MODELLING RESEARCH GROUP

UNIVERSITY OF SOUTHERN CALIFORNIA
1880
PALMAM QUI MERUIT FERAT

DEPARTMENT OF ECONOMICS
UNIVERSITY OF SOUTHERN CALIFORNIA
UNIVERSITY PARK
LOS ANGELES, CALIFORNIA
90089-0152

MODELLING RESEARCH GROUP

BIASED PREDICTORS, RATIONALITY AND THE EVALUATION OF FORECASTS

ARNOLD ZELLNER*

MRG WORKING PAPER #M8612

DEPARTMENT OF ECONOMICS
UNIVERSITY OF SOUTHERN CALIFORNIA
UNIVERSITY PARK
LOS ANGELES, CALIFORNIA
90089-0152

BIASED PREDICTORS, RATIONALITY AND THE EVALUATION OF FORECASTS

ARNOLD ZELLNER*

MRG WORKING PAPER #M8612

March 1986

ABSTRACT

Optimal, rational forecasts are often biased and thus the empirical finding that actual forecasts are biased is not necessarily evidence of irrational behavior. Using an asymmetric loss function and a regression forecasting problem, these points are explicitly demonstrated.

*H.G.B. Alexander Research Foundation, Graduate School of Business, University of Chicago, visiting the University of Southern California.

Biased Predictors, Rationality and the Evaluation of Forecasts

by

Arnold Zellner*
University of Chicago

Abstract

Optimal, rational forecasts are often biased and thus the empirical finding that actual forecasts are biased is not necessarily evidence of irrational behavior. Using an asymmetric loss function and a regression forecasting problem, these points are explicitly demonstrated.

It has long been recognized that biased predictors and estimators may be better than unbiased predictors and estimators in terms of expected loss or risk relative to various loss functions--see e.g. Stein (1955), Efron and Morris (1973), Judge et al. (1985, Chs. 3,4), and Zellner (1963, 1985). Indeed, Bayesian estimators and predictors, based on proper informative priors, are generally biased but are known to be admissible and to minimize average or Bayes risk. In view of these well known facts, it is surprising that many studies that attempt to evaluate forecasts concentrate attention on whether forecasts are unbiased and interpret a departure from unbiasedness as evidence of a departure from rationality and/or

---

*Research financed in part by the National Science Foundation and by income from the H.G.B. Alexander Endowment Fund, Graduate School of Business, University of Chicago. The author's address is: Graduate School of Business, U. of Chicago, 1101 East 58th Street, Chicago, Illinois 60637.

the rational expectations hypothesis. For a review of a large portion of this literature, see Zarnowitz (1985) who, at my suggestion, mentioned that optimal forecasts may be biased. Earlier, Grossman (1975) recognized that the rational expectations hypothesis does not imply that the anticipated value of a variable is equal to its mathematical expectation in all circumstances.

The purpose of this note is to provide an explicit regression example in which a biased predictor is optimal relative to an asymmetric loss function and to indicate the implications for usual "tests of rationality." Also, it will be shown how to estimate the shape parameter of an asymmetric loss function that determines its degree of asymmetry.

Let the model for the $n\times 1$ observation vector $\underset{\sim}{y}$ be a standard normal multiple regression model,

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{u} \tag{1}$$

where X is a given nonstochastic $n\times k$ matrix of rank k, $\underset{\sim}{\beta}$ is a $k\times 1$ vector of regression parameters and $\underset{\sim}{u}$ is an $n\times 1$ vector of disturbance terms that are assumed independently drawn from a normal distribution with zero mean and variance σ^2. Further, let a future value of the dependent variable, y_f be given by

$$y_f = \underset{\sim}{x}_f'\underset{\sim}{\beta} + u_f \tag{2}$$

where $\underset{\sim}{x}_f$ is a given $k\times 1$ vector and u_f is a future error term drawn independently of the errors in (1) from a $N(0,\sigma^2)$ distribution. Our problem is to forecast y_f given the data, the model in (1) and the vector $\underset{\sim}{x}_f$. From results in the literature, it is known that the least squares predictor (LSP),

$$\hat{y}_f^{LS} = \underset{\sim}{x}_f'\hat{\underset{\sim}{\beta}} \tag{3}$$

where $\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$ is a minimum variance, unbiased predictor of y_f under the normality assumption introduced above. That is $\hat{y}_f^{LS}$ is optimal relative to a squared error loss function, $L_s = c(\tilde{y}_f - y_f)^2$, with $c > 0$ and $\tilde{y}_f$ a point predictor, subject to the side condition that $\tilde{y}_f$ be unbiased, that is $E(\tilde{y}_f - y_f) = 0$. The restriction that $\tilde{y}_f$ be unbiased can be costly in terms of say, mean squared error (MSE). Note that MSE $= \mathrm{var}(\tilde{y}_f - y_f) + [E(\tilde{y}_f - y_f)]^2$ = variance plus squared bias. As is well known, a predictor with a small variance and a small bias can have lower MSE than the LSP in (3)--see, e.g. Zellner (1963) for examples. Also a Bayesian predictor based on a proper prior distribution for the regression parameters will in general be biased but has lower average risk relative to a squared error loss function than the LSP. Thus even with use of a symmetric squared error loss function, various good or optimal predictors are biased.

Instead of a symmetric squared error loss function, consider use of the following asymmetric LINEX loss function, introduced by Varian (1975),

$$L(\Delta) = b[e^{a\Delta} - a\Delta - 1] \qquad \begin{array}{l} a \neq 0 \\ b > 0 \end{array} \tag{4}$$

where $\Delta = \tilde{y}_f - y_f$, the forecast error associated with the use of the point forecast $\tilde{y}_f$. It is seen that $L(0) = 0$, the minimal value of L, and when $a > 0$, loss rises almost exponentially for $\Delta > 0$ and almost linearly when $\Delta < 0$. Thus with $a > 0$, over-forecasting results in higher loss than with equal under-forecasting and the reverse is true for $a < 0$. When $|a|$ is small in value, $L(\Delta) \doteq a^2\Delta^2/2$, a squared error loss function, as can be seen by expanding $e^{a\Delta} \doteq 1 + a\Delta + a\Delta^2/2$ and inserting this expression in (4).

Given that we have a predictive pdf for y_f, $p(y_f|D)$, where D denotes given sample and prior information, it can be used to compute expected loss as follows,

$$EL(\Delta) = b[e^{a\tilde{y}_f}Ee^{-ay_f} - a(\tilde{y}_f - Ey_f) - 1].$$

On minimizing this expression with respect to $\tilde{y}_f$, as in Zellner (1985), the result is

$$\tilde{y}_f^* = (-1/a)\log[Ee^{-ay_f}]. \tag{5}$$

In the case of the regression model (1) with a diffuse prior pdf, the predictive pdf given σ^2 is normal with mean $\hat{y}_f^{LS} = \underset{\sim}{x}_f'\hat{\underset{\sim}{\beta}}$ and variance $v = [1 + \underset{\sim}{x}_f'(X'X)^{-1}\underset{\sim}{x}_f]\sigma^2$ --see e.g. Zellner (1971, p.72). Thus, from (5),

$$\tilde{y}_f^* = \hat{y}_f^{LS} - av/2, \tag{6}$$

a special application of (5) presented in Varian (1975). This predictor is biased but as shown in Zellner (1985) it has uniformly lower risk than the LSP, $\hat{y}_f^{LS}$, and thus the latter <u>is inadmissible relative to the LINEX loss function in (4)</u>. This is also true if v in (6) is replaced by $\hat{v} = [1 + \underset{\sim}{x}_f'(X'X)^{-1}\underset{\sim}{x}_f]s^2$, where $s^2 = (\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})/(n-k)$.

Now suppose that a forecaster uses (6) with $v = \hat{v}$ to forecast y_f. Then his forecast error is e_f given by

$$e_f = \tilde{y}_f^* - y_f = \hat{y}_f^{LS} - y_f - a\hat{v}/2 \tag{7}$$

and his mean error or bias is

$$Ee_f = -av/2 = -a[1 + \underset{\sim}{x}_f'(X'X)^{-1}\underset{\sim}{x}_f]\sigma^2/2. \tag{8}$$

Thus the algebraic sign of the bias depends on the value of the LINEX parameter a. The larger $|a|$, the greater the asymmetry of the loss function and the greater is the bias in (8). However, the forecaster is rational in the sense that he is minimizing expected loss by using the biased predictor $\tilde{y}_f^*$ in (6).

If we have a sequence of forecast errors, $e_{fi} = \tilde{y}_{fi} - y_{fi}$, $i = 1,2, \ldots, m$, where $y_{fi} = \underset{\sim}{x}'_{fi}\underset{\sim}{\beta} + u_{fi}$ and $\tilde{y}^*_{fi} = \hat{y}^{LS}_{fi} - a\hat{v}_i/2$, with $\hat{y}^{LS}_{fi} = \underset{\sim}{x}'_{fi}\hat{\underset{\sim}{\beta}}$, $\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$, $\hat{v}_i = [1 + \underset{\sim}{x}'_{fi}(X'X)^{-1}\underset{\sim}{x}_{fi}]s^2$, and $s^2 = (\underset{\sim}{y}-X\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y}-X\hat{\underset{\sim}{\beta}})/(n-k)$, then the expectation of the mean forecast error $\bar{e}_f = \sum_{i=1}^{m} e_{fi}/m$ is

$$E\bar{e}_f = -a\bar{v}/2 \tag{9}$$

where $\bar{v} = \sigma^2 \sum_{i=1}^{m} [1 + \underset{\sim}{x}'_{fi}(X'X)^{-1}\underset{\sim}{x}_{fi}]/m$. From (9), the algebraic sign of the "bias," will be determined by the algebraic sign of the LINEX parameter a. Also, from (9), the following is an estimator of a

$$\hat{a} = -2\bar{e}_f/\hat{\bar{v}} \tag{10}$$

where $\hat{\bar{v}}$ is $\bar{v}$ with σ^2 replaced by s^2. Since $\bar{e}_f \equiv \bar{e}^{LS}_f + a/2\hat{\bar{v}}$, where $\bar{e}^{LS}_f = \sum_{i=1}^{m} e^{LS}_{fi}/m$, $\bar{e}^{LS}_f$ and s^2 are independent, and $E\bar{e}^{LS}_f = 0$, $E\hat{a} = a$, that is the estimator in (10) is unbiased. By a direct calculation $\mathrm{Var}(\hat{a}) = 4\nu[1 + \bar{\underset{\sim}{x}}'_f(X'X)^{-1}\bar{\underset{\sim}{x}}_f]/m\sigma^2c^2(\nu-2)$, where $\nu = n-k$, $\bar{\underset{\sim}{x}}_f = \sum_{i=1}^{m} \underset{\sim}{x}_{fi}/m$ and $c = \sum_{i=1}^{m} [1 + \underset{\sim}{x}'_{fi}(X'X)^{-1}\underset{\sim}{x}_{fi}]/m$.

It will be noted from Zarnowitz (1985, Table 3, p. 301) that his mean forecast errors tend to differ from zero. As he explains, "Table 3 shows that almost all forecasters underestimated inflation....In contrast, real growth [of GNP] was predominately overestimated...." (pp. 299-300). Whether his non-zero mean forecast errors are (a) peculiar to the period studied, (b) evidence of irrationality and/or (c) due to forecasters' use of optimal biased forecasts cannot, of course, be ascertained from just study of past forecast errors' properties. However, under assumption (c) and an assumed use of a LINEX loss function, it appears from (10) that the parameter a may be positive for inflation forecasting and negative for forecasting real growth of GNP. To make such inferences secure, further research is needed to determine the forms of forecasters' loss functions and how they use them, explicitly or implicitly, to generate forecasts.

References

Efron, B. and C.F. Morris (1973), "Stein's Estimation Rule and Its Competitors--An Empirical Bayes Approach," Journal of the American Statistical Association, 68, 117-130.

Grossman, S. (1975), "Rational Expectations and the Econometric Modelling of Markets Subject to Uncertainty: A Bayesian Approach," Journal of Econometrics, 3, 255-272.

Judge, G. et al. (1985), The Theory and Practice of Econometrics, 2nd ed., New York: John Wiley & Sons, Inc.

Stein, C. (1955), "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, Berkeley: University of California Press, 197-206.

Varian, H.R. (1975), "A Bayesian Approach to Real Estate Assessment," in S.E. Fienberg and A. Zellner (eds.), Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage, Amsterdam: North-Holland Publishing Co., 195-208.

Zarnowitz, V. (1985), "Rational Expectations and Macroeconomic Forecasts," Journal of Business and Economic Statistics, 3, 293-310.

Zellner, A. (1963), "Decision Rules for Economic Forecasting," Econometrica, 31, 111-130.

__________ (1971), An Introduction to Bayesian Inference in Econometrics, New York: John Wiley & Sons, Inc.

__________ (1985), "Bayesian Estimation and Prediction Using Asymmetric Loss Functions," Technical Report, H.G.B. Alexander Research Foundation, Graduate School of Business, University of Chicago, to appear in the Journal of the American Statistical Association.

REPORTS IN THIS SERIES

M8519: TIMUR KURAN: Anticipated Inflation and Aggregate Employment: The Case of Costly Price Adjustment.

M8520: LUIGI L. PASINETTI: Sraffa's Circular Process and the Concepts of Vertical Integration.

M8521: SUBAL C. KUMBHAKAR: Econometric Models for Input-Specific Technical and Allocative Inefficiency: U.S. Class 1 Railroads.

M8522: DAMAYANTI GHOSH: Maximum Likelihood Estimation of the Dynamic Shock-Error Model.

M8523: JEAN-JACQUES LAFFONT and JEAN TIROLE: The Dynamics of Incentive Contracts.

M8524: DEAN C. MOUNTAIN and CHENG HSIAO: Peak and Off-Peak Industrial Demand for Electricity: The Hopkinson Rate in Ontario, Canada.

M8525: RICHARD H. DAY: Policy, Prediction and Intrinsic Business Fluctuations.

M8526: TIMUR KURAN: The Economic System in Contemporary Islamic Thought: Interpretation and Assessment.

M8527: DENNIS J. AIGNER and DENZIL G. FIEBIG: Sample Design Considerations for Telephone Time-of-Use Pricing Experiments with an Application to OTC-Australia.

M8528: REZAUL KHANDKAR: The Effect of Temporary Versus Permanent Job Separation on the Reservation Wage and Job Search Strategy of the Unemployed.

M8529: ROBERT KALABA and LEIGH TESFATSION: Automatic Differentiation of Functions of Derivatives.

M8530: MICHAEL J. P. MAGILL and WAYNE J. SHAFER: Equilibrium and Efficiency in a Canonical Asset Market Model.

M8531: DENNIS J. AIGNER and PIETRO BALESTRA: On the Optimal Experimental Design for Error Commponents Models.

M8532: ROGER GUESNERIE: Stationary Sunspot Equilibria in an N Commodity World.

M8601: JOHN M. VEITCH: Repudiations and Confiscations by the Medieval State.

M8602: MARIO D. TELLO: Imperfect International Markets, Multinational Enterprises and Manufactured Exports from Developing Countries.

M8603: JEFFREY B. NUGENT: Arab Multinationals: Problems, Potential and Policies.

M8604: RICHARD H. DAY and T.-Y. LIN: Irregular Fluctuations and Comparative Policy Dynamics in the Keynesian Regime.

M8605: RICHARD M. GOODWIN: The Economy as an Evolutionary Pulsator.

M8606: Z.M. BERREBI and J. SILBER: Interquantile Differences, Income Inequality Measurement and the Gini Concentration Index.

M8607: JOSEPH G. HIRSCHBERG: A Comparison of Three Flexible Cost Functions Using Establishment Level Electricity Use Data.

M8608: IAN E. NOVOS and MICHAEL WALDMAN: Complementarity and Partial Nonexcludability: An Analysis of the Software/Computer Market.

M8609: ARNOLD ZELLNER: Further Results on Bayesian Minimum Expected Loss (MELO) Estimates and Posterior Distributions for Structural Coefficients.

M8610: ARNOLD ZELLNER: Bayesian Analysis in Econometrics.

M8611: ARNOLD ZELLNER: A Tale of Forecasting 1001 Series: The Bayesian Knight Strikes Again.

M8612: ARNOLD ZELLNER: Biased Predictors, Rationality and the Evaluation of Forecasts.

M8613: ARNOLD ZELLNER: Bayesian Estimation and Prediction Using Asymmetric Loss Functions.