FORECASTING PROCEDURES FOR SUCCESS

Suzy V. Landram, University of Northern Colorado
Frank G. Landram, West Texas A&M University
Chris Furner, West Texas A&M University

ABSTRACT

This study brings an awareness of five mistakes frequently found in forecasting. In doing so, a least squares model is derived that predicts new homes sold. This model employs current predictive innovations that have consistently out-performed other forecasting models. By comparing the out-of-sample accuracy of this model with others, frequent errors are noted and evaluated. Hence, our objective is to promote an understanding and awareness of common mistakes made in forecasting. We also bring a summary of current innovations used to enhance the accuracy of predictive models. **JEL Classification:** C53

INTRODUCTION

In this study, new homes sold are forecasted revealing the predictive economic status of our US economy. New homes are newly constructed homes that have never been occupied. (http://research.stlouisfed.org/fred2/data /HSN1FNSA.TXT). However, we begin with a discussion concerning the innovative least squares, time series decomposition model that is superior to traditional decomposition operations. This model also consistently outperforms other forecasting models with regard to both in-sample and out-of-sample accuracy (Landram, 2008a). Hence, attention is directed to this extraordinarily accurate model. Next the five common mistakes made in forecasting are discussed. In an effort to fix ideas and clarify concepts the new homes forecast is then given. This leading indicator helps predict the depths and duration of our economic decline. The example also illustrates the differences between in-sample and out-of-sample statistical measures. A discussion and concluding remarks follow.

MODEL DERIVATION AND EVALUATION

Consider the following least squares time series model

$$\hat{Y}_t = b_0 + b_1T_t + b_2C_t + b_3S_j + b_4T_tC_tS_j$$
(1)

where C_t and S_j represent cyclical and seasonal components obtained from the decomposition of a time series. The trend component is represented by any polynomial form of $T_t = b_0 + b_1X_t$, where X_t, = (1, 2, ..., n) and $T_tC_tS_j$ represents (T_t*C_t*S_j.)--the multiplicative form of the decomposition Observe the accuracy of (1) is superior to the traditional time series decomposition since it contains both additive and multiplicative variables placed in a least squares model. Further accuracy is usually obtained by employing judgmental modification-event modeling operations (Bunn and Wright, 1991), (Lawrence, et al., 1986). Enhancements are also obtained by combining forecasts. It is widely believed that judgmental modification and combined forecasts are essential elements in forecasting (Bates and Granger (1969), Batchelor and Dua. (1995), Fang, (2003). For more on judgmental modification see Blattberg and Hoch (1990), Manganelli (2007).

The inclusion of an event modeling variable D_j, which indicates periods of economic contraction, is also included together with the interaction term D_jT_t:

$$\hat{Y}_t = b_0 + b_1T_t + b_2C_t + b_3S_j + b_4T_tC_tS_j + b_5D_t + b_6D_t\,T_t$$
(2a)

Observe (2a) represents two forecasting equations. For example, starting in 2005 let $D_t = 1$, otherwise $D_t = 0$. Note, (2a) becomes equation (1) before 2005. Starting in 2005, (2b) is realized:

$$\hat{Y}_t = [b_0 + b_5] + [b_1 + b_6]T_t + b_2C_t + b_3S_j + b_4T_tC_tS_j.$$
(2b)

Notice the intercepts differ as do the slopes for T_t. in (2b) and (1).

Common Mistakes In Forecasting

1. Statistical measures used in explanatory models are incorrectly used in predictive models. Shmueli (2010) lists a host of journal articles that incorrectly employ statistical measures used in explanatory models to measure the accuracy of predictive models. Certainly, goodness of fit does not guarantee goodness of predictions. In forecasting new home sold below, we use PRESS, P^2, and Mallows Cp statistics. These statistics are explained below. Often, adjusted R^2 is incorrectly used to measure a model's out-of-sample accuracy.

2. Statistical models are employed where cause-effect relationships are incorrectly implied. This may happen when an explanatory variable is included in the model. Correlation does not imply causation brings to mind the age-old illustration that for young grade 1 children X=foot size and Y=achievement scores are highly correlated. This correlated does not imply that large feet cause

high achievement scores. The latent variable in this study is age. In general, first grade children near seven or eight rather than barely six years old have larger feet and higher achievement scores. The high correlation coefficient between foot size and achievement scores should direct attention to the underlying cause. Knowledge of the subject matter is then needed to discover the true cause-effect relationship existing between age and achievement scores.

3. Conditional error occurs when an X variable is predicted and then used in predicting Y. This error inflates interval predictions but is often ignored. In the example below, we show a method of measuring interval predictions. This method should be employed when conditional error exists. The mistake is the lack of awareness that conditional error inflates interval predictions.

4. Many believe that multicollinearity is not harmful to forecasts. Makridakis et al. (1998) incorrectly state, "Multicollinearity will not affect the ability of the model to predict." Multicollinear and unrelated variables cause overfitting, which, in turn, inflates out-of-sample interval predictions. In the Prediction Variance section below, we derive a statistic that measures the width of interval predictions.

5. Autocorrelation can occur by incorrectly omitting a relevant variable, say Z. If Z causes X and Y then when Z is omitted, the omission of Z often results in the errors being correlated. Most textbooks state misspecified models are the major cause of autocorrelation. They then plunge into the differencing method of correction. This should be the last resort, one should first try finding the omitted variable.

PREDICTION VARIANCES

Consider the sample regression equation

$$\hat{\mathbf{Y}} = \mathbf{Xb}, \quad (3)$$

where **X** is an n *x* p data matrix of full rank and **b** is the p x 1 vector of estimated regression coefficients. The HAT matrix is defined as

$$[h_{ij}] = \mathbf{H} = \mathbf{X}(\mathbf{X^TX})^{-1}\mathbf{X^T}, \quad (4)$$

where **H** represents the n x n HAT matrix (Hoaglin and Welsch, 1978). Vector **b** in (3) is defined as

$$\mathbf{b} = (\mathbf{X^TX})^{-1}\mathbf{X^TY}. \quad (5)$$

By employing (6) and (7) prediction variances are computed for $E[Y_i]$ and Y_i sometime referred to as Y-average and Y-individual:

$$\hat{V}(\hat{Y}_i) = (MSE_p)h_{ii} \quad (6)$$

$$\hat{V}(Y_i - \hat{Y}_i) = MSE_p(1+h_{ii}). \quad (7)$$

The mean square error is represented as

$$MSE_p = SSE_p / (n-p), \quad (8)$$

where SSE_p is the sum-of-squares error, n is the number of observations, and p is the number of regression coefficients in the model. The HAT diagonal h_{ii} is defined in (4).

Prediction intervals. Interval predictions for Y-average and Y-individual are

$$\hat{Y}_i \pm t_{\alpha/2}[(MSE)\, h_{ii}]^{1/2}, \quad (9)$$

$$\hat{Y}_i \pm t_{\alpha/2}[MSE\,(1+h_{ii})]^{1/2}, \quad (10)$$

where $MSE(h_{ii})$ and $MSE(1+h_{ii})$ are defined above, and $t_{\alpha/2}$ is an appropriate value from the t-distribution. Since an individual Y_i value is assumed to differ from its respective mean $E[Y_i]$ in a random manner, $\hat{Y}_i$ is the best estimate for both Y-average $E[Y_i]$ and Y-individual.

Measuring prediction variances. If there are n observations in the sample, there are n point and interval predictions. These n interval predictions may differ in size (heterskedasticity); therefore, a preferred method of measurement is to sum all n prediction variances:

$$PV(p) = \sum_{i=1}^{n} \hat{V}(\hat{Y}_i) = (MSE_p)\sum_{i=1}^{n} h_{ii} = (MSE_p)\,p, \quad (11)$$

where MSE_p is the mean square error from the model with p regression coefficients. Hocking (2003) shows

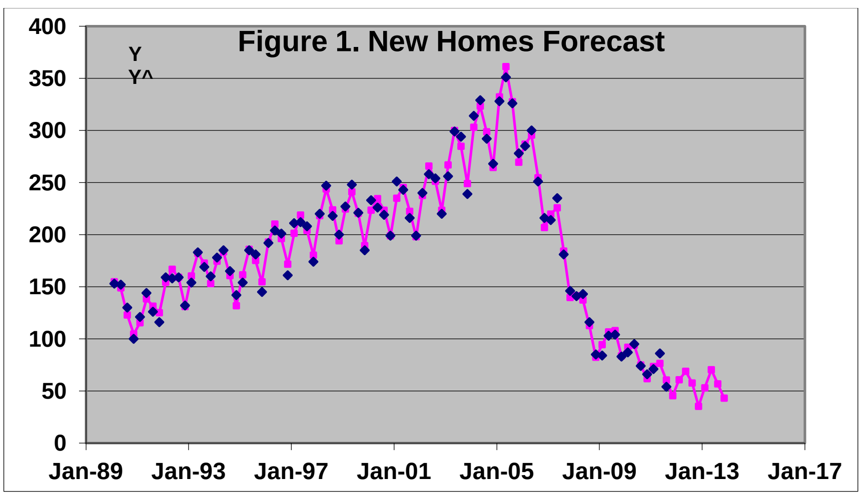
$$\sum_{i=1}^{n} h_{ii} = p. \quad (12)$$

NEW HOMES SOLD FORECAST

Figure 1 is the visualization of new homes sold (http://research.stlouisfed.org/fred2/data /HSN1FNSA.TXT). Data ranged from quarter 1 of 1990 to quarter 3 of 2011. This figure shows the actual value, fitted values, and out-of-sample predictions. Projections were from quarter 4, 2011 to quarter 4, 2013.

$$\hat{Y}_t = b_0 + b_1X_t + b_2C_t + b_3S_j + b_4T_tC_tS_j + b_5D_t + b_6D_tX_t \quad (13)$$

Although similar, (13) differs from (2); X_t rather than the equation $T = b_0 + b_1X_t$ is in the model. Both X_t and T_t increase each quarter by a constant amount. The variable TCS_t is the combined forecast values produced from traditional decomposition software. It is not from multiplying the time series components X_t, C_t, and S_j found in this data set. As shown above with (1) and (2b), the event modeling variable D_t converts (13) into two equations.



DW = 1.98			n	87	
R^2	0.992		MSE	40.663	
Adj R^2	0.992		P*MSE	284.64	
P^2	0.780		PRESS	91,404.30	
ANOVA					
	df	*SS*	*MS*	*F*	*p-value*
Reg	6	412806.89	68801.15	1691.97	0.000
Error	80	3253.06	40.66		
Total	86	416059.95			
	Coef	*S(b)*	*t Stat*	*P-value*	*Partial F*
Intercept	-247.255	39.793	-6.213	0.00000	38.607
X	0.626	0.155	4.029	0.00013	16.233
S	41.163	13.248	3.107	0.00261	9.655
C	231.010	35.913	6.433	0.00000	41.377
TCS	0.764	0.056	13.762	0.00000	189.395
D	212.878	48.826	4.360	0.00004	19.009
XD	-3.140	0.750	-4.190	0.00007	17.553

Table 1 reveals that (13) is extremely accurate for in-sample fitted values with an $R^2 = 0.992$. However, the coefficient of prediction $P^2 = 0.780$ (explained below) is considerably lower revealing once again that goodness of fit does not guarantee goodness of predictions.

PREDICTIVE MODELS

The PRESS, prediction sum of squares (Hoaglin and Welsch 1978), and P_p^2 statistics (Landram, et al., 2008b) are used to evaluate predictive accuracy:

$$P_p^2 = 1 - (\text{PRESS} / \text{SST}); \quad (14)$$

where

$$\text{PRESS} = \sum_{i=1}^{n} (Y_i - \hat{Y}_{(i)})^2 = \sum_{i=1}^{n} e_{(i)}^2 \quad (15)$$

and SST = $\Sigma(Y_i - \bar{Y})^2$. The subscript in P_p^2 indicates the number of regression coefficients in the model. Since each Y_i value in (15) is not used in deriving the model that computes the associated prediction $\hat{Y}_{(i)}$, the independence of Y_i and $\hat{Y}_{(i)}$ is established. Remarkably, these residuals are computed in a direct manner:

$$e_{(i)} = e_i / (1-h_{ii}), \quad (16)$$

where e_i are ordinary residuals and h_{ii} are diagonal elements of the hat matrix **H** defined by (4) above. By dividing PRESS by SST and subtracting the ratio from one, P_p^2 is similar to the determination coefficient R_p^2 both are relative measures. However, P_p^2 indicates quality of fit while P_p^2 indicates quality of predictions. Again, since each $\hat{Y}_{(i)}$ in (15) is independent of $Y_{i.}$, PRESS residuals, PRESS, and P_p^2 are considered out-of-sample statistics and use to indicate the accuracy of out-of-sample predictions.

DISCUSSION AND CONCLUDING REMARKS

As stated above, Shmueli (2010) lists a host of journal articles that incorrectly employ statistical measures used in explanatory models to measure the accuracy of predictive models. In the example above, the new homes sold data possesses an extremely high R^2 and adjusted R^2 of 0.9922 and 0.9916, respectively. However, for predictions, out-of-sample measures such as PRESS, P^2 and other model selection statistics must be used along with judgmental modification variables. Certainly, in an election year with a congress many say is dysfunctional, judgmental modification is essential. Do not use goodness of fit statistics to infer prediction accuracy.

Conditional error occurs when the value of an independent X-value must be predicted and then used in predicting Y. Although time series variables are usually better than econometric variables, they still contain a degree of conditional error. Innovations such as moving seasonal indices help. Never the less, one should be cognizant that conditional error results in inflated prediction variances.

Multicollinearity is harmful to forecasts. The inclusion of unrelated and multicollinear variables cause overfitting and therefore inflates interval predictions. In nested models, use (11) to measure the size of the interval predictions.

Inflated Prediction Variances Rationale. It is important to remember when an additional variable is included in the regression model, h_{ii} in (6) and (7) can not decrease and SSE in (8) can not increase. With this in mind, as overspecification in the model increases, the MSE in (6) and (7) will not decrease significantly and may even increase when irrelevant variables enter the model. These factors cause $PV(p) = (MSE_p)\, p$ in (11) to increase thereby indicating inflated prediction variances.

Concluding Remarks. The objective of this study is to promote an understanding and awareness of mistakes frequently found in forecasting. The derivation of the new homes model not only gives a practical prediction concerning the state of our economy but also provides a review of the latest innovations used in predictive modeling. Indeed, the merits of employing judgmental modification, combined forecasts, and other innovations provide researchers with a head start in making reliable forecasts.

REFERENCES

Batchelor, R., P. Dua. (1995). “Forecast Diversity and the Benefits of Combining Forecasts,” *Management Science,* 41: 68-75.

Bates, J. and C. Granger (1969). "The Combination of Forecasts", *Operations Research Quarterly,* 20: 451-469.

Blattberg, R., S. Hoch. (1990). “Database Models and Managerial Intuition: 50% Model and 50% Managers,” *Management Science,* 36: 887-899.

Bunn, D. and G. Wright (1991). "Interaction of Judgmental and Statistical Forecasting Methods," *Management Science* 37, 501

Clemen, R. (1989). "Combining Forecasts: A review and Annotated Biography (with discussion)," *International Journal of Forecasting*, 5: 559-583.

Data can be obtained from http://research.stlouisfed.org/fred2/data/HSN1FNSA.TXT

Fang, Y. (2003). "Forecasting Combination and Encompassing Tests, International *Journal of Forecasting*, 19: 87-94.

Hoaglin, D, and Welsch, R. (1978) The HAT Matrix in Regression and ANOVA, *The American Statistician*, 32: 17-22.

Hocking, R. (2003) *Methods and Applications of Linear Models*, New York: John Wiley

Landram,F., R. Pavur, and B. Alidaee (2008a), "Combining Time Series Components for Improved Forecasts", *Decision Sciences :Journal of Innovative Education*, 5: 2.

Landram,F., R. Pavur, and B. Alidaee (2008b), "An Algorithm for Enhancing Spreadsheet Regression with Out-of-Sample Statistics," *Communication in Statistics: Simulation and Computation*, 32: 1578-1592.

Lawrence, M. J., Edmundson, R. H. and O'Connor M. J. (1986). "The Accuracy of Combining Judgmental and Statistical Forecasts," *Management Science*, 32: 1521-1532.

Makridakis, S., S. Wheelwright, R. Hyndman (1998) *Forecasting: Methods and Applications*, 3rd ed., Wiley, New York

Mallows, C.L. (1966). Choosing a Subset Regression. Joint Statistical Meetings, Los Angeles, CA

Manganelli, S. (2007), "Forecasting with Judgment," *Fifth European Central Bank Workshop on Forecasting Techniques: Forecast Uncertainty in Macroeconomics and Finance*.

Sanders, N. and K. Manrodt (2003). "The efficacy of using judgmental versus quantitative forecasting methods in practice," *Omega* 31: 511-522

Shmueli, G, (2010). "To Explain or to Predict?" *Statistical Science*, 25: 289-310