ISSN 1032-3813
ISBN 0 7326 0792 2

MONASH UNIVERSITY

AUSTRALIA

BUSINESS FORECASTING WITH EXPONENTIAL SMOOTHING:

COMPUTATION OF PREDICTION INTERVALS

Ralph D. Snyder and Simone Grose

Working Paper 11/96
September 1996

DEPARTMENT OF ECONOMETRICS

Business Forecasting with Exponential Smoothing

Computation of Prediction Intervals

Ralph D. Snyder
and
Simone Grose

Department of Econometrics
Monash University
Clayton, Victoria 3185
Australia

Correspondence
A/Professor R. D. Snyder, Department of Econometrics, Monash University, Clayton, Australia 3168. Telephone ++ 613 9905 2366. Fax ++ 613 9905 5474
E-mail: Ralph.Snyder@BusEco.Monash.Edu.Au.

Acknowledgments
This research was financed by a Faculty of Business and Economics research grant. We are thankful to our colleagues Max King and Brett Inder for their helpful comments on an earlier draft of the paper.

C:\ARALPH\EXPO11.DOC 16 August, 1996

Abstract

The problem considered in this paper is how to find reliable prediction intervals with simple exponential smoothing and trend corrected exponential smoothing. Methods for constructing prediction intervals based on linear approximation and bootstrapping are proposed. A Monte Carlo simulation study, in which the proposed methods are compared, indicates that the most reliable intervals can be obtained with a parametric form of the bootstrap method. An application of the method to predicting Malaysian GNP per capita is considered.

Key Words

Forecasting, exponential smoothing, prediction intervals, bootstrap method

1. INTRODUCTION

The success of the exponential smoothing methods (Brown, 1959; Holt 1967, Winters, 1960) for short term forecasting in business has tended to hide the fact that their implementation in computer business systems is done without proper regard for their statistical foundations. Consider inventory applications, for example. Prediction error variances of total demand over a delivery lead time, generated for safety stock determination, are typically based on the assumption of independent and identically distributed demands. Being incompatible with the intertemporal dependencies implicit in the exponential smoothing methods, this assumption gives rise to quite sizeable errors (Johnston & Harrison ,1986; Harvey and Snyder,1990). Prediction variances turn out to be too small, a problem leading to poor customer service levels in the inventory context.

Appropriate methods for the determination of prediction error variances may have largely been ignored in practice because of the relative complexity of analytical approaches to the problem. If this is the case, bootstrap methods (Efron, 1979; Horowitz, 1995) offer a relatively simple alternative now that cheap high powered computers have become commonly available . Bootstrap methods have been successfully applied in a time series context to autoregressive processes (McCullough, 1994). The time is now ripe to consider them in the context of the exponential smoothing methods with a focus on the determination of prediction intervals.

2. MODELS AND MAXIMUM LIKELIHOOD ESTIMATION

2.1 Invariant State Space Model

The general framework used here for the exponential smoothing methods is taken from Snyder (1985). y_t, the value of a time series in period t, is shaped by forces

from the past, reflected in a k-vector of state variables $\mathbf{x}_{t-1}$, and forces from the present represented by a random disturbance e_t. The relationship, in more specific terms, is given by the measurement equation

$$y_t = \mathbf{h}'\mathbf{x}_{t-1} + e_t, \tag{2.1}$$

$\mathbf{h}$ being a fixed k-vector. The evolution of the state variables over time is governed by the first-order recurrence relationship

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \alpha e_t, \tag{2.2}$$

$\mathbf{F}$ being a fixed kxk transition matrix and α a fixed k-vector of so-called smoothing parameters. The e_t are NID$(0,\sigma^2)$ random disturbances, independent of the so-called seed state vector $\mathbf{x}_0$. $\mathbf{F},\mathbf{h},\alpha$ are all potentially functions of an unknown parameter p-vector θ.

Example 1: Local Level Model

When the state at the beginning of period t is represented by a single number a_{t-1}, called the local level, the measurement equation becomes

$$y_t = a_{t-1} + e_t. \tag{2.3}$$

Changes in the underlying level are governed by

$$a_t = a_{t-1} + \alpha e_t. \tag{2.4}$$

Here $\theta = \alpha$.

Example 2: Local Trend Model

The local level in example 1 can be supplemented with a local growth rate, denoted at the beginning of period t by b_{t-1}. The measurement equation then becomes

$$y_t = a_{t-1} + b_{t-1} + e_t, \tag{2.5}$$

the changes in the state variables governed by the level and rate equations:

$$a_t = a_{t-1} + b_{t-1} + \alpha_1 e_t \tag{2.6}$$

$$b_t = b_{t-1} + \alpha_2 e_t. \tag{2.7}$$

Here $\theta = (\alpha_1 \quad \alpha_2)'$.

2.2 Exponential Smoothing and the Decomposition of the Likelihood Function

The point predictions in the study are obtained from the above models fitted to time series data using maximum likelihood methods. We show, using the theory of conditional probability, that the likelihood function can be written quite simply in terms of the one-step ahead prediction errors. We then demonstrate that minimising the sum of squared one-step ahead prediction errors also yields the maximum likelihood estimates.

To obtain an expression for the likelihood function, begin by considering the situation at the *beginning* of a typical period t. In these circumstances the old series values $y_1, y_2, \cdots, y_{t-1}$ will be known from past observations of the process. y_t, however, still being unknown at this point of time, is uncertain. The following argument is built on the inductive hypothesis that given trial values of the seed vector $\mathbf{x}_0$ and the parameter vector θ, the state vector $\mathbf{x}_{t-1}$ may be treated as *fixed* and known. Equation (2.1) then implies that $y_t | y_1, \cdots, y_{t-1}, \mathbf{x}_0, \theta, \sigma$ has the same distribution as e_t. Using $\phi(\bullet)$ as a generic symbol for a probability density function, and recognising

that the Jacobian of the transformation between $y_t|y_1,\cdots,y_{t-1},\mathbf{x}_0,\theta,\sigma$ and e_t is unity, the associated density is

$$\phi(y_t|y_1,\cdots,y_{t-1},\mathbf{x}_0,\theta,\sigma)=\left(2\pi\sigma^2\right)^{-1/2}\exp\left(-e_t^2/2\sigma^2\right). \tag{2.8}$$

Moving to the *end* of period t, y_t is observed. A fixed value for the error, the one-step prediction error, is calculated with

$$e_t = y_t - \mathbf{h}'\mathbf{x}_{t-1}. \tag{2.9}$$

Then a fixed value of $\mathbf{x}_t$ can be computed with equation (2.2), thus confirming the use of the inductive hypothesis about the fixed nature of $\mathbf{x}_{t-1}$. Together, these steps at the end of period t, define the most general linear form of exponential smoothing - see Box & Jenkins, 1976. Some important special cases are considered in the following two examples.

Example 3 (Simple Exponential Smoothing)

For the local level model (Example 1) the general form of exponential smoothing reduces to

$$e_t = y_t - a_{t-1} \tag{2.10}$$

$$a_t = a_{t-1} + \alpha e_t \tag{2.11}$$

This corresponds to the case of simple exponential smoothing (Brown, 1959). Note that a_t is now conditioned on the sample $y_1, y_2,\ldots,y_t$ and the seed a_0. In contrast to the model in Example 1 where it was random, a_t is now fixed and known.

Example 4 (Trend Corrected Exponential Smoothing)

For the local trend model (Example 2) we obtain

$$e_t = y_t - a_{t-1} - b_{t-1} \tag{2.12}$$

$$a_t = a_{t-1} + b_{t-1} + \alpha_1 e_t \tag{2.13}$$

$$b_t = b_{t-1} + \alpha_2 e_t$$

This is conventional trend corrected exponential smoothing (Holt, 1957). Again it should be noted that being conditioned on $y_1, y_2,...,y_t$ and the seeds a_0 and b_0, both a_t and b_t are fixed. This should be contrasted with their counterparts in the local trend model (Example 2), where both a_t and b_t were random.

The joint density of a sample of size n governed by the dynamic model (2.1) - (2.2), conditional on $\mathbf{x}_0, \theta$ and σ, is denoted by $\phi(y_1, y_2, \cdots y_n|\mathbf{x}_0, \theta, \sigma)$. By the theory of conditional probability, it can be written in the recursive form

$$\phi(y_1, y_2, \cdots y_n|\mathbf{x}_0, \theta, \sigma) = \phi(y_n|y_1, y_2, \cdots y_{n-1}, \mathbf{x}_0, \theta, \sigma)\phi(y_1, y_2, \cdots y_{n-1}|\mathbf{x}_0, \theta, \sigma) \quad (2.14)$$

and resolved into the product of conditional densities

$$\phi(y_1, y_2, \cdots y_n|\mathbf{x}_0, \theta, \sigma) = \prod_{t=1}^{n} \phi(y_t|y_1, y_2, \cdots y_{t-1}, \mathbf{x}_0, \theta, \sigma). \quad (2.15)$$

More specifically

$$\phi(y_1, y_2, \cdots y_n|\mathbf{x}_0, \theta, \sigma) = (2\pi\sigma^2)^{-n/2} \exp\left(-\sum_{t=1}^{n} e_t^2 \Big/ 2\sigma^2\right) \quad (2.16)$$

If the time series were stationary, the seed state vector $\mathbf{x}_0$ would have a well-defined density. This could then be used in conjunction with (2.11) to obtain the unconditional joint density $\phi(y_1, y_2, \cdots y_n|\theta, \sigma)$. The time series are non stationary in the cases of simple and trend corrected exponential smoothing. In these circumstances, because the distribution of $\mathbf{x}_0$ does not exist, the unconditional density of the sample cannot be obtained.

The density (2.12) effectively summarises all that can be known about the sample generated by a non-stationary stochastic process and the definition of likelihood must be based upon it. The likelihood, a function of $\mathbf{x}_0$ together with the parameters θ and σ, is defined as:

$$\ell(\mathbf{x}_0, \theta, \sigma|y_1, y_2, \cdots y_n) = (2\pi\sigma^2)^{-n/2} \exp\left(-\sum_{t=1}^{n} e_t^2 \Big/ 2\sigma^2\right) \quad (2.17)$$

The rationale for the inclusion of $\mathbf{x}_0$ is that it summarises the effects of the past. Yet the past provides no information on $\mathbf{x}_0$ To the extent that past forces have shaped the sample, any information about the past must be gleaned from the sample alone. It therefore makes sense to estimate $\mathbf{x}_0$ from the sample, hence the need to include $\mathbf{x}_0$ in the definition of the likelihood function.

The likelihood (2.17) looks remarkably like the prediction error decomposition of the likelihood function associated with Kalman filtering (Schweppe,1965; Harvey, 1991). It differs, however, in that the one-step prediction errors used in (2.17) are obtained with exponential smoothing rather than the Kalman filter. Unlike the Kalman filter, the one-step ahead prediction errors from exponential smoothing are homoscedastic. The conditioning on $\mathbf{x}_0$ eliminates the need to deal with the heteroscedastic one-step ahead prediction errors.

2.3 Estimation Method

The maximum likelihood estimate of σ is given by

$$\hat{\sigma} = \sqrt{\sum_{t=1}^{n} e_t^2 \Big/ n} \qquad (2.18)$$

On concentrating σ^2 out of the likelihood function (2.17) with (2.18), it may be established that maximising the likelihood is equivalent to minimising the sum of squared errors function

$$S(\mathbf{x}_0, \theta) = \sum_{t=1}^{n} e_t^2 , \qquad (2.19)$$

the e_t being the one-step ahead prediction errors from the exponential smoothing routine which are conditional on $\mathbf{x}_0$ and θ.

Those values of $\mathbf{x}_0$ and θ which minimise the sum of squared errors function (2.19), denoted by $\hat{\mathbf{x}}_0$ and $\hat{\theta}$, are obtained by numerical optimisation, possibly with a suitably adapted version of the Newton-Raphson method. Special transformations

may be needed to ensure that parameters such as α_1 and α_2 are restricted to non negative values. The number of variables to be optimised in $\mathbf{x}_0$ and θ can become quite sizeable in some models, particularly those involving seasonality. It is anticipated, however, that the model's linearity in $\mathbf{x}_0$ helps to maintain computational loads at relatively low levels.

2.4 Predictions

Predictions can be obtained, for $t = n+1, \cdots n+h$, with the equations

$$\hat{y}_t = \mathbf{h}'\hat{\mathbf{x}}_{t-1} \tag{2.20}$$

$$\hat{\mathbf{x}}_t = \mathbf{F}\hat{\mathbf{x}}_{t-1} \tag{2.21}$$

It is convenient to combine the *h* predictions into the *h*-vector $\hat{\mathbf{y}}$.

3. PREDICTION INTERVALS

3.1 Conditional Hessian Method (CHS)

The future values of the time series, denoted by the *h*-vector $\mathbf{y}$, ultimately depend on $\mathbf{x}_0$ and θ. The relationship, which can be written as

$$\mathbf{y} = \mathbf{f}(\mathbf{x}_0, \theta) + \mathbf{e} \tag{3.1}$$

where $\mathbf{e}$ is the *h*-vector of future disturbances, has a Jacobian with respect to $\mathbf{x}_0, \theta$ which, when evaluated at the optimal solution $\hat{\mathbf{x}}_0, \hat{\theta}$, will be denoted by the matrix **J**. The Hessian of the sum of squares function (2.19), evaluated at the optimal solution, will be denoted by **H**. If $\mathbf{y}$ is a *h*-vector of future values of the time series, then

$$E(\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})' \cong \hat{\sigma}^2(\mathbf{J}\mathbf{H}^{-1}\mathbf{J}' + \mathbf{I}) \tag{3.2}$$

The square roots of the diagonal elements of the resulting RHS matrix in (3.2) may then be used as approximate root mean squared prediction errors required to establish the prediction intervals.

A potential difficulty with this strategy is that the elements of θ are constrained in our context to be non negative. The establishment of prediction intervals when inequality

restrictions apply to the parameters is a non trivial problem. To avoid the associated complexities, we only find the Hessian with respect to those elements of θ which are *not* at their lower bound of zero in the optimal, feasible solution. The other elements are fixed at their optimal value of zero and are treated as though their sampling distribution is entirely concentrated at this point. The effect is to understate the risk and hence underestimate the root mean squared prediction errors. To the extent that the estimates are consistent, the errors from this form of conditioning are likely to be insignificant in all but small samples.

3.2 Bootstrap Methods

Bootstrap methods (Horowitz, 1995) represent an alternative where inequality restrictions are handled relatively simply by imposing constraints on the optimisation used to obtain the estimates. It is assumed that estimates $\hat{\mathbf{x}}_0, \hat{\theta}$ of the model have been obtained from the time series data and that the associated errors are represented by an n-vector $\hat{\mathbf{e}}$. The bootstrap strategy, in our context, is to simulate from a model with seed vector $\mathbf{x}_0^*$ and parameter vector θ^* determined by the approximation

$$\mathbf{x}_0^* = \hat{\mathbf{x}}_0 \text{ and } \theta^* = \hat{\theta}. \tag{3.3}$$

3.2.1 Conventional Bootstrap Method (CBS)

The typical trial τ of the conventional bootstrap method, in our context, consists of the following steps:

B1 Generate $n+h$ disturbances $e_1^\tau, \cdots e_{n+h}^\tau$ using random selection with replacement from the elements of $\hat{\mathbf{e}}$.

B2 Calculate the enlarged time series sample $y_1^\tau, \cdots y_{n+h}^\tau$ with the relationships in the model (2.1)-(2.2) for the seed vector and parameter vector (3.3).

B3 Find the constrained maximum likelihood estimates $\hat{\mathbf{x}}_0^\tau, \hat{\theta}^\tau$ for the sample $y_1^\tau, \cdots y_n^\tau$. Then generate the corresponding predictions $\hat{y}_{n+1}^\tau, \cdots \hat{y}_{n+h}^\tau$ with (2.20) and (2.21).

B4 Compute the associated prediction errors

$\hat{e}_t^{\tau} = y_t^{\tau} - \hat{y}_t^{\tau} \quad (t = n+1, \cdots n+h)$,

storing them in the τ^{th} row of a matrix $\hat{\mathbf{E}}$.

These steps are repeated a total of m times. On completion, the typical column j of the result matrix $\hat{\mathbf{E}}$ corresponds to a sample of size m of the j-step ahead prediction errors. The (1-P)/2 and (1+P)/2 fractiles of this sample are used as the limits of a prediction error interval, P being the nominated confidence level.

3.2.2 Parametric Bootstrap Method (PBS)

Another strategy is to again rely on the approximation (3.3) but to generate the disturbances from a continuous distribution presumed to be a good approximation to the empirical distribution of the residuals $\hat{\mathbf{e}}$. The possibility used here is an $N(\mathbf{0}, \hat{\sigma}^2)$ distribution, $\hat{\sigma}$ being the estimate of the standard deviation computed from the vector $\hat{\mathbf{e}}$ with (2.18). This parametric form of the bootstrap method is essentially the same as the conventional bootstrap method, the only difference being that step B1 above is replaced by random selection from an $N(\mathbf{0}, \hat{\sigma}^2)$ distribution.

4. MONTE CARLO SIMULATION

The three approaches to prediction interval determination potentially yield different results. Monte Carlo simulation studies were therefore undertaken to gauge the differences under controlled conditions.

4.1 Monte Carlo Design

The following steps were repeated 1000 times in each simulation:

S1 Generate an extended sample of disturbances $e_1, \cdots, e_{n+h}$ from an $N(0, \sigma^2)$ distribution.

S2 Use the model (2.1) to (2.2), in conjunction with the disturbances from step 1, to generate the extended time series $y_1, \cdots, y_n, y_{n+1}, \cdots, y_{n+h}$.

S3 Obtain the prediction intervals for a nominated confidence level P with one of the three methods applied to the sample $y_1, \cdots, y_n$. ($m = 1000$ in the case of the standard bootstrap and parametric methods).

S4 Update coverage counters for lead times $\ell = 1,2,\ldots h$ with

$$\kappa_{n+\ell} \leftarrow \begin{cases} \kappa_{n+\ell} + 1 & \text{if } y_{n+\ell} \text{ is in prediction interval for period } n+\ell \\ \kappa_{n+\ell} & \text{otherwise} \end{cases} . \quad (4.1)$$

On completion of the simulation compute the coverage indexes:

$$\gamma_{n+\ell} = \frac{\kappa_{n+\ell}}{1000P} \quad \text{for } \ell = 1,2,\ldots h \quad (4.2)$$

Ideally, when expressed as percentages, these indexes should be 100 percent. Any significant deviation from 100 percent is an indicator of serious problems with the prediction intervals.

The computer programs for the simulation, written in Fortran 77, were run on a DEC Alpha 7000 - 610 computer under the VMS operating system. Optimisation and random number generation were undertaken with standard subroutines from the NSWC subroutine library (Morris, 1993). Since the optimiser was designed for problems without constraints, non negative conditions on the parameters were enforced by replacing each non negative quantity by the square of a quantity permitted to range over the entire real line.

The simulation was applied under a range of experimental conditions to the models in examples 1 and 2. In the case of the local level model, all simulations began with $a_0 = 200$. The smoothing parameter α was varied, being assigned any of the values 0.0, 0.5 and 1.0. In all cases of the local trend model $a_0 = 200$ and $b_0 = 3$. The smoothing parameters $(\alpha_1 \quad \alpha_2)'$ were assigned values of $(0 \quad 0)$, $(0.8 \quad 0.5)$ and $(1 \quad 1)$. The standard deviation σ of the disturbances was set to 5 and 10. The simulations were undertaken for sample sizes n = 30, 50, 200. In the case of the local trend model we report results for a sample size of 20. To test the robustness of the results, all the simulations were repeated with a t-distribution in place of the normal

distribution in step S1 above. More specifically, random variates were generated from a *t*-distribution with 5 degrees of freedom. These variates were then scaled by a quantity of appropriate size to give the required disturbances with the nominated standard deviations of 5 or 10. The objective was to determine the effect of fatter tails on the prediction intervals.

4.1 Monte Carlo results

Tables 1 and 2 contain cross tabulations of the results expressed in percentage terms. Most sections of the tables refer averages of the coverage indexes. Two refer to the standard deviations of the coverage indexes. The following notation is used:

CBS	standard bootstrap method
CHS	conditional Hessian method
PBS	parametric bootstrap method
Smpl Size	sample size
Nominal CL	nominal confidence level

From the tables it can be seen that all methods produced prediction intervals that were too small on average. The parametric method is best amongst the three proposed methods, its performance being close enough to the 100 percent mark to justify its use in practice.

The conditional Hessian method out performed the conventional bootstrap method, despite previous findings indicating that, in many applications, bootstrap statistics are more accurate in small samples than first-order approximations (Horowitz, 1995).

It may have been anticipated, given the dependence of the other methods on the normal distribution, that the conventional bootstrap method would have worked best in the case of the *t*-distribution. This preconception, however, is not supported by the results. Interestingly, for the local trend model, the prediction intervals proved to be

more reliable in the case of the *t*-distribution. The resulting prediction intervals were presumably slightly wider than those for data generated from a normal distribution and hence had less of a downward bias.

More detailed findings are as follows:

- The prediction intervals improve with sample size. The parametric bootstrap method could be relied upon for most applications for sample sizes in excess of 30.
- The prediction intervals become less reliable with increases in the prediction horizon. Again the parametric approach seems to produce satisfactory results for the five periods tested.
- The simulations were undertaken for two nominal confidence levels (0.9 and 0.95). The results were remarkably similar for both the cases considered.
- The results for the various parameter values are more complex. Both small and large values of the parameters were associated with a strong performance. Intermediate values were associated with a diminution in the reliability of the intervals.

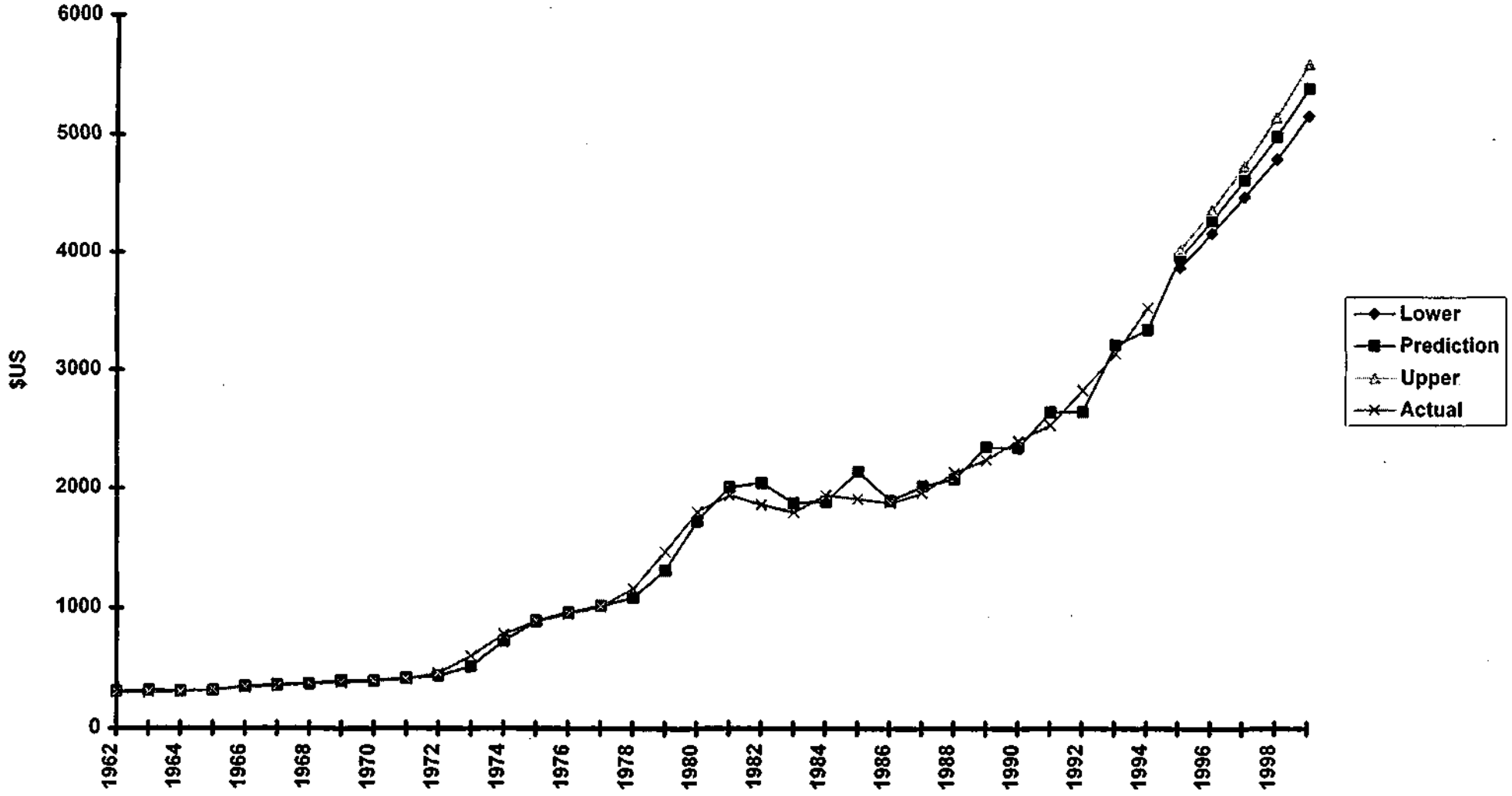
Tables summarising the results for the two standard deviations of 5 and 10 are not shown in the paper. The results for the prediction intervals were very close for both cases. This may reflect an invariance property with respect to the scale parameter σ.

5. APPLICATION

The parametric method for prediction interval determination is illustrated with an application to current GNP per capita of Malaysia. The data, available in annual

TO BE INSERTED INTO WORKING PAPER 11/96
R.D. SNYDER & S. GROSE

Figure 1. Malaysia: Current GNP Per Capita



terms, was taken from the World Bank World Tables (released: 22-April-96), extending from 1962 to 1994. This data is reproduced in Table 3.

Because of the presence of a long term exponential trend, the entire series was initially transformed with the natural logarithm function. Trend corrected exponential smoothing was applied to the transformed series in conjunction with a numerical optimiser to find the best smoothing parameters. Point forecasts were generated for the period 1995-1999 in the 'log space' and converted back to original terms with the exponential function. The limits of the 90 percent prediction intervals, expressed as percentage deviations from the point predictions in Table 4, were obtained directly from the results of the parametric bootstrap in 'log-space' based on 1000 replications. The results appear to be plausible and are illustrative of the potential that this approach holds in practice.

6. CONCLUSIONS

The bootstrap approach provides a relatively simple solution to the problem of establishing prediction intervals for the exponential smoothing methods. Despite the associated approximations, this study indicates that the parametric form of the bootstrap method produces sufficiently accurate prediction intervals for them to be used in practice. The computational loads associated with this method are quite high on traditional standards. With the recent advances in modern desktop computers, however, it has reached a point where the associated calculations can be done in a relatively short period of time, in the range of 30 to 60 seconds for the 1000 iterations of the bootstrap on the DEC Alpha computer. The parametric bootstrap method is now an attractive proposition for many business forecasting applications.

REFERENCES

Box, G. E. P. and G. M. Jenkins 1976, Time series analysis: forecasting and control (Holden-Day, San Francisco).

Brown, R. G. 1959, *Statistical forecasting for inventory control* (McGraw Hill, New York).

Efron, B. 1979, Computers and the theory of statistics: thinking the unthinkable, *SIAM Review* 21, 460-480.

Harvey, A. C. and R. D. Snyder, 1990, Structural time series models in inventory control, *International Journal of Forecasting 6,* 187-198.

Harvey, A. C., 1991, *Forecasting, structural time series models and the Kalman filter* (Cambridge University Press, Cambridge).

Holt, C. E. 1957, Forecasting trends and seasonals by exponentially weighted averages, *O.N.R. Memorandum No. 52,* Carnegie Institute of Technology, Pittsburgh.

Johnston, F. R. and P. J. Harrison, 1986, The variance of lead time demand, *Journal of the Operational Research Society* 37, 303-308.

Horowitz, J.L., 1995, Bootstrap methods in econometrics: theory and numerical performance, Working Paper Series #95-10, The University of Iowa.

McCullough, B. D., 1994, Bootstrapping forecast intervals: an application to AR(p) models, *Journal of Forecasting* 13, 51-66.

Morris, A. H. 1993, *NSWC library of mathematics subroutines* (Naval Surface Warfare Center, Dahlgren).

Schweppe, F. 1965, Evaluation of likelihood functions for Gaussian signals, *IEEE Transactions on Information Theory* 11, 61-70.

Snyder, R. D. 1985, Recursive estimation of dynamic linear models, *The Journal of the Royal Statistical Society: Series B* 47, 272-276.

Winters P. R., 1960, Forecasting sales by exponentially weighted moving averages, *Management Science* 6, 324-342.

	Method		
Factor	**CBS**	**CHS**	**PBS**
Distribution	*Effect on Average Coverage*		
Normal	97.6	98.5	99.1
t-Dist	97.1	98.4	98.8
Distribution	*Effect on Std Dev of Coverage*		
Normal	1.8	1.4	1.2
t-Dist	2.4	1.6	1.5
Smpl Size	*Effect on Average Coverage*		
30	96.4	97.8	99.0
50	97.3	98.3	98.8
200	99.1	99.4	99.5
Lead Time	*Effect on Average Coverage*		
1-Step	99.0	99.6	100.2
2-Step	97.6	98.5	99.0
3-Step	97.5	98.1	98.8
Nominal CL	*Effect on Average Coverage*		
0.9	97.4	98.3	99.0
0.95	97.9	98.7	99.1
Alpha	*Effect on Average Coverage*		
0	98.6	99.0	99.6
0.5	96.9	97.9	98.5
1	97.4	98.6	99.1

Table 1. Simulation of Prediction Intervals: Local Level Model

		Method		
	Factor	**CBS**	**CHS**	**PBS**
	Distribution	*Effect on Average Coverage*		
	Normal	95.8	96.1	98.2
	t-Dist	96.4	97.2	98.7
	Distribution	*Effect on Std Dev of Coverage*		
	Normal	3.9	3.7	2.5
	t-Dist	2.8	2.4	1.6
	Smpl Size	*Effect on Average Coverage*		
	20	84.5	85.7	90.7
	30	93.8	94.2	97.6
	50	96.4	96.8	98.6
	Lead Time	*Effect on Average Coverage*		
	1-Step	97.9	98.0	99.8
	2-Step	96.2	96.5	98.4
	3-Step	95.3	95.6	97.7
	4-Step	95.1	95.7	97.7
	5-Step	94.3	95.0	97.2
	Nominal CL	*Effect on Average Coverage*		
	0.9	95.4	95.8	98.1
	0.95	96.1	96.5	98.2
Alpha1	***Alpha2***	*Effect on Average Coverage*		
0	**0**	96.9	96.8	99.1
0.8	**0.5**	93.8	94.3	96.7
1	**1**	96.9	97.6	99.0

Table 2. Simulation of Prediction Intervals: Local Trend Model

Year	1966	1967	1968	1969	1970
NGNP	340	350	370	380	390
Year	1971	1972	1973	1974	1975
NGNP	410	460	600	780	890
Year	1976	1977	1978	1979	1980
NGNP	950	1010	1160	1470	1800
Year	1981	1982	1983	1984	1985
NGNP	1940	1860	1800	1940	1910
Year	1986	1987	1988	1989	1990
NGNP	1880	1960	2130	2240	2400
Year	1991	1992	1993	1994	
NGNP	2530	2830	3140	3520	

Table 3. Current Gross National Product per Capita of Malaysia ($US)
Source: World Bank World Tables (released: 22-April-96)

Year	Prediction	Lower Limit	Upper Limit
1995	3955	-4.3%	3.7%
1996	4273	-4.7%	4.3%
1997	4617	-5.4%	5.1%
1998	4989	-6.3%	5.8%
1999	5390	-7.3%	6.7%

Table 4. Predictions and 90 Percent Prediction Intervals
Current GNP per Capita of Malaysia

TO BE INSERTED INTO WORKING PAPER 11/96
R.D. SNYDER & S. GROSE

Figure 1. Malaysia: Current GNP Per Capita