of optimal forecasts the following conditions hold:

$$a_0 = 0; \quad FM = AM; \quad UM = 0$$

$$a_1 = 1; \quad FSD = R.ASD; \quad UR = 0$$

UD should therefore tend to unity. If these conditions did not hold, the forecasts’ accuracy could in principle be improved by a simple linear correction. In Table 3 $IN = \hat{a}_0$ and $SL = \hat{a}_1$, respectively the intercept and slope parameters in equation (1), and an asterisk denotes $\hat{a}_0 \neq 0$ or $\hat{a} \neq 1$ at the 5 percent significance level.

A well-known property of an optimal forecast is that it should understate the variability of outcomes. The actual variance of a stochastic economic process includes, in addition to a systematic component replicated by its optimal predictor, the variance of random disturbances which post-date the forecast. We therefore compute the ratio of the standard deviations:

$FASD = FSD/ASD$

In Table 3, an asterisk indicates that the *FASD* value is less than unity at the 5 per cent significance level.

A simple test of whether accuracy has improved or deteriorated during the sample period involves the regression

$$|E_t \ / \ A_t| = b_0 + b_1.TREND + u_t \qquad (5)$$

TREND is a time trend. In Table 3, *TIME* = $\hat{b}_1$, and an asterisk denotes $\hat{b}_1 \neq 0$ at the 5 per cent significance level.

For some purposes, such as the successful timing of changes in the direction of policy, it may be more harmful to make a smaller prediction error yet mis-forecast the direction of change, than to make a larger, directionally correct error. Therefore the overall prevalence of turning point errors is calculated:

$TPE = NTPE / N$

NTPE is the number of turning-point errors, that is the number of pairs of observations for which the sign of F_t differs from that of A_t.

So far, our criteria for judging whether or not HP proxy forecasts are rational stem from Muth's long-standing and commonly used definition that rational expectations "are essentially the same as predictions of the relevant economic theory. In particular, the [rational expectations] hypothesis asserts that the economy generally does not waste information, and that expectations depend specifically on the structure of the entire system" (Muth, 1961, p. 315). The upshot is the analysis of the properties of the quantitative forecast errors which we have described above. However, a more recent approach to rationality compares the *direction* of forecasts with that of the corresponding outcomes. The groundwork for non-parametric tests on the direction of forecasts was developed by Merton (1981) and Henriksson and Merton (1981) in the context of whether a market-timing forecast, that is a forecast of when stocks will outperform bonds, or *vice versa*, would have value to an investor. Thus for Merton (1981, p. 384), "a forecast is said to be *rational* if, given the forecast, no investor would modify his prior [distribution for the return on the market] in the opposite direction of the forecast."

Let $p_1(t)$ denote the probability of a directionally correct forecast conditional upon an actual downturn at t; let $p_2(t)$ denote the probability of a directionally correct forecast, conditional upon no actual downturn at t. Merton then shows that a necessary and sufficient condition for the forecast to be rational is that $p_1(t) + p_2(t) \geq 1$. A test of directional rationality for the HP forecasts therefore examines the null hypothesis that $p_1(t) + p_2(t) \geq 1$ against the alternative that $p_1(t) + p_2(t) < 1$. Estimates of probabilities $p_1(t)$ and $p_2(t)$ are obtained from our sample data. Henriksson and Merton (1981) demonstrate that the conditional distribution of these estimates is given by the hypergeometric distribution. We then use our data to calculate the probability of being in the tail of this hypergeometric distribution, and then test against the 5 percent significance level.

Rational forecasts may or may not be useful. Following Merton (1981) and Henriksson and Merton (1981), Stekler (1994, p. 495) defines a macroeconomic forecast as having value "if it could change the user's prior distribution about the direction of change of the economy". Merton (1981) shows that a necessary and sufficient condition for a prediction to have no value is that $p_1(t) + p_2(t) = 1$, and, *assuming directional rationality*, a sufficient condition for positive value is that $p_1(t) + p_2(t) > 1$. (The larger is $p_1(t) + p_2(t)$, the more valuable are the forecasts. In the limit, forecasts which are always directionally correct have $p_1(t) = p_2(t) = 1$, so $p_1(t) + p_2(t) = 2$.) When the null of rational forecasts cannot be rejected, Henriksson and Merton therefore test the hypothesis that the forecasts have no value, i.e. $p_1(t) + p_2(t) = 1$, against the alternative that the forecasts are of positive value, i.e. $p_1(t) + p_2(t) > 1$, proceeding in a way similar to the rationality test.

We follow Stekler, and form the following contingency table to test for the independence of the predicted and actual changes, using two procedures: the χ^2 test and Fisher's Exact Test (Fisher, 1941) denoted in the table by FE. Fisher's Exact Test is the uniformly most powerful unbiased test for independence, and is identical to Henriksson and Merton's test for predictive value.

		Forecast	
		< 0	≥ 0
Actual	< 0	$P_1(t)$	$1 - p_1(t)$
	≥ 0	$1 - p_2(t)$	$P_2(t)$

Pesaran and Timmerman (1992) have also developed a non-parametric test on the correct prediction of the signs of actuals and forecasts. They test for a significant difference between the observed, sample estimate of the probability of a correctly signed forecast, and the estimate of what that probability would be under the null of independence between forecasts and outcomes. We denote their test statistic by S_n^2. When tabulating the results of all three tests, an asterisk denotes that the null hypothesis "H_0: the forecasts and outcomes are independent" is rejected at the five percent level: had they been made at the time, these proxy forecasts would have had value to hypothetical users.

4. Results

4.1 Rationality Tests.

Table 2 reports the results of 'Muthian' rationality tests on the quantitative errors in the HP proxy forecasts. The HP inflation series generally pass both tests for bias: that is, HP data could stand as unbiased proxy forecasts of the inflation rate. First differencing the inflation rate yields data on acceleration and deceleration in prices, presents a more challenging test of a forecast's performance, and leads here to a different result. The HP series on acceleration and deceleration in the GDP deflator (but not consumer prices) are biased: they fail, at the 5 percent significance level, the joint test of $\alpha_0 = 0,\ \alpha_1 = 1$ in equation (1) over both the long 1878-1996 period and the post-World War II sub-sample. Inefficiency, however, is pervasive. Usually the HP series fail both Box-Ljung tests for autocorrelated errors, irrespective of whether the series is the GNP deflator or the Consumer Price Index, measured as the inflation rate or price acceleration/deceleration, and calculated over both long and shorter time periods. In addition, the HP GNP deflator fails the test of $\beta_0 = \beta_1 = 0$ in equation (4): successive forecast errors are linearly dependent. All told, were these HP proxy forecasts in fact authentic *ex ante* predictions, they would not incorporate efficiently all information contained in recent known forecast errors.

Our results for the Livingston Survey data confirm earlier findings that these expectations are not rational either (see Carlson, 1977 and Figlewski and Wachtel, 1981). Like the HP data, there is evidence of inefficiency, though here it is the orthogonality test which is failed rather than the Box-Ljung autocorrelation tests, which provide the primary evidence of HP inefficiency. Unlike the HP 'forecasts', the Livingston series are also clearly biased, failing both tests at the 5 percent significance level.

Finally, we examine whether HP data on output is rational. For the long series, from 1878 to 1996 the answer is no: the data are unbiased but inefficient, according to both Box-Ljung statistics. Post-World War II however, the HP proxy for a forecast of GDP is rational, passing all tests for unbiasedness and efficiency.

4.2 Accuracy and Error Diagnostics.

Table 3 presents inequality coefficients and reveals the sources of error in the HP data, treating these data as if they were genuine forecasts. It should be noted at the outset that the Theil inequality coefficient, U, and its variants are descriptive, and cannot be used to test hypotheses *per se*.

All HP 'forecasts' record inequality coefficients less than unity, indicating a better performance in predicting changes (in prices and output) than that of a very naïve model, the assumption of no change from the previous period. As one might expect, the HP filter 'predicts' the inflation rate more accurately than it does acceleration or deceleration in prices. Similarly, given likely structural changes in the US macroeconomy, inflation and output are more accurately predicted by HP post-World War II than over the long sample of observations for 1878-1976. Also, for the shorter, more recent period, the HP filter outperforms the Livingston survey of consumer price expectations. Table 4 provides further evidence on the relative accuracy of the HP data: for comparison it shows the inequality coefficients calculated for actual one year ahead OECD forecasts of inflation and output for both the USA and all G7 economies - Canada, France, Italy, Germany, UK and USA. Too much should not be made of this comparison, not least because the sample periods for HP and OECD are different. However, the HP data are not markedly inferior to the actual forecasts; indeed the HP filter is clearly the more accurate 'predictor' of acceleration and deceleration in both price indices.

Judged by the size of the inequality proportions shown in Table 3, HP forecasting error is predominantly non-systematic. The UD random proportion of mean squared error is usually well above 90 percent, and appreciably larger than the corresponding values for the Livingston Survey. Four HP series fail the optimality test; that is, in equation (1) either $\vec{\alpha}_0 \neq 0$ or $\vec{\alpha}_1 \neq 1$ or both. A simple linear correction would then in principle improve their accuracy. Three of these four cases underline our earlier finding of bias in the HP series for price acceleration/deceleration. However, sub-optimality of this sort is most apparent in the long output series which the rationality tests show to be bias free. Significant errors in the intercept and slope parameters of equation (1) offset each other, illustrating Holden and Peel's (1990) demonstration that testing the joint hypothesis that $\vec{\alpha}_0 = 0$ and $\vec{\alpha}_1 = 1$ is a sufficient, but not a necessary, condition for unbaisedness. As one would expect - indeed, as one requires of an optimal predictor - the smooth filtered HP series have a smaller variance than the corresponding raw data: FASD values are invariably less than unity, and usually significantly so.

We have already remarked on the serial correlation of HP errors. Here we note that the accuracy of both the HP filter and the Livingston survey improve in the later years of the sample period. All but two of the time-trend coefficients are negative, significantly so in three cases: the GDP deflator over the long period, and the Consumer Price Index post war for both HP and Livingston. Not surprisingly, turning-point errors are most frequent when predicting acceleration or deceleration in prices: at least a quarter of both HP and Livingston forecasts for these series are wrong signed.

4.3 Directional Analysis.

Table 5 presents the results of our non parametric tests of direction. These tests are not defined when there are no downturns in a series, which is indeed the case for the GDP deflator, post-World War II. Otherwise the results for directional rationality are unanimous: all HP series as well as the Livingston survey are rational. There is greater ambiguity about the usefulness of these series for predicting the direction of change. In Table 4 we report results for Fisher's Exact Test (FE) and the χ^2 test for independence of predicted and actual changes, as well as Pesaran and Timmerman's S_n^2 test. All three tests show that four of the HP 'forecasts' are useful, in the present context: the long GDP inflation series; the first difference in this series over both long and short sample periods; and the output series, but only for the period after the Second World War. This last result can be seen as complementing recent research by Canova (1999): he concludes that the HP filter was one of two - out of twelve - detrending methods whose performance best mimics NBER and Department of Commerce post-war US business cycles. The Livingston series for consumer prices also passes our tests for value, whereas the corresponding HP series does not.

5. Conclusions

The main question examined in our paper is this: might an HP filtered series stand as an *ex post* proxy for corresponding *ex ante* rational expectations? From our analysis of US inflation data the answer is a highly qualified 'yes'. In contrast to Livingston survey data on inflation expectations, our results show the HP series for the GDP deflator and for the Consumer Price Index to be generally unbiased. On the other hand, like the Livingston data, they are very clearly inefficient: were they authentic *ex ante* predictions, they would not incorporate efficiently all information contained in the recent forecast errors. So while the HP series are not fully rational in the sense of Muth (1961) and most later authors, they do meet the criterion of 'weak rationality' recently proposed by Grant and Thomas (1999). They are also rational proxy predictors of direction for, following Merton (1981), agents would not change their prior in the opposite direction to these 'forecasts'. Our finding that the HP series are unbiased but have serially correlated errors suggests two directions for future research: reduce the size of the smoothness parameter λ, and/or explicitly model the serial correlation of HP errors. We are at present undertaking experiments along these lines, and hope in due course to report our results.

Whether or not HP series are rational, they may anyway be used as proxy forecasts. OECD researchers, Orr *et al*. (1995) and Martins and Scarpetta (1999), have already done so. It is pertinent therefore to inspect their accuracy and carry out diagnostic checks on their errors. We find that the accuracy of the HP series is comparable with that of genuine *ex ante* inflation forecasts, and more accurate than both the very naïve no-change prediction and the Livingston Survey data. In general they have the properties of an optimal predictor, for example error is predominantly non-systematic. There is greater ambiguity about the usefulness of the HP series for predicting the direction of change but, as Ash *et al*. (1998) and Stekler (1994) have shown, this is a common finding too for *ex ante* forecasts of inflation and output.

Most of our analysis has concerned price expectations. However we have also evaluated the HP series for GDP over both a long, 119 year sample period and post-World War II. Over the long period the familiar conclusion holds: the HP filter is only weakly rational, being unbiased but inefficient. But for the shorter, more recent sub-period the HP series is unbiased *and efficient*. It is also directionally rational and useful. Now as we noted earlier the HP filter is often used to measure *potential* output. The validity of this practice is called into question if the HP series were also to proxy a strong rational expectation of *actual* output. For both procedures to be consistent, the output gap would have to be a purely random phenomenon.

TABLE 1: JOHANSEN RANK CONDITIONS

Series	Lags	Null hypothesis	Johansen trace stat.
LIV	2	$m \leq 1$	1.988
		$m = 0$	17.801*
HP PC	2	$m \leq 1$	3.454
		$m = 0$	33.989*

* Denotes significance at 0.05 level.

Lags are determined using both the Akaike Information Criterion and Schwartz-Bayes Criterion.

m Denotes the rank of the matrix of actual and respective expectations series.

TABLE 2: RATIONALITY TESTS

Series		XB	TM	BLE	BLR	XO	T
P	*1878-1996*	2.006	0.000	38.254*	38.458*	12.243*	119
DP	*1879-1996*	11.067*	0.095	37.019*	19.137*	11.982*	118
P	*1946-1996*	2.402	0.000	19.136*	17.858*	13.317*	51
DP	*1947-1996*	6.596*	-0.729	11.241*	8.611	10.847*	50
PC	*1947-1997*	0.472	0.000	20.120*	20.491*	4.694	50
DPC	*1948-1997*	4.211	-0.028	20.569*	13.501*	5.541	49
LIV	*1947-1997*	11.600*	2.879*	4.127	4.826	11.425*	50
DLIV	*1948-1997*	9.177*	2.695*	5.463	5.148	11.129*	49
Y	*1878-1996*	5.685	0.000	23.503*	21.916*	2.589	119
Y	*1946-1996*	2.008	0.000	3.229	1.892	0.535	51

TABLE 3: ACCURACY AND ERROR DIAGNOSTICS

Series		U	UM	UR	UD	IN	SL	R	FM	AM	FASD	TIME	TPE	T
P	*1878-1996*	0.684	0.000	0.017	0.983	-0.449	1.177	0.653	2.532	2.532	0.554*	-0.442*	0.092	119
DP	*1879-1996*	0.832	0.000	0.086	0.914	0.045	0.713*	0.605	0.037	0.071	0.849	-0.045	0.254	118
P	*1946-1996*	0.376	0.000	0.046	0.954	-0.783	1.175	0.825	4.476	4.476	0.702*	-1.485	0.000	51
DP	*1947-1996*	0.728	0.011	0.108	0.881	-0.243	0.749*	0.720	-0.229	-0.415	0.960*	-0.082	0.340	50
PC	*1947-1997*	0.391	0.000	0.010	0.990	-0.335	1.087	0.773	3.868	3.868	0.711*	-0.024*	0.040	50
DPC	*1948-1997*	0.825	0.000	0.081	0.919	-0.011	0.725*	0.611	-0.013	-0.021	0.843*	0.133	0.429	49
LIV	*1947-1997*	0.464	0.142	0.047	0.811	1.403*	0.818	0.731	3.012	3.868	0.894*	-0.011*	0.060	50
DLIV	*1948-1997*	0.959	0.129	0.029	0.842	0.595	0.745	0.474	-0.827	-0.021	0.637*	4.186	0.388	49
Y	*1878-1996*	0.770	0.000	0.046	0.954	-2.461*	1.707*	0.467	3.483	3.483	0.274	-0.035	0.218	119
Y	*1946-1996*	0.654	0.000	0.039	0.961	-0.882	1.321	0.633	2.747	2.747	0.479*	-0.009	0.157	51

TABLE 4: INEQUALITY COEFFICIENTS FOR HP DATA, 1946-1996, AND OECD FORECASTS 1967-1987

		P	DP	PC	DPC	Y
Hodrick-Prescott,	USA	0.376	0.782	0.391	0.825	0.654
OECD						
	USA	0.234	1.209	0.291	0.993	0.513
	G7 average	0.305	0.987	0.306	0.969	0.653

Source: Ash *et al*. (1990).

TABLE 5: NON PARAMETRIC TESTS OF DIRECTION

Series		N_1	N_2	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}$	s	FE	χ^2	S_n^2
P	*1878-1996*	20	99	0.750	0.848	1.598	1.0000	0.0000*	28.51*	31.888*
DP	*1879-1996*	53	65	0.736	0.723	1.459	1.0000	0.0000*	22.82*	24.83*
P	*1946-1996*	0	51				ND	ND	ND	ND
DP	*1947-1996*	25	25	0.640	0.680	1.320	0.9949	0.0232*	3.93*	5.23*
PC	*1947-1997*	2	48	0.000	1.000	1.000	1.0000	1.0000	0.00	0.00
DPC	*1948-1997*	26	23	0.500	0.652	1.152	0.9140	0.2166	0.62	1.18
LIV	*1947-1997*	2	48	1.000	0.938	1.938	1.0000	0.0082*	9.78*	19.13*
DLIV	*1948-1997*	26	23	0.769	0.435	1.204	0.9664	0.1122	1.48	2.36
Y	*1878-1996*	27	92	0.074	0.989	1.063	0.9893	0.1286	1.31	3.42
Y	*1946-1996*	10	41	0.300	0.976	1.276	0.9992	0.0205*	5.07*	8.62*

N_1 number of outcomes which are negative.

N_2 number of outcomes which are non-negative.

$\hat{p}_1$ estimate of $p_1(t)$

$\hat{p}_2$ estimate of $p_2(t)$

$\hat{p}$ $= \hat{p}_1 + \hat{p}_2$

s significance level testing H_0: $p_1(t) + p_2(t) \geq 1$ against H_1: $p_1(t) + p_2(t) < 1$.

FE Fisher's Exact Test.

χ^2 Chi-square test of independence of forecasts and outcomes.

S_n^2 Pesaran-Timmerman test.

ND Test not defined when $N_1 = 0$ or $N_2 = 0$.

* Null hypothesis rejected at 5% level.

References

Ash, J.C.K. and D.J. Smyth, (1973) *Forecasting the UK Economy*, Farnborough, Hants and Lexington, Mass: D.C. Heath.

Ash, J.C.K., D.J. Smyth and S.M. Heravi (1990) The Accuracy of OECD Forecasts of the International Economy: Demand, Output and Prices. *International Journal of Forecasting*, **6**, 379-392.

Ash, J.C.K., D.J. Smyth and S.M. Heravi (1998) Are OECD Forecasts Rational and Useful?: A Directional Analysis. *International Journal of Forecasting*, **14**, 381-391.

Batchelor, R. and P. Dua (1989) Household Versus Economist Forecasts of Inflation: A Reassessment, *Journal of Money, Credit and Banking*, **21**, 539-544.

Bryan, M. and W. Gavin, (1986) Models of Inflation Formation: A Comparison of Household and Economist Forecasts, *Journal of Money, Credit and Banking*, **18**, 539-544.

Canova, F. (1999) Does Detrending Matter for the Determination of the Reference Cycle and the Selection of Turning Points? *The Economic Journal*, **109**, 126-150.

Carlson, J. (1977) A Study of Price Forecasts, *Annals of Economic and Social Measurement*, **6**, 27-56.

Cogley, T. and J. Nason (1995) Effects of the HP Filter on Integrated Series, *Journal of Economic Dynamics and Control*, **19**, 253-278.

Cukierman, A. and Meltzer, A.H. (1982) What do Tests of Market Efficiency Show in the Prescence of the Permanent-Transitory Confusion, *mimeo Tel-Aviv and Carnegie-Mellon University, Pittsburgh, PA.*

Dow, J.C.R. (1965) *The Management of the British Economy 1945-60*, London: Cambridge University Press.

Figlewski, S. and P. Wachtel, (1981) The Formation of Inflationary Expectations, *Review of Economics and Statistics*, **63**, 1-10.

Fisher, R.A. (1941) *Statistical Methods for Research Workers*, London: Oliver and Boyd.

Gordon, R.J., (1997) *Macroeconomics,* Boston: Little, Brown and Co. USA.

Gramlich, E. (1983) Models of Inflation Expectations Formation, *Journal of Money, Credit and Banking*, **15**, 155-173.

Grant, A.P. and Thomas, L.B. (1999) Inflationary Expectations and Rationality Revisited, *Economic Letters*, **62**, 331-338.

Granger, C.W.J. and P. Newbold (1973) Some Comments on the Evaluation of Economic Forecasts, *Applied Economics*, **5**, 35-47.

Harvey, A.C. and A. Jaeger (1993) Detrending, Stylized Facts and the Business Cycle, *Journal of Applied Econometrics*, **8**, 231-247.

Henriksson, R.D. and R.C. Merton (1981) On Market Timing and Investment Performance 2: Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business*, **54**, 513-533.

Hodrick, R.J. and E.C. Prescott, (1997) Postwar US Business Cycles: An Empirical Investigation, *Journal of Money, Credit and Banking*, **29**, 1-16.

Holden, K., Peel, D.A. and J.L. Thompson (1985) *Expectations: Theory and Evidence*, London: Macmillan.

Holden, K. and D.A. Peel (1990) On Testing for Unbiasedness and Efficiency of Forecasts, *The Manchester School*, **58**, 120-127.

Johansen, S. (1988) Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, **12**, 231-254.

King, R. and S. Rebelo (1993) Low Filtering and the Business Cycle, *Journal of Economic Dynamics and Control*, **17**, 207-231.

Laster, D., P. Bennett and I.S. Geoum (1999) Rational Bias in Macroeconomic Forecasts, *Quarterly Journal of Economics*, **114**, 293-318.

Laxton, D., G. Meredith and D. Rose (1995) Asymmetric Effects of Economic Activity on Inflation, *IMF Staff Papers*, **42**, 344-374.

Martin, J. and S. Scarpetta (1999) The Levels and Cyclical Behaviour of Mark-Ups Across Countries and Market Structures, *OECD Economics Department Working Paper, No. 213.*

Mullineaux, D. (1978) On Testing for Rationality: Another Look at the Livingston Price Expectations Data, *Journal of Political Economy*, **86**, 329-336.

McNees, S.K. (1978) The 'Rationality' of Economic Forecasts, *American Economic Review*. **68**, 301-305.

Merton, R.C. (1981) On Market Timing and Investment Performance 1: An Equilibrium Theory of Value for Market Forecasts, *Journal of Business*, **54**, 363-406.

Mullineaux, D. (1978) On Testing for Rationality: Another Look at the Livingston Price Expectations Data, *Journal of Political Economy*, **86**, 329-336.

Muth, J.F. (1961) Rational Expectations and the Theory of Price Movements, *Econometrica*, **21**, 315-335.

Orr, A., M. Edey, and M. Kennedy, (1995) The Determination of Real Long Term Interest Rates: 17 Country Pooled-Time-Series Evidence. *OECD Economics Department Working Paper, No. 155.*

Paquet, A. (1992) Inflationary Expectations and the Theory of Price Movements, *Economic Letters*, **40**, 303-308.

Pesaran, M.H. and A.G. Timmerman (1992) A Simple Non-Parametric Test of Predictive Performance, *Journal of Business and Economic Statistics*, **10**, 461-465.

Rich, R. (1989) Testing Rationality of Inflation Forecasts from Survey Data: Another Look at the SRC Expected Price Change Data, *Review of Economics and Statistics*, **71**, 682-686.

St.-Amant, P. and S. van Norden, (1997) *Measurement of the Output Gap: A Discussion of Recent Research at the Bank of Canada*, Bank of Canada, Technical Report, No. 79.

Smant, D.J.C. (1998) Modelling Trends, Expectations and the Cyclical Behaviour of Prices, *Economic Modelling*, **15**, 151-161.

Smyth, D.J. (1992) Measurement Errors in Survey Forecasts of Expected Inflation and the Rationality of Inflation Expectations, *Journal of Macroeconomics*, **14**, 439-448.

Stekler, H.O. (1994) Are Economic Forecasts Valuable? *Journal of Forecasting*, **13**, 495-505.

Stekler, H.O. and R. Fildes, (1999) *The State of Macroeconomic Forecasting*, George Washington University, Center for Economic Research, Discussion Paper No. 99-

04, and the Management School, Lancaster University, Discussion Paper No. EC2/99.

Theil, H. (1966) *Applied Economic Forecasting*, Amsterdam: North-Holland.

Theil, H. (1961) *Economic Forecasts and Policy*, Amsterdam: North-Holland.

Zarnowitz, V. (1967) *An Appraisal of Short-Term Economic Forecasts*, New York: National Bureau of Economic Research, and Columbia University Press.