Journal of Forecasting, Volume 15, Number 1, 1996, Pages 19-36

Assessing the Historical Accuracy of Regional Economic Forecasts*

Carol Taylor West
Thomas M. Fullerton, Jr.
Department of Economics and
Bureau of Economic & Business Research
University of Florida
221 Matherly Hall
Gainesville, FL 32611-2017
Telephone: 904-392-0171
Facsimile: 904-392-4739

* An earlier version of this paper was presented at the annual meetings of the American Statistical Association, Toronto, Canada, August 1994.

Assessing the Historical Accuracy of Regional Economic Forecasts

Carol Taylor West is Professor of Economics and Forecasting Program Director at the College of Business Administration, University of Florida. Her work has appeared in outlets such as **Journal of Economic Theory**, **Review of Economics & Statistics**, **Regional Science & Urban Economics**, and **Journal of Urban Economics**.

Thomas Fullerton is Senior Economist at the Bureau of Economic and Business Research at the University of Florida, where he also teaches Latin American Political Economy. His work has appeared in outlets such as **Applied Economics**, **International Journal of Forecasting**, **Journal of Policy Modeling**, and **Public Budgeting & Finance**.

Assessing the Historical Accuracy of Regional Economic Forecasts

Abstract

This study addresses for the first time systematic evaluation of a widely used class of forecasts, regional economic forecasts. Ex ante regional structural equation model forecasts are analyzed for 19 metropolitan areas. One-quarter to ten-quarter ahead forecasts are considered and the seven year sample spans a complete business cycle. Counter to previous speculation in the literature, (1) dependency on macroeconomic forecasting model inputs does not substantially erode accuracy relative to univariate extrapolative methodologies and (2) stochastic time series models do not on average yield more accurate regional economic predictions than structural models. Similar to findings in other studies, clear preferences among extrapolative methodologies do not emerge. Most general conclusions, however, are subject to caveats based on step-length effects and region-specific effects.

Key Words: Econometric models, Univariate time series analysis, Regional forecast evaluation

Regional economic forecasts are grist for state and local policy making and budgeting mills around the country. Despite the widespread use of regional economic forecasts, they have remained curiously exempt from accuracy evaluation. For over two decades there has been on-going monitoring of the accuracy of national macroeconomic forecasts (Nelson, 1972 and 1984; McNees, 1981 and 1992, Ashley, 1983 and 1988). Recent work in demography, including regional demography, has admitted that population "projections" are in fact "forecasts" and has begun addressing questions of forecast accuracy (Ahlburg and Land, 1992). Along with national macroeconomic forecasts and population forecasts, regional economic forecasts are commonly utilized, but unlike their macroeconomic and demographic counterparts, their accuracy has not been evaluated.

Recent literature has heightened interest in the issue, but has not addressed it. Ashley (1983) clearly questioned the efficacy of the popular regional modeling methodology of constructing area econometric models which operate as "satellites" to national macroeconomic models. Broad claims that regional stochastic time series models (henceforth RSTIMS) are more accurate and cheaper than regional structural equation models (henceforth RSEMS) abound in the regional forecasting literature (Anderson, 1979; Weller, 1989 and 1990; Lesage, 1990; Shoesmith, 1990 and 1992), but the assertions have not been tested. Regional demography has intensely debated the relative merits of causal and extrapolative methodologies, but again direct comparisons are not cited (see for example, the Smith (1987a), Beaumont and Isserman (1987), Smith (1987b) exchange).

The current research addresses these issues by evaluating the accuracy of ex ante RSEMS forecasts for 19 metropolitan areas. One-quarter to ten-quarter ahead forecasts are considered and the seven year sample set spans a complete business cycle. General results include: (1) Ashley's (1983) direst prognostications for the ex ante accuracy of RSEMS using forecast macroeconomic variables as inputs are not realized when a large number of regions and forecast periods are considered--i.e., the RSEMS forecasts are on average more accurate than projecting the regional variable at its historical mean (but not unanimously over all regions and time horizons); (2) univariate ARIMA RSTIMS are not on average more accurate than RSEMS and relative accuracy generally exhibits the horizon effects observed by Nelson (1984) for macroeconomic forecasts--i.e., accuracy of RSEMS relative to ARIMA declines with an increase in time horizon; (3) there are clear region/methodology interactions analogous to the interaction of product/market situation and methodology accuracy observed in sales forecasting (e.g., Thomas, 1993), but unfortunately the most accurate methodology for a particular area is not readily determined from basic economic/demographic characteristics which are commonly used to distinguish types of regions.

Background

The "rolling recession" of the 1980s highlighted the differential growth paths of many regional economies compared with the national average, increasing demand for projections specific to local markets. Despite the expansion of regional forecasting activities, and RSEMS forecasting in particular, evaluation of ex ante RSEMS forecasts is an almost non-existent literature. To the authors' knowledge, only one article addresses the question. Conway (1990) evaluates one ex ante long-term RSEMS forecast for one state. Results are difficult to interpret, however, because no comparative norm is presented. Several studies consider the relative accuracy of RSEMS and RSTIMS. Two (Anderson, 1979, and Lesage, 1990) are crude, yielding no useful information. They compare outcomes of model A for region R over time period Y with those of model B for region S over time period Z. Not only are time horizon and time period effects ignored, but early work on regional forecasting models established that alternative methodologies cannot be compared when the methodology and the region are simultaneously changed (Taylor, 1982b). Like different economic variables, some regions are more "forecastable" than others.

Less crude comparisons of RSEMS and RSTIMS exist (Schmidt, 1979; Taylor, 1982a; Fullerton, 1989; Glennon et al., 1987), but each is highly restricted in scope. None is based on ex ante forecasts. Only one (Fullerton, 1989) considers time horizon effects and only one (Taylor, 1982a) analyzes more than one region. None tests over varying phases of the business cycle. To summarize briefly, in forecasting metropolitan area employment, Taylor(1982a) finds RSEMS outperform RSTIMS in forecasting total jobs and Glennon et al.(1987) find the same for nonmanufacturing employment, but the opposite for manufacturing employment. In forecasting state sales taxes and retail sales, Fullerton (1989) finds RSEMS outperform RSTIMS and Schmidt (1979) finds the opposite. To determine whether these conflicting results are reflective of strong region-specific factors or the limited nature of the individual studies requires more comprehensive analysis.

Lack of evaluation undoubtedly derives from a combination of three factors. First, on-going RSEMS forecasting has a shorter history than its national macroeconomic counterpart. Only recently have a large sample of ex ante forecasts become available for any region. Second, some of the gap must be attributed to regional forecaster reluctance to perform the evaluation. Even long-standing regional modeling systems such as NRIES and REMI still eschew the opportunity to present evaluation of ex ante forecasts (Lienesch and Kort, 1992, and Treyz, et al., 1992). Third, regional data are subject to large revisions over time which confound comparative analysis. McNees (1988) has shown how difficult it is to ex post incorporate these revisions into the evaluative process and in a recent article (McNees, 1992), he illustrates how choice of "actual" can significantly effect conclusions on forecast accuracy.

The more fundamental problem with revisions derives from Nelson's (1984) observation that forecasts from two methodologies are genuinely comparable only when they are made ex ante and at the same time. Ex ante and at the same time imply (1) the same complete lack of knowledge of the forecast period and (2) utilization of the same data set as it existed ex ante to the forecast period. Condition (1) can be replicated ex post only with univariate extrapolative methodologies. The frequent and substantial revisions to regional economic data series often render ex post replication of condition (2) virtually impossible. To the authors' knowledge, no regional forecasting systems have made simultaneous forecasts with distinct methodologies and hence it is difficult to make meaningful comparisons or evaluations of forecasting techniques in this field.

Such a study is feasible for Florida, however, at the current time because of the existence of a set of ex ante RSEMS economic forecasts for 19 metropolitan areas which span one to ten quarter projection horizons and jointly cover the various phases of the business cycle. Equally important, the data set for the variable which is central to most regional forecasting models, employment, can be replicated as it existed at any historical point in time for this particular series of forecasts. Consequently, comparative univariate extrapolative models can be constructed ex post. This unique situation arises because of an anomalous choice for the measurement of regional employment in these RSEMS models. The commonly used series in regional forecasting is the Bureau of Labor Statistics monthly employer survey data (BLS 790 data) which is subject to large revisions for up to two years (Coomes, 1992). In contrast, the present set of regional forecasts is based on employment data from Unemployment Insurance reports (ES 202 data) which are not revised over time. They are in fact the data which are used to generate the sequence of BLS 790 revisions. Note that the lack of revision to ES 202 data not only permits historical reconstruction of data sets, it simultaneously avoids all problems of trying to account for the revisions in the evaluation process.

The RSEMS employment forecasts are ex ante projections published quarterly for 19 metropolitan statistical areas (MSAs) of Florida in The Florida Outlook, Volume 10, No. 1 through Volume 16, No. 3 (see, e.g., West and Fullerton, 1994). Although all the sample metropolitan areas are within Florida, this fourth largest state encompasses a variety of economic regions. Some salient metropolitan area features are summarized in Exhibit 1. Size, growth rates, income levels and demographic composition span a wide range. Not shown in Exhibit 1 are the varied economic bases of the metropolitan areas. Some key components are well known--tourism in Orlando, state government in Tallahassee, retirees and "snowbirds" in Sarasota, space and defense manufacturing and research in Brevard County (the Melbourne MSA which includes Kennedy Space Center and Cape Canaveral), the state university in Gainesville. Some less familiar ones include the historical phosphate mining and manufacturing bases of Lakeland, the dominance of military bases in Pensacola and Ft. Walton Beach, the shifting economic base of Jacksonville from pulp/paper and food manufacturing to finance, insurance and international port services, and the urban/rural mix of the West Palm Beach MSA (besides being Florida's wealthiest and sixth largest MSA, Palm Beach County ranks number one among Florida's 67 counties in agricultural sales).

Place Exhibit 1 here

The forecast sample period from 1985.4 through 1992.2 covers rapid growth of mid-1985 to 1988, deceleration in 1989 and early 1990, decline into recession in late 1990 and 1991 and recovery from recession in the early part of 1992. Employment data prior to 1985.4 (and forecasts based on these data) were not generated by the same methodology and, hence, are not comparable. Forecasts beyond 1992.2 were excluded for two reasons. First, the landfall of Hurricane Andrew in 1992.3 introduced outliers in the time series data for southeastern Florida and indeed the entire state. Extrapolative methodologies are not reasonable benchmarks in such situations. Second, in the ensuing quarter, the new geographic definitions of Florida metropolitan statistical areas released by the U.S. Census Bureau were incorporated into the models and projections of The Florida

Outlook. Consequently, regional consistency was disrupted at that point for several important metropolitan areas.

Methodology

The Florida MSA forecasting system is specified in a standard satellite arrangement to both a state econometric model and a national macroeconomic model (see Bolton, 1985). Each metro model is of the smaller simultaneous equation variety described in Taylor (1982a) and includes stochastic equations for total employment (nonagricultural, civilian), unemployment rate, two components of personal income (adjusted labor earnings and non-labor income), net migration, taxable sales, single and multifamily housing starts. For each quarterly forecast, projections of national exogenous variables were from the most recently available control scenario of DRI/McGraw-Hill.

For purposes of constructing the alternative extrapolative models, quarterly historical data series for the 19 MSAs were recreated for 27 different sample periods, 1975.1 through 1985.3, 1975.1 through 1985.4, . . . , 1975.1 through 1992.1. Seasonally adjusted total nonagricultural employment for each MSA is the sum of nine seasonally adjusted components, calculated using the common Census X-11 procedure. ES-202 data prior to 1975 did not include public sector employment and hence could not be used. Major changes in ES-202 coverage subsequent to that time are confined to the agricultural sector and did not affect nonagricultural employment estimates.

Four extrapolative methodologies were considered. The simplest is the Makridakis and Hibon (1979) Naive 1 model which is also used by McNees (1992) in evaluating national macroeconomic forecasts. It is the common no-change assumption and in this case is applied to growth rates--i.e., at quarter T, forecast growth in employment between T and T+1 is that which occurred between T-1 and T, forecast two-period growth, T to T+2, is that which occurred between T-2 and T, etc.

Ashley (1988) proposed simple extrapolation of historical trends and this is similar to the LINE methodology used in regional demography (Smith, 1987a). Ashley's(1983 and 1988) arguments are of particular relevance to the present case. The derivation is specific to the situation in which national macroeconomic forecasts are required as inputs to other forecasts, as is the case with the RSEMS models. Furthermore, the RSEMS outputs in turn are often input into regional sales, budgeting and revenue models and hence the same question can be asked about their usefulness as inputs that Ashley (1983 and 1988) asked of the macroeconomic forecasts. This recursive process suggests scope for error interaction effects and highlight the importance of examining whether the regional RSEMS forecasts are better than historical mean extrapolation. Following the conditions set forth by Ashley (1983), the extrapolation was applied to historical mean quarterly growth, not the levels of employment directly.

Based on its generally good performance in the M competition (Makridakis, et al., 1982), Holt-Winters exponential smoothing was used as the more sophisticated deterministic time series extrapolative model. Finally, analogous to Nelson (1984), univariate ARIMA models were developed for each series using the well-known steps of identification, estimation and diagnostic checking. Similar to Shoesmith (1992), the visually obvious nonstationarity of the levels was formally confirmed with an augmented Dickey-Fuller test using the first four lags of the first differences and then eliminating lags not significant at the .05 level. A second test generally confirmed stationarity of the series differenced once although in some isolated instances, second order differencing was required.

All extrapolative models were updated each quarter before a new set of forecasts was generated. Note that the ARIMA models did not simply rotely update the initial model derived from the first estimation period. Closely competing models and the length of the early sample periods raised the possibility that a forecaster at the time would change model structure as more information emerged on the time series properties for the MSA employment data in the second half of the 1980s and early 1990s. Historically, the RSEMS models were regularly updated and revised based on standard statistical tests, the economic plausibility of the coefficients and emerging evidence on the relative forecast accuracy of different exogenous variables. The RSTIMS models were similarly reviewed each quarter, re-testing for stationarity and changing form as either previously significant variables became insignificant and/or the statistical evidence in favor of closely competing specifications shifted. Four MSAs (Fort Walton Beach, Gainesville, Miami and Tallahassee) retained unchanging RSTIMS specifications throughout the development of the 27 models, two (Ft. Pierce and Sarasota) underwent five or more respecifications and in general, one to two specification changes were required for other areas.

The published RSEMS data and the extrapolative simulations generated 27 sets of forecasts for each methodology. Data availability associated with the former set restricted the analysis to one- to ten-quarter ahead projections. For each methodology, there were 27 one-quarter ahead forecasts, 26 two-quarter ahead forecasts, ..., and 18 ten-quarter ahead forecasts. Accuracy was compared by compiling modified Theil U statistics, calculated as the ratio of the root mean square error (RMSE) of the RSEMS forecasts to that of each alternative extrapolative methodology for every area and step-length.

Empirical Results

The relative RMSE comparisons are summarized in Exhibit 2 by area. Exhibits 1A through 4A in the appendix present the relative RMSE for each of the areas and each of the ten steps individually. Exhibit 3 presents the (unweighted) mean of the relative RMSE across areas for each horizon length. On average, the RSEMS forecasts clearly are more accurate than those of the naive model as the RMSE for RSEMS averages 26 percent below that of the naive across all areas and step lengths. Of the 190 comparison points (10 forecast steps, 19 areas), only 38 (20 percent) have values greater than .85 and only 14 (7.4 percent) reach equality. In contrast, 77 (40.5 percent) are less than .70.

Place Exhibit 2 here

In evaluating macroeconomic forecasts, McNees (1992, p. 30) found the naive model to be a "surprisingly stringent" standard, but that finding is not replicated here. The erratic short-term growth paths of these metro areas reduce accuracy of the naive model when applied a few steps out. The fact that an entire business cycle is spanned in the test set sharply reduces accuracy at longer steps--i.e., high recent growth in the late 1980s was projected into the subsequent recession period.

The averages in Exhibit 3 show little effect of forecast step length on relative accuracy. The same is true for 12 of the 19 MSAs and in three, the average relative accuracy of the RSEMS deteriorates by at least 10 percentage points between the first five quarters and the last five. The opposite occurs in four other MSAs.

Place Exhibit 3 here

There are, however, clear regional effects impacting relative accuracy. Of the 38 points with relative RMSE greater than .85, 35 are contributed by five MSAs--Daytona Beach, Fort Pierce, Jacksonville, Naples, Tallahassee and West Palm Beach. Of the 77 points with relative RMSE less than .70, 56 are accounted for by six areas--Lakeland, Melbourne, Orlando, Panama City, Pensacola and Tampa. Referring to Exhibit 1, there are no obvious features which distinguish these two groups. Both contain large and small MSAs, both contain high growth and slower growth MSAs and both contain relatively diversified and relatively non-diversified MSAs.

Columns 3 and 4 in Exhibit 2 and row 2 in Exhibit 3 show that the historical mean projection is a more stringent test for the RSEMS forecasts than is the naive model. Averaging across all areas and step lengths, the RMSE of the RSEMS forecast improves upon that of the historical mean forecast by 15 percent instead of almost 25 percent. Most striking is the emergence of clear step length effects. On average, the relative RMSE of RSEMS is .78 that of the historical mean forecast for the first five periods, but that relative accuracy deteriorates to .92 in the last five quarters of forecast and is at 1.0 for the tenth quarter out. Fifteen of the nineteen MSAs similarly exhibit at least a .1 increase in relative RMSE between the two sets of step lengths. The four which defy this trend do form a logical grouping--Fort Walton Beach, Panama City, Pensacola and Tallahassee. In all four of these panhandle metro areas, state and/or federal government play significant roles in their economic bases.

Data in Exhibit 2 do not confirm Ashley's (1983) speculations regarding the accuracy of RSEMS using national macroeconomic forecast inputs. RSEMS remain more accurate than extrapolation of the historical mean more than "a couple of quarters ahead" (Ashley, 1988, p. 363). Three factors may contribute to this outcome. First, the national forecasts are drawn from a different period than those analyzed by Ashley (1983, 1988) and forecast accuracy may have improved. Indeed, such improvement in accuracy is suggested by at least one recent study (McNees, 1992). Second, as noted above, historical prediction accuracy of macroeconomic input variables was used in selecting RSEMS equation specifications and less reliable forecasts may have been excluded. Third, Ashley's study specifically considered ignoring national macroeconomic forecasts. Projecting regional employment growth at its historical mean is equivalent to "ignoring" forecasts for all independent variables in the employment regression. It cannot be ruled out that ignoring the macroeconomic forecasts only might have improved regional accuracy. Although the relative accuracy of RSEMS compared with the historical mean projection do ultimately deteriorate for most areas, the deterioration is not as rapid as might have been expected from Ashley's (1988) study.

In confirmation of the Makridakis et al. (1982) observation that increased sophistication among extrapolative methodologies does not necessarily improve accuracy, the Holt-Winters forecasts do not provide a more challenging evaluative standard for RSEMS than the simple projection of historical mean. Similar to the naive model and in contrast to the historical mean model, there is little evidence of step-length impacts on relative accuracy--neither in the summary measures of Exhibit 3 nor the individual MSA results of Exhibit 2. Strong regional effects, however, clearly characterize the Holt-Winters comparison. Averaging the relative RMSE across all ten step lengths for each MSA, the resultant series of averages has the highest variance of any of the four comparisons. For five areas, the average RSEMS RMSE is 35 percent or more below that of the Holt-Winters forecast (Fort Lauderdale, Lakeland, Orlando, Sarasota and Tampa). In contrast, the Holt-Winters forecasts are more accurate than the RSEMS for Jacksonville and Panama City and startlingly so for West Palm Beach. Again, casual perusal of Exhibit 1 does not suggest any clear factor defining these groups. Interestingly, five of the six largest metropolitan areas are split between the two extremes.

Comparison of RSEMS with the ARIMA alternative is very similar to the historical mean comparison. This is not surprising since the ARIMA forecasts tend toward the latter as the steps are lengthened. Substantial accuracy gains characterize the RSEMS relative to the ARIMA in the first six periods, but relative accuracy then steadily declines. Nelson (1984) observed this same general phenomenon in evaluating macroeconomic forecasts, but in that case the deterioration of the structural model forecasts relative to the ARIMA forecasts began almost immediately in the second quarter out. Both these results contrast with Fullerton's (1989) finding of few forecast horizon effects.

It should be noted that Nelson's observations about the first quarter relative accuracy also apply here. It is somewhat surprising that the ARIMA forecasts do not fare better than the RSEMS for the first quarter forecast. However, in the RSEMS forecasts, the "first quarter out" for regional employment data typically includes preliminary estimates for some national exogenous variables as well as for some other regional variables. Hence, the RSEMS "forecast" for the first quarter out embodies considerable available information about the period and this phenomenon may account for its unexpected accuracy. The same consideration does not, however, apply to the subsequent quarters when the RSEMS forecasts actually improve relative to the ARIMA. Although the ARIMA and historical mean comparisons are generally similar, notable exceptions are provided by the ARIMA forecasts for Fort Pierce and Naples. Both of these are clearly more accurate than the RSEMS forecasts, outcomes which did not characterize the historical mean forecasts for those two

MSAs.

While Exhibits 2 and 3 summarize RSEMS forecasting errors relative to each of the four extrapolative methodologies in turn, Exhibit 4 briefly examines all five simultaneously. Column A in Exhibit 4 reports the percent of the 190 area-step length projections that each method produced the most accurate forecast. Clearly, the RSEMS methodology dominates with almost two-thirds of the observations. For seven of the 19 areas, it was the most accurate at all step lengths (Fort Lauderdale, Fort Myers, Lakeland, Miami, Orlando, Sarasota and Tampa) and it was the most accurate at all step lengths except one in two others (Gainesville and Pensacola). Dominance of the historical mean is clearly a forecast length phenomenon. Eight of the 11 times it outperformed other methodologies occurred at quarter lengths 8 to 10 periods out. In contrast, the Holt-Winters dominance was clearly a region-specific phenomenon. Twenty three of the 26 times it produced the best forecast, it was a forecast in one of three areas -- Jacksonville, Panama City and West Palm Beach. Less clear step-length and/or region patterns characterize the times the ARIMA method produced the lowest forecast error.

Place Exhibit 4 here

Relative root mean square errors calculated in this study cannot be formally tested for significant difference from one. Mizrach (1992) presents critical values if the forecast errors are correlated (and many are in the sample utilized in this study). However, the test does not allow for bias or autocorrelation in the errors. Some bias in the forecasts was noted, but more restrictively, autocorrelation is inevitable in sequential multi-step forecasts which share periods in common. Less formally, we can examine "noticeably" if not "significantly" better. Column B in Exhibit 4 reports the percent of the 190 comparisons that a method which was most accurate had a root mean square error at least 10 percent below that of all other methodologies. Note that the figures do not sum to 100 percent since in almost 40 percent of the cases, one method did not dominate all others by at least 10 percent greater accuracy. Generally, only RSEMS achieved "noticeable" improvement.

More negatively, column C in Exhibit 4 reports the percent of the 190 cases in which a methodology was dominated by all other methods. It highlights the relatively poor performance of the naive model which was less accurate than all other methods in over 50 percent of the comparisons. There was a strong step-length impact here. For 45 of the 57 comparisons for step lengths 8 to 10, the naive model was the least accurate. The poorest performance of the historical mean was also a step-length phenomenon, but in the opposite direction. Twenty-six of the 27 times it was dominated by all other methodologies occurred in the first five quarters of forecast. Again, the Holt-Winters outcome was strongly region-specific. Twenty five of the 44 times it was dominated by all other methodologies occurred across a spectrum of time periods in three metro areas--Lakeland, Sarasota and Tallahassee. Less clear patterns characterized the fewer times the ARIMA methodology was dominated and the RSEMS forecasts were never poorer than all four extrapolative techniques.

In summary, despite recent qualms in the literature, ex ante RSEMS forecasts can often achieve higher accuracy than simpler extrapolative methodologies widely recognized as the appropriate bases for evaluation. The outcome in this study is "often," not "always" and indeed, "always" would be a suspect outcome and suggestive of either absurdly naive extrapolative alternatives and/or an insufficiently diversified set of regions. If "best" method is defined as lowest average RMSE (averaged over all ten forecast horizons) and lowest RMSE for at least half of the ten horizons individually, then all 19 areas have a well-defined best method. It is Holt-Winters for three--Jacksonville, Panama City and West Palm Beach. It is ARIMA for two--Fort Pierce and Naples. It is RSEMS for the remaining 14.

Concluding Remarks

This paper has addressed a long-standing gap in the forecasting literature, evaluation of ex ante regional forecasts. In its choice of extrapolative alternatives, it has also contributed to the debate initiated by Ashley (1983) on the efficacy of RSEMS which depend on inputs from macroeconomic forecasting models. Again in its choice of extrapolative alternatives, it has addressed the untested claims that RSTIMS are a more accurate alternative to RSEMS for regional forecasters. To conclude briefly, (1) the ex ante RSEMS forecasts analyzed in this study generally are more accurate than a range of extrapolative alternatives; (2) forecast error interaction due to satellite modeling arrangements do not seem severe and; (3) RSTIMS forecasters would do well to consider the information available from RSEMS.

To expand upon point (1), accuracy is only one aspect to be considered in choice of regional economic forecasting methodology. RSEMS produce forecasts which are internally consistent for far more than one variable. Furthermore, as Glennon et al. (1987) have noted, the flexibility of RSEMS in conducting alternative scenarios is valuable. Perhaps the question should not be "Are RSEMS more accurate?," but "Do they unreasonably sacrifice accuracy in order to retain alternative scenario flexibility?" Results of the current study would suggest that there is little sacrifice of accuracy and in a number of cases, some clear gain. Whether the extent of gain is worth the additional modeling effort depends of course upon how much the additional accuracy is valued and the degree to which all other forecasts being simultaneously generated in the RSEMS system are necessary. If no alternative scenarios were desired and only employment was of interest, then the marginal gains of some MSA RSEMS over alternative extrapolative techniques could not be justified. In contrast, some MSAs--most particularly in this set Fort Lauderdale, Gainesville, Orlando, Sarasota and Tampa--so consistently defied extrapolative techniques that even if no alternative scenarios were required and employment was the only variable of interest, RSEMS would still be the efficient methodology.

With respect to point (2), it should be re-iterated that the RSEMS models in this study were constructed taking into account the relative accuracy of different national macroeconomic forecasts. Notoriously poorly forecast variables were avoided even if their inclusion would have improved within-sample estimation fits. The results of the current study do not suggest that Ashley's (1983) warning was unimportant. Rather, they suggest that judiciously following that warning permits construction of useful RSEMS.

With respect to point (3), it can be argued that only the simplest class of RSTIMS has been considered in this study and hence no conclusions can be drawn with respect to RSEMS and RSTIMS in general. In this debate, it should be noted that the regional time series literature to date provides no absolute basis for preferring multivariate time series methodologies over univariate techniques and no consistent preferences within the almost unlimited class of multivariate models (see, for example, Hoehn et al., 1984, and Kinal and Ratner, 1986).

For many metropolitan areas, multivariate methodologies requiring extensive time series are precluded. Employment is in fact typically the series with the most available data points. As noted earlier, the ES-202 data used in this paper are restricted to 1975.1 to the present. Returning to the BLS 790 data does not significantly change the situation. Indeed, the BLS 790 data have a longer history for states and the nation, but they are not revised to reflect the changing geographic boundaries of metropolitan areas. Consequently, if consistently defined employment data are needed back just to 1970.1, then it will be available for only approximately a quarter of today's 315 currently defined MSAs.

Ideally, methodological choice is not restricted to a single technique. Rather, the forecasts complement each other and results of the current research clearly suggest room for beneficial combination of forecasts. But some guidance is needed in that effort. What stands out most clearly from the present study is the inability to readily link regional characteristics to forecasting methodology. It would have been difficult to predict that Holt-Winters extrapolation would outperform all other methodologies in Jacksonville, Panama City and West Palm Beach, yet prove to be highly inaccurate for Lakeland, Sarasota and Tallahassee. Even the set of areas noted above which defied all extrapolative methodologies have no obviously unifying traits in common. The same comment applies to almost all MSA groupings which emerged in the discussion of empirical results. Linking regional characteristics to best methodology (or combination of methodologies) obviously requires more complex, multivariate analysis and it is a topic which should be high on the agenda for regional forecasting research. Prerequisite to such investigation, however, is evaluation of ex ante regional forecasts. It is hoped that the current paper encourages more rigorous analysis of this large class of widely used forecasts.

Appendix

Place Exhibit 1A here

Place Exhibit 2A here

Place Exhibit 3A here

Place Exhibit 4A here

References

Ahlburg, D.A. and Land, K.C., "Population forecasting: Guest Editors' introduction," International Journal of Forecasting, 8 (1992), 289-99.

Anderson, P.A., "Help for the regional economic forecaster: vector autoregression," Federal Reserve Bank of Minneapolis Quarterly Review, 3 (1979), 2-7.

Ashley, R., "On the usefulness of macroeconomic forecasts as inputs to forecasting models," Journal of Forecasting, 2 (1983), 211-23.

Ashley, R., "On the relative worth of recent economic forecasts," International Journal of Forecasting, 4 (1988), 363-76.

Bolton, R., "Regional econometric models," Journal of Regional Science, 25 (1985), 495-520.

Conway, R.S., "The Washington projection and simulation model: a regional interindustry econometric model," International Regional Science Review, 13 (1990), 141-65.

Beaumont, P. M. and Isserman, A.M., "Comment," Journal of the American Statistical Association, 82 (1987), 1004-9.

Coomes, P.A., "A Kalman filter formulation for noisy regional job data," International Journal of Forecasting, 7 (1992), 473-81.

Fullerton, T.M., Jr., "A composite approach to forecasting state government revenues: Case study of the Idaho sales tax," International Journal of Forecasting, 5 (1989), 373-80.

Glennon, D., Lane, J. and Johnson, S., "Regional econometric models that reflect labor market relations," International Journal of Forecasting, 3 (1987), 299-312.

Hoehn, J.G., Gruben, W.C. and Fomby, T.B., "Time series forecasting models of the Texas economy: a comparison," Federal Reserve Bank of Dallas Economic Review, (1984), 11-23.

Kinal, T. and Ratner, J., "A VAR forecasting model of a regional economy: its construction and comparative accuracy," International Regional Science Review, 10 (1986), 113-26.

Lesage, J.P., "Forecasting metropolitan employment using an export-base error-correction model," Journal of Regional Science, 30 (1990), 307-23.

Lienesch, T. and Kort, J.R., "The NRIES II multiregional macroeconomic model of the United States," International Regional Science Review, 14 (1992), 255-74.

Makridakis, S. and Hibon, M., "Accuracy of forecasting: an empirical investigation," Journal of the Royal Statistical Society A, 142 (1979), 97-145.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J. Parzen, E. and Winkler, R., "The accuracy of extrapolation (time series) methods: results of a forecasting competition," Journal of Forecasting, 1 (1982), 111-53.

McNees, S.K., "How accurate are macroeconomic forecasts?" New England Economic Review, (1988), 15-36.

McNees, S.K., "How large are economic forecast errors?" New England Economic Review, (1992), 25-33.

Mizrach, B., "The distribution of the Theil U-statistic in bivariate normal populations," Economics Letters, 38 (1992), 163-7.

Nelson, C.R., "The prediction performance of the FRB-MIT-PENN model of the U.S. economy," The American Economic Review, 52 (1972), 902-17.

Nelson, C.R., "A benchmark for the accuracy of econometric forecasts of GNP," Business Economics, 19 (1984), 52-8.

Schmidt, J.R., "Forecasting state retail sales: econometric vs. time series models," The Annals of Regional Science, 13 (1979), 91-101.

Shoesmith, G.L., "The forecasting accuracy of regional Bayesian VAR models with alternative national variable choices," International Regional Science Review, 13 (1990), 257-69.

Shoesmith, G.L., "Co-integration, error correction and improved medium-term regional VAR forecasting," Journal of Forecasting, 11 (1992), 91-109.

Smith, S.K., "Tests of forecast accuracy and bias for county population projections," Journal of the American Statistical Association, 82 (1987a), 991-1003.

Smith, S.K., "Rejoinder," Journal of the American Statistical Association, 82 (1987b), 1009-1012.

Taylor, C.A., "Econometric modeling of urban and other substate areas: an analysis of alternative methodologies," Regional Science and Urban Economics, 12 (1982a), 425-48.

Taylor, C. A., "Regional econometric model comparisons: what do they mean?" The Annals of Regional Science, 16 (1982b), 1-15.

Thomas, R.J., "Method and situational factors in sales forecast accuracy," Journal of Forecasting, 12 (1993), 69-77.

Treyz, G.I., Rickman, D.S. and Shao, G., "The REMI economic-demographic forecasting and simulation model," International Regional Science Review, 14 (1992), 221-53.

Weller, B.R., "National indicator series as quantitative predictors of small region monthly employment levels," International Journal of Forecasting, 5 (1989), 241-47.

Weller, B.R., "Predicting small region sectoral responses to changes in aggregate economic activity: a time series approach," Journal of Forecasting, 9 (1990), 273-81.

Weller, B.R. and Kurre, J.A., "Applicability of the transfer function approach to forecasting employment levels in small regions," The Annals of Regional Science, 21 (1987), 34-43.

West, C.T. and Fullerton, T.M., Jr., "Forecast summary: Florida economy," The Florida Outlook, 18 (1994), 11-28.