Predicting the Signs of Forecast Errors*

by

Nazaria Solferino and Robert Waldmann**

University of Rome "Tor Vergata"

Abstract: The signs of forecast errors can be predicted using the difference between individuals' forecasts and the average of earlier forecasts of the same variable. It is possible to improve forecasts without worsening any. It is difficult to reconcile this result with the rational expectations hypothesis, because the average of earlier forecasts is in the information set of the forecasters

Keywords: Rational Expectations, Panel, Loss Function, Forecast, Interest Rate..
JEL Classification: G14, E47

* We would like to thank Steve Peterson for encouragement, useful e-mail discussions and a computer readable version of the Bluechip Forecast data set
** Corresponding author: Robert J. Waldmann, Professor of Economics, University of Rome "Tor Vergata", Department of Economics, Via Columbia 2, 00133 Rome Italy
Phone: +390672595740
E-mail robert.waldmann@gmail.com

1.Introduction.

In this paper we attempt to determine whether apparent rejection of the rational expectations hypothesis by analysis of survey data is the result of auxiliary assumptions about loss functions. In empirical work with survey data, it is generally assumed that survey participants have quadratic loss functions or at least that they report the mean of their subjective probability distribution (see for instance the recent survey by Elliot and Timmermann [2008] which summarizes empirical studies of rationality for inflation and output forecasts). Without any assumptions about agents' aims it is possible to reconcile any behavior with full rationality, but a quadratic loss function seems restrictive. Several authors have considered the theoretical implications of asymmetric loss functions including Granger [1968,1999], Varian [1974], Zellner[1986], Weiss [1996], Christofersen and Deibold [1997], Batchelor and Peel [1998], Granger and Pesaran [2000], Pesaran and Skouros [2002]. They all note that the evidence against the joint hypothesis rationality and quadratic losses is perfectly consistent with rationality and asymmetric losses if disturbances are heteroskedastic. They also note that it is difficult to derive implications from relatively weak assumptions about loss functions without imposing strong assumptions about the data generating process. Recently, Patton and Timmmermann [2006] have derived testable implications of the joint hypotheses of rational expectations, a homogenous loss function and a data generating process such that the conditional distribution of the variable being forecast is a function of the conditional mean and the conditional variance.

In empirical work, quadratic loss functions have been generalized to finite dimensional parametric classes of loss functions whose parameters are estimated by GMM in an initial stage. Using this approach Elliott, Komunjer and Timmermann [2005,2008]] have found weaker evidence against rationality in survey data on forecasts of nominal output growth that was obtained assuming a quadratic loss function.

We consider an implication of rational expectations which requires almost no assumptions about loss functions except the assumption that losses decrease as forecasts move closer to realizations.

In particular it is not necessary for the loss function to be homogeneous or to belong to a known parametric family of loss functions. Nor it is necessary to assume that different forecasters have the same loss function nor that this function is time invariant for any one forecaster. Thus it is not necessary to assume that losses are a function of forecast errors alone. Finally we do not need to make strong assumptions about the data generating process. It is not required that disturbances to the variable being forecast are homoskedastic or even belong to a finite dimensional paremetric class of distributions with time varying parameters.

The only additional assumption that must be made about the data generating process and the loss function (as) is that they do not create peso problems, that is rare events which may not be observed in a finite sample which are nonetheless very important to forecasters in expected value.

Our results imply that it is possible to recommend revisions to forecasts which lower losses assuming only that losses are reduced if a forecast is changed slightly in the direction of the outcome. The striking pattern noted in this paper is that whenever forecasts of quarterly averages of annualized yields on 30 year treasury bonds are very far from the average of older forecasts of the same variable, they are too far from this lagged average. That is forecasts far higher (lower) than the lagged average forecast are always higher (lower) than the outcome.

In this paper we define "far" from the lagged average forecast as far compared to the root mean forecast error of the lagged average forecast. All 1,100 forecasts which are at least 2.6 times this root mean squared error from the lagged average forecast are too far from the lagged average forecast. This result is striking because forecasts are published each month so the lagged average forecast is known by all forecasters.

Even the assumption that forecasters' losses are reduced when forecast errors are reduced is somewhat restrictive and some loss function would reconcile the data with the rational expectations hypothesis. Furthermore it is impossible to rule out peso problems – survey participants' predictions may reflect rational expectations of the probability of rare extreme events which are not observed in the sample. Nonetheless, the results presented here appear to be striking evidence against the rational expectations hypothesis. They are also very difficult to reconcile with the view that forecasters or economic agents in general exhibit herd behavior, that is understate the difference between their opinion and the conventional wisdom. If anything the data support the view that agents overstate this difference if the lagged average is taken to represent the conventional wisdom.

The paper is organized as follows. In section II we briefly describeour dataset, while the section II describes our methodology used to build a measure of a too "far " distance between analysts' forecasts and average lagged values of the interest rate on 30 year treasury bonds for the same period, that can be usefully used to predict the signs of forecasts error and so to improve them. In section IV we discuss our results which show that it is possible to improve 1,100 forecasts without worsening any. Concluding remarks are contained int he section V.

II. Yield forecasts data set.

This paper analyses the large data set collected in S.Peterson (2001) from the Blue Chip Financial Survey. It concerns interest rates over eight different maturities (3 month, and six month US Treasury bill yields and one, two, five-, seven-, ten- and thirty-year US Treasury bond yields) over the period 1987-1996. Therefore, this sample is very rich in information and lets us determine whether our method works for a period which includes a recession (1990-1991). The master database is an unbalanced panel including more than 28000 forecasts of professional economists from banks, financial firms, prominent corporations and academia recorded in the Blue Chip Financial Survey. A new time series is automatically created for any change in affiliation, name, or composition of a forecaster or group of forecasters.

Each month, participants in the survey submit forecasts of the average quarterly yield of each of the eight maturities for each of the next four quarters. Therefore analysts' forecasts consist of a multistep-ahead set including three periodic revisions for one-quarter-through four-quarter-head horizons. To measure forecasts errors we compute the difference between the n-step-ahead quarterly forecast and the quarterly realized yield obtained as the monthly average from the Federal

Reserve Board of Governors release on the constant maturity yield.

We focus on forecasts of interest rate yields on 30 year treasury bonds. Our approach is not invariably successful for interest rates at shorter maturities. We briefly discuss our results for the shorter maturities. It is important to consider the fact that the record of successful predictions of the signs of forecast errors reported below concerns only one of the bonds which we analyzed.

Waldmann (1995) had a perfect record forecasting the signs of forecast errors in the quarterly average of the annualized yield on 91 day treasury bills using a similar method. Waldmann analyzed a much smaller data set with only 506 forecasts in total, thus his perfect record is much less impressive than the not quite perfect record reported here.

III. Methodology.

Define r_t as the average interest rate yield on 30 year treasury bonds in the secondary market in quarter t. Define f_{itj} as forecast of the ith forecaster of the average interest rate of maturity m in the t'th quarter in the sample based on information available j months before the quarter ended. Define I_{itj} as an indicator variable which indicates non-missing f_{itj}. Define $f_{t,j}$ as the average of f_{itj} across forecasters. Note that our data include "forecasts" of the current quarter e.g. f_{it1}-- the forecast of the annualized interest rate on 30 year treasury bonds made by the i'th forecaster at the beginning of the third month of the same quarter.

We use an extremely simple technique to test the claim that whenever forecasts are much higher (lower) than the average of lagged forecasts they are higher (lower) than the outcome. To be precise first define σ_{tk}^2 as

$$\sigma_{t,k}^2 = \frac{\sum_{s=1}^{t-1-B}(r_s - f_{s,k})^2}{t-1-B} \qquad 2$$

Where B is the largest integer less than or equal to k/3. Note that, so long as k is greater than j, $f_{s,k}$ and $\sigma_{t,k}^2$ are calculated with information available when forecast f_{itj} is made.

IV. Results

Our first result is that for k equal to j+1, j+2, j+3 or j+4

if $f_{imtj}-f_{m,t,k} > 2.6\sigma_{m,t,k}$ then $f_{imtj} > r_{mt}+0.05\%$
and
if $f_{imtj}- f_{m,t,k} <- 2.6\ \sigma_{tk}$ then $f_{imtj} < r_{mt}-0.05\%$.

2.6 was chosen ex post as the smallest value such that the result holds. Thus we snooped enough to estimate one parameter. (fn in contrast with his much smaller data set Waldmann chose the critical level a priori). To restate the result, if the difference between a forecast and the average of forecasts lagged one, two, three or four months is greater than 2.6 times the mean squared error of that lagged average calculated with past data, then the forecast can be improved by moving it 10 basis points closer to that lagged average.

Notice that this result implies that some forecasters have not chosen a forecasting rule which minimizes any function of forecast errors which is reduced as the forecast errors are reduced. For any such function, the modified rule in which forecasters calculate their forecasts then reduce forecasts by the smallest allowed change (0.1 percent). This is true even if the loss function is asymmetric and disturbances are heteroskedastic. This is even true if the loss function is time varying, and if each forecaster has a different loss function.

To restate the result, if the difference between a forecast and the average of forecasts lagged one, two, three or four months is greater than 2.6 times the mean squared error of that lagged average calculated with past data, then the forecast can be improved by moving it 10 basis points closer to that lagged average.

In our sample, all 1,082 forecasts which are above at least one of the four intervals calculated for averages lagged one two three and four months are above the outcome, and all 18 forecasts which are below one of the four intervals are below the outcome. This means that the technique improves 1,100 forecasts without worsening any. Notably relatively few (80) forecasts are clearly wrong based on the difference with the average forecast lagged just one month. In contrast large numbers of forecasts are detectably wrong based on information lagged 2, 3 and 4 months. Average forecasts approach the truth relatively slowly, so the intervals are not all that much larger when older lagged averages are used.

In contrast, forecasters act as if they receive valuable information every month so many forecasts are much further from the 2 month lagged average than from the one month lagged average. Using information available to the forecasters it is possible to determine that many of these forecasts are much too far from the 2 month lagged average. Results are shown in Table I.

Table I Improvements of Predictions of forecasts of the 30 year interest rate

Average lagged	Predictably too high	Predictably too low
One month	80	0
Two months	359	2
Three months	619	9
Four months	338	8
Total	1082	18

Notes: The total is not equal to the sum, because the signs of some forecast errors can be predicted using different lagged averages.

The gaps between lagged average forecasts and the forecasts we correctly predict are too far from the lagged average range from 339 to 347 basis points depending on the lag. This is important because the one key issue is whether our result could be due to a peso problem where a possible extreme event which does not happen to occur in the sample period has an important effect on expected values. Certainly September 2008 has included extreme events related to interest rates, yet over the period from 9/2/08 to 9/29/08 the yield on 30 year treasury bonds varied only from a high of 4.41 to a low of 4.13 (U.S Treasury 2008). The large event would have to be much more dramatic than the events of September 2008 to rationalize the forecasts in our sample

Table II classifies improved forecasts such by the time from the forecast to the realization, that is the number of months from the forecast to the end of the quarter whose average 30 year Treasury bond yield was forecast. Unsurprisingly it becomes more difficult to predict the signs of forecast errors as this interval becomes longer. Strikingly, it is possible to predict the sign of 156 forecasts more than a year before the realization.

Table II: Forecasts Improved by Full Months from The Forecast to the End of the Quarter To Which the Forecast Refers

Months	Too High	Too Low	Improved
1	108	0	108
2	38	0	38
3	43	0	43
4	112	0	112
5	36	0	36
6	72	0	72
7	114	0	114
8	45	0	45
9	105	1	107
10	95	0	95
11	68	1	69
12	90	5	95
13	90	0	90
14	51	0	51
15	15	10	25
Total	1082	18	1100

Table III classifies improved forecasts by the quarter to which the improved forecast refers. At least one forecast with an error of detectable sign refers to each quarter from the first quarter of 1987 until the first quarter of 1997, that is from the first quarter in the data set to the 5th from the last. Relatively few forecasts refer to outcomes from the 2nd quarter of 1997 through the 1st quarter of 1998 and none of these are improved. Our ability to improve forecasts is not based on our procedure happening to flag forecasts which refer to a few quarters in which something unusual happened to interest rates.

Table III. Forecasts Improved by Quarter to Which Forecast Refers.

Quarter of Outcome	Forecasts Improved	Forecasts Worsened	Total Forecasts
1987:1	1	0	93
1987:2	7	0	239
1987:3	37	0	383
1987:4	40	0	518
1988:1	33	0	655
1988:2	20	0	703
1988:3	20	0	704
1988:4	30	0	708
1989:1	32	0	718
1989:2	11	0	721
1989:3	9	0	721
1989:4	12	0	721
1990:1	22	0	712
1990:2	12	0	706
1990:3	10	0	710
1990:4	16	0	712
1991:1	14	0	697
1991:2	8	0	699
1991:3	12	0	711
1991:4	12	0	719
1992:1	14	0	723
1992:2	10	0	723
1992:3	11	0	722
1992:4	5	0	723
1993:1	14	0	728

Table III continued

Quarter of Outcome	Forecasts Improved	Forecasts Worsened	Total Forecasts
1993:2	13	0	720
1993:3	13	0	725
1993:4	10	0	732
1994:1	31	0	729
1994:2	27	0	724
1994:3	24	0	727
1994:4	19	0	727
1995:1	55	0	722
1995:2	32	0	720
1995:3	32	0	729
1995:4	40	0	731
1996:1	124	0	723
1996:2	71	0	722
1996:3	61	0	730
1996:4	78	0	730
1997:1	58	0	618
1997:2	0	0	474
1997:3	0	0	336
1997:4	0	0	192
1998:1	0	0	46
Total	1100	0	28726

Table IV classifies improved forecasts by the quarter in which the forecasts were made. Since the series of forecasts are monthly we have summed over the three months in a quarter to make the table smaller. The majority of the improved forecasts were made in the first quarter of the year. This would correspond to a large number of forecasts made in the first quarter which are much higher than the lagged average. It appears that early in each year, many forecasters predict that future 30 year treasury bond yields will be much higher than generally predicted in the past. In our sample, they are wrong 100% of the time.

Table IV Forecasts of 30 year T-Bond rates improved by Quarter in which the Forecast was made

Quarter of Forecast	Number Improved	Number Worsened	Number of Forecasts
1987:1	120	0	703
1987:2	3	0	730
1987:3	0	0	719
1987:4	7	0	655
1988:1	83	0	699
1988:2	7	0	711
1988:3	0	0	733
1988:4	2	0	741
1989:1	42	0	724
1989:2	4	0	693
1989:3	0	0	711
1989:4	0	0	725
1990:1	55	0	691
1990:2	2	0	711
1990:3	3	0	686
1990:4	0	0	715
1991:1	38	0	710
1991:2	2	0	719
1991:3	2	0	740
1991:4	0	0	720
1992:1	35	0	724
1992:2	0	0	705
1992:3	2	0	716
1992:4	0	0	730
1993:1	54	0	730
1993:2	0	0	739
1993:3	1	0	729
1993:4	0	0	734
1994:1	115	0	723
1994:2	1	0	712
1994:3	0	0	720
1994:4	0	0	740
1995:1	182	0	717
1995:2	1	0	724
1995:3	0	0	725
1995:4	0	0	744
1996:1	338	0	723
1996:2	1	0	707
1996:3	0	0	728
1996:4	0	0	720

Table V shows that the vast majority of forecasting teams make at least one forecast which can be improved using lagged data. Only 17 forecasting teams make no such predictable errors and each of them makes relatively few forecasts. Only 424 forecasts were made by teams such that we can improve none of the team's forecasts.

Table V Forecasting Teams Which Make no Predictable Forecast Errors

Name	Number of Forecasts Made.
Anthony Cham	45
Berson	40
Grace Ortiz	20
James M. Griffin/Christine M	65
Jerry L. Jordan/Lynn Reaser	15
Joel L. Naroff/Veronika Whit	10
John Park	20
John Ryding	3
John Tuccillo/Robert Barr	5
Kathleen Camilli	46
Martin Regalia	25
Mitchel Held/Schindewolf	45
Paul Casperson	20
Paul Goulekas/James M. Griff	30
Richard Rippe	5
Richard Rippe/Michelle Laug	20
William Helman	10
Total	424

.

Many members of the 17 teams with perfect records made improvable forecasts as part of other teams. A total of 11 forecasters made no improvable forecasts whether alone or as part of a team. These forecasters made a total of 254 forecasts.

Such striking evidence of irrationality is not found for lower maturities. Many fewer predictions fall out of the calculated intervals so we propose many fewer changes – a total of 115 for the 7 maturities shorter than 30 years. One of the changes we propose, a reduction by the minimum 10 basis points of Richard Berner and Russell Sheldon's prediction made in May 1994 of the average yield on 5 year treasury bonds in second quarter of 1994 worsens that forecast, so our overall record is 1214 forecasts improved and one forecast worsened. Results are shown in Table VI.

Table VI Number of Forecasts Modified For Different Maturities

Maturity	Number Decreased	Number Worsened By Decrease	Number Increased	Number Worsened By Increase
30 Years	1082	0	18	0
10 Years	11	0	5	0
7 Years	11	0	2	0
5 Years	6	1	3	0
2 Years	7	0	3	0
1 Year	3	0	6	0
6 Months	5	0	10	0
3 Months	14	0	29	0

This means that our record of 1,100 right out of 1,100 predictions of the signs of forecast errors is based, in part, on two ex post choices. It applies to the interest rates of 30 year treasury bills which are one of 8 securities which we investigated and it relies on the parameter 2.6 which was chosen ex post. Aside from that, our approach used information available to forecasters when they made the 1,100 forecasts which we can improve without worsening any.

V. Conclusions.

The forecasts of quarterly average 30 year treasury bond yields in data set collected by S.Peterson (2001) from the Blue Chip Financial Survey (1987-1996), contain errors which can be predicted using information available to forecasters. Using a simple natural definition of "far", we find that all forecasts which are far above the average of one month lagged forecasts are too high and that all forecasts which are far below the average of one month forecasts are too low. This means that it is possible to reduce losses using only lagged information for any loss function which increases if forecasts are further from the truth.

References

Batchelor R. and Peel D.A (1998), "Rationality Testing under Asymmetric Loss", Economics Letters 61, 49-54.

Christoffersen P.F. and Diebold F.X. (1997), "Optimal prediction under asymmetric loss", *Econometric Theory* 13, 808-817.

Ehrbeck T. (1993), "Essays on the Rationality of Expectations", Phd Dissertation European University Institute.

Ehrbeck, T. and R. Waldmann (1996), " Why Are Professional Forecasters Biased? Agency Vs. Behavioral Explanations", *Quarterly Journal of Economics*, 111, 21-40.

Elliott, G; I. Komunjer and A Timmermann (2005) "Estimation and Testing of Forecast Rationality under Flexible Loss", *Review of Economic Studies*, 72, 1107-25.

Elliott, G.; I. Komunjer, and A. Timmermann, A (2008) "Biases in Macroeconomic Forecasts: Irrationality or Asymmetric Loss?", *Journal of the European Economic Association*, 6, 122-57

Elliott, G. and A. Timmermann (2008) “Economic Forecasting”, *Journal of Economic Literature,* 46,, 3-56.

Granger C.W.J. (1968), “Prediction with a generalized cost function”, OR, 20, 199-207.

Granger C.W.J. (1999), “Outline of Forecast Theory Using Generalized Cost Functions”, Spanish Economic Review 1, 161-173.

Granger C.W.J. and. Pesaran M.H (2000), “Economic and Statistical Measures of Forecast Accuracy”, Journal of Forecasting 19, 537-560.

Patton A. and Timmermann A. (2007), “Properties of Optimal Forecasts under Asymmetric Loss and Nonlinearity”, Journal of Econometrics 140, 884-918.

Pesaran M.H. and Skouras S. (2002), “Decision-based Methods for Forecast Evaluation”, in Clements, M.P. and D. F. Hendry (Eds.), A Companion to Economic Forecasting. Blackwell, Oxford.

Peterson S.P. (2001), “Rational Bias in Yield Curve Forecasts”, *The Review of Economics and Statistics*, Vol.3, N.3 , pp.457-464.

U.S. Treasury (2008) http://www.ustreas.gov/offices/domestic-finance/debt-management/interest-rate/yield.shtml

Varian H. R. (1974), “A Bayesian Approach to Real Estate Assessment”, in Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage, eds. S.E. Fienberg and A. Zellner, Amsterdam: North Holland, 195-208.

Waldmann, R. (1995) “Predicting the Signs of Forecast Errors” European University Institute Working Paper ECO 95/22.

Weiss A.A. (1996), "Estimating Time Series Models Using the Relevant Cost Function", Journal of Applied Econometrics 11, 539-560.

Zellner A. (1986), "Bayesian Estimation and Prediction Using Asymmetric Loss Functions", Journal of the American Statistical Association, 81, 446-451.