REAL -TIME OR CURRENT VINTAGE: DOES THE TYPE OF DATA MATTER FOR FORECASTING AND MODEL SELECTION?

Hui Feng

Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

August, 2005

Abstract

In this paper we investigate the impact of data revisions on forecasting and model selection procedures. A linear ARMA model and nonlinear SETAR model are considered in this study. Two Canadian macroeconomic time series have been analyzed: the real-time monetary aggregate M3 (1977-2000), and residential mortgage credit (1975-1998). The forecasting method we use is multi-step-ahead non-adaptive forecasting.

Author Contact:

Hui Feng, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: hui@uvic.ca; Phone: (250) 472-4418; FAX: (250) 721-6214

1. Introduction

When economists are trying to build a good economic model, their aim is not only to gain a good fit in terms of within-sample performance, but also to get highly accurate out-of-sample forecasts. However, accurate out-of-sample forecasting is heavily affected by the forecasting model selected, and the data that are used to fit the model are also essential to the quality of the forecasting results. Economic data such as real output, the money stock and aggregate consumption, which are routinely published by official government agencies, are subject to revisions and redefinitions as new information becomes available. The information set available at a particular date is called a 'vintage'. Real-time data specifically refer to the information set available to economic researchers who are making forecast in real-time. Due to the fact that information may creep into forecasts of current vintage data, the forecasting results for real-time data and current vintage data are likely to be different to some degree.

In this paper, to further explore the problem of data revisions, we focus our attention on two issues: the impact of data revisions on a model's forecasting performance; and the impact on the problem of choosing between the linear Autoregressive Moving Average (ARMA) model and the nonlinear Self-Excited Threshold Autoregressive (SETAR) model. The rest of the paper is organized as follows: Section 2 provides some insights into the history of this data revision problem. In Section 3, we discuss the ARMA and SETAR models. The forecasting techniques considered in this paper are discussed in Section 4. Section 5 deals with the comparison between the forecast performance of real-time and current vintage models, and some robustness tests are carried out in Section 6. Section 7 concludes for the paper.

## 2. Developments in Data Revisions

The early research in this field focused on the use of preliminary rather than final data. Cole (1969) provided evidence that certain types of forecasts are strongly affected by data revisions. She found that "the use of preliminary rather than revised data resulted in a doubling of the forecast error." Diebold and Rudebusch (1991) looked at the importance of data revisions for the index of leading indicators. They suggested that the use of real-time data was crucial, as the variables included in the index of leading indicators were chosen *ex post*. In real-time, the leading indicators neither indicate, nor lead!

1

Many other studies have been undertaken on the consequences of data revisions. Denton and Kuiper (1965) agreed with Cole's conclusion that the use of the preliminary rather than final data leads to large forecast errors, but Trivellato and Rettore (1986) found effects that were rather more modest. Croushore and Stark (2001) suggested that, when evaluated over long periods, forecast-error statistics are not sensitive to the distinction between real-time and currently available data, even though the forecasts for isolated periods can diverge. Croushore and Stark (2000) also found that standard measures of forecast accuracy can be vintage-sensitive when constructed from short spans of data, as are sometimes used by researchers for forecast evaluation.

Another problem relating to the data revision issue is model selection. Feng and Liu (2003) studied the difference between the standard linear ARMA model and the widely used nonlinear SETAR model using current vintage data, and found that for one-step non-adaptive recursive forecasts, the SETAR model is preferred to the ARMA model. However, for multi-step non-adaptive recursive forecasting, there is no significant difference between the two. It would be interesting to see if this result still holds if we use real-time data.

Hence, two aspects of data revisions will be analyzed in this paper. First, we consider the impact of using real-time data, as opposed to current vintage data, on forecasting accuracy and model specification within a family of ARMA and SETAR models respectively. Second, we evaluate the impact of using real-time data on model selection associated with choosing between the ARMA and SETAR models. Monetary aggregates series are often revised to adjust for various factor such as the inflation rate of each year. Figure 1 plots the Canadian monetary aggregate M3 at the first quarter of 1989 reported from 1990 to 2002. The substantial changes in the value of M3 in 1989Q1 illustrates the impact that revisions can have on the raw data. The question is how these changes impact on the use of the data.. Two quarterly Canadian series are used in this study: the monetary aggregate M3 (1977-1998), and residential mortgage credit (1975-1998). For forecasting, given that the two time series are quarterly, the four-step-ahead non-adaptive rolling forecasting scheme will be used.

Figure 1. Canadian M3 for 1989Q1 Reported from year 1990 to 2002
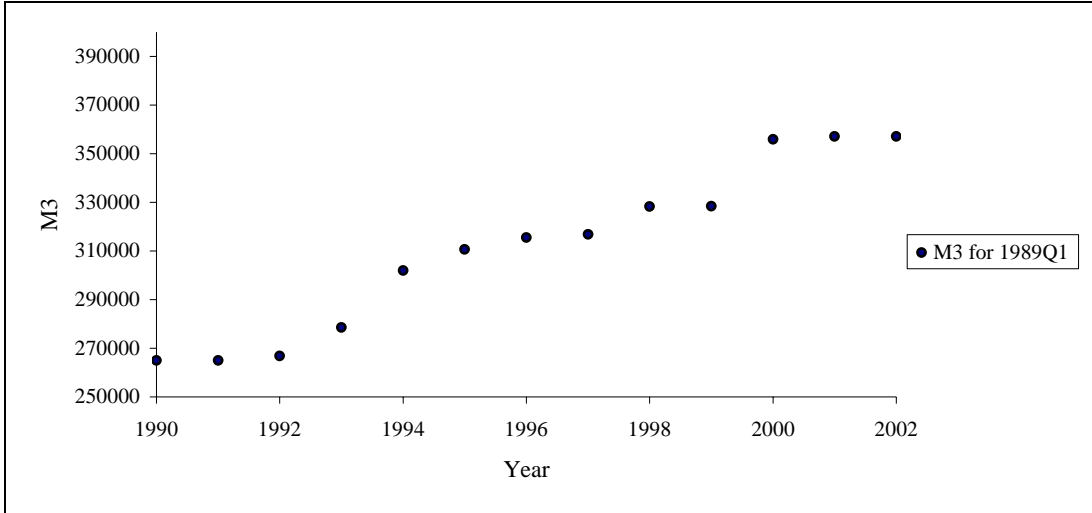
3. Introduction of the Two Models: ARMA and SETAR

3.1 SETAR (Self-Excited Threshold Autoregressive)

The SETAR model belongs to the family of regime switching models and was first proposed by Tong (1978). It assumes that the regime that occurs at time *t* can be determined by the lagged value of the time series itself relative to a threshold value, which we denote as *c*. The particular setting enables the model to be linear within a regime, and also liable to moving between regimes as the process crosses the threshold: see Tong (1983, 1990) and Hansen (1997, 2000). In our case, we assume that there are only two regimes, and the process is a p_i-th order autoregression in the *i*-th regime, so we denote this by writing SETAR $(2; p_1; p_2)$, and algebraically the model is:

$$ y_t = \left(\phi_{0,1} + \phi_{1,1} y_{t-1} + \cdots + \phi_{p,1} y_{t-p_1} \right) I\left[y_{t-d} \leq c \right] + \left(\phi_{0,2} + \phi_{1,2} y_{t-1} + \cdots + \phi_{p,2} y_{t-p_2} \right) I\left[y_{t-d} > c \right] + \varepsilon_t $$

(1)

where *I* [A] is an indicator function with *I* [A] = 1 if the event A occurs and *I* [A] = 0 otherwise. In particular, in the 2-regime SETAR model, y_t will lie within the first regime if the value of y_{t-d} is smaller than or equal to the threshold variable *c*, otherwise, y_t will lie within the second regime.

More compactly, this SETAR model can be rewritten as,

$$y_t = \phi_1' x_t I[y_{t-1} \le c] + \phi_2' x_t I[y_{t-1} > c] + \varepsilon_t \qquad (2)$$

where $\phi_j = (\phi_{0,j}, \phi_{1,j}, \cdots, \phi_{p_j,j})'$, $j = 1,2$, $x_t = (1, y_{t-1}, \cdots, y_{t-p_j})'$ and ε_t is well-behaved normally distributed error term. In the context of SETAR models we restrict the candidate threshold variable to be the lagged endogenous variable y_{t-d}, for some positive integer d. The estimators of the nuisance parameters, namely the threshold value c and the lag-length d, can be determined by minimizing the residual variance, or optimizing some other such criterion, among a reasonable range of choices of c and d (say, $c \in C$ and $d \in \{1, \cdots, d*\}$ for some upper bound $d*$). The set of allowable threshold values, C, should be such that each regime contains enough observations for the estimator defined above to produce reliable estimates of the autoregressive parameters. A popular choice of C is to allow each regime to have at least a fraction π of the observations, that is,

$$C = \left\{ c \,\middle|\, y_{([\pi(n-1)])} \le c \le y_{([(1-\pi)(n-1)])} \right\} \qquad (3)$$

A safe choice for this fraction appears to be 0.15 (Franses and van Dijk, 2000). For an optimization criterion we follow standard practice and use Akaike's Information Criterion (AIC) (Akaike, 1974). Then, identification of the appropriate threshold value, c, and lag order, d, can be chosen from the model that minimizes the AIC value. Tong (1990) defines the AIC for a 2-regime SETAR model as the sum of the AICs for the AR models in the two regimes, which is:

$$AIC(p_1, p_2) = n_1 \ln \hat{\sigma}_1^2 + n_2 \ln \hat{\sigma}_2^2 + 2(p_1 + 1) + 2(p_2 + 1) \qquad (4)$$

where, $\hat{\sigma}_j^2$ $j = 1,2$, is the variance of the residuals in the j-th regime and in our case, we set $p = p_1 = p_2$ (Feng and Liu, 2003).

3.2 ARMA (Autoregressive Moving Average)

Generally, an ARMA (p,q) model takes the following form:

$$y_t = c + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + ... + \gamma_p y_{t-p} + \varepsilon_1 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ... - \theta_q \varepsilon_{t-q} \qquad (5)$$

where $\varepsilon_t \sim$ iid N($0, \sigma^2$).

The values of the *p* and *q* are determined by using the General to Specific method. According to the characteristics of the data, and also the sample size, we could pre-assign the maximum values for *p* and *q*. Gradually, the two values are then reduced in a systematic way. The AIC values from all possible combinations of *p* and *q* are to be compared in order to get the model's correct specification. The AIC is defined as:

$$\text{AIC} = \text{LOG}(\hat{\varepsilon}^2) + 2 * (p+1) / T \qquad (6)$$

where $\hat{\varepsilon}^2$ is the sum of the squared residuals, ($p+1$) is the number of regressors, and *T* is the sample size.

4. Out-of-Sample Forecasts

In this paper, we consider four-step-ahead non-adaptive rolling forecasts. Non-adaptive rolling forecasting means that, for every forecast stage, the model specification will not be changed and the sample size will be fixed.

There is some difference between real-time forecasts and current vintage forecasts: for each real-time forecast, we have to change the data set completely, whenever the data have been revised. For instance, if we want to forecast M3 for the four quarters of the year 1999, we have to use the data that were first made available in the first quarter of 2000 to construct the out-of-sample forecasts, given that there is always a one-quarter-delay in the reporting of the data by the Bank of Canada. Without changing the model specification, we carry on to forecast four quarters of the year 2000. This time the data that were released in the first quarter of the year 2001 are used to fit the model.

To compare the various forecasting results, we use the Mean Absolute Percentage Error (MAPE) and the Root Mean Squared Percentage Error (RMSPE):

$$\text{MAPE} = [1/m \sum_{i=1}^{m} abs((\hat{y}_t - y_t)/y_t)]*100 \tag{7}$$

$$\text{RMSPE} = \sqrt{1/m \sum_{i=1}^{m} [(\hat{y}_t - y_t)/y_t]^2} *100 \tag{8}$$

where *m* is the number of forecast periods. Both criteria can be used to compare nested and non-nested models. The S-test for model selection (Diebold and Mariano, 1995) will be used as well. This test is used to test the hypothesis that there is no difference between the true forecasting performances of two models. The test statistic is:

$$S = (\frac{m}{4})^{-\frac{1}{2}} \left[\sum_{i=1}^{12} d_i - \frac{m}{2} \right] \tag{9}$$

where *m* is the number of forecast periods; $d_i = 1$ if the squared prediction error for the SETAR model exceeds that for the ARMA model, and is zero otherwise. Under the null hypothesis of equal performances, *S* is asymptotically standard normal.

5. Comparison of Real-time and Current Vintage Models and Model Selection

The quarterly Canadian aggregate money supply M3 and mortgage credit data are obtained from the Bank of Canada *Review*. Due to the limited degrees of freedom, in this paper both the ARMA and SETAR models are assigned a maximum lag length of 6. We will use the Canadian M3 (1977-1998) to carry out the usual test about the effect of the data problem on the models' within-sample fit as well as forecasting performance. In section 6, we will use the same M3 series but with a different sample (1975-1996) as well as the Canadian residential mortgage credit series to test for the results' robustness.

We use the levels of the time-series to fit the models. The unit root in the data has been neglected. This is because recent research has suggested that unit root tests have low power in discriminating against trend stationarity and other alternatives in typical macro-econometric contexts (Christiano and Eichenbaum, 1990; Sowell, 1992). In particular, the power of such tests can be dramatically diminished and a unit root mistakenly identified in the presence of threshold

determined regime switching. The data are seasonally adjusted, but we find that both models achieve a better fit by including seasonal factors.

5.1 Forecasts from the ARMA Model

The general ARMA (6,6) model is:

$$y_t = c + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + ... + \gamma_6 y_{t-6} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - ... - \theta_6 \varepsilon_{t-6} \; . \tag{10}$$

As well as the autoregressive and moving average processes, we also consider seasonal autoregressive and moving average factors in order to obtain a fair comparison with the SETAR model. The general terms for seasonal autoregressive and moving average are added from (4,4) to (3,3), to (2,2) and also (1,1) for convenience. We use the AIC to perform model selection, as shown below in Table 1, with the ten smallest AIC values for the real-time and current vintage M3 series:

Table 1.Model Selection for M3 for Real-Time and Current Vintage ARMA Model

Real-time		Current Vintage	
Model specification	AIC	Model specification	AIC
(6,2,4,4)*	22.119	(1,5,0,4,4)**	15.529
(1,5,4,4)	22.235	(1,4,0,2,2)	16.556
(3,2,4,4)	22.277	(4,4,0,3,3)	16.569
(4,2,3,3)	22.318	(1,6,0,3,3)	16.576
(1,2,4,4)	22.328	(4,6,0,3,3)	16.583
(1,2,4,4)	22.328	(2,5,0,3,3)	16.593
(1,2,3,3)	22.337	(1,6,0,1,1)	16.611
(5,4,4,4)	22.351	(3,5,0,1,1)	16.615
(3,6,4,4)	22.356	(1,2,0,4,4)	16.629
(2,2,4,4)	22.369	(6,1,0,4,4)	16.635

*(6,2,0,4,4) denotes model specification for the ARMA is AR=6, MA=2, SAR=4, SMA=4
**(1,5,0,4,4) denotes model specification for the ARMA is AR=1, MA=5, SAR=4, SMA=4

As a result, the best ARMA model for real-time M3 is ARMA (6,2,4,4), and it is ARMA (1,5,4,4) for the current vintage M3 series. Table 2 provides the within-sample MAPE and RMSPE for the

7

two preferred ARMA models. For later use, we also report the results for the SETAR model within-sample comparison between current vintage and real-time data.

Table 2. Within-Sample MAPE & RMSPE: Real-time vs. Current Vintage

	ARMA		SETAR	
	MAPE	RMSPE	MAPE	RMSPE
Real-time	0.8399	1.1416	0.592	1.057
Current Vintage	0.813	1.0414	0.741	1.029

Both criteria indicate that there is no significant difference between real-time and current vintage ARMA models for within sample fit. However, the data revisions alter the model specification to some degree.

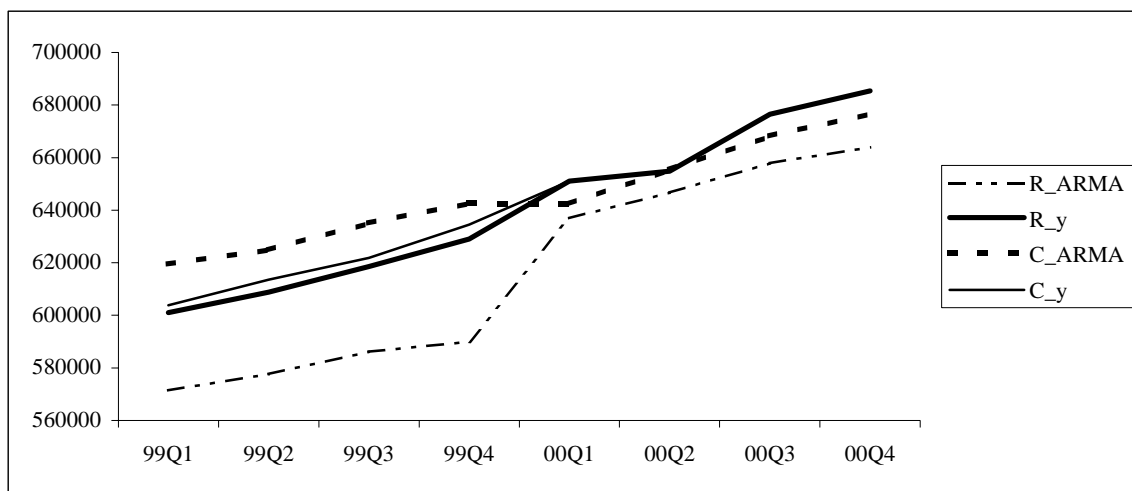Figure 2. Out-of-Sample Forecasting Comparison: Real-time vs. Current Vintage

Figure 2 shows out-of-sample forecasts for the period from the first quarter of 1999 to the fourth quarter of 2000 models. For later use, the results for the SETAR out-of-sample forecast are also reported for the current vintage and real-time series. The message from the data is clear: the structural change between 1999Q4 to 2000Q1 is captured well by the real-time model. However, the current vintage ARMA model forecasts better than the real-time forecast model. Another finding to note is that in this case, the real time ARMA model always under-forecasts M3, but the forecasts from the current vintage ARMA model are mixed: for the first four periods, it over-

8

forecasts; for the second four periods, it under-forecasts most of the time. Finally, for every forecasting stage, the current vintage forecast is closer to the realized value than is the corresponding real-time forecast. In the left section of Table 3, we find that the current vintage model is better than the real-time model under the criteria of MAPE and RMSPE:

Table 3. Forecasting MAPE and RMSPE Comparison: Real-time vs. Current Vintage

		ARMA		SETAR	
Criteria	Period	Real-time	Current Vintage	Real-time	Current Vintage
MAPE	Stage 1	5.377	1.965	5.118	0.545
	Stage 2	2.324	0.979	4.319	2.983
	Over all	3.85	1.472	4.718	1.764
RMSPE	Stage 1	5.4	2.021	5.244	0.636
	Stage 2	2.429	1.103	4.461	3.231
	Over all	3.915	1.562	4.852	1.934

At this stage it is possible to reach some conclusions based on the ARMA model: the current vintage model is better than the real-time model in terms of forecasting. More specifically, the current vintage model forecasts better than the real-time model during every forecasting period. Notice that in both cases, the ARMA model tends to improve its forecasting accuracy as we move forward in time. We now wish to see whether the non-linear SETAR model would support the same conclusion or not.

5.2 Forecasts from the SETAR Model

The SETAR model is a typical regime-switching model, and we assume two regimes throughout this research. In order to achieve a fair comparison, we modify the SETAR model from the general case of 6 lags to the lower order cases, using the levels of the data. In addition, the critical value for the model to decide which regime the data are in also starts from the 6th lag. We also consider seasonal factors in this case. As a result, seasonal dummy variables have been added to both the real-time and current vintage models. Table 4 reports the AIC values for the real-time SETAR model.

Table 4. Model Selection for M3 for Real-Time SETAR Model

c\AR	1	2	3	4	5	6
1	21416040	22026940	24441600	20856250	**19980600**	23867080
2	22112480	22925050	24715850	25524210	25233860	24564990
3	25023730	25999180	26382770	28550350	29766280	29744550
4	27927560	28394840	29982370	27101820	29327780	27132230
5	27440990	28148760	29612430	28141240	29750180	31297700
6	26963010	27624210	29014930	27285370	28301080	28830500

Under the model selection criterion of minimizing the AIC value, we chose the model SETAR (2,5,1). That is, there are two regimes, five autoregressive terms, and the first lagged value of the series will be used to compare against the threshold value to determine the regime, in the preferred model specification. We carried out exactly the same procedure for the current vintage SETAR model and found that SETAR (2,1,3) is preferred. Table 5 provides AIC report for the current vintage M3 model:

Table 5. Model Selection for M3 for Current Vintage SETAR Model

c\AR	1	2	3	4	5	6
1	28102780	29685720	30279040	29773590	**32232040**	33270260
2	26788750	28306780	29613560	29586830	30751080	33784820
3	25666300	27150760	28854790	31596120	31395310	32879020
4	27071670	28486710	30272850	30091940	30337280	33326110
5	27456960	28629930	30380860	30873840	30834490	32144900
6	28084270	28537000	29988930	30619480	31583350	32657920

From the second section of Table 2, we notice that, in terms of the impact of the data, there is little difference between the non-linear SETAR model and the linear ARMA model. In this case, the MAPE shows that the real-time model is better than the current vintage model, but the RMSPE shows that there is little difference between the two. Over all, the within sample fit is similar for the models. Another point to note is that the data revisions change the preferred model from the SETAR (2,5,1) to SETAR (2,3,1) from the real-time case to the current vintage case, which is the same as the ARMA model case. Figure 3 shows the results for two four-step-ahead forecasts for M3. We find similar results compared with the ARMA model— the real-time

forecasts have a similar *pattern* to the realized values, but the current vintage forecasts are *closer* to the realized values.

Figure 3. Out-of-Sample Forecasting Comparison: Real-time vs. Current Vintage
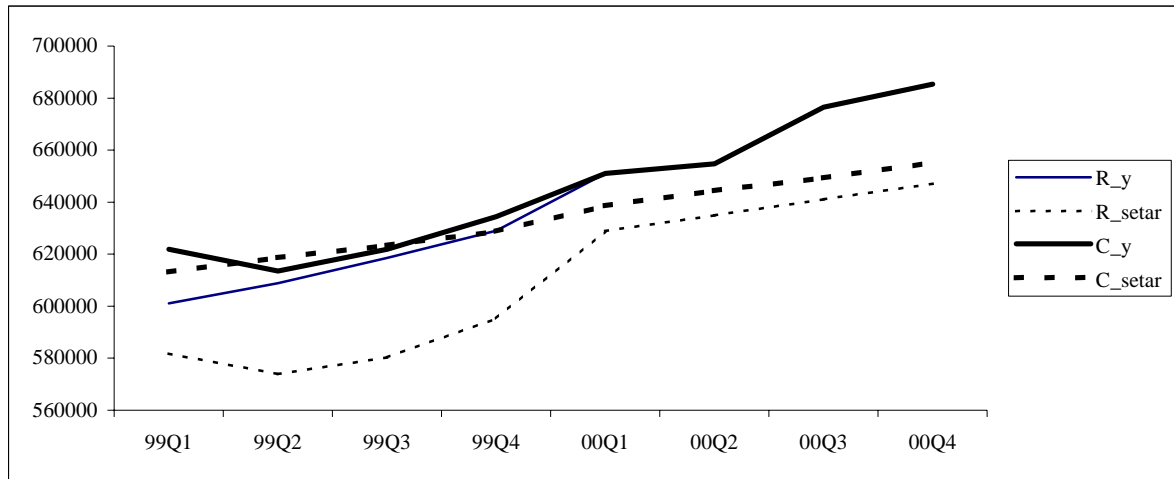
The right section of Table 3 reports the forecasts of the SETAR model for the real-time case and current-vintage case, which coincides with that of the ARMA model: both MAPE and RMSPE support the current vintage SETAR model. Furthermore, the real-time SETAR model tends to improve its forecasting accuracy as we roll four periods ahead, but this is not true for the current vintage SETAR.

5.3 Comparison between the ARMA Model and SETAR Model

In this section, we compare the within-sample performance and out-of-sample forecasts across the two models. For current vintage data, we find that the preferred model specifications are ARMA (1,5,4,4) and SETAR (2,1,3). For real-time data the preferred model specifications are ARMA (6,2,4,4) and SETAR (2,5,1). The results in Tables 6 and 7 show that the forecasts from the ARMA model tend to improve as the forecasting period increases, but this is not true for the current vintage SETAR model, which is consistent with arguments advanced in Feng and Liu (2003). For the first four forecasting periods, in the current vintage case, the SETAR model is better than the ARMA model, but the results are mixed for the real time case - in the last four

11

forecasting periods, the ARMA model is better than the SETAR model in both cases, which can clearly be seen in Figures 4 and 5.

Table 6. Forecasting Percentage Error for the ARMA and SETAR Models: M3

	Current Vintage		Real-time	
Period	ARMA	SETAR	ARMA	SETAR
99Q1	2.591	0.403	-4.935	-3.2
99Q2	1.85	0.806	-5.119	-5.741
99Q3	2.129	0.075	-5.233	-6.176
99Q4	1.29	0.896	-6.22	-5.352
00Q1	-1.328	-1.918	-2.173	-3.414
00Q2	0.102	-1.593	-1.244	-3.028
00Q3	-1.222	-4.028	-2.758	-5.251
00Q4	-1.264	-4.395	-3.12	-5.584

** The first 4 forecasts belong to the first forecast stage where we use the same set of data to forecast, while the second 4 forecasts belong to the second forecast stage where we roll the real-time data sample forward one year.

Table 7. Forecasting MAPE and RMSPE for the SETAR and ARMA models: M3

		Current Vintage		Real-time	
		MAPE	RMSPE	MAPE	RMSPE
ARMA	Stage 1	1.965	2.021	5.377	5.400
	Stage 2	0.979	1.103	2.324	2.429
	Sum	1.472	1.628	3.850	4.187
SETAR	Stage 1	0.545	0.636	5.118	5.244
	Stage 2	2.983	3.231	4.319	4.461
	Sum	1.764	1.934	4.718	4.852

Figure 4. Real-Time Forecast Comparison between ARMA and SETAR
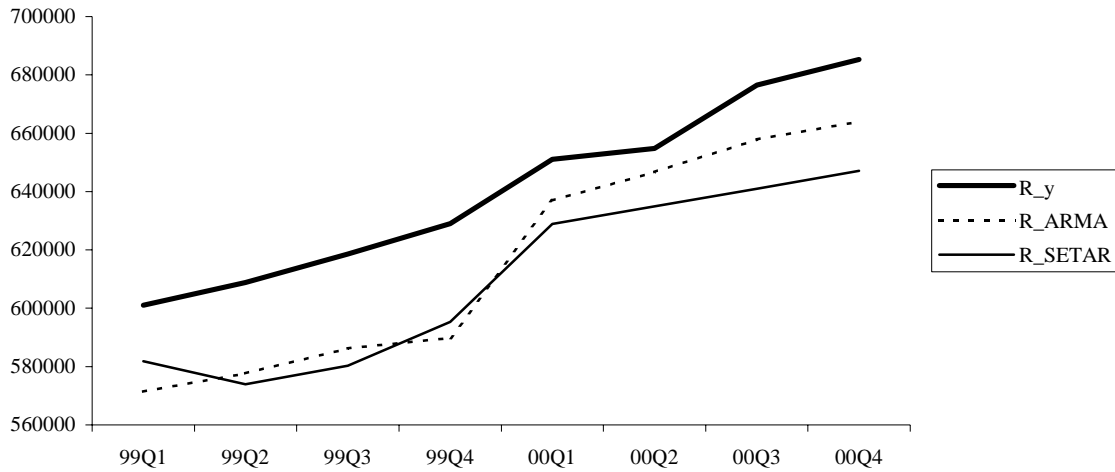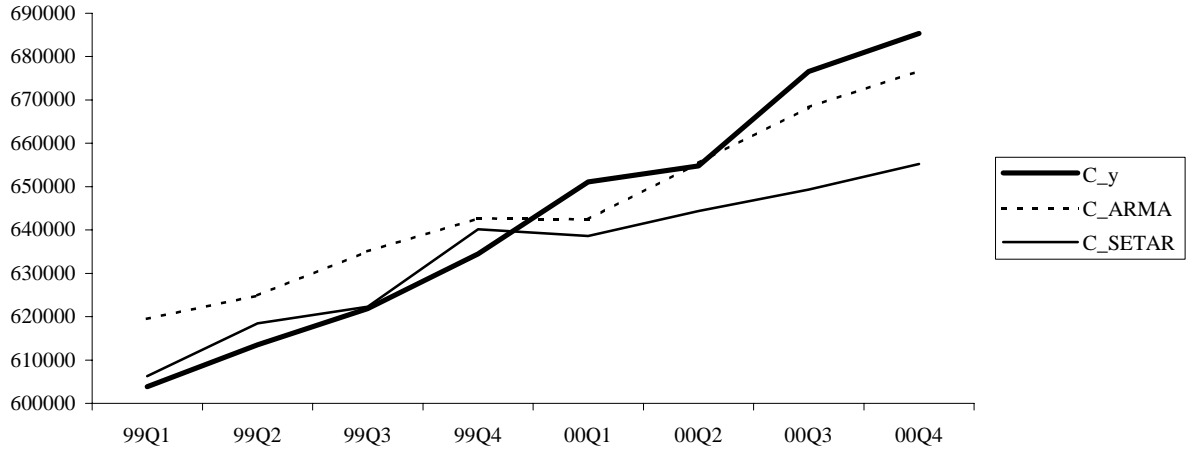

Figure 5. Current Vintage Forecast Comparison between ARMA and SETAR

Here we are interested in a general comparison. Both MAPE and RMSPE suggest that the ARMA model forecasts better than the SETAR model in real-time and with current vintage data. Over all of the eight forecasting periods, we apply the S-test with the null hypothesis that the forecasting ability for the SETAR model and the ARMA models are equal. For the current vintage case the S-statistic is zero, indicating support for the null hypothesis, and again corroborating the results of

1

Feng and Liu (2003). In the case of the real-time model the S-statistic is –1.414. Thus we cannot reject the hypothesis that the two models are equal in terms of forecast performance, at the 5% level of significance.

To sum up, in terms of forecasting performance over all forecasting periods, the SETAR model essentially performs as well as the ARMA model in both real-time and current vintage, even though both the MAPE and RMSPE are slightly in favor of the ARMA model. However, we need to check the robustness of this conclusion against other data.

6. Robustness Analysis

Often, results can be highly sensitive to sample features. We focus on two types of robustness tests here. First, we keep the same series in the model, but fit in a different time period: 1975-1996. Second, we consider Canadian residential mortgage credit in place of M3. The forecasting performance and model selection between the ARMA model and the SETAR model are re-valuated in this light. In this section, we will not go through data results in a detailed manner as earlier, but we will focus primarily on the forecasting comparison between the real-time and current vintage models, as well as making comparisons between the SETAR and ARMA models.

Table 8 shows the real-time and current vintage forecasting comparison for the ARMA and SETAR models for M3 (1975-1996) and Canadian mortgage credit (1975-1996). Table 9 provides the S-test results for the models for these two cases. These two tables demonstrate the following findings. First, the current vintage model is always better than the real-time model for ARMA and SETAR respectively. Second, for both series, the MAPE and RMSPE values support the SETAR model in the current vintage case, but these two criteria favor the ARMA model in the real-time case. From the S-test results, we find that the two models are equivalent in terms of forecasting performance, except for current vintage Canadian mortgage credit. In that case the SETAR model forecasts out-perform those of the ARMA model. Third, both models tend to improve their forecasting performance in the case of real-time as we roll four periods ahead. Finally, we also find that the real-time forecast is not always lower than the current vintage forecast.

Table 8. Forecasting Comparison for ARMA and SETAR: Real-Time vs. Current Vintage

			Current Vintage		Real-time	
			ARMA	SETAR	ARMA	SETAR
M3 (1975-1996)	MAPE	Stage1	2.096	0.479	4.288	4.367
		Stage 2	1.534	1.017	1.192	2.532
		Over all	1.815	0.748	2.74	3.45
	RMSPE	Stage 1	2.236	0.6	4.387	4.386
		Stage 2	1.577	1.219	1.245	2.649
		Over all	1.907	0.91	2.816	3.518
Mortgage Credit (1975-1996)	MAPE	Stage 1	1.71	1.144	0.892	1.267
		Stage 2	2.001	1.221	0.53	1.221
		Over all	1.856	1.182	0.711	1.244
	RMSPE	Stage1	1.897	1.488	1.115	1.494
		Stage 2	2.069	1.49	0.615	1.49
		Over all	1.985	1.489	0.9	1.492

Table 9. The S-test for Model Selection between SETAR and ARMA

		S-statistics	Accept/Reject
M3 (1975-1996)	Real-time	-1.414	Accept
	Current Vintage	1.414	Accept
Mortgage Credit (1975-1996)	Real-time	-1.414	Accept
	Current vintage	2.828	Reject

We find that the results for M3 in both time periods are the same. The two models give equivalent forecasts both for real-time and current vintage data, which indicates that the results are quite robust for the same data series. For Canadian mortgage credit, in the current vintage case, the SETAR model outperforms the ARMA model in forecasting, but the two are equivalent in real-time forecast using S-test, which is not supported by the MAPE and RMSPE results.

7. Conclusions

As macroeconomic time-series data are revised, the information contained in the time-series becomes increasingly rich. The results we have obtained, which demonstrate that the current vintage model forecasts better than the real-time model, are apparently robust. Equally robust is our finding that, for real-time data, the non-linear SETAR model and the linear ARMA models are equivalent in terms of forecasting quality. In the current vintage case, for the same series (M3

in our case), the conclusion that the models are equivalent holds even if we change the sample range and the forecasting periods. However, if we switch to Canadian residential mortgage credit series, this result may not necessarily hold. Finally, the linear ARMA and non-linear SETAR models tend to improve their forecasting performance as data revisions take place, but in the current vintage case, the SETAR model's data-sensitive characteristics offset the improvement gained from data revisions.

In future research it would be interesting to see a similar study involving higher frequency time-series data. High frequency series are less prone to data revisions, but they are more likely to exhibit more volatility. Accordingly, the modeling and forecasting abilities of the nonlinear SETAR model might be expected to be superior to those of the linear ARMA model, both in real time and current vintage data forms.

References:

Akaike H. (1974) "A new Look at Statistical Model Identification," *IEEE Transactions on Automatic Control,* 19, 716.

Cole, R. (1969) "Data Errors and Forecasting Accuracy," in J. Mincer (ed.), *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance,* New York: National Bureau of Economic Research, 47-82.

Christiano, L. J., and Eichenbaum, M. (1992) "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations," *American Economic Review,* 86, 430-450.

Croushore, D., and Stark T. (1999) "Does Data Vintage Matter for Forecasting?" Federal Reserve Bank of Philadelphia Working Paper 99-15.

Croushore, D., and Stark T. (2000) "A Real-time Data Set for Macroeconomists: Does the Data Vintage Matter?" Federal Reserve Bank of Philadelphia Working Paper 00-06.

Croushore, D., and Stark T. (2001) "A Real-time Data Set for Macroeconomists," *Journal of Econometrics,* 105, 111-130.

Denton, F. T., and Kuiper J. (1965) "The Effect of Measurement Errors on Parameter Estimates and Forecasts: A Case Study Based on the Canadian Preliminary National Accounts," *Review of Economics and Statistics*, 47, 198-206.

Diebold, F. X., and Mariano R. S., (1995) "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253-63.

Feng, H and Liu, J, (2003) "A SETAR Model for Canadian GDP: Non-linearities and Forecast Comparisons," *Applied Economics*, 18, 1957 – 1964.

Franses, P.H., and van Dijk, D. (2000) *Non-Linear Time Series Models in Empirical Finance,* Cambridge University Press, Cambridge.

Gallegati, M., and Mignacca, D. (1995) "Nonlinearities in Business Cycle: SETAR Model and G7 Industrial Production Data," *Applied Economics Letters*, 2, 422-427.

Hansen, B.E. (1997) "Inference in TAR Models," *Studies in Nonlinear Dynamics & Econometrics*, Berkeley Electronic Press, 2, 1-14.

Hansen, B.E. (2000) "Sample Splitting and Threshold Estimation," *Econometrica*, 68, 575-604.

SHAZAM (1997) *SHAZAM Econometrics Computer Program, Version 8.0: User's Reference Manual*, McGraw-Hill, New York.

Sowell, F. (1992) "Modeling long-Run Behaviour With the Fractional ARIMA Model," *Journal of Monetary Economics*, 29, 277-302.

Swanson, N. R. (1996) "Forecasting Using First-Available Versus Fully Revised Economic Time-Series Data," *Studies in Nonlinear Dynamics and Econometrics*, Berkeley Electronic Press, 1, 47-64.

Swanson, N. R., Ghysels E., and Callan M., (1999) "A Multivariate Time Series Analysis of the Data Revision Process for Industrial Production and the Composite Leading Indicator," in R. F. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting*. Oxford: Oxford University Press.

Swanson, N. R., and White H., (1997) "A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks," *Review of Economics and Statistics,* 79, 540-50.

Tong, H. (1978) "On a Threshold Model," in C. H. Chen (ed.), *Pattern Recognition and Signal Processing*, Sijhoff and Noordoff, Amsterdam.

Tong, H. (1983) *Threshold Models in Non-Linear Time Series Analysi*s, Springer-Verlag, New York.

Tong, H. (1990) *Non-linear Time Series. A Dynamical System Approach*, Clarendon Press, Oxford.

Tong, H., and Lim, K. S. (1980) "Threshold Autoregression, Limit Cycles and Cyclical Data," *Journal of the Royal Statistical Society*, B, 42, 245–292.

Trivellato, U. and Rettore E., (1986) "Preliminary Data Errors and Their Impact on the Forecast Error of Simultaneous-Equations Models," *Journal of Business and Economic Statistics*, 4, 445-53.