The GSBGM Working Paper Series

WORKING PAPER 5/00

On log-transformations, vector autoregressions and empirical evidence

John Haywood

ISSN 1173-4523
ISBN 0-475-11556-2

The GSBGM Working Paper Series 5/00, September 2000.
*John Haywood, 'On log-transformations, vector autoregressions and empirical evidence'

Please address enquiries concerning this paper direct to the author*.

Other queries to:
Angela Dolan
Manager Administration Services
Graduate School of Business and Government Management
Victoria University of Wellington
PO Box 600
Wellington
New Zealand

Telephone: 64-4-463 5366
Facsimile: 64-4-463 5454
E-mail: angela.dolan@vuw.ac.nz

*John Haywood
School of Economics and Finance
Victoria University of Wellington
PO Box 600
Wellington
New Zealand

On log-transformations, vector autoregressions and empirical evidence

John Haywood*

Abstract

This paper re-analyses a bivariate US macroeconomic time series, previously used to demonstrate the need for bias corrections in the forecasts of levels of variables, modelled and originally forecast using a VAR after log-transformation. It is demonstrated that claims previously made for these data, concerning improvements in forecast accuracy following bias correction, were not well founded. Simple univariate forecasting procedures are shown to be more successful for these data than a cointegrated VAR, with or without bias correction. In the light of previous empirical work, such findings could have been expected. This further reinforces the call for an answer to why well-motivated theoretical advances in time series analysis often do not lead to noticeable improvements in out-of-sample forecast accuracy.

Keywords

Bias correction; Empirical evidence; Forecast comparison; Log-transformation; Simple forecasting procedures; VAR models.

* School of Economics and Finance, Victoria University of Wellington, PO Box 600, Wellington, New Zealand.
E-mail: John.Haywood@vuw.ac.nz

1. Introduction

The use of data transformations in empirical time series analysis is ubiquitous. Of such transformations, the natural logarithm is certainly one of the most commonly applied. In a recent paper, Arino & Franses (2000) – hereafter AF, considered the common situation of forecasting an untransformed (vector) time series, after the series have been modelled (and forecast) following log-transformation.

Granger & Newbold (1976) originally presented theoretical expressions for the autocorrelation properties of stationary and integrated univariate transformed series, which included the inverse logarithmic (exponential) transform as a particular (but commonly used) case. They then derived optimal forecasts for the transformed series and hence gave expressions for the losses incurred by the use of naive forecasts, which simply applied an appropriate transformation to the forecasts made on the modelled series. AF extended the work of Granger & Newbold in an important direction: while limiting themselves to autoregressive models (unlike Granger & Newbold), they considered vector time series. Particularly since the influential papers of Sims (1980) and Engle & Granger (1987), there has been a huge interest in non-structural (VAR) modelling and in cointegrated systems, with estimation of both the long run relationships between time series and their short run dynamics. Hence the extension to multivariate systems which AF made is certainly important and seems relevant for much empirical work.

In this paper I reconsider the two-dimensional time series which AF analysed as an empirical illustration of their proposed bias correction. AF presented evidence which, they claimed, demonstrated the superiority of forecasts following bias correction. They also suggested, on the basis of their empirical example, that bias correction becomes more important as the forecast lead time increases. As I show in Section 2, AF's results owe far more to the behaviour of the data in the particular out-of-sample period used than they do to the bias correction. In fact, it is easy to find out-of-sample periods in the same data for which the proposed bias correction will definitely make forecasts worse than with no correction, irrespective of the cost function chosen to assess the forecasts.

In Section 3, I suggest that if only these two series are considered, there are no apparent benefits for forecasting from using AF's cointegrated VAR model over simple univariate procedures, even after bias correction. There is, however, evidence that log transformation is sensible for these data (certainly in comparison to no transformation), so the issue of biased forecasts following transformation is still relevant. I relate this particular example to several other papers in Section 4 and suggest that the findings here could have been expected. Section 5 concludes.

2. Data revisited: the cointegrated VAR of AF

The data which AF analysed were a two-dimensional time series $(X_1(t), X_2(t))'$ of quarterly observations, recorded from 1947Q1 to 1988Q1, where X_1 is US real GNP and X_2 is US real gross domestic investment (both in billions of 1982 US dollars). The data are listed in Pindyck & Rubinfeld (1991, chapter 12) and are plotted in Figures 1 and 2.

Figure 1

US real GNP, 1947Q1 to 1988Q1

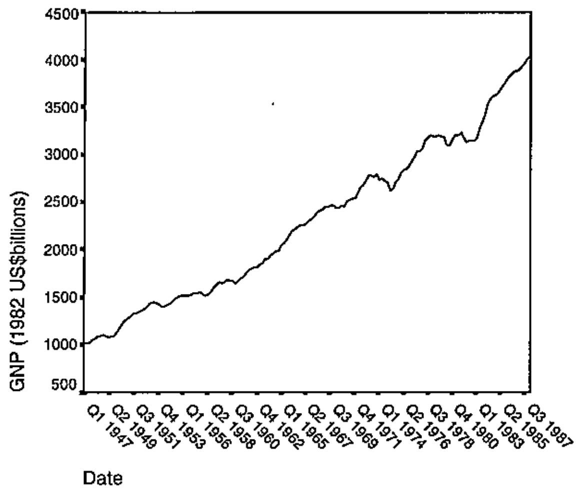
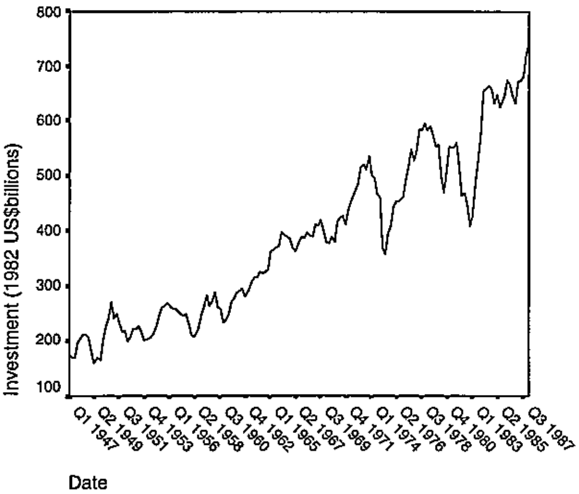


Figure 2

US real Investment, 1947Q1 to 1988Q1



AF estimated a VAR(3) on data up to 1980Q4 and withheld the remaining 29 observations for out-of-sample forecast evaluation. Restrictions were imposed on the VAR in order to impose a simple cointegrating relationship between GNP and Investment (in logs). No details were provided of the particular software which AF used, or of any notable features of the method of estimation. However, given the equations for the estimated VAR (AF, p.114) it is a simple matter to reproduce their naive forecasts, which are plotted along with the out-of-sample data in Figures 3 and 4. These figures well illustrate the true reasons for some of the conclusions that AF drew from their analysis.

AF presented a range of summary statistics for naive forecasts (as plotted in Figures 3 and 4) and unbiased forecasts, produced using their proposed correction applied to the naive forecasts (see AF, Table 1, p. 115). They also presented the same summary statistics, calculated for just the last 20 of the 29 out of sample forecasts (see AF, Table 2). It is well known that forecast error variability (usually) increases with increasing lead time. Thus one of the most striking aspects of AF's paper was the massive decrease in MSE (for example) when forecasts for lead times 1 to 9 were excluded (e.g. for GNP, the naive MSE was reported as dropping from 12,248 to 2,511). Similarly striking was the change in sign of mean error (for GNP), from negative when all 29 errors were considered, to positive.

Figure 3
Data with 29 out-of-sample forecasts from a cointegrated VAR(3) model (as estimated by AF): US real GNP, 1981Q1 to 1988Q1

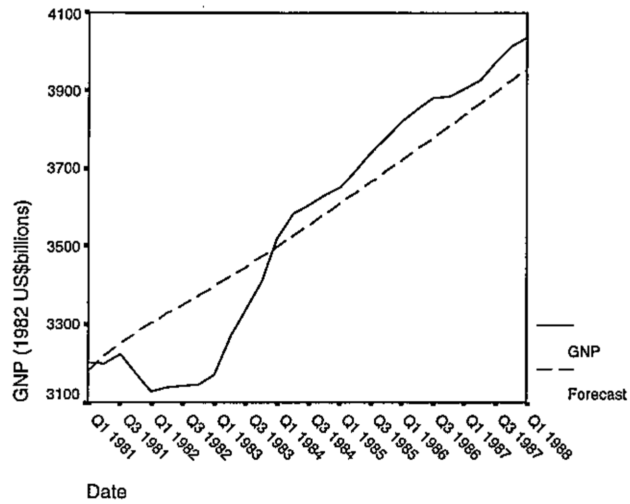
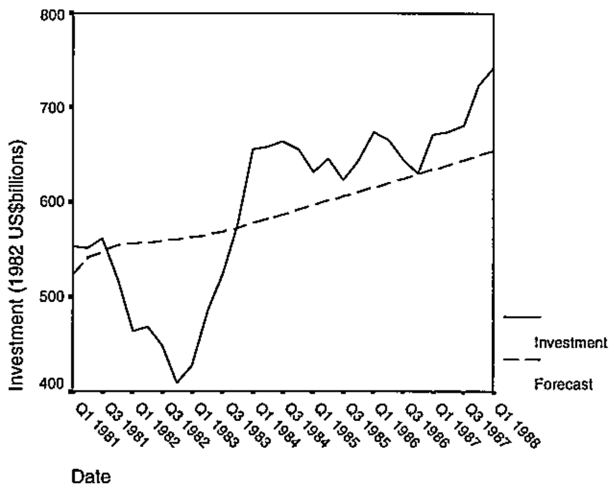


Figure 4
Data with 29 out-of-sample forecasts from a cointegrated VAR(3) model (as estimated by AF): US real Investment, 1981Q1 to1988Q1



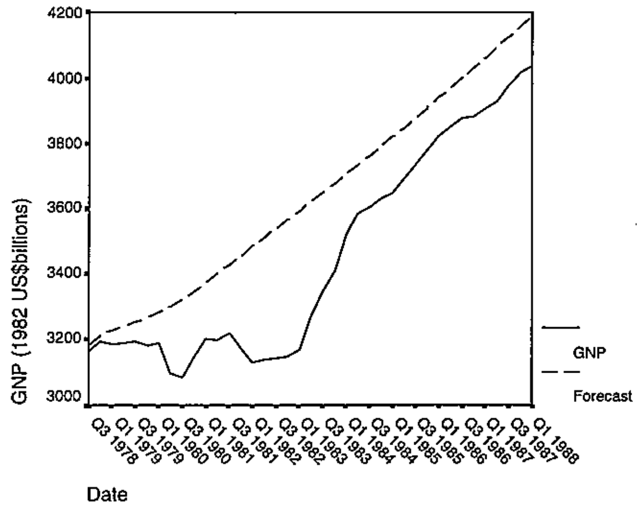
Figures 1 and 3 reveal the very obvious reason for these 'unusual' movements in summary statistics (but see also the further discussion of this point early in Section 3). The US recession of 1981/2 was not forecast at all by AF's VAR, so all but one of the first 12 errors were negative, and many of those were particularly large in absolute value (causing the change in sign in mean error noted above).

The form of bias correction which AF proposed can be considered either as an additive correction, added to a variable in log form before exponentiation, or as a multiplicative factor, applied to the naive forecasts. The additive form of correction is always non-negative (as in Granger & Newbold, 1976), while the multiplicative factor is always at least unity, and is guaranteed to be monotonically non-decreasing. Hence applying the bias correction to the last 17 out-of-sample forecasts plotted in Figure 3, all of which under-forecast GNP, is certain to improve forecast performance, provided the correction is of a 'sensible' size. In fact, AF noted this particular feature as evidence of the superiority of their bias-corrected forecasts, particularly when applied to long lead time forecasts. Conversely, the 11 forecasts from 1981Q2 to 1983Q4, which are all overestimates, will definitely deteriorate following positive bias correction (of any magnitude), irrespective of the cost function used to assess them. AF did not suggest that this implied their bias correction was inappropriate for short lead time forecasts.

It is not surprising therefore, that when considering only the last 20 out-of-sample forecasts, AF found an improvement using their bias correction (see AF, Table 2). Similar comments can be made about the behaviour of US real Investment and the summary statistics which AF presented: see Figures 2 and 4. To show that those results were due simply to features of these data, rather than to bias corrections which become more important as forecast lead time increases, consider Figure 5. Using AF's model equations, predicted values can easily be calculated starting from before 1980Q4, but continuing as far as is desired. Recall that 1980Q4 was the end of AF's estimation period, so such additional predictions include data that was actually used in the model-fitting procedure. Figure 5 shows the resulting (naive) predicted values of US real GNP, starting in 1978Q3 (i.e. an extra 10 forecasts than were calculated originally by AF, giving 39 in total).

Figure 5

Data with 39 forecasts from a cointegrated VAR(3) model (as estimated by AF): US real GNP, 1978Q3 to 1988Q1 inclusive. See text for further details



It is clear that, since every observation is over-predicted, any bias correction that increases the value of the forecasts (as AF's must) will inevitably worsen all the forecasts and any summary cost function of the forecast errors. This demonstrates that the claims AF made, concerning their bias correction and this *particular* data set, are not well founded.

3. Further data analysis

As noted in Section 2, the massive reduction in MSE when excluding the forecast errors from 1981Q1 to 1983Q1 inclusive was surprising, but understandable to some extent, given the behaviour of these data. However, a closer look at AF's Table 2 suggested that the figures may not be entirely reliable, since the mean error (ME) for GNP (X_1) is a larger positive value than the corresponding mean absolute error (MAE). For comparison with the summary statistics for additional models presented later, Table 1 presents a corrected version of the summary measures for the naive forecasts presented in AF's Table 2. (Note that AF's Table 1 was correct, as were the MAE and mean average percentage error (MAPE) rows in AF's Table 2.) Without re-estimating their model exactly, it is impossible to reproduce AF's bias-corrected forecasts and the associated summary measures.

Table 1

Evaluation of 20 quarters out-of-sample forecasting performance from 10 to 29 periods ahead for naive forecasts for levels from a cointegrated VAR(3) model for log-transformed series (as estimated by AF) [a]

	Log-model X_1	Log-model X_2
	Naive forecast	Naive forecast
ME	43.752	35.179
MAE	75.529	47.385
MAPE	2.04%	7.37%
MSE	6,468	2,882
RMSE	80.4	53.7

[a] The forecast errors are defined as the true value minus the predicted value. Forecast evaluation criteria are: mean error, ME; mean absolute error, MAE; mean absolute percentage error, MAPE; mean squared error, MSE; and root mean squared error, RMSE.

It is clear that while there is a reduction in MSE (from AF's Table 1 to Table 1 here), the reduction is nowhere near as large as that originally recorded in AF. For example, MSEs for the 20 out-of-sample forecasts from 10-29 periods ahead are approximately 50-65% of those for the full 29 periods, rather than the 18-20% recorded in AF. Perhaps of more general interest is the increase in absolute size of mean error, following the exclusion of short lead time forecasts: compare the naive MEs in AF's Table 1 with those presented in Table 1 here. A suggestion which AF (or others) may like to pursue is the possible relevance of bias corrections for log-transformed data, applied to estimation procedures which explicitly minimise functions of multi-step errors; see, for example, Haywood & Tunnicliffe Wilson (1997), Tiao & Xu (1993), and Weiss (1991).

There is a large literature which has demonstrated that simple forecasting methods often perform as well as, or better than, more complex procedures; see Fildes *et al.* (1998), Makridakis & Hibon (1979), Makridakis *et al.* (1982), and additional references therein. While those papers considered just univariate methods and a huge variety of time series, Fildes & Stekler (1999) focussed specifically on US and UK real GNP and inflation. Fildes & Stekler considered a range of structural (multivariate) macroeconomic forecasts and models, and summarised comparisons of such published forecasts with those produced by univariate and VAR/BVAR time series models. While noting that different studies found evidence that conflicted to some extent, among the points which Fildes & Stekler made are the following. Firstly, univariate time series methods generally provided real US GNP forecasts with accuracy comparable to that of structural model forecasts. Secondly, VAR forecasts of real US GNP were also comparable to structural model forecasts, except that in certain studies structural models had a clear advantage. Finally, most forecasts (structural or otherwise) failed to predict recessions in advance and sometimes even failed to recognize them contemporaneously.

Given the findings of Fildes & Stekler (1999), it is instructive to contrast AF's cointegrated VAR forecasts of X_1 and X_2 with those produced by some simple univariate methods. Table 2 gives a summary of the 29 out-of-sample naive forecast errors from AF's cointegrated VAR (as in AF, Table 1) along with a summary of the errors from simple exponential smoothing with drift (EWMA) and from the Robust Trend method, described in Fildes *et al.* (1998, Appendix A). Unbiased forecasts from AF's VAR (taken from AF, Table 1) and from the EWMA are also included.

Before discussing Table 2 further, I will elaborate slightly on the simple univariate methods that are summarised there. As is well known, simple exponential smoothing is an optimal forecasting procedure for data generated as a random walk signal plus noise, or alternatively as an ARIMA(0,1,1) process. For a stochastic trend to well approximate either of the (logged) series considered here, a drift term is essential. Hence the form of EWMA that was used can be viewed as the simplest exponentially weighted forecasting scheme which could be consistent with the behaviour of the data. The in-sample MSE was minimised to choose the discount factor, and for both series a value of unity was selected: therefore future values (at lead times 1, 2, ..., k) were forecast simply as the most recent observed value, plus an appropriate multiple (1, 2, ..., k) of the drift term. The drift term for each series was estimated as the slope coefficient from a simple regression of the (logged) data on a linear time trend, up to 1980Q4.

Table 2

Evaluation of 29 quarters out-of-sample forecasting performance for naive forecasts of levels from a cointegrated VAR(3) model (as estimated by AF), a univariate Robust Trend approach, and simple exponential smoothing with drift (EWMA), all for log-transformed series. Unbiased forecasts for the VAR and EWMA are also evaluated

	Log-model X_1 forecasts					Log-model X_2 forecasts				
	Naive			Unbiased		Naive			Unbiased	
	VAR (from AF)	Robust Trend	EWMA	VAR (from AF)	EWMA	VAR (from AF)	Robust Trend	EWMA	VAR (from AF)	EWMA
ME [a]	-9.720	-74.461	-29.781	-15.159	-33.808	4.906	-20.279	9.617	0.557	-6.116
MAE	93.474	76.304	66.680	89.587	64.200	55.652	43.865	53.654	53.659	44.421
MAPE	2.72%	2.28%	2.00%	2.62%	1.94%	10.20%	8.56%	9.74%	9.96%	8.50%
MSE	12,248	11,633	9,028	11,873	9,039	4,475	3,493	3,987	4,395	3,423
RMSE	110.7	107.9	95.0	109.0	95.1	66.9	59.1	63.1	66.3	58.5

[a] These statistics are defined in Table 1.

The unbiased forecasts for the EWMA were calculated in a standard way for integrated log-transformed series, following Granger & Newbold (1976). It is worth noting here that the bias correction of Granger & Newbold requires the assumption of a parametric model form, from which expressions for forecast error variances up to k steps ahead naturally follow. The EWMA is however a forecasting procedure, not a time series model. Given the comments made later in Section 4, that absence of an underlying model may be beneficial. However, here I assume an ARIMA(0,1,1) model is appropriate and construct the bias correction accordingly. While the correction does generally improve the EWMA forecasts, the improvements are not large and in some cases the performance actually deteriorates (e.g. ME and MSE for X_1). In that regard the comments made later, concerning the bias correction which AF applied to their cointegrated VAR, could also apply here.

The Robust Trend method was designed for series that follow a random walk with drift. Hence with that method, forecasts of future values *always* add an appropriate multiple of the estimated drift term to the most recent observed value, irrespective of the nature of the data. The robustness comes from the estimation of the drift parameter, which is achieved using the median first difference, itself augmented by a nonlinear function of the median differences: see Fildes *et al.* (1998) for further details. Forecasting using either of these univariate approaches requires only simple calculations, which can be implemented easily in a spreadsheet (as was the case here). As in AF, estimation was carried out using data up to 1980Q4.

Returning to Table 2, with the exception of the mean error, the out-of-sample performances of EWMA (with or without bias correction) and of Robust Trend are noticeably better for all summary measures and for both time series than the performance of AF's cointegrated VAR, even after their bias correction.

The reason for the smaller mean error from AF's VAR is clear from Figures 6 and 7. Figure 6 shows AF's (naive) forecast for X_1 along with the (naive) EWMA forecast, while Figure 7 displays AF's (naive) forecast for X_2 along with the (naive) Robust Trend forecast. None of the methods forecast the recession of 1981/2 (evident in both series), which is perhaps not surprising given the conclusions of Fildes & Stekler (1999), a point discussed further in Section 4. However, ignoring that recession and the subsequent recovery, the simple methods more accurately capture the long run trend in both X_1 and X_2. It is the underestimate of that long run trend by AF's VAR which offsets to some extent the large negative errors caused by not modelling the recession. Following AF's line of argument (AF, p. 115), note that the (naive) EWMA forecasts for X_1 outperform AF's (naive) forecasts 24 times to 5. Similarly, the (naive) Robust Trend forecasts for X_2 outperform AF's (naive) forecasts 18 times to 11.

Figure 6
Data with 29 out-of-sample forecasts from a cointegrated VAR(3) model (as estimated by AF) and from simple exponential smoothing with drift (EWMA): US real GNP, 1981Q1 to 1988Q1

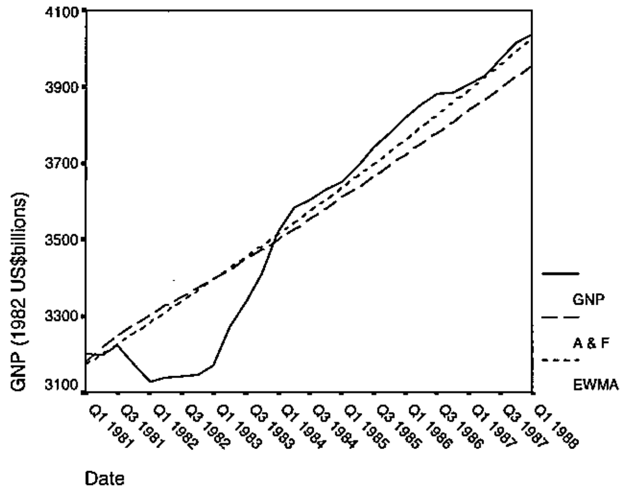
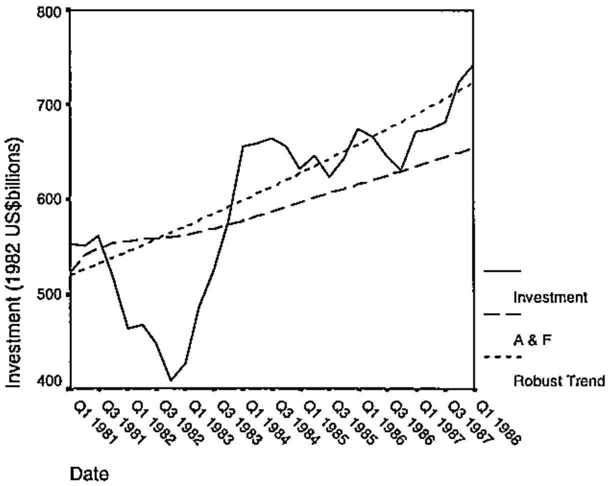


Figure 7
Data with 29 out-of-sample forecasts from a cointegrated VAR(3) model (as estimated by AF) and from a Robust Trend approach: US real Investment, 1981Q1 to 1988Q1



While the analysis of Section 2 suggests that there is little point in focussing solely on long lead time forecasts with just these out-of-sample data, for completeness consider Table 3. The summary statistics of Table 2 are repeated, but now the first nine forecasts are excluded (so only lead times 10 to 29 are considered, as in AF, Table 2). Note that unbiased forecasts for AF's VAR could not be included, due to the errors in their Table 2, first mentioned above.

Table 3

Evaluation of 20 quarters out-of-sample forecasting performance from 10 to 29 periods ahead for naive forecasts of levels from a cointegrated VAR(3) model (as estimated by AF), a univariate Robust Trend approach, and simple exponential smoothing with drift (EWMA), all for log-transformed series. Unbiased forecasts for the EWMA are also evaluated (those presented by AF for their VAR are not included: see text)

	Log-model X_1 forecasts				Log-model X_2 forecasts			
	Naive			Unbiased	Naive			Unbiased
	VAR (from AF)	Robust Trend	EWMA	EWMA	VAR (from AF)	Robust Trend	EWMA	EWMA
ME[a]	43.752	-50.462	8.381	3.081	35.179	-4.089	35.357	14.620
MAE	75.529	50.619	42.355	38.244	47.385	29.605	46.703	31.772
MAPE	2.04%	1.40%	1.18%	1.08%	7.37%	4.87%	7.26%	5.09%
MSE	6,468	4,837	3,106	2,945	2,882	1,402	2,801	1,631
RMSE	80.4	69.5	55.7	54.3	53.7	37.4	52.9	40.4

[a] These statistics are defined in Table 1.

The respective performance of the forecasting methods is unchanged from Table 2. Again with the exception of the mean error, the out-of-sample performances of EWMA (with or without bias correction) and of Robust Trend are noticeably better for all summary measures and for both time series than the (naive) performance of AF's cointegrated VAR. Even the exception of the mean error to that general superiority is marginal: for both time series, two of the three simple methods outperform the VAR on that criterion also.

The appropriateness of logarithmic transformations has been debated in the forecasting literature for some time, and in fact AF closed by noting the general difficulty in determining the empirical relevance of such a transformation. Box and Cox (1964) originally suggested the use of more general power transformations, selected by maximum likelihood, and including the log-transform as a particular (limit) case. An early example of the types of problem associated with data transformations and forecasting was given in a case study presented and extensively discussed in Chatfield & Prothero (1973a, 1973b), Wilson (1973) and Box & Jenkins (1973). Chatfield & Prothero (1973a) originally selected a logarithmic transformation for some sales data, the use of which was challenged as over-transformation by both Wilson and Box & Jenkins (who each favoured a power transformation). In their reply, Chatfield & Prothero (1973b, p. 347) concluded that log transforms should be avoided in general and in fact they suggested that their (new) approach, "would always be to analyse the untransformed observations, except possibly in exceptional circumstances".

More recently, Makridakis & Hibon (1997) re-analysed the 1001 time series originally used in the M-Competition (Makridakis *et al.* (1982)). They found that logarithmic or power transformations (applied as appropriate, to achieve stationarity in the variance) did result in improvements in out-of-sample forecasts (as measured by MAPE). They noted that the improvements were small, but consistent, and also noted that their findings were in contrast to previous ones (including their own; see Makridakis & Hibon (1979)).

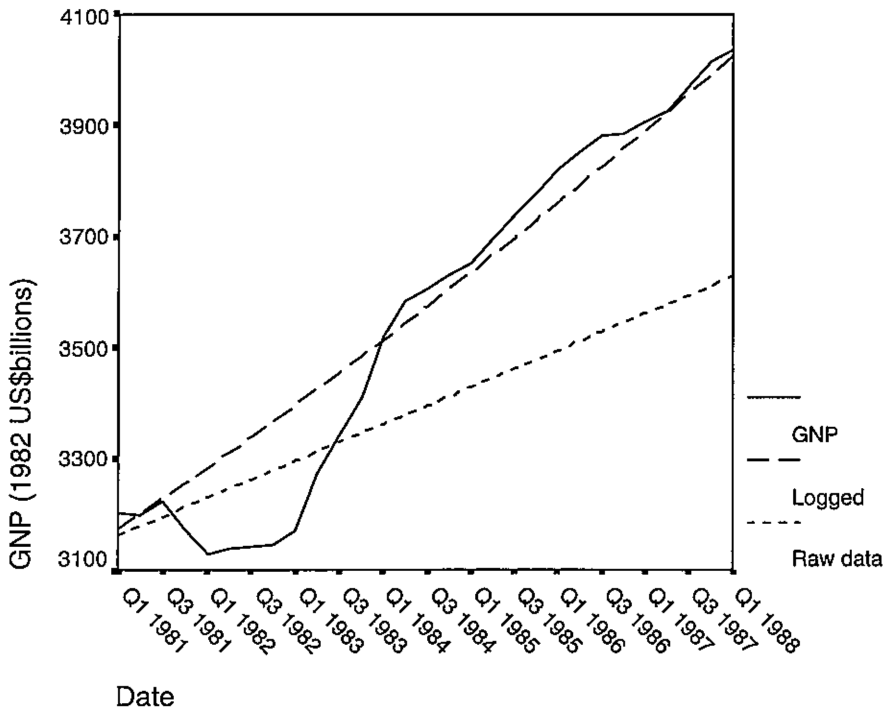
For the data analysed here, if attention is restricted to simple extrapolation procedures of the type discussed above, then the dichotomous choice between log-transformation and no transformation is relatively straightforward: the use of logarithms is clearly preferable for out-of-sample forecast accuracy reasons. Given the linearity of the forecast function used in EWMA and in Robust Trend, Figures 1 and 2 (which clearly depict nonlinear growth in both series) suggest that those methods will not be appropriate for either series without transformation. That is obviously confirmed, without the need for summary statistics, by Figure 8. There, 29 out-of-sample forecasts (1981Q1 onwards) are displayed for X_1, from an EWMA optimised on the raw data (by in-sample MSE), and from an EWMA optimised for the logged series (as already plotted in Figure 6). The tuned discount factor is unity in both cases and the more appropriate long-run trend in the forecasts from the logged data is very clear; a similar result is found for X_2.

4. Lessons from empirical evidence

The theoretical idea presented in AF, of bias correction for forecasts of levels made on log-transformed multivariate data, is certainly important. As they demonstrated, it can also improve the empirical performance of VAR models. However, as this paper shows, such improvements can be very much sample-dependent. More particularly, with the data considered here, by simply changing the forecast origin one can find examples where such bias correction makes forecasts uniformly worse. The validity of the cointegrated VAR estimated by AF for these data must be questioned by such findings. As demonstrated in Section 3, simple univariate procedures do in fact perform better out-of-sample than AF's VAR, even after bias correction. However, I now suggest that such a finding could have been anticipated.

Figure 8

Data with 29 out-of-sample forecasts from simple exponential smoothing with drift, optimised over both log-transformed and raw data: US real GNP, 1981Q1 to 1988Q1 inclusive



Simple procedures have repeatedly been shown to be at least comparable with more complex methods, as already noted in Section 3. Fildes and Makridakis (1995) pointed out that such empirical evidence has consistently been ignored, at least by academic forecasters. They argued that more attention should be paid to reliable empirical evidence, as provided (for example) by comparative forecasting accuracy studies. Fildes and Makridakis found, using classification by key words, that within-sample issues (such as model fitting and hypothesis testing) had accounted for the vast majority of articles with a time series focus published over a 21 year period (1971-1991 inclusive). They also noted that comparisons with other models (or methods) were very rare. So it is perhaps not surprising that, although AF did look at out-of-sample performance, they did not note the better performance of more simple methods. Yet such a result surely questions any conclusions drawn from their multivariate model.

Of course, the arguments for systems models in applied economics are overwhelming, since no one believes economic time series to be independent. The study of business cycles (the majority of which have been with US data, as analysed here) was historically concerned with identifying peaks and troughs in the level of variables that together reflect economic activity; see Harding & Pagan (1999a, 1999b) and additional references therein. Policy decisions concerning the business cycle inevitably involve many macroeconomic variables. However, if sensibly specified models that are designed to capture long-run movements can not forecast out-of-sample as well as the simplest univariate projection methods, how sensible can the models be? Improving the performance of an inappropriate model is not difficult, but such improvements are unlikely to be robust. Perhaps more attention needs to be devoted to different measures of model 'appropriateness'.

For example, Fildes *et al.* (1988) stressed the need for (out-of-sample) forecast comparisons based on multiple time origins. Such comparisons were absent from AF. Presumably though, AF simply wished to present an illustration of their new theoretical advance. However, the conclusions they drew from their empirical evidence were quite general, and yet clearly dependent on a particular time origin. Much of the improvement in forecast performance exhibited by the simple methods in Section 3 (over that of AF's model) was attributed to forecast functions that more closely follow the long run trend in the data. I conveniently 'ignore' the recession of 1981/2, which is indeed helpful, since simple procedures with linear forecast functions can not hope to follow such data movements out-of-sample. Yet the cointegrated VAR of AF does not pick up the recession (and subsequent recovery) either and that is not uncommon for multivariate methods, as Fildes & Stekler (1999) noted.

Harding & Pagan (1999a, 1999b) showed that a simple random walk with drift model, when combined with appropriate cycle-dating procedures, could simulate most of the features of the classical US (and for that matter UK and Australian) business cycles. They also showed that using nonlinear models of the Markov-switching type, as suggested by Hamilton (1989), did not generate business cycles that were 'more similar' to the (log) data, even though such models did produce cycles quite different to those from a random walk with drift; see Harding & Pagan (1999b). Interestingly, Harding & Pagan (1999b, p. 28) also noted that such nonlinear models were, "chosen over random walk models by statistical tests"; a finding that could be interpreted as a further warning against extensive reliance on in-sample statistics (*cf.* Fildes & Makridakis (1995)).

Chauvet & Potter (2000) concluded that the probability of a recession in the US business cycle has reduced dramatically since 1984. Using Gibbs sampling and a Bayesian analysis, Chauvet & Potter estimated the posterior median time to recession, post 1984, as 18 years. This compares with an estimated posterior median time to recession of 4.6 years prior to 1984, the date identified by

Chauvet & Potter as that of a structural break (specifically a reduction in volatility) in the common factor identified as driving US macroeconomic series.

Thus the success of the very simple forecasting procedures in Section 3 was perhaps to have been expected, since recent evidence suggests that time series models for which such procedures are optimal (random walk with drift) do in fact well approximate US macroeconomic series. Further, the frequency of recessions has been shown to have diminished markedly, so linear forecast functions may well prove even more appropriate with more recent data than that analysed here.

The success of any time series model in predicting the future obviously relies on the assumption of constancy; patterns identified in the past and accurately modelled, will continue into the future. While the date of the structural break identified by Chauvet & Potter (2000) was outside the sample of data used for estimation in this paper, there is other evidence of non-constancy before 1981Q1. For example, the sample autocorrelation (acf) and sample partial autocorrelation (pacf) functions of the first 17 years of differenced, log-transformed data (67 observations from 1947Q2 to 1963Q4) show the importance of the seasonal lag in both time series. For real Investment, the fourth lag is in fact the first that is significant in either the acf or pacf. However, repeating that exercise for the last 17 years prior to 1981Q1 (i.e. 1964Q2 to 1980Q4), the fourth lag is now insignificant in the acf and pacf, for both time series. Thus, particularly for economic data, procedures like exponentially weighted predictors, designed specifically to robustly track the underlying structure of a variety of time series, are likely to fare well in comparison to particular (selected) models whose success relies far more on the assumption of constancy.

In Section 3 it was shown that the use of logarithmic transformations with these data did lead to improved out-of-sample forecast accuracy, when compared to not transforming the data at all. Other (power) transformations were not considered though. However, such an omission is 'standard practice', especially in applied economics, where it is almost mandatory to take logs as a first step in any data analysis. There are good reasons for this, of course: most economic data are non-linear and can be thought of as consisting of components (for example, 'trend-cycle', 'seasonal', and 'irregular') which combine in a multiplicative fashion. That empirical fact is not surprising when one considers the multiplicative relationships in, for example, compound growth rates. However, the majority of time series models in widespread use are linear, so transformation via the use of logarithms, from a multiplicative to an additive combination of time series components, is desirable from a (linear) modelling point of view.

So the issue of bias in naive forecasts of levels, made by transforming forecasts made in logarithms, is still very relevant. However, the approaches to correcting that bias given in Granger & Newbold (1976) and in AF are dependent on a number of assumptions, as both papers readily noted. A crucial one is the normality of the innovations (univariate or multivariate, as appropriate). It is also necessary that the forecasts of the (log) transformed variable(s) are optimal in the sense of quadratic loss. Now it is obvious from the analysis in Section 3 that the forecasts from AF's cointegrated VAR were far from optimal for the logged data, since the naive VAR forecasts of X_1 and X_2 were inferior to those produced by simple procedures. That the bias corrected forecasts were also inferior to those from the simple procedures (themselves with or without bias correction) is then not too surprising, given that the correction used was not appropriate. Thus while theoretically relevant, the practical benefits of AF's work may well depend far more on an increased attention to out-of-sample errors at the time of model selection.

Although the use of logarithmic transformations (and therefore the need for bias corrections of forecasts) seems compelling, there are alternative approaches to modelling multiplicative data, and alternative approaches to bias correction, following transformation. For example, Ozaki & Thomson (2000) propose a nonlinear dynamic model for multiplicative seasonal time series. Ozaki & Thomson extend the X-11 paradigm in a parametric modelling direction, which explicitly avoids the use of logarithmic transformations, partly due to the inherent bias problem. In related work, Thomson & Ozaki (2000) present a range of bias correction procedures, appropriate for data that are to be both transformed and seasonally adjusted, before presentation on the original scale. Rather than restricting themselves solely to logarithmic transformations, Thomson & Ozaki consider a general class of power transformations of the type originally proposed by Box & Cox (1964).

5. Conclusions

In this paper I have re-analysed a bivariate US macroeconomic time series, originally used to demonstrate the need for bias corrections in the forecasts of levels of variables, modelled (and originally forecast) following log-transformation. I demonstrated that the claims previously made, concerning improvements in forecast accuracy following bias correction for these data, were not generally well founded. I also showed that, for these data, simple univariate forecasting procedures were more successful than a cointegrated multivariate model (with or without bias correction). It was noted that in the light of previous empirical work, such findings could have been expected. This paper further reinforces the call for an answer to why well-motivated theoretical advances in time series analysis often do not lead to noticeable improvements in out-of-sample forecast accuracy.

Acknowledgements

Partial financial support from the Victoria University of Wellington, FCA Research and Study Leave Grant: 7001-23077-7001-000 is gratefully acknowledged. This work was carried out while the author was visiting the School of Economic Studies, University of Manchester. Their hospitality, and in particular that of my academic host, Professor Denise Osborn, is very much appreciated.

References

Arino, M.A. & Franses, P.H. (2000). Forecasting the levels of vector autoregressive log-transformed time series, *International Journal of Forecasting*, **16**, 111-116.

Box, G.E.P. & Cox, D.R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society B*, **26**, 211-252.

Box, G.E.P. & Jenkins, G.M. (1973). Some comments on a paper by Chatfield and Prothero and on a review by Kendall, *Journal of the Royal Statistical Society A*, **136**, 337-345.

Chatfield, C. & Prothero, D.L. (1973a). Box-Jenkins seasonal forecasting: Problems in a case study (with discussion), *Journal of the Royal Statistical Society A*, **136**, 295-336.

Chatfield, C. & Prothero, D.L. (1973b). A reply to "Some comments on a paper by Chatfield and Prothero and on a review by Kendall", by G.E.P. Box & G.M. Jenkins, *Journal of the Royal Statistical Society A*, **136**, 345-352.

Chauvet, M. & Potter, S. (2000). Recent changes in the US business cycle. Conference paper given at: Growth and Business Cycles in Theory and Practice, University of Manchester, July 2000. Available from: http://www.ses.man.ac.uk/cgbcr/programm.htm

Engle, R.F. & Granger, C.W.J. (1987). Cointegration and error correction: Representation, estimation and testing, *Econometrica*, **55**, 251-276.

Fildes, R. & Makridakis, S. (1995). The impact of empirical accuracy studies on time series analysis and forecasting, *International Statistical Review*, **63**, 289-308.

Fildes, R., Hibon, M., Makridakis, S. & Meade, N. (1998). Generalising about univariate forecasting methods: Further empirical evidence, *International Journal of Forecasting*, **14**, 339-358.

Fildes, R. & Stekler, H. (1999). The state of macroeconomic forecasting. Discussion paper 99-04, Department of Management Science, Lancaster University. Available from: http://www.lums.lancs.ac.uk/mansci/Staff/fildes.htm

Granger, C.W.J. & Newbold, P. (1976). Forecasting transformed series, *Journal of the Royal Statistical Society B*, **38**, 189-203.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, **57**, 357-384.

Harding, D. & Pagan, A.R. (1999a). Knowing the cycle. Mimeo. Available from: http://econrsss.anu.edu.au/staff/adrian/

Harding, D. & Pagan, A.R. (1999b). Dissecting the cycle: A methodological investigation. Mimeo. Available from: http://econrsss.anu.edu.au/staff/adrian/

Haywood, J. & Tunnicliffe Wilson, G. (1997). Fitting time series models by minimizing multistep-ahead errors: A frequency domain approach, *Journal of the Royal Statistical Society B*, **59**, 237-254.

Makridakis, S. & Hibon, M. (1979). Accuracy of forecasting: An empirical investigation (with discussion), *Journal of the Royal Statistical Society A*, **142**, 97-145.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, P. & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition, *Journal of Forecasting*, **1**, 111-153.

Makridakis, S. & Hibon, M. (1997). ARMA models and the Box-Jenkins methodology, *Journal of Forecasting*, **16**, 147-163.

Ozaki, T. & Thomson, P. (2000). A nonlinear dynamic model for multiplicative seasonal-trend decomposition, *Journal of Forecasting*, in press. Available from: http://www.statsresearch.co.nz/

Pindyck, R.S. & Rubinfeld, D.L. (1991). *Econometric Models and Economic Forecasts (3rd ed.)*, New York: McGraw-Hill.

Sims, C.A. (1980). Macroeconomics and reality, *Econometrica*, **48**, 1-48.

Thomson, P. & Ozaki, T. (2000). Transformation and seasonal adjustment. Submitted for publication. Available from: http://www.statsresearch.co.nz/

Tiao, G.C. & Xu, T. (1993). Robustness of maximum likelihood estimates for multi-step predictions: The exponential smoothing case, *Biometrika*, **80**, 623-641.

Weiss, A.A. (1991). Multi-step estimation and forecasting in dynamic models, *Journal of Econometrics*, **48**, 135-149.

Wilson, G.T. (1973). Contribution to discussion of "Box-Jenkins seasonal forecasting: Problems in a case study", by C. Chatfield & D.L. Prothero, *Journal of the Royal Statistical Society A*, **136**, 315-319.

WORKING PAPER SERIES

The main purpose of this series is to reach a wide audience quickly for feedback on recently completed or in progress research. All papers are reviewed before publication.

A full catalogue with abstracts and details of other publications is available, for enquires and to be included in our distribution list, write to:

Angela Dolan
Manager Administration Services
GSBGM, Victoria University of Wellington,
PO Box 600, Wellington, New Zealand
Tel: (04) 463 5366; Fax: (04) 463 5454

Code in bold denotes order number, eg: WP 1/99

--- Group denotes the author's academic discipline Group (note this does not necessarily define the subject matter, as staff's interests may not be confined to the subjects they teach).

WP 1/95 **Management Group**
Gilbertson, D.K., Wright, H., Yska, G, Gilbertson, D.W. and 1994 Students of MGMT 306 'Kiwi entrepreneurs: A study.'

WP 2/95 **Management Group**
Cavana, R. 'Policy issues related to coastal and international shipping in New Zealand'
Shipping policy issues. *Transportant: The Journal of the Chartered Institute of Transport in New Zealand* 1995, Vol 25, No 2, 17-19.

WP 3/95 **Information Systems Group**
Bonner, Marcus 'On seeing information systems as bridges'

WP 4/95 **Management Group**
Cavana, Bob, Rob Crozier, Barrie Davis and Perumal Pillai 'A survey of academic staff attitudes towards the system of academic titles used in New Zealand universities'
Attitudes Towards Reclassifying Academic Titles in New Zealand Universities. *The Journal of Higher Education Policy and Management 1996, Vol 18, No 1.*

WP 5/95 **Econometrics Group**
Krawczyk, J.B. and G. Zaccour 'Pollution management through levies and subsidies'

WP 6/95 **Marketing Group**
Ashill, Nicholas and Malcolm Wright 'Marketing information systems - A review and reconceptulisation'

WP 7/95 **Information Systems Group**
Casey, Mary-Ellen 'An exploratory study into the use of information technology as an important enabler of organisational differentiation in the financial sector'

WP 8/95 **Economics Group**
Boles de Boer, David and Lewis Evans 'The economic efficiency of telecommunications in a deregulated market: the case of New Zealand'

WP 9/95 **Management Group**
Mabin, Victoria J. 'Using spreadsheet optimisation facilities as a decision aid within the theory of constraints framework'

WP 10/95 **Economics Group**
M. Khaled, M.D. Adams and M. Pickford 'Estimates of scale and scope economies in the New Zealand life insurance industry.'

WP 11/95 **Economics Group**
John A. Carlson and Robert A. Buckle 'Price duration with two-sided pricing rules'
In Karl Heinrich Oppenlaender and Guenter Poser (Eds.) *Business Cycle Surveys: Forecasting Issues and Methodological Aspects.* Aldershot, Avebury. 1996: 101-118.

WP 12/95 **Economics Group**

Ganesh Nana 'Developing a multi-sectoral CGE model of the New Zeland economy.'

WP 13/95 **Money and Finance Group and Economics Group**

Stephen Burnell, Lewis Evans and Shuntian Yao 'Network games: The optimal network contract and the efficiency of bypass in oligopolistic network industries under light regulation'

WP 14/95 **Economic History Group**

Gordon Boyce 'The Nickel Syndicate, 1901 - 1939'

WP 15/95 **Money and Finance Group**

Jan Whitwell 'Monetary disinflation with inflation inertia: Central bank autonomy in an open economy'

WP 16/95 **Economics Group**

Emery, H. Daniel V. Gordon and Doug McClintock 'On the efficacy of construction site safety inspections.'

WP 17/95 **Economics Group**

Ganesh Nana 'An inter-temporal CGE model with fiscal and sector balances'

WP 18/95 **Economics Group**

Jie Zhang 'Government debt, human capital, and endogenous growth'

WP 19/95 **Accountancy Group**

Zahirul Hoque and Manzurul Alam 'Quality management and accounting in a New Zealand service organisation: Towards an institutional perspective on management accounting'

WP 20/95 **Economics Group**

Paul Calcott 'Can we test for supplier-induced demand by comparing informed with uninformed consumers?'

WP 1/96 **Management Group**

M.W. Lee, J. Bennett, R.J. Taylor and R.Y. Cavana 'A dynamic simulation model for possum and gorse control on a farm woodlot.'

> R.Y.Cavana, M.W.Lee, J.Bennett & R.J.Taylor. 1996. Possum and gorse control on a farm woodlot in New Zealand: A system dynamics analysis. *Asia-Pacific Journal of Operational Research*, Vol 13, No. 2, pp. 181-207.

WP 2/96 **Economics Group**

Jie Zhang 'Optimal public investments in education, and endogenous growth'

WP 3/96 **Economics Group**

Paul Tompkinson 'The benefits of tariff reductions in the presence of psychological adjustments costs.'

WP 4/96 **Economics Group**

Shuntian Yao 'A note on the decision of a sales maximizer in response to the increase of per unit cost.'

WP5/96 **Economics Group**

Karen Palmer and Margaret Walls 'Optimal policies for solid waste disposal: taxes, subsidies and standards'

WP6/96 **Economics Group**

Margaret Walls and Jean Hanson 'Distributional impacts of an environmental tax shift: the case of motor vehicle emission taxes.'

WP7/96 **Economics Group**

Lloyd S Shapley and Shuntian Yao 'Dynamic properties of the Nash equilibrium.'

WP 8/96 **Accountancy Group**

Kamran Ahmed, Kazi Feroz Alam and Manzurul Alam 'An empirical study of factors affecting accounting students' career choice in New Zealand'

WP9/96 **Accountancy Group**

Andrew MC Smith and Paul V Dunmore 'Tax Avoidance and the Financial Structures of Non -Resident Controlled Companies in New Zealand'

WP10/96 **Econometrics Group**

L Fraser Jackson. Relative Prices and Inflation.'

WP11/96 **Public Policy Group**

Robert Gregory. 'Reserve Bank independence, political responsibility, and the goals of anti-democratic policy: A political 'Cri de Coeur' in response to an economist's perspective.

WP12/96 **Economics Group**

Robert A. Buckle and John A. Carlson. 'Inflation and asymmetric price adjustment.'

WP13/96 **Econometrics Group**

J.B. Krawczyk, O. Pourtallier and M. Tidball. 'Modelling and solution to the municipal effluent management problem.

WP14/96 **Economics Group**

Viv B. Hall and David Rae. 'Fiscal expansion, interest rate risk premia, and wage reactions: Some illustrative simulations with NBNZ-DEMONZ.'

WP15/96 **Economics Group**

Viv. B. Hall, Kunhong Kim and Robert A. Buckle. 'Pacific rim business cycle analysis: Dating, volatility and synchronisation.'

WP16/96 **Management Group**

S. Davenport, C. Grimes and J. Davies. 'Collaboration and organisational learning: A study of a New Zealand collaborative research programme.'

WP1/97 **Information Systems Group**

Marcus Bonner and Mary-Ellen Casey. 'The informated individual - a consideration of importance to the business organisation.'

WP2/97 **Management Group**

R.Y. Cavana, I.G. Harrison, F.E.B. Heffernan and C.C. Kissling. 'Freight transport industry in New Zealand.'

WP3/97 **School of Economics and Finance**

Steffan Berridge and Jacek B Krawczyk. 'Relaxation Algorithams in finding Nash Equilibria.'

WP 4/97 **School of Economics and Finance**

Margaret Walls and Karen Palmer. 'Upstream Pollution, Downstream Waste Disposal, and the Design of Comprehensive Environmental Policies'

WP 1/98 **School of Economics and Finance**

John Haywood and Granville Tunnicliffe Wilson. 'Selection and Estimation of Trigonometric Component Models for Seasonal Time Series'

WP 2/98 **School of Economics and Finance**

Roger J Bowden, 'CAPM and Empirical Embedding: When is 'near enough', good enough?'

WP 3/98 **School of Economics and Finance**

Roger J Bowden, 'The Discrete Time Beta'

WP 4/98 **School of Economics and Finance**

Viv B Hall, 'Assessing Structural Adjustment and Economic Reform: the Case of New Zealand'

WP 5/98 **School of Economics and Finance**

Jacques Poot and Jacques J Siegers, 'The Macroeconomics of Fertility in Small Open Economies: A test of the Becker-Barro Model for the Netherlands and New Zealand'

WP 6/98 **School of Business and Public Management**

Nicholas J. Lee and John W. Cadogan, "Factors Influencing Sales Manager Effectiveness: Preliminary Findings and Future Research Directions"

WP 7/98 **School of Business and Public Management**

Nicholas J. Ashill, Lisa J. Casagranda and Peter M. Stevens, "Creating Competetive Advantage Using the Internet in Primary Sector Industries"

WP 1/99 **School of Business and Public Management**

Robert Y. Cavana, Paul J. Carroll, Mark B. Lockhart and Ronald J. Peake, "Hawkes Bay Apple and Pear Growers' Attitudes towards the proposed Deregulation of the Pipfruit Industry in New Zealand"

WP 2/99 **School of Economics and Finance**

Olivier Jeanne and Andrew K. Rose, "Noise Trading and Exchange Rate Regimes"

WP 3/99 **School of Business and Public Management**

Robert Y. Cavana and Leslie V.Clifford, "The Matrix, the Spiderweb and the Influence Diagram: Developments in Systems Thinking at the New Zealand Customs Service".

WP 4/99 **School of Economics and Finance**

Jacques Poot, 'Reflections on Local and Economy-wide Effects of Territorial Competition'

WP 5/99 **School of Economics and Finance**

John Haywood and Granville Tunnicliffe Wilson, 'An improved state space representation for cyclical time series'

WP6/99 **School of Business and Public Management**

Nicholas J. Ashill and David Jobber, 'A Conceptual Framework of MkIS Design: The Impact of Environmental Uncertainty Perceptions, Decision-Maker Characteristics and Work Environment Factors on the Perceived Usefulness of Marketing Information Characteristics'

WP 1/00 **School of Economics and Finance**

Philip S. Morrison, Kerry L. Papps and Jacques Poot, 'Local Labour Markets, Migration and Wage Determination: Theory and Evidence for the Wage Curve in New Zealand'

WP 2/00 **School of Business and Public Management**

Robert Stephens, Paul Frater and Charles Waldegrave, 'Below the Line: an Analysis of Income Poverty in New Zealand, 1984 – 1998'

WP 3/00 **Graduate School of Business and Government Management**

Paul M. Dickie, 'Corporate Debt/Equity Ratios and the Korean Financial Crisis'

WP4/00 **School of Business and Public Management**

Nicholas J. Ashill, John Davies and Anthony Joe, 'Sponsorship Scale Development: A Study of a National Sports Event in New Zealand'.

WP5/00 **School of Economics and Finance**

John Haywood, 'On log-transformations, vector autoregressions and empirical evidence'