JRAP (1999) 29: 2

The Journal of Regional Analysis & Policy

A Research Note on the Accuracy of Simple Methods for Updating Untimely Data**

Jon R. Miller and Julia M. Dickson*

Abstract. In this paper we address an irritating aspect of applied economics in regional science, the untimeliness of broad measures of regional economic performance. Our purpose is to assess the accuracy of simple methods for updating untimely regional income data. We examine three simple methods here, the naive model, percent projection, and exponential smoothing. Forecast accuracy is assessed on the basis of out-of-sample percent forecast errors. Not surprisingly, one-year forecasts are more accurate than two-year forecasts. For large regions, in tranquil times of regional growth, percent projection and exponential smoothing methods have similar accuracy, and both out-perform the naive model. But when regional economic times become volatile, as is the case in some time periods for large regions, and is commonplace for many counties studied here, exponential smoothing methods are superior in terms of accuracy.

1. Introduction

In the broadest sense, information on regional economic performance answers three fundamental questions about the economy: 1) Where are we? 2) Where have we been? and, 3) Where are we going? As regional scientists we do pretty well on the second question, by separating regional time series data into various temporal, sectoral, and spatial aggregations, and examining patterns over time. Our response to

*Professor of Economics, University of Idaho Moscow, Idaho 83844-3240, Tel: (208) 885-7506 FAX: (208) 885-8939 e-mail: jrmecon@uidaho.edu and Analyst, Process Competency Group, Andersen Consulting LLP, Suite 900, 801 Second Avenue, Seattle, Washington 98104, Tel: (206) 623 - 8950 FAX: (206) 623 - 8615, respectively.

**A previous version of this paper was presented at the thirty-eighth annual meeting of the Western Regional Science Association in Ojai, California, in February 1999. Much of the research on this paper occurred when Ms. Dickson was a senior economics and finance major in the College of Business and Economics at the University of Idaho.

the third question is, at the least, reasonably well-defined, as regional forecasting is an established sub-discipline. Surprisingly, it is the first question where we often have difficulty.

In this paper we address an irritating aspect of applied economics in regional science. Unless we are willing to measure regional economic performance in employment units, waiting for data is often like waiting for Godot. Broad measures of economic performance, such as annual personal income, appear with a lag approaching a year and a half. In order to discuss "current conditions" in the regional economy, regional science practitioners must forecast the present with past data. While the technical expertise of academic regional scientists is certainly high enough to apply sophisticated forecasting methods, the same may not be true for a staff member of, say, a city or county planning commission. Likewise, budgets may not be sufficient to contract for this kind of analysis.

The purpose of this paper is to assess the accuracy of simple methods for updating untimely regional income data. For the purposes of this research, we define a simple method to have three characteristics. First, it must use readily available, univariate time series data. This requirement excludes regression models, for example, which require the maintenance of multiple time series. Second, the method or model must not be overly data-hungry. Autoregressive integrated moving average (ARIMA) models would be precluded here, as long time series are required to estimate forecasting equations with these methods. Finally, a simple method must not be analytically or technically complex, requiring extensive technical training. Ordinary regression models, ARIMA models, and other time series techniques, such as vector autoregression, error correction models, and cointegration analysis would be similarly precluded.

We examine three simple methods here, the naive model, percent projection, and exponential smoothing. With the naive model we assume future values of a variable will be equal to the last reported value. Percent projection requires two time periods of data, and extrapolates the most recent percent change into the future. Our final simple method is exponential smoothing, or exponentially-weighted moving average models, which we define more formally in the next section. Admittedly, we are adding some complexity here. But the methods are easy to understand, are univariate, require very little data, and can be applied with only a modicum of training. We teach public university undergraduate business students to develop smoothing models in approximately three to five hours of class time, and software packages exist which are extremely user-friendly. Exponential smoothing methods also have a history of accurate short term forecasts, often beating more complex methods in forecasting competitions (e.g., see Makradakis, et al. 1982).

We apply these simple methods and assess the accuracy of one and two-year updates of real personal income in seven large economic regions of the Northwest United States. The economic regions are those areas defined by the Bureau of Economic Analysis (Johnson 1995). For one of these large regions, the Spokane Economic Area, we perform a similar analysis for the component counties of the region. We also check the sensitivity of the analysis to length of historic period and to the possible uniqueness of the most current period.

2. Exponential Smoothing Models

The models applied here are what are commonly referred to as Holt exponential smoothing models, which incorporate a forecast level and linear trend, but no seasonality. Models are estimated with the software package Forecast Pro for Windows (see Stellwagen and Goodrich 1994). Models are defined as follows:

$$Y_t(m) = L_t + mT_t \tag{1}$$

where the notation is as follows:

$Y_t(m)$	Forecast for time t+m from base t
L_t	Smoothed level at end of time t
m	Forecast lead time
T_t	Smoothed trend at end of time t

Level and trend values are obtained from the following equations:

$$L_t = aY_t + (1-a)(L_{t-1} + T_{t-1}) \tag{2}$$

$$T_t = g(L_t - L_{t-1}) + (1-g)T_{t-1} \tag{3}$$

where a and g are smoothing parameters for level and trend, respectively. To estimate model parameters, the Forecast Pro program uses an iterative search to minimize the sum of squared errors over the historic data.

3. Measure of Forecast Accuracy

Forecast accuracy is measured on the basis of out-of-sample forecasts of annual data for 1995 and 1996. For example, from a time series of real personal income, by region, running from 1969 through 1996, we "hold out" the 1995 and 1996 observations. Exponential smoothing models are then estimated on the 1969 -1994 series, and used to make forecasts for 1995 and 1996 real personal income. Forecasts are also developed from the naive and percent projection models. Forecasted values are then compared to actual 1995 and 1996 values to determine the accuracy of the method. Percent error is the accuracy measure.

4. Data

Data are derived from nominal personal income for counties and economic areas from U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1997), and U.S. Department of Commerce, Bureau of Economic Analysis (1998). Nominal personal income is converted to 1996 constant dollars with the CPI-U index from U.S. Bureau of Labor Statistics (1998).

5. Results

In Table 1 we show the value of smoothing constants and standard in-sample goodness-of-fit measures for exponential smoothing models applied to a 1969 - 1994 time series of real personal income for regions of the Northwest U.S. The specific regions appear in the leftmost column.

Table 1. Smoothing Constants And Goodness Of Fit Information For Real Personal Income Exponential Smoothing Forecasting Models For Northwest Economic Regions. (Sample: 1969-1994)

Region[a]	Smoothing Constant		R^2	In-sample MAPE
	Level	Trend		
Seattle	.99	.58	.99	1.9
Portland	.99	.69	.98	2.3
Spokane	.99	.06	.97	2.3
Eugene	.99	.10	.96	3.1
Tri-Cities	.99	.06	.97	2.9
Boise	.99	.78	.99	2.6
Pendelton	1.0	-	.82	3.4

[a]Regions defined as BEA economic areas (Johnson 1955)..

For all regions but Pendelton, the model is of the form described in equation (1) above. The values in the "smoothing weight" columns for "Level" and "Trend" correspond to the "a" and "g" parameters in equations (2) and (3), respectively. A high value indicates that more recent observations have a greater influence on the forecasted value of level or trend. Lower values mean more distant past observations influence the forecast. Because no trend is apparent in the Pendelton series, simple exponential smoothing is selected for that region.

Furthermore, the value of the smoothing constant is 1.0, indicating that in this case the exponential smoothing model produces forecasts identical to those of the naive model.

As can be seen in the R^2 column, the overall in-sample fits of the models are good. The rightmost column of Table 1 shows the in-sample mean absolute percent error (MAPE), the percent, on average, that the in-sample forecasted value differs from the actual value.

Table 2 shows percent errors for one and two-year forecasts of regional real personal income for the three simple forecasting methods. Because Northwest U.S. regions have been experiencing annual growth since the early 1980s, it is not surprising that both the percent projection and exponential smoothing models out-perform the naive model. The accuracy of the percent projection model is slightly better than that of exponential smoothing, but only slightly. The percent projection model is more accurate in four regions, and the exponential smoothing model is more accurate in three. The MAPE across regions is very similar for the two methods, 1.0 and 1.2 for the 1995 forecast, and 2.9 and 3.1 for the 1996 forecast.

Table 2. A Comparison of the Accuracy of Alternative Simple-Model One and Two-Year Real Personal Income Forecasts for Northwest Regions[a]

	1995 Percent Error[b]			1996 Percent Error[b]		
Region[c] Model[d]	Naive	%Projection	Exp. Smoothing	Naive	%Projection	Exp. Smoothing
Seattle	-3.2	-1.0	-1.1	-8.3	-4.1	-4.3
Portland	-4.9	-0.9	-1.3	-10.0	-2.2	-3.2
Spokane	-2.8	0.4	-0.9	-5.7	0.6	-2.1
Eugene	-3.5	-0.2	-1.5	-8.5	-2.2	-4.8
Tri-Cities	-1.6	-0.5	-0.3	-4.8	-2.6	-1.0
Boise	-5.1	0.4	0.3	-9.4	1.3	-0.9
Pendelton	-2.8	-3.6	-2.8	-5.5	-7.1	-5.5
MAPE (across regions)	3.4	1.0	1.2	7.5	2.9	3.1
Accuracy Ranking Score[e]	1.5	10.0	9.5	1.5	10.0	9.5

[a]Actual and forecasted values in 1996 constant dollars.
[b]Calculated as ((Forecast-Actual)/Actual) x 100.
[c]Regions defined as BEA economic areas (Johnson 1955).
[d]Forecast magnitude with the Naive model is the 1994 actual value. The % projection model forecasts by projecting forward the 93-94 % change in the actual value. For exponential smoothing models, see Table 1.
[e]Calculated by assigning 2 points, 1 point, and 0 points for the most, next, and least accurate regional forecast, respectively, then summing across regional forecasts for each method. Ties receive one-half the sum of both rank scores.

The last row of Table 2 contains an accuracy ranking score. For each regional forecast, two points are awarded to the model with the lowest percent forecast error, one point for the next most accurate, and no points for the least accurate. Ties receive one-half the sum of the two appropriate ranking scores. In both the 1995 and 1996 cases, the scores are 10 and 9.5 for the percent projection and exponential smoothing methods, respectively.

While judgements about accuracy level are subjective, these results suggest that either percent projection or exponential smoothing models have acceptable levels of one-year forecast accuracy when applied to large Northwest regions experiencing reasonably tranquil economic times. Two-year forecasts are more problematic, with percent forecast errors roughly two to three times larger than those for the one-year forecasts.

Before turning to the county forecasts, we present some "sensitivity" information about alternative historic samples and end years. The results are presented for the Seattle region in Table 3, and for the Spokane Region in Table 4. In each table we show one and two-year forecasts from the end years, 1976, 1981, 1986, and 1992. The end years were chosen partly to test the models in time periods which were more volatile than the recent past, such as the "double dip" national recession of the early 1980s. Also, for the end years 1981 and 1986, we estimate the exponential smoothing model over historic samples of different lengths.

Table 3. A Comparison of the Accuracy of Alternative Simple-Model One and Two-Year Real Personal Income Forecasts for the Seattle Region for Different Historic Samples and End Years[a]

	One-Year Forecast % Error[b]			Two-Year Forecast % Error[b]		
Sample Model[c]	Naive	% Projection	Exp. Smoothing	Naive	% Projection	Exp. Smoothing
1969-76	-4.8	1.5	-0.4	-12.7	-0.8	-4.5
1969-81	0.2	1.9	2.0	-1.8	1.5	1.6
1974-81	-	-	1.8	-	-	1.2
1974-86	-3.1	2.4	2.0	-7.9	2.9	1.9
1979-86	-	-	2.1	-	-	2.0
1979-92	-0.9	3.6	2.5	-3.1	5.9	3.6
MAPE	2.3	2.3	1.7	6.4	2.8	2.9
Accuracy Ranking Score[d]	4.0	3.0	5.0	2.0	5.0	5.0

[a]Actual and forecasted values in 1996 constant dollars.
[b]Calculated as ((Forecast-Actual)/Actual) x 100.
[c]Forecast magnitude with the Naive model is the 1994 actual value. The % projection model forecasts by projecting forward the 93-94 % change in the actual value. For exponential smoothing models, see Table 1.
[d]Calculated by assigning 2 points, 1 point, and 0 points for the most, next, and least accurate regional forecast, respectively, then summing across regional forecasts for each method. Ties receive one-half the sum of both rank scores.

Table 4. A Comparison of the Accuracy of Alternative Simple-Model One and Two-Year Real Personal Income Forecasts for the Spokane Region for Different Historic Samples and End Years[a]

	One-Year Forecast % Error[b]			Two-Year Forecast % Error[b]		
Sample Model[c]	Naive	% Projection	Exp. Smoothing	Naive	% Projection	Exp. Smoothing
1969-76	-2.5	2.5	0.2	-9.7	-0.1	-4.6
1969-81	4.3	3.5	6.4	-1.9	-3.4	2.1
1974-81	-	-	4.3	-	-	-1.9
1974-86	0.6	2.9	0.6	-0.8	3.7	-0.9
1979-86	-	-	-0.3	-	-	-1.7
1979-92	-3.8	1.1	-1.5	6.8	2.8	-1.7
MAPE	2.8	2.5	2.2	4.8	2.5	2.3
Accuracy Ranking Score[d]	4.0	4.5	4.5	4.0	3.0	5.0

[a]Actual and forecasted values in 1996 constant dollars.
[b]Calculated as ((Forecast-Actual)/Actual) x 100.
[c]Forecast magnitude with the Naive model is the 1994 actual value. The % projection model forecasts by projecting forward the 93-94 % change in the actual value. For exponential smoothing models, see Table 1.
[d]Calculated by assigning 2 points, 1 point, and 0 points for the most, next, and least accurate regional forecast, respectively, then summing across regional forecasts for each method. Ties receive one-half the sum of both rank scores.

As seen in Tables 3 and 4, forecast errors differ with sample length, but the differences are reasonably minor. In only one instance, the forecasts from 1981 in the Spokane region, do absolute percent forecast errors differ by more that one percentage point. This is comforting from an estimation standpoint, as shorter time series are always more available and easier to maintain than longer series.

In the less tranquil time periods examined in Tables 3 and 4, the supremacy of the percent projection and exponential smoothing models over the naive model either vanishes or falls significantly. The exponential smoothing models have the lowest across-region MAPE for the one-year forecasts in both regions, and the two-year forecast in the Spokane region, but the other methods are not far behind. This is also shown in the similar accuracy ranking score, where exponential smoothing ties for first in two cases, and wins in two cases, but again not by much. These results suggest that the moderately superior performance of the percent projection method seen in the 1995 and 1996 forecasts shown in Table 2, likely depends on the tranquil times of steady growth found preceding these years in most Northwest U.S. regions. In more volatile economic times, the percent projection model's accuracy falls, relative to the other simple models.

The effect of time series volatility on the forecast accuracy of simple methods is easily seen in the forecasts of county real personal income. In Table 5 we show estimation results like those presented in Table 1 for the

Table 5. Smoothing Constants And Goodness Of Fit Information For Real Personal Income Exponential Smoothing Forecasting Models For Component Counties of the Spokane Region. (Sample: 1969-1994)

	Smoothing Constant			
County	Level	Trend	R^2	In-sample MAPE
Spokane	1.00	0.35	.98	1.9
Kootenai	0.99	0.84	.99	3.0
Nez Perce	0.99	0.05	.93	2.6
Whitman	0.48	-	.02	4.8
Stevens	1.00	0.35	.96	4.2
Latah	0.88	0.05	.90	3.9
Bonner	0.99	0.06	.97	3.9
Asotin	1.00	0.31	.96	3.0
Idaho	0.92	-	.55	4.1
Shoshone	1.00	-	.82	.47
Lincoln	0.55	-	.23	8.9
Pend O'Reille	0.96	0.05	.96	3.8
Benewah	0.54	0.04	.76	5.9
Clearwater	1.00	-	.65	3.9
Boundary	0.90	0.05	.87	4.3
Ferry	0.99	0.05	.95	4.3
Lewis	0.24	-	.05	8.5
Garfield	1.00	-	.67	12.4

larger regions. Here the spatial units are the component counties of the Spokane Region. Note the much wider variety in smoothing constants for level in Table 5, and the generally lower R^2 values. Seven of the counties exhibit no trend in real income over the sample period. While not presented here for space considerations, graphs of the time series are much more likely to look like a saw blade, rather than the knife edge characterizing recent large-region growth.

Table 6 contains a comparison of the forecasts from the three simple models for the eighteen component counties of the Spokane region. Note that very large forecast errors appear for some counties, with the percent

Table 6. A Comparison of the Accuracy of Alternative Simple-Model One and Two-Year Real Personal Income Forecasts for Counties in the Spokane Region[a]

	1995 Percent Error[b]			1996 Percent Error[b]		
County Model[c]	Naive	% Projection	Exp. Smoothing	Naive	% Projection	Exp. Smoothing
Spokane	-2.4	0.8	0.6	-4.5	1.9	1.3
Kootenai	-4.2	4.0	3.1	-8.6	7.5	5.2
Nez Perce	-1.7	0.8	-0.1	-5.2	-0.4	-2.3
Whitman	-4.8	-8.6	-4.1	-10.8	-17.7	-10.1
Stevens	-1.4	2.3	2.3	-3.2	4.1	4.0
Latah	-3.7	-1.1	-2.2	-7.1	-2.1	-4.0
Bonner	-2.2	4.1	0.5	-6.2	6.2	-1.1
Asotin	-1.4	2.0	2.1	-4.4	2.2	-2.2
Idaho	0.3	0.5	0.3	-2.5	-2.2	-2.5
Shoshone	-1.6	1.8	-1.6	-7.5	-1.0	-7.5
Lincoln	-14.9	-26.3	-10.5	-17.7	-38.3	-13.5
Pend O'Reille	-3.6	0.9	-1.6	-6.7	2.3	-2.7
Benewah	2.1	3.5	-3.2	-5.2	5.9	-4.8
Clearwater	-1.0	2.2	-1.0	-5.2	1.0	-5.2
Boundary	-1.5	2.3	0.0	-7.0	0.4	-4.0
Ferry	-1.0	1.9	1.0	-2.9	9.0	6.9
Lewis	-3.4	-6.3	-4.3	-3.5	-9.3	-4.5
Garfield	-6.0	-24.2	-6.0	-19.0	-47.4	-19.0
MAPE	3.2	5.2	2.5	7.1	8.8	5.6
Accuracy Ranking Score[d]	19.5	9.5	25.0	11.5	18.0	24.5

[a]Actual and forecasted values in 1996 constant dollars.
[b]Calculated as ((Forecast-Actual)/Actual) x 100.
[c]Forecast magnitude with the Naive model is the 1994 actual value. The % projection model forecasts by projecting forward the 93-94 % change in the actual value. For exponential smoothing models, see Table 1.
[d]Calculated by assigning 2 points, 1 point, and 0 points for the most, next, and least accurate regional forecast, respectively, then summing across regional forecasts for each method. Ties receive one-half the sum of both rank scores.

projection method having the largest error, e.g., under-forecasting 1995 real personal income by 24.2% and 26.3% for Garfield and Lincoln Counties, respectively. The other methods, while also greatly errant in these cases, are "less bad." In fact, the cross-region MAPE for the 1995 forecast is a respectable 2.5% for the exponential smoothing model. The two-year forecast errors are, on average, like their counterparts in the large region case, two to three times larger than the one-year forecast errors.

The exponential smoothing model is clearly superior to the other simple models for updating county personal income data. The across-region MAPE is the lowest for both 1995 and 1996, and the accuracy ranking score is larger by a wide margin.

6. Conclusions

The main conclusion from this research is that among a set of simple forecasting methods which includes the naive model, percent projection, and exponential smoothing, the latter is likely to yield the most accurate forecasts, on average. In tranquil times of continuous regional economic growth, the advantage of exponential smoothing over percent projection does not hold. But if volatility in real personal income exists, as it does for some time periods for large Northwest regions, and over most time periods for many of the counties studied here, exponential smoothing models are superior to the other methods. The "benefits" of increased forecast accuracy with exponential smoothing would appear to outweigh the "costs" of the modest increase in technical complexity the methods create.

While the acceptability of any level of forecast error is subjective, we suggest that using simple methods, especially exponentially smoothing, to generate one-year forecasts to update untimely regional real personal income data has a reasonable level of accuracy for large regions and counties, as well. The accuracy of two-year updates with simple methods is more suspect, and more caution is warranted.

References

Johnson, K.P. 1995. Redefinition of the BEA economic areas. *Survey of Current Business* 75(February): 75-81.

Makradakis, S., A. Anderson, R. Carbone, R. Fildes, M. Hibron, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler. 1982. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting* 1(1982): 111- 153.

Stellwagen, E.A., and R.L. Goodrich. 1994. *Forecast Pro for Windows*. Belmont, MA: Business Forecast Systems, Incorporated.

U.S. Department of Commerce. 1997. Regional Economic Information System 1969-95: Personal Income and Employment estimates for all counties and metropolitan areas in the United States. Washington, D.C.: USDC, Economics and Statistics Administration, Bureau of Economic Analysis.

U.S. Department of Commerce. 1998. Website: www.bea.doc.gov/bea/dr1.html. Washington, D.C.: USDC, Bureau of Economic Analysis.

U.S. Department of Labor. 1998. Website: www.bls.gov/top20.html. Washington, D.C.: USDC, Bureau of Labor Statistics