Agricultural Economics, 4 (1990) 365–380
Elsevier Science Publishers B.V., Amsterdam

Predictive performance of the World Bank's commodity price projections

Peter G. Warr
Department of Economics, Research School of Pacific Studies, Australian National University, GPO Box 4, Canberra, A.C.T. 2601 (Australia)

(Accepted 23 August 1989)

ABSTRACT

Warr, P.G., 1990. Predictive performance of the World Bank's commodity price projections. *Agric. Econ.*, 4: 365–380.

The World Bank's commodity price projections are widely used for various planning purposes. Two aspects of the Bank's projections of relative prices are studied in this paper. The first is whether the forecasts make efficient use of the information available at the time the forecast is made. The second is whether the forecasts predict future prices with greater accuracy than alternative forecasting methods. These matters are studied by comparing the World Bank's past price projections with the actual prices that were subsequently observed. The results show that, overall, the World Bank forecasts do not pass either test. First, the World Bank forecasts are informationally inefficient. Prediction error (projection minus actual price) tends to be positively correlated with the projections themselves. Although the direction of future price movements tends to be correctly predicted, the magnitude of these movements tends to be overpredicted. Second, the World Bank forecasts do not perform well even compared with the simplest of alternative forecasting methods – the prediction of no change.

1. INTRODUCTION

The prices of primary commodities are notoriously volatile, but there are circumstances in which forecasts must be attempted. The ex ante evaluations of public and private sector investment projects inherently involves forecasts of relative price movements, often many years into the future. Likewise, institutions' projections of borrowing countries' balance of payments, necessary for assessment of the future capacity of these countries to repay, are also often highly sensitive to assumptions about the future prices of a few impor-

0169-5150/90/$03.50

tant export commodities.[1] Sometimes, as in the case of petroleum, they are sensitive to the prices of particular import commodities as well. In countries dependent on exports of primary commodities, virtually any long-term planning activities inherently involve some assumptions about the future prices of these crucial commodities.

The forecasting methodologies in common use fall into two main categories. The first method can loosely be called 'naive' projections. This includes use of the relative prices observed in a base year, simple extrapolation of past price trends, simple autoregressive time series models, and more sophisticated univariate statistical procedures such as Box–Jenkins methods, spectral analysis, etc. All these approaches are 'naive' in that they involve projecting future values of a price series solely from its own past and present values. The second method of projection entails *economic modelling.* It rests on an attempt to describe the economic forces determining commodity prices and to use this information to predict prices.

The World Bank has invested considerable resources in research reflecting the economic modelling approach, and produces regular forecasts of future prices for a large number of commodities. Many national and international organizations – including the World Bank itself and most regional development banks and bilateral aid agencies – now rely upon these World Bank forecasts in their own ex ante appraisal of investment projects. In recent years, the World Bank's projections have become both influential and controversial, replacing naive projections in many organizations' planning activities.

World Bank projections for several commodities have been published since 1972 and since 1976 for more than 40 commodities. It is now possible to review the performance of the projections to date. The future is inherently uncertain, and *any* attempt to forecast volatile commodity prices will necessarily entail errors. But some forecasting methods are presumably superior to others. Moreover, when price forecasts are being used, it is helpful to have some knowledge of the degree of precision that can be expected from these forecasts. Other series of commodity price forecasts exist, but are either confined to very short-term forecasts of less than 2 years, or are available only on a commercial basis and inaccessible for this research, or have not been produced long enough for an assessment of their performance to be possible. The World Bank's forecasts are uniquely suitable for this statistical exercise and, because their use is so widespread, an evaluation of their performance is of practical importance.

Section 2 of this paper compares observed prices with the World Bank's earlier projections by means of statistical tests of the unbiasedness and effi-

[1]Since only relative prices are important for project evaluation, the problem is one of projecting the prices of these commodities relative to others; it is real rather than nominal prices that must be projected.

ciency across a wide range of commodities. Section 3 compares the accuracy of the projections with those obtained from some 'naive' projection methods.

2. TESTS OF BIAS AND EFFICIENCY

2.1 Data

With the kind assistance of staff of the Commodity Studies and Projections Department of the World Bank, time series data have been assembled for 18 commodities of interest. For each of these commodities, the data are of two kinds:

- actual prices, assembled on an annual basis, from 1950 to 1985 where available, or beginning in a later year in the case of some commodities; and
- projected prices, also on an annual basis, commencing on the first year for which World Bank projections are available.

All prices in this data set, actual and projected, are expressed in *real* terms. Actual prices have been deflated by the observed values of the World Bank's manufacturing unit value index (MUV) expressed in US$.[2] The World Bank projects commodity prices in both real and nominal terms. The real forecasts relate to the movement of individual commodity prices relative to the MUV. The nominal price projections, expressed in US$, may be viewed as having two components: the real forecasts, just mentioned; and forecasts of the US rate of inflation, as reflected in the US$ value of the MUV. Only the real price projections have been used in this study. Nominal price projections have been ignored because forecasts of U.S. rates of inflation, or of exchange rates between the US$ and other currencies, are of no interest for this study.

Projected prices are classified in two ways: the year in which the forecast was made (superscript) and the year to which it relates (subscript). The notation P_t^{t-r} is used to denote the projection made in year $t-r$ about the price that will hold in year t. Obviously, r indicates the number of years ahead that the projection is made. For notational simplicity identifiers indicating the particular commodity under discussion have been suppressed. In our data, r, the period of the forecast, ranges from one to six years. The notation A_t will subsequently be used to denote the actual price observed in year t.

Table 1 summarizes the data set for each of the 18 commodities. It shows the number of observations of actual and projected prices which the data set contains. The year in which the actual price series begins for a given commod-

[2]The Manufacturing Unit Value (MUV) index is a c.i.f. index of US$ prices of industrial countries' manufactured exports (SITC 5-8) to the developing countries. It is compiled by the World Bank to serve as an international price deflator relevant for developing countries dependent on exports of primary commodities. It is published in World Bank, *Price Prospects for Major Primary Commodities*, various issues.

TABLE 1

Data set: number of observations of actual and projected prices

	Commodity																	
	Petroleum	Copper	Tin	Coffee	Tea	Sugar	Rice	Wheat	Maize	Palm oil	Copra	Rubber	Logs	Sawn wood	Cotton	TSP	Urea	Phosphate rock
	Actual prices																	
	36	36	36	31	35	36	36	36	36	36	36	36	31	28	33	19	23	31
	Projections																	
1 year	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	5	9	11
2 years	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	4	8	10
3 years	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	2	6	9
4 years	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	0	4	6
5 years	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	0	0	4
6 years	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	0	2

Source: Commodity Studies and Projections Department, World Bank, Washington, DC.

ity can be found by subtracting the number of annual data points listed from 1986. For example, the rice series begins in 1950. It is apparent that the number of data points for which longer term predictions are available is too small to allow meaningful statistical tests of the quality of these predictions at the individual commodity level. Pooling of commodities will thus be necessary, and this will be explained in the later discussion.

2.2 *Statistical test*

The econometric literature on forecasting has generally entailed separate tests of 'bias' and 'efficiency'. Forecasting bias is tested by estimating an equation of the form:

$$A_t = a_0 + a_1 P_t + e_t \tag{1}$$

and testing the joint hypothesis that $(a_0, a_1) = (0, 1)$. Forecasting efficiency is tested by estimating an equation such as:

$$A_t - P_t = bA_{t-r} + f_t \tag{2}$$

The left side of (2) is the forecasting error and testing the hypothesis that $b = 0$ is equivalent to testing whether forecasting error is orthogonal to past values of the actual price.

The similarity of these two tests was pointed out by Ravallion (1985, p. 179). It is revealed by manipulating (1) to give:

$$A_t - P_t = a_0 + (a_1 - 1)P_t + e_t \tag{3}$$

and then comparing this with (2). We now see that testing whether $a_1 = 1$ (i.e. $a_1 - 1 = 0$), amounts to asking whether the forecasting error is orthogonal to the forecasts themselves. Both past values of the actual price (as in (2)) and the present value of the forecast (as in (3)) belong to the information sets available to the forecaster at the time the forecast is made. The tests for 'bias' and 'efficiency' are thus seen to be each special forms of the more general question of whether ex post forecasting error is correlated with information available at the time of the forecast is made. Whether any such test is called a test for 'bias' or 'efficiency' seems somewhat arbitrary. In subsequent discussion only tests of forecasting efficiency are referred to, meaning the joint testing for what the earlier literature has called 'bias' and 'efficiency'.

The two tests can be combined to give an equation of the form:

$$A_t = c_0 + c_1 A_{t-1} + c_2 P_t + c_3 P_{t-1} + c_4 t + u_t \tag{4}$$

The term in t is included here to allow for the possibility of time dependence of forecasting errors.

This type of regression equation has recently been shown to have some convenient properties. Hendry and Mizon (1978) show that if Sargan's Com-

TABLE 2

Results of analysis of efficiency: one-year projections

	Commodity							
	Petrol-eum	Copper	Tin	Coffee	Tea	Sugar	Rice	Wheat
d_0	2.419	9.629	4.371	4.401	7.440	6.097	3.794	2.446
	(2.214)	(5.036)	(0.266)	(2.707)	(3.849)	(1.488)	(2.381)	(1.758)
d_1	0.144	−0.244	0.589	0.290	−0.351	0.059	0.379	0.522
	(0.365)	(−1.019)	(0.290)	(1.059)	0.989	(0.100)	(1.587)	(2.151)
$1\text{-}d_1$	0.856	1.244	0.411	0.710	1.351	0.941	0.621	0.478
	(2.160)	(5.192)	(0.202)	(2.591)	(3.805)	(1.593)	(2.599)	(1.972)
d_2	0.048	−0.047	−0.115	−0.032	−0.001	−0.148	−0.030	0.001
	(1.145)	(−2.421)	(−1.424)	(−1.122)	(−0.004)	(−1.211)	(−1.011)	(0.067)
c_1	0.651	0.254	0.128	−0.096	0.184	0.403	0.013	0.163
	(2.843)	(0.872)	(0.427)	(−0.319)	(0.622)	(0.461)	(0.043)	(0.548)
SEE	0.173	0.118	0.757	0.317	0.241	0.454	0.172	0.117
R^2	0.760	0.585	0.213	0.190	0.071	0.593	0.715	0.595
D–W statistic	1.984	1.742	1.218	1.696	1.774	1.618	1.956	1.767
F-test[a] $(d_0, d_1) = (0, 1)$	2.452	18.077**	0.398	4.009*	7.445*	1.700	7.874*	7.652*
F-test[a] $(d_0, d_1, d_2) = (0, 1, 0)$	1.636	13.119**	1.026	2.673	5.077*	1.407	5.815*	5.710*

[a]F-test results: *denotes significant at 5% level; **significant at 1% level.

TABLE 3

Results of analysis of efficiency: two-year projections

	Petrol-eum	Copper	Tin	Coffee	Tea	Sugar	Rice	Wheat
d_0	2.646	10.602	−34.161	5.904	7.882	5.317	6.629	3.589
	(4.245)	(5.486)	(−1.191)	(3.155)	(3.736)	(1.741)	(2.050)	(2.188)
d_1	0.069	−0.366	4.700	0.074	−0.424	0.122	−0.064	0.296
	(0.296)	(−1.504)	(1.541)	(0.232)	(−1.083)	(0.277)	(−0.131)	(1.032)
$1-d_1$	0.931	1.366	−3.700	0.926	1.424	0.878	1.064	0.704
	(4.019)	(5.610)	(−1.213)	(2.899)	(3.640)	(2.006)	(2.175)	(2.454)
d_2	0.070	−0.056	−0.190	−0.069	−0.011	−0.098	−0.069	0.007
	(2.344)	(−3.592)	(−2.647)	(−2.615)	(−0.447)	(−0.935)	(−1.399)	(0.316)
SEE	0.120	0.092	0.655	0.207	0.232	0.415	0.198	0.098
R^2	0.882	0.752	0.485	0.590	0.218	0.592	0.556	0.316
D–W statistic	1.746	1.837	0.9773	2.184	2.278	1.975	1.627	2.432
Ftest[a] $(d_0, d_1) = (0, 1)$	9.146*	24.750**	3.074	8.833*	7.673*	4.429	5.689*	21.420**
F-test[a] $(d_0, d_1, d_2) = (0, 1, 0)$	8.806*	26.991**	2.790	6.025*	5.173*	3.635	6.565*	16.850**

[a]See Table 2.

Maize	Palm oil	Copra	Rubber	Logs	Sawn wood	Cotton	TSP	Urea	Phosphate rock
3.584	7.401	7.114	7.384	9.606	14.411	7.362	4.539	4.262	1.915
(2.092)	(5.580)	(4.509)	(3.485)	(5.617)	(.409)	(4.803)	(3.760)	(3.312)	(1.403
0.287	−0.121	−0.146	−1.508	−1.007	−1.628	−0.210	0.083	0.206	0.511
(0.895)	(−0.619)	(−0.600)	(−1.185)	(−2.821)	(−5.198)	(−0.753)	(0.391)	(0.937)	(1.662
0.713	1.121	1.146	1.508	2.007	2.628	1.2105	0.917	0.794	0.489
(2.226)	(5.714)	(4.698)	(3.519)	(5.623)	(8.390)	(4.332)	(4.308)	(3.616)	(1.590
−0.020	−0.028	−0.001	−0.020	−0.056	0.081	−0.064	−0.002	0.024	−0.018
(−0.067)	(−1.891)	(−0.099)	(−0.812)	(4.177)	(6.388)	(−5.093)	(−0.080)	(−0.974)	(−0.530
0.273	0.287	0.103	0.473	0.032	−0.544	−0.506	0.462	0.068	−0.008
(0.943)	(0.109)	(0.344)	(1.782)	(0.106)	(−2.098)	(−1.760)	(1.646)	(−0.228)	(−0.026
0.137	0.158	0.293	0.167	0.113	0.103	0.0880	0.136	0.192	0.246
0.440	0.287	0.049	0.272	0.693	0.737	0.821	0.190	0.423	0.520
1.567	2.101	1.831	1.390	1.890	1.722	1.836	1.594	1.782	1.605
4.424	17.923**	11.684**	6.230*	15.807**	35.457**	152.555**	16.606**	11.484**	2.415
4.542*	13.636**	7.899*	4.392*	12.461**	24.769**	384.27**	11.391**	9.346**	1.965

Maize	Palm oil	Copra	Rubber	Logs	Sawn wood	Cotton	TSP	Urea	Phosphate Rock
8.138	11.466	9.547	5.813	6.319	11.673	11.481	5.391	8.106	4.069
(2.898)	(12.412)	(4.942)	(1.093)	(2.206)	(5.627)	(4.998)	(3.270)	(8.220)	(4.604)
−0.579	−0.736	−0.501	−0.155	−0.289	−1.088	−0.975	−0.058	−0.483	−0.042
(−1.107)	(−5.251)	(−1.672)	(−0.142)	(−0.489)	(−2.897)	(−2.316)	(−0.187)	(−2.835)	(−0.204)
1.579	1.736	1.501	1.155	1.289	2.088	1.975	1.058	1.483	1.042
(3.019)	(12.386)	(5.005)	(1.056)	(2.180)	(5.561)	(4.693)	(3.417)	(8.705)	(5.118)
−0.055	−0.043	−0.038	−0.036	0.026	0.042	−0.086	−0.017	−0.056	−0.030
(−1.706)	(−8.312)	(−1.679)	(2.713)	(1.575)	(2.951)	(−6.709)	(−0.879)	(−3.521)	(−1.352)
0.132	0.087	0.217	0.129	0.119	0.134	0.064	0.121	0.119	0.111
0.230	0.808	0.517	0.559	0.516	0.583	0.889	0.443	0.575	0.664
2.334	1.957	2.551	1.308	2.2062	1.243	2.338	1.330	1.751	1.880
8.059*	77.241**	12.868**	3.707	2.941	17.681**	322.955**	7.892*	62.107**	21.850**
7.881*	77.895**	8.926**	2.818	4.078	15.418**	898.263**	8.963*	50.906**	17.912**

mon Factor (COMFAC) restriction holds, namely that $c_1c_2+c_3=0$, then (4) can be reduced to the much simpler equation[3]:

$$A_t = d_0 + d_1 P_t + d_2 t + v_t \tag{5}$$

where

$$v_t = c_1 v_{t-1} + u_t \quad |c_1| < 1 \tag{6}$$

The correspondence between the parameters of these models is now:

$$d_0 = (c_0 - c_1 d_2)/(1 - c_1) \tag{7}$$

$$d_1 = c_2 \tag{8}$$

and

$$d_2 = c_4/(1 - c_1) \tag{9}$$

Provided the COMFAC restriction applies and the autoregressive error structure is appropriately handled, the model based on (5) and (6) provides a more powerful test for forecasting efficiency than does (4). Equations (5) and (6) entail fewer parameters to be estimated, and considering the limited number of degrees of freedom the data set makes available for estimation in the present study, this is an important advantage. This is the procedure adopted here.[4]

2.3 Results

Tables 2 to 4 present the results of estimating equations (5) and (6) using a cross-sectionally correlated and time-wise autoregressive package. The results can be summarized as follows.

One-year forecasts. Tests of the COMFAC restrictions were accepted for all commodities. Estimates of d_0 were positive for all commodities and significantly positive at the 5% level for twelve of the 18 commodities and for eight commodities at the 1% level. Every estimate of d_1 is smaller than unity. Testing the null hypothesis that $d_1=1$ (i.e. that $1-d_1=0$), the hypothesis is rejected at the 5% level for twelve of the 18 commodities (the same twelve as above) and at the 1% level for nine commodities (including the above eight). The joint restriction that $(d_0, d_1)=(0, 1)$ is rejected by an F-test at the 5% level for twelve of the 18 commodities, etc., in precisely parallel fashion.

The commodities for which the hypothesis is rejected are: copper, tea, rice, wheat, palm oil, copra, rubber, logs, sawn wood, cotton, TSP, and urea. Those

[3]For expositions of this point, see Hendry and Mizon (1978, pp. 550–555).
[4]See also Ravallion (1985, pp. 178–180).

TABLE 4

Results of analysis of efficiency: three-year projections

	Commodity group						
	Metals	Beverages	Grains	Oil	Wood	Petroleum and Rubber	Fertilizer
d_0	−0.139	0.723	−0.475	−0.186	0.019	−0.175	−1.217
	(−0.830)	(3.352)	(−6.064)	(−0.686)	(0.147)	(−0.680)	(−4.121)
d_1	0.896	0.325	0.883	0.381	1.013	−0.098	0.136
	(1.5537)	(0.921)	(3.279)	(1.447)	(2.942)	(−0.233)	(0.503)
$1-d_1$	0.104	0.675	0.117	0.619	−0.013	1.098	0.864
	(0.180)	(1.913)	(0.435)	(2.349)	(−0.037)	(2.611)	(3.196)
d_2	0.002	−0.105	0.026	−0.073	0.024	0.075	0.042
	(0.049)	(−2.336)	(1.977)	(−1.960)	(0.728)	(1.396)	(0.811)
Degrees of freedom	11	11	18	11	9	11	15
SEE	0.847	0.845	0.780	0.783	0.808	0.852	0.762
R_2	0.2325	0.4751	0.5107	0.4126	0.5995	0.220	0.079
D–W statistic	1.2504	1.1293	1.4986	1.8203	1.958	1.378	1.114
F-test[a] $(d_0, d_1) = (0, 1)$	0.410	5.918*	33.833**	4.629*	0.013	3.573*	8.524*
F-test[a] $(d_0, d_1, d_2) = (0, 1, 0)$	0.804	4.208*	67.601**	4.464*	1.478	2.584*	6.087*

[a]See Table 2.

for which the hypothesis is not rejected are: petroleum, tin, coffee, sugar, maize and phosphate rock.

These results imply that forecasting error, defined as projected minus actual, is positively correlated with projected prices. This is seen by rearranging (5) to give:

$$P_t - A_t = -d_0 + (1-d_1)P_t - d_2 t - v_t \tag{10}$$

Alternatively, writing the lagged value of (5) and subtracting:

$$A_t - A_{t-1} = d_1(P_t - P_{t-1}) + v_t - v_{t-1} \tag{11}$$

When $0 < d_1 < 1$, the expected value of projected price changes exceeds actual price changes; there is a tendency for projected price changes to overestimate the magnitude of actual price changes.

Two-year forecasts. Results here are similar to those above, except that the joint hypothesis is rejected for 14 of the 18 commodities, the additional commodities being petroleum, coffee, maize and phosphate rock, whereas the hypothesis is no longer rejected for rubber and logs.

Three and four-year forecasts. To study three-year forecasts, the number of observations was such that pooling of commodities into groups was required. The following seven pooled groups were constructed:

Pooled group	Commodities
Metals	Copper, tin
Beverages	Coffee, tea
Grains	Rice, wheat, maize
Vegetable oils	Palm oil, copra
Wood	Logs, sawn wood
Petroleum and rubber	Petroleum, rubber
Fertilizer	TSP, urea, rock phosphate

This list includes 16 of the 18 commodities. Sugar and cotton did not seem to belong to any natural pooled group and were dropped from the sample. To form each composite commodity data set, all prices were normalized by dividing by the 1974 observed price of that commodity and then treated as independent observations.

The results for three-year forecasts are summarized in Table 4. The null hypothesis that $(d_0, d_1) = (0, 1)$ is rejected for five of the seven groups, the exceptions being the two groups metal and wood. This analysis was repeated for four-year forecasts. Since the results are very similar, they need not be presented separately, except to note that the above null hypothesis fails to be rejected for the petroleum and rubber group.

Six-year forecasts. A composite pool of all 18 commodities was formed to study the efficiency of the longest term forecasts available. The results (with 29 degrees of freedom) were (with t-statistics shown in parentheses):

$d_0 =$	0.384	$d_1 =$	0.916	$1-d_1 =$	0.084
	(16.717)		(30.123)		(2.777)

The hypotheses that $d_0=0$ and $d_1=1$ and the joint hypothesis that $(d_0, d_1) = (0, 1)$ were each rejected by an F-test at the 5% level of significance.

3. COMPARSION WITH NAIVE PREDICTORS

3.1 Two naive predictors

Statistical tests of 'bias', 'efficiency', etc. provide one approach to studying the performance of commodity price forecasts. Another, perhaps more interesting exercise, is to compare the performance of these forecasts with alternative forecasts. One such alternative forecasting method is obviously to use past and present price information available at the time the World Bank forecast is made. This kind of forecast is 'naive' in that it is based only on the past and present values of the variable being projected. Unlike the World Bank's forecasts there is no structural model of supply and demand being estimated at all. It seems reasonable to expect the World Bank's forecasts to out-perform any such naive forecasts.

Two kinds of naive forecasts are considered:

(a) The actual price observed in the year in which the forecast is made:

$$P_t^{t-r} = A_{t-r} \quad r = 1, 2, \ldots, 6 \tag{12}$$

(b) Simple autoregressive time series forecasts based on estimated regression equations of the form[5]:

$$A_t = a_0 + a_1 A_{t-1} + a_2 A_{t-2} + a_3 A_{t-3} + a_4 t + w_t \tag{13}$$

Type (a) is clearly the most 'naive' of all possible forecasts – the assumption that there will be no change. We shall subsequently call this the 'base period' forecast. Type (b) entails fitting a simple autoregressive time series model to actual price data available at the time the forecast is made. This estimated equation can then be used to generate a sequence of forecasts. First a one-year forecast is made, then a two-year forecast is produced by feeding the one-year forecast back into the estimated equation, etc. This is subsequently referred to as the 'autoregressive' forecast. For the purpose of this study naive forecasts of both types were generated so that their predictive performance could be compared with that of the World Bank forecasts.

3.2 *Statistical test*

Predicted and actual prices are compared by computing the mean squared error, defined as (Madalla, 1977, p. 344):

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^{n} \left(\frac{P_t^{t-r} - A_t}{A_{t-r}} \right)^2 \tag{14}$$

where r is the period of the forecast, as before. An F-test, taking account of the appropriate degrees of freedom, is then used to test the significance of the observed differences. The results are presented in Tables 5 to 7.

In all three tables the results shown represent *ratios* of mean squared error (MSE) obtained from either the 'base period' or 'autoregressive' forecasts relative to the MSE obtained from the corresponding World Bank projection. That is, for each World Bank price projection data point contained in the data

---

[5]This particular autoregressive form was selected by regression analysis using actual price data for each commodity for the period 1950–1972. Five related autoregressive functional forms were tried and were compared according to the resulting R^2 statistic adjusted for degrees of freedom. The superiority of this form over the four others tested was surprisingly robust across commodities.

set two rival forecasts were generated – a 'base period' and an 'autoregressive' forecast. The differences between these forecasts and actual prices are then aggregated across time periods to give mean squared errors as indicated in (14) above. A number greater than unity in the tables means that the rival forecasting method ('base period' or 'autoregressive') generates a higher MSE than the World Bank's forecasts. Whether this ratio is significantly greater than unity at the 5% or 1% significance level, as revealed by an F-test, is indicated by the presence of '+' or '+ +' superscripts, respectively. Whether a number shown is significantly smaller than unity is indicated correspondingly by '–' or '– –' superscripts.

3.3 *Results*

Table 5 depicts individual commodity results for one-year and two-year forecasts. The available degrees of freedom are too small to make this exercise

TABLE 5

Forecasting errors of naive models vs. World Bank projections: individual commodities (ratios of mean squared error)

	1-year forecasts		2-year forecasts	
	Base period/ World Bank	Autoregressive/ World Bank	Base period/ World Bank	Autoregressive/ World Bank
Petroleum	0.695	8.203^{++}	1.000	22.111^{++}
Copper	0.796	1.810	1.041	3.163^{+}
Tin	0.124^{--}	0.150^{--}	0.148^{--}	0.167^{--}
Coffee	0.835	1.626	1.471	1.576
Tea	0.796	1.133	0.825	1.301
Sugar	0.931	1.489	2.112	1.726
Rice	1.149	3.089^{+}	1.338	4.811^{++}
Wheat	0.656	0.781	1.140	1.372
Maize	0.532	0.745	0.527	0.986
Palm Oil	0.514	0.450	0.549	0.2332^{-}
Copra	0.840	0.829	0.993	0.469
Rubber	0.000	2.751	1.955	8.955^{++}
Logs	0.981	0.750	2.500	1.500
Sawn wood	0.547	1.113	1.580	1.706
Cotton	0.532	0.583	1.014	1.645
TSP	1.017	0.233	2.352	0.519
Urea	0.784	1.946	0.960	0.950
Phosphate rock	0.776	3.724^{+}	1.375	3.875^{+}

F-test results: $^{+}$, significantly greater than unity at 5% level.
$^{++}$, significantly greater than unity of 1% level.
$^{-}$, significantly smaller than unity 5% level.
$^{--}$, significantly smaller than unity at 1% level.

useful for longer forecasts. The Table shows that in almost all cases base period forecasts generate a lower MSE than the autoregressive forecasts, and in the case of one-year forecasts, they result in a lower MSE than the World Bank's projections for 15 of the 18 commodities. Nevertheless, this difference between the MSE resulting from base period and World Bank forecasts is statistically significant for only one of these commodities. The results are more mixed for two-year forecasts. The results do not indicate a single commodity for which the World Bank's price projections outperform base period forecasts at the 5% level of significance, either for one-year or two-year forecasts.

Table 6 shows similar results for three-year and four-year forecasts using the pooled commodity groups described in Section 2.3 above. These results show that for three-year forecasts there is one commodity group, metals, for which both base period and autoregressive forecasts provide significantly smaller forecasting errors than the World Bank's forecasts, and three groups, wood, petroleum and fertilizer, for which the reverse is true. For four-year forecasts the results are similar, except that the World Bank's forecasts also outperform both kinds of naive forecasts for the beverages group.

Table 7 displays pooled results for all commodities for three-year, four-year, five-year and six-year forecasts. To show the degree to which the extreme volatility of petroleum prices is affecting the results, the lower half of the Table shows the results when petroleum is deleted from the pooled results. In the case of six-year forecasts the World Bank's projections outperform both kinds of naive forecasts whether petroleum is included in the results or not. For shorter periods the results are less impressive. When petroleum is deleted from the pooled data set, it is only in the case of the six-year forecasts that the World Bank's projections are significantly superior to either type of naive forecast.

TABLE 6

Forecasting errors of naive models vs. World Bank projections: pooled commodity groups (ratios of mean squared error)

	3-year forecasts		4-year forecasts	
	Base period/ World Bank	Autoregressive/ World Bank	Base period/ World Bank	Autoregressive/ World Bank
Metals	0.216^{--}	0.421^{-}	0.111^{--}	0.105^{--}
Beverages	1.391	1.736	3.060^{+}	4.150^{+}
Grains	0.744	1.637	0.889	2.163
Vegetable oils	1.250	1.018	2.115	1.662
Wood	2.375^{+}	2.141	3.241^{+}	2.165
Petroleum	2.917^{+}	79.390^{++}	1.082	78.354^{++}
Fertilizer	5.654^{++}	4.014^{++}	5.477^{++}	4.148

See Table 5.

TABLE 7

Forecasting errors of naive models vs. World Bank projection: pooled results (ratios of mean squared error)

	3-year forecasts	4-year forecasts	5-year forecasts	6-year forecasts
Including petroleum				
Base period/World Bank	1.324^{+}	0.904	0.783	1.701^{++}
Autoregressive/ World Bank	9.302^{++}	15.677^{++}	20.158^{++}	3.779^{++}
Excluding petroleum				
Base period/World Bank	1.141	0.880	0.713^{-}	2.325^{++}
Autoregressive/ World Bank	1.406	1.279	1.007	3.552^{++}

See Table 5.

4. CONCLUSIONS

This paper has reported analyses of the efficiency with which the World Bank's price forecasts utilize available information and of the magnitude of ex post prediction errors. The results are generally unfavourable with respect to the performance of these forecasts. Our statistical tests of 'bias and efficiency' (Section 2) suggest that the World Bank's projections do not make efficient use of the information available at the time the forecasts are made. Prediction error (projection minus actual price) tends to be positively correlated with the projections themselves and the hypothesis that the World Bank's projections are statistically efficient ($d_1 = 1$) was rejected in most cases. The results further imply that although the direction of price movements tended to be correctly predicted ($d_1 > 0$), the magnitude of these movements was overpredicted ($d_1 < 1$). The latter result is consistent with the observation that speculative market behaviour is not adequately captured in the World Bank's projection models. The effect of such speculation in diminishing the magnitude of price movements ex post is thus partly overlooked.

Comparison of the magnitude of the ex post prediction errors resulting from the World Bank's projections with those resulting from 'naive' forecasting methods (Section 3) suggests that for short-term forecasts – of 1 and 2 years – the World Bank's projections fail to out-perform naive methods. For longer-term forecasts, the World Bank's projections perform relatively better. When compared with naive forecasts, such as the prediction of no price change ('base period' forecasts), the performance of the World Bank's projections is best in the case of the longest period of projection, 6 years, and worst in the case of the shortest periods, one and two years. This suggests that the operation of market forces is such that the current market price of a commodity tends to incorporate the information presently available about the likely values of that price one or two years into the future, at least as effectively as do the World Bank's econometric models. It is in the case of longer term forecasts that the supply and demand models used in the World Bank's projections become relatively more fruitful.

The technical quality of the World Bank's work on price forecasting is impressive. The underlying research has contributed to our understanding of the markets for many commodities. Nevertheless, the task of economic modelling for prediction purposes is expensive and the results of this study's analysis of the output of that effort to date are not encouraging. It must be hoped that the predictive performance of the World Bank's commodity price projections will improve. Otherwise, if the World Bank's rigorous principles of benefit–cost analysis were applied to its own work on price forecasting, based on its past predictive performance the activity would surely be discontinued.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the outstanding research assistance of Corrine Condie and Lorraine Dearden. Generous cooperation was received from the staff of the Commodity Studies and Projections Department of the World Bank, especially from Ronald Duncan, Donald Mitchell and Suan Tan. Helpful comments were also received from Graeme Dorrance, Ronald Duncan, Ulrich Hiemenz, and the Editor-in-Chief and a referee of this journal. Discussions with Norman Gemmell, D.T. Nguyen, Martin Ravallion and R.M. Sundrum were also helpful. The author is responsible for the views presented and any errors.

REFERENCES

Hendry, D.F. and Mizon, G.E., 1978. Serial correlations as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the Bank of England. Econ. J., 88: 549–563.

Maddala, G.S., 1977. Econometrics. McGraw-Hill, New York.

Ravallion, M., 1985. The information efficiency of traders' price expectations in a Bangladesh rice market. Oxford Bull. Econ. Stat., 47: 171–184.