PRELIMINARY ESTIMATES OF GROSS STATE PRODUCT

James R. Schmidt and Scott A. Loseke*

I. Introduction

During the past two decades, gross state product (GSP) has become a major indicator of state economic activity. GSP has great conceptual appeal with its definition being "the gross market value of the goods and services attributable to labor and property located in a State", (Renshaw, et.al., 1988, p. 30). Until the late 1980s, GSP was not an official indicator released by a government agency. Rather, it was constructed by approximation techniques traceable to the work of Kendrick and Jaycox (1965). Since there was no official standard of computation, analysts were on their own to compute GSP accounts for their locale as they saw fit. Naturally, a variety of nuances were added to the original computation method. Weber (1979) covers several of these in his survey. Despite the lack of official status, GSP product accounts became very popular and found their way into several branches of research in regional economics. Perhaps the most intensive use of the accounts has occurred in state econometric models. Examples of models built on the cornerstone of GSP include the Mississippi model of Adams, Brooking, and Glickman (1975); the Ohio multiregional model of Baird (1983); and the recent New Jersey model of Weber (1986).

In 1985, the Bureau of Economic Analysis (BEA) released a staff paper describing methods that had been used to create "experimental" estimates of GSP by industry in all states for selected years in the 1963-1977 span (BEA Staff, 1985).[1] In the spring of 1988, a more complete set of GSP estimates for all years in the 1963-1986 span were released (Renshaw, et.al., 1988). At the time of this writing, almost three years have elapsed since the terminal year of the official GSP estimates. This paper analyzes the issue of timeliness in the GSP estimates and proposes, plus evaluates, a strategy for preparing preliminary GSP for years beyond the end of the official reporting period.

The GSP initiative within BEA establishes a new environment for analyses and research efforts dealing with industry level outputs. The new "official" estimates will become the prevailing data for gross product and will supplant the traditional "synthetic" estimates from the Kendrick-Jaycox method or its various extensions. Comments, and perhaps pointed criticism, concerning BEA's measurement methods may appear but the fact remains that the supplier of the driving component in past synthetic GSP estimates, industry earnings, has moved regional accounting forward by preparing a set of official estimates. Competing GSP estimates and measurement methods may continue to be used, but the official BEA estimates have probably already become the standard of comparison. However, as we demonstrate below, there is still a complementary role for synthetic GSP estimates to play in the future.

Official GSP estimates for 1987 and 1988 are being prepared in conjunction with the BEA regional projections program and are to be released in mid-1990. The reporting lag for the 1988 estimates implied by this timetable roughly matches the lag associated with the release of the 1963-1986 estimates made in the spring of 1988. Both releases will have been made approximately one and one-half years after the final year to which the estimates refer. Of course, the reporting lag for 1987 is quite severe at two and one-half years. These reporting lags, while not unprecedented in regional economic data, may seem quite long to data users who need timely estimates. Potential users of the official GSP estimates might find great appeal in the comprehensiveness of the GSP concept and be eager to base analyses and discussions of contemporary economic conditions upon it. However, the prospect of long reporting lags will temper such enthusiasm. For example, suppose a public official was interested in compiling an economic profile of a region and wanted the most recent and broadest measures of state economic activity to be the cornerstone of the profile. If such an effort were to be pursued at the time of this writing, imagine the chagrin on the part of the official when it is discovered that the latest official estimates conclude with 1986.

In addition to general public reporting, the issue of timeliness is also critical to the school of regional econometric modelers that has traditionally been attracted to the GSP accounts. These modelers will face the same problems of timeliness if their models are respecified to use the official GSP estimates. For example, consider a GSP based state econometric model estimated with the currently

available data that ends in 1986. If the model was used to create projections for 1990, the projection horizon is four periods beyond the end of the estimation period. A method that provides preliminary estimates of GSP for years beyond the terminal year of the official data would allow the econometric model's projection horizon to be shortened.

II. A Method for Preliminary GSP Estimates

Many methods for creating preliminary estimates of the GSP accounts could be entertained so some guidelines need to be established. First, the method should be based upon publicly available information that has a reasonable reporting lag. A reporting lag of four months, which matches the reporting lag of data on industry earnings, is likely to be the shortest that can be attained. However, the available set of information having relevance to the GSP accounts may be considered as too incomplete just four months past the end of the year. Other information, with reporting lags longer than four months, may enhance the accuracy of the preliminary estimates. In that event, the decision to sacrifice accuracy gains for timeliness would have to be made. Second, the method should be dependent upon the patterns in the official GSP estimates. This criterion rules out the use of synthetic GSP estimates as the preliminary estimates. However, exploiting the historical differences between official and synthetic GSP estimates remains a possibility. Third, the method should be relatively easy to apply but should also retain an element of empirical rigor. For example, a state econometric model with a GSP equation block could be used to predict GSP values past the terminal year of the estimated equations. However, such a model may not be available, there might not be an interest in creating one, and they are not necessarily easy to work with.

The method proposed here for creating preliminary GSP estimates uses predictions from single equation models that relate official GSP estimates to either the synthetic GSP estimates of the corresponding industry or to selected components of the synthetic GSP estimates. There should be a fairly strong relationship between the official and synthetic GSP estimates for an industry in any given state. Recalling the usual definition of synthetic GSP, it is the ratio of gross domestic product (GDP) and national earnings in the industry multiplied by the state earnings in the industry. Earnings is a dominant component in official GSP so close covariation of the official and synthetic estimates is likely. In addition, GDP and the official GSP of an industry may show similar patterns of movement through time.

While the relationships between official and synthetic GSP can expected to be strong, they will not be

perfect. Synthetic GSP embodies the restrictive assumption that the "State's share of the Nation's nonearnings components of gross product is the same as its share of the Nation's earnings" (BEA Staff, 1985, p. 1). Alternatively stated, the proportions of GSP and GDP in an industry attributable to nonearnings components are constrained to be identical. This constraint cannot be expected to hold but the difference between the nonearnings proportions of synthetic and official GSP may remain relatively stable or change in a systematic way over time. Under such circumstances, synthetic GSP could be a useful predictor of official GSP.

Four prediction models for official GSP will be examined, one pertaining to current dollar GSP and the remaining three pertaining to constant dollar GSP. This division of attention between current and constant dollar magnitudes is made to allow for the uncertain future of reporting lags in industry GDP and deflator data. In each model, official GSP is a function of contemporaneous values of publicly available information having shorter reporting lags. The first model pertains to current dollar GSP and uses state earnings as the functional argument:

$$GSPC_i = f(SEC_i) \qquad (1)$$

where $GSPC_i$ is the official GSP estimate for industry i in current dollars and SEC_i is state earnings in current dollars. The second model uses state earnings as the argument, the third uses the synthetic GSP estimate, and the fourth uses two of the three components of synthetic GSP, state earnings and national GDP:

$$GSP_i = f(SE_i) \qquad (2)$$
$$GSP_i = f(SGSP_i) \qquad (3)$$
$$GSP_i = f(SE_i, NGDP_i) \qquad (4)$$

where, for industry i, GSP_i is the official GSP estimate, SE_i is state earnings, $SGSP_i$ is synthetic GSP_i, and $NGDP_i$ is national GDP, all of these variables being in constant dollar terms. Synthetic GSP contains both of the arguments in (4) but combines them in a nonlinear fashion which may be an unreasonable constraint in terms of eventual predictive efficiency. Each model could be estimated using data from 1963 through the final year of official GSP availability. The model's out of sample predictions generated by contemporaneous data on the argument variables would then serve as the preliminary GSP estimates.

Model (1) contains only the state earnings variable, in current dollars, for which data is reported with a lag of four months. Thus, predictions of official GSP in current dollars could be made for the most recent year just four months after its conclusion. In contrast, models (2), (3), and (4) involve either GDP in constant dollars or an implicit industry output price deflator as well as earnings. GDP and the industry output price deflators have traditionally been available with a reporting lag of around one-

half year, so there has been only a minor delay in obtaining constant dollar predictions from models (2), (3), or (4) as compared to obtaining current dollar predictions from (1). However, the release of GDP and the deflators pertaining to 1988 did not follow the traditional schedule. Indications are that the data for 1988 will be released in the spring of 1990. Thus, projections of GSP from models (2), (3), and (4) will be delayed by around nine months and carry a reporting lag of over one year. The next release of official GSP is anticipated in mid-1990 so using the projections from the model(s) as preliminary estimates of GSP would fill a several month void of GSP information pertaining to 1988.

The reporting lag for industry GDP and the deflators will also impact the timing of official GSP releases. Up to six months are required to complete official GSP estimates after industry GDP figures for the corresponding year(s) are finished. If GDP reporting returns to its traditional one-half year lag, then the official GSP figures could attain a reporting lag of one year or slightly less were an annual release program to be put in place. However, if GDP reporting incorporates the recent delays on a permanent basis, then the reporting lag for official GSP would stretch to between one and one-half and two years under an annual release program.

An overview of the release date situation and timing of the respective modeling strategies is given in Table 1. The table is for illustration purposes and in no way claims to represent definitive intentions or schedules of official releases. Release schedules will no doubt continue to change. However, the schedules shown are sufficient to reveal the problems and strategies surrounding the creation of preliminary GSP estimates. A four-year reporting span of 1989-1992 is segmented into quarters for portrayal of release dates. The span could be further segmented into months but the use of quarters simplifies the exposition. Numbers under columns two through five refer to the year(s) for which data are being released during the particular quarter (row). The table reflects an optimistic assumption that the annual release of industry GDP and deflator data will occur in the third quarter from 1990 onward. The first number of each list under the last column also refers to the year for which data are being released. Trailing numbers in the lists under column six refer to the model types that would be providing the GSP predictions and to the prediction horizon (one or two years) being encountered, respectively. Two release schedules for official GSP, annual and biennial, are presented in response to the uncertain future of the release frequency. Again, we stress that the schedules are not official and are for illustration purposes only. The table shows that prediction models based solely on earnings can provide preliminary GSP estimates, in current dollar terms, in the second quarter of the succeeding year. Models like (2), (3), and (4), being based on GDP information as well as earnings, can provide preliminary constant dollar GSP estimates in the third quarter of the succeeding year.

The prediction horizons faced are dependent upon the release frequency adopted for official GSP. Under an annual release frequency, the prediction horizon for the models can be kept to one year. For example, official GSP for 1987 and 1988 and earnings data through 1989 would be available in the third quarter of 1990. Any of the models (1) through (4) could be estimated immediately upon receipt of official GSP and used to create the one-step ahead prediction (preliminary estimate) of GSP for 1989. It follows that the preliminary GSP estimates so obtained would lead the release date of official 1989 GSP by one-half year. Under a biennial release frequency for official GSP, the prediction horizon becomes a mixture of one and two years. In the second quarter of 1991, the latest available year for official GSP will be 1988 and the earnings data for 1990 will have been released. Model (1) could then be used to compute the two-step ahead prediction (preliminary estimate) of current dollar GSP for 1990. Models (2), (3), and (4) could not be used to make the two-step ahead predictions of constant dollar GSP for 1990 until the third quarter of 1991. The reporting frequencies of the various series in the prediction models may well change over time, causing some realignment of the timing patterns in Table 1. However, the creation of preliminary GSP estimates will continue to be a prediction problem having horizons of one year and two years, the latter being relevant in the event that a biennial release frequency is adopted for official GSP.

III. Accuracy of Preliminary GSP Estimates

The prediction accuracy of models (1) through (4) will be assessed in this section. Four states that comprise a diverse set of regional economies have been selected for analysis: Florida, Nebraska, Pennsylvania, and Texas. For each state, the models are estimated at the one-digit industry level for a sequential set of seven time spans, all beginning with the initial year for which official GSP estimates are available, 1963. The time spans end in 1979 through 1985, respectively. The models estimated over the six time spans ending in 1980 through 1985 are used to create six predictions of the one year ahead variety in each industry. Mean absolute percentage error (MAPE) is computed from the six predictions and used here as the evaluation measure for the prediction model. The models estimated over the six time spans ending in 1979 through

Table 1
Releases of Annual Data and Timing
of Preliminary GSP Estimates

Release Date (Quarter)	Earnings (Current $)	Gross Domestic Product	Gross State Product (Annual)	Gross State Product (Biennial)	Preliminary Gross State[a,b,c] Product
89I	—	—	—	—	—
89II	88	—	—	—	88[*], 1, 2
89III	—	—	—	—	—
89IV	—	—	—	—	—
90I	—	88	—	—	88, 2, 2
90II	89	—	—	—	89[*], 1, 3
90III	—	89	87, 88	87, 88	89, 2, 1
90IV	—	—	—	—	—
91I	—	—	89	—	—
91II	90	—	—	—	90[*], 1, 1 or 2
91III	—	90	—	—	90, 2, 1 or 2
91IV	—	—	—	—	—
92I	—	—	90	89, 90	—
92II	91	—	—	—	91[*], 1, 1
92III	—	91	—	—	91, 2, 1
92IV	—	—	—	—	—

[*]Current dollars.

[a]The first number in the list is the final year for which preliminary GSP estimates can be prepared.

[b]The second number indicates the method by which preliminary estimates are made. Method 1 uses earnings exclusively while method 2 uses earnings and gross domestic product.

[c]The third number indicates the prediction horizon of the number of years ahead from the end of the official GSP estimates.

1984 are used to create six predictions of the two year ahead variety. MAPE is then computed for each set of six predictions.

Farming is included in the set of one-digit industries to be modeled. However, the traditional method of computing farm GSP has used a value added accounting framework with data from the U.S. Department of Agriculture. The necessary data has usually been available with a reporting lag of less than a year so the traditional method could continue to be used to prepare preliminary GSP estimates for farming in advance of official BEA releases.

Each of the four models are specified as linear but all variables are subjected to the transformation of Box and Cox [1964] with the exception of the unit variable to which the constant term applies. For example, model (4) becomes:

$$GSP_{it}(\lambda) = \beta_1 + \beta_2 * SE_{it}(\lambda) + \beta_3 * NGDP_{it}(\lambda) + e_{it},$$

where each variable receives the transformation:

$$Y_{it}(\lambda) = (Y_{it}^{\lambda} - 1)/\lambda \qquad \lambda \neq 0$$
$$\ln Y_{it} \qquad \lambda = 0.$$

If λ equals one, then a model that is linear in the variables is implied. Thus, the standard linear models in the levels

and logarithms of the variables are special cases. Use of the Box-Cox transformation allows us to abstract from concerns over competing functional forms for the predictive models. The transformation has been used in other regional econometric applications, for example in revenue forecasting by Chang (1979). Allowance is made for first-order autocorrelation in the error terms of the models by using the maximum likelihood algorithm of Savin and White (1978) for estimation. Estimation results such as parameters, significance tests, and model fit are not presented due to the large number of models estimated. Each of the one-digit industries in a given state had models (1) through (4) estimated for it using seven different sample periods. Model fit was excellent in the vast majority of estimations, with coefficients of determination above ninety-five percent being the rule rather than the exception.

Table 2 contains the MAPE values for the one year ahead and two year ahead prediction errors of current dollar GSP, by state and by industry, from model (1). Recall that the model relies exclusively upon current dollar earnings as the predictor, so the specification is quite

Table 2
Mean Absolute Percentage Error of Current Dollar GSP
Predictions From Model (1), 1981-1986*

	One Year Ahead				Two Years Ahead			
	Florida	Nebr.	Penn.	Texas	Florida	Nebr.	Penn.	Texas
Farming	9.13%	8.62%	11.35%	8.64%	9.20%	11.43%	12.21%	12.29%
Ag. Services	1.69	10.58	3.24	3.68	1.90	11.58	5.23	3.43
Mining	24.28	4.95	6.15	34.32	27.33	5.93	6.45	20.03
Construction	4.82	2.53	2.29	2.41	8.72	4.52	4.66	5.17
Manuf. Durables	2.68	4.94	1.89	5.36	3.96	5.45	3.38	2.07
Manuf. Nondurables	4.21	2.47	2.67	3.86	7.43	3.78	3.61	6.09
Trans., Comm., Util.	3.34	4.26	8.24	12.04	7.16	6.29	8.28	11.70
Wholesale Trade	2.80	1.77	1.44	2.23	2.89	2.50	1.74	2.86
Retail Trade	2.38	1.97	0.87	3.41	1.72	2.70	1.13	3.34
Fin., Ins., Real Est.	14.23	3.26	9.05	3.72	9.74	6.40	12.15	5.91
Services	0.77	1.03	0.81	0.71	0.76	0.77	1.05	0.52
Fed. Government	0.57	4.66	0.57	0.67	1.13	13.45	0.96	1.34
S & L Government	0.58	0.81	0.42	0.49	1.11	1.71	0.83	0.95
Total Nonfarm	1.26	1.60	1.94	1.62	2.51	2.82	3.15	2.50
Total Nonfarm (Sum)	3.85	1.81	2.46	4.22	3.52	3.11	2.89	2.37

*Model (1) uses earnings, in current dollars, as the predictor variable of current dollar GSP.

cryptic. Predictions of nonfarm GSP were prepared in two ways: from the models of the nonfarm GSP aggregate and, second, by summing the predictions from the nonfarm industry models.

The MAPE values in Table 2 show a mixed performance of the predictions from model (1). There is wide variation in the degrees of accuracy across states and across industries. Large errors are made in several instances with mining having particularly large errors in Florida and Texas during the evaluation period. Gross product in mining has been quite erratic during the early portion of the 1980s and difficult to predict. Predictions in the wholesale and retail trade industries perform relatively well across the states but there is certainly room for improvement. The services industry is a bright spot with the MAPE values falling below one percent in three of the four states. State and local government predictions also perform well. In six of the eight cases, nonfarm GSP is predicted more accurately as an aggregate from its own of model than from summing the predictions of the respective nonfarm industries. Accuracy deteriorates in most industries as we move to the two year ahead predictions. Yet, there are some curious exceptions. For example, prediction accuracy in Texas for the nonfarm industry total obtained by summing industry predictions is substantially improved versus the one year ahead predictions. Such anomalies are not unexpected when dealing with short estimation and evaluation periods for the models. On balance, it would appear that the use of

earnings to prepare preliminary estimates of current dollar GSP, by industry, for the year concluded just four months previously is a risky venture at best. However, preparing estimates of the nonfarm total for current GSP on this schedule does hold promise.

The performance of predictions for constant dollar GSP from model (2) is shown in Table 3. Recall that model (2) is the constant dollar version of model (1), using only earnings as a functional argument. Several of the state-industry cases that had rather large MAPE values under model (1) now have more moderate values under model (2). Mining and TCU in Texas, FIRE in Florida, and TCU in Pennsylvania are cases showing substantial improvement. However, prediction accuracy deteriorates for several state-industry cases. MAPE for agricultural services increases for all four states and there are other scattered cases of poor performance. Accuracy in the total nonfarm GSP from summing industry predictions improves in three states with Nebraska showing less accurate predictions in the move from current to constant dollar GSP.

Table 4 contains the MAPE values from model (3) that uses synthetic GSP as the lone functional argument. Use of synthetic GSP allows the predictions to account for swings in the ratio of official GSP to GDP that accompany reallocations of GDP across states. Reallocation may occur slowly over time or may be abrupt in response to major movements in the shares of national industrial capacity from one state to another. For example, opening of

The Review of Regional Studies

Table 3
Mean Absolute Percentage Error of Constant Dollar
GSP Predictions from Model (2), 1981-1986*

	One Year Ahead				Two Years Ahead			
	Florida	Nebr.	Penn.	Texas	Florida	Nebr.	Penn.	Texas
Farming	8.50%	10.52%	5.98%	9.75%	9.58%	14.28%	9.39%	14.67%
Ag. Services	4.45	11.29	7.88	6.38	5.87	15.72	7.92	8.01
Mining	17.97	4.49	6.10	4.55	26.29	6.20	10.11	5.88
Construction	1.92	1.77	1.75	1.06	4.61	2.05	3.59	1.76
Manuf. Durables	3.34	5.70	2.51	3.18	5.48	6.74	5.25	3.27
Manuf. Nondurables	6.35	4.00	2.64	2.99	9.36	7.66	3.21	4.29
Trans., Comm., Util.	2.78	4.23	3.79	6.08	5.60	5.80	7.01	10.28
Wholesale Trade	1.92	1.88	1.38	2.17	2.73	2.86	1.38	2.69
Retail Trade	1.13	1.64	3.48	0.76	1.45	2.86	3.11	1.42
Fin., Ins., Real Est.	7.08	3.52	6.24	2.42	8.24	6.97	10.30	3.92
Services	0.81	1.16	0.79	0.68	0.89	0.99	1.10	0.46
Fed. Government	0.96	6.54	1.94	0.90	1.59	15.75	3.89	1.60
S & L Government	0.58	0.66	0.58	0.55	0.92	1.26	0.89	0.86
Total Nonfarm	0.96	1.72	1.86	1.25	1.76	3.03	3.04	2.18
Total Nonfarm (Sum)	2.41	2.15	1.84	0.83	3.64	3.69	2.64	1.09

*Model (2) uses earnings, in constant dollars, as the predictor variable of constant dollar GSP.

Table 4
Mean Absolute Percentage Error of Constant Dollar GSP Predictions from Model (3), 1981-1986*

	One Year Ahead				Two Years Ahead			
	Florida	Nebr.	Penn.	Texas	Florida	Nebr.	Penn.	Texas
Farming	8.77%	5.06%	6.63%	13.18%	9.59%	5.52%	7.17%	18.33%
Ag. Services	2.46	8.56	3.49	2.43	2.78	11.75	3.06	2.47
Mining	21.51	4.75	5.41	4.46	30.06	10.69	9.93	7.39
Construction	2.12	1.29	1.93	1.24	4.81	1.63	4.04	2.16
Manuf. Durables	2.65	3.17	2.12	1.25	2.79	5.11	3.44	2.08
Manuf. Nondurables	1.94	2.00	0.59	2.88	3.88	3.28	0.80	5.83
Trans., Comm., Util.	0.89	1.87	3.45	3.09	1.76	2.27	2.97	5.61
Wholesale Trade	1.45	0.89	1.09	1.12	2.99	1.47	1.30	1.74
Retail Trade	0.36	0.67	0.73	0.50	0.63	0.57	1.30	0.81
Fin., Ins., Real Est.	3.27	2.43	5.04	6.33	4.38	4.39	7.55	10.23
Services	0.75	0.57	0.54	0.60	0.93	0.59	0.64	0.77
Fed. Government	0.61	6.39	1.36	0.73	1.20	15.21	2.60	0.55
S & L Government	0.48	0.46	0.47	0.50	0.49	0.98	0.65	0.45
Total Nonfarm	0.30	1.73	0.97	1.39	0.44	1.59	1.47	2.83
Total Nonfarm (Sum)	1.10	0.85	0.94	1.13	1.86	1.46	0.97	1.91

*Model (3) uses synthetic GSP, in constant dollars, as the predictor variable of constant dollar GSP.

the Alaska pipeline caused a rather abrupt reallocation of mining GDP across states. If the official portion of GDP accruing to a state suddenly changes, it may well be accompanied by a corresponding change in the ratio of state to national earnings. Thus, the value of synthetic GSP would change and the model's predictions would be moved in the appropriate direction.

When the values in Table 4 are compared with those of Table 3, in most cases there is noticeable improvement in predictive accuracy from using model (3) versus model (2). Thus, incorporation of national information in the nonlinear fashion dictated by the synthetic GSP formula seems to contribute positively, in general, to the prediction effort. There are a few exceptions in the comparison where accuracy deteriorates. FIRE in Texas and mining in Florida stand out in particular in the one year ahead predictions while federal government in Pennsylvania, and mining in both Florida and Nebraska, show accuracy losses in the move from earnings to synthetic GSP. These exceptions are far overshadowed by the improvement in accuracy in many of the state-industry cases. Selected cases enjoying dramatic improvement include both the manufacturing durables and nondurables industries in all states except Texas and for both horizons, the TCU industry in all states and for both horizons (with the exception of Pennsylvania in the one year horizon), and retail trade in all states and for both horizons. Total nonfarm GSP is predicted more accurately by model (3) than model (2),

using either the model of the aggregate or the summed industry predictions, with the exceptions of Texas, and Nebraska at the one year horizon. The MAPE measures for nonfarm GSP are quite low in several of the cases, reaching 0.30% and 0.44% for the one year and two year ahead predictions of the Florida aggregate, respectively. For the summed prediction of nonfarm GSP, the minimum MAPE measures are 0.85% and 0.97%, applying to Nebraska and Pennsylvania, respectively.

The results in Table 4 demonstrate the usefulness of incorporating national information into the prediction effort. However, model (3) incorporates the information in a nonlinear fashion as dictated by the synthetic GSP formula. The nonlinear restriction embedded in the information might seem a bit peculiar from the point of view of building a predictive model. Thus, model (4) was specified to include GDP as a separate functional argument, freed from the nonlinear restriction involving state earnings. National earnings was not included as a third explanatory variable since it is highly correlated with state earnings and GDP.

Table 5 contains the MAPE values for the predictions from model (4). Prediction accuracy in many of the state-industry cases is roughly comparable to that shown by model (3) in Table 4 although in some cases there is noticeable deterioration in accuracy. For example, manufacturing durables in Florida, Pennsylvania, and Texas have much higher MAPE values at both horizons. Yet,

Table 5
Mean Absolute Percentage Error of Constant Dollar
GSP Predictions from Model (4), 1981-1986*

	One Year Ahead				Two Years Ahead			
	Florida	Nebr.	Penn.	Texas	Florida	Nebr.	Penn.	Texas
Farming	5.36%	6.32%	4.06%	3.80%	5.36%	9.40%	6.38%	3.32%
Ag. Services	4.47	8.72	6.48	4.85	5.90	10.26	6.92	5.99
Mining	20.38	5.92	11.19	1.30	27.94	7.03	9.81	2.80
Construction	1.59	1.95	2.52	3.88	5.40	1.70	4.64	7.78
Manuf. Durables	5.48	2.40	2.82	1.91	8.35	2.27	4.87	2.47
Manuf. Nondurables	4.09	2.64	1.58	1.92	5.93	5.47	1.62	3.79
Trans., Comm., Util.	1.23	1.46	1.20	3.84	1.93	1.52	2.81	6.68
Wholesale Trade	1.79	1.61	1.20	2.48	2.69	3.15	2.42	4.54
Retail Trade	1.44	0.50	1.12	2.22	1.58	1.09	2.16	2.39
Fin., Ins., Real Est.	2.61	2.44	2.90	3.88	3.58	4.65	4.69	4.68
Services	0.81	0.78	0.56	0.92	0.83	0.61	0.52	1.27
Fed. Government	1.28	6.84	1.83	0.76	2.28	16.01	3.65	1.39
S & L Government	0.59	0.77	0.65	0.52	0.91	1.51	0.69	0.79
Total Nonfarm	0.96	0.86	1.40	1.08	1.78	1.50	2.22	1.65
Total Nonfarm (Sum)	1.31	0.51	0.59	0.64	2.23	0.89	1.00	1.20

*Model (4) uses earnings and gross domestic product, both in constant dollars, as the predictor variables of constant dollar GSP.

there are counterbalancing cases of accuracy improvement in moving to model (4) from (3). FIRE in all the states except Nebraska and at both horizons are such cases. As concerns nonfarm GSP, a pattern of superiority of model (3) over model (4) does emerge for Florida at the one year horizon, and for Florida and Pennsylvania at the two year horizon. Model (4) is the superior predictor of nonfarm GSP for Nebraska and Texas at the one year horizon, and for Texas at the two year horizon.[2]

Application of the above methods to the two-digit industrial level is an obvious extension to consider in light of the successes in the one-digit industries. As industry aggregation moves from the one-digit to two-digit level, more variation in output levels usually occurs and the ratio constraints explained earlier involving nonearnings components of synthetic GSP become more suspect. Thus, predictions for two-digit industries may not achieve the accuracy levels of the one-digit industries. We confine our work to model (4) at the two-digit level and have chosen to experiment with ten industries from the mining, manufacturing, and FIRE group at the one-digit level. Table 6 contains the MAPE values of the predictions. Several state-industry cases show extremely poor performances with oil and gas extraction in Florida being the worst. The official GSP time series for this case behaves very erratically during the estimation period with several year-to-year changes being above fifty percent. Several of the industries turn in reasonably good prediction performances at the one year horizon, notably food and kindred products and insurance carriers, although there is deterioration in accuracy when moving to the two year horizon. On balance, the performance of the two-digit predictions falls short of that achieved by the one-digit predictions.

IV. Summary

The need for preliminary estimates of GSP stems from the present and future reporting lags of official GSP. These lags can be over two years in length. Observers of state economies and public policy makers need information that is broad in coverage and that can provide measurement of an economy's position with minimal delay. No measure is available that provides a more broad description of a state economy's status than GSP. Thus, methods to produce GSP estimates in advance of official releases are worthy of study and, if used with care, are in the public interest.

We have proposed a strategy to create preliminary GSP estimates from the predictions of models based upon state and national industry earnings data and GDP. These publicly available series are released at a shorter reporting lag than GSP and can be used to form predictions of GSP in advance of official releases. Our strategy causes the current dollar GSP reporting lag to match the state earnings reporting lag (four months) and the constant dollar GSP reporting lag to match the GDP reporting lag. Several model specifications were entertained and their accuracies assessed. While predictions of nominal GSP can be created with a reporting lag of just four months, caution is urged. Several industries in the four states chosen here showed rather large prediction errors in the current dollar GSP model. Accuracy of the predictions improved in the constant dollar GSP models that employ national earnings and earnings in tandem with GDP data. Even so, certain industries such as mining and FIRE proved to be formidable challenges for the prediction models. Overall, the prediction accuracy of the models for constant dollar GSP

Table 6

Mean Absolute Percentage Error of Constant Dollar GSP Predictions
for Two-Digit Industries From Model (4), 1981-1986*

	One Year Ahead				Two Years Ahead			
	Florida	Nebr.	Penn.	Texas	Florida	Nebr.	Penn.	Texas
Oil and Gas Extraction	32.04%	10.42%	10.24%	2.43%	50.97%	15.97%	17.09%	6.38%
Nonmetallic Minerals	6.57	10.08	4.93	10.74	11.30	15.78	8.65	19.74
Fabricated Metals	5.46	4.44	4.21	9.35	10.89	4.07	7.00	17.02
Electric Equipment	4.62	9.55	4.10	7.50	6.92	13.41	6.63	8.40
Food/Kindred Products	2.58	3.93	.55	3.18	4.40	4.91	1.05	6.12
Chemicals	3.94	4.62	1.94	4.67	7.74	5.00	3.64	9.16
Apparel	2.95	8.82	4.25	6.87	5.41	11.07	7.43	12.13
Printing and Publishing	3.63	4.29	1.62	5.30	5.82	5.27	3.14	13.96
Banking	5.67	5.20	3.08	4.23	7.92	9.37	4.65	5.60
Insurance Carriers	2.59	1.71	2.39	3.80	4.15	3.19	2.80	7.29

*Model (4) uses earnings and gross domestic product, both in constant dollars, as the predictor variables of constant dollar GSP.

based on both state and national information shows that our approach is promising for the majority of one-digit industries and the nonfarm GSP total. Further extensions may prove worthwhile as well. For example, others may wish to expand the model specifications, perhaps by including lags of the involved variables or entertaining other variables, in hope that prediction accuracy might be enhanced.

Notes

[1]GSP is the sum of earnings (employee compensation and proprietors' income), indirect business tax and nontax charges, rental income of persons, corporate profits with inventory valuation adjustment, net interest, business transfer payments, subsidies less current surplus of government enterprises, and capital consumption allowances.

[2]At the suggestion of a referee, a model relating GSP in constant dollars to work force employment (job counts) was estimated for each of the one-digit nonagricultural industries of Nebraska. The MAPE values for the one and two year ahead predictions across the industries roughly matched the performance achieved by model (2) but were higher than the values registered by models (3) and (4). Thus, predictive models based on employment do not offer improvements over the models based on earnings and GDP, at least in the time period analyzed here.

References

Adams, F. Gerard, Carl G. Brooking, and Norman J. Glickman. "On the Specification and Simulation of a Regional Econometric Model: A Model of Mississippi." Review of Economics and Statistics. 57 (1975). 286-298.

Baird, Catherine Ann. "A Multiregional Econometric Model of Ohio." Journal of Regional Science. 23 (1983). 501-516.

Box, George E. P. and David R. Cox. "An Analysis of Transformations." Journal of the Royal Statistical Society, Series B. 26 (1964). 211-243.

BEA Staff. "Experimental Estimates of Gross State Product by Industry." Bureau of Economic Analysis Staff Paper 42. National Technical Information Service. (May, 1985).

Chang, Hui S. "Functional Forms and the Estimation of a State's Sales Tax Revenue: Tennessee as a Case Study." Annals of Regional Science. 14 (1979). 8-18.

Kendrick, John W. and C. Milton Jaycox. "The Concept and Estimation of Gross State Product". Southern Economic Journal. 32 (1965). 153-168.

Renshaw, Vernon, Edward A. Trott, Jr., and Howard L. Friedenberg." Gross State Product by Industry, 1963-1986". Survey of Current Business. 68 (May 1988). 30-46.

Savin, Neil E., and Kenneth J. White. "Estimation and Testing for Functional Form and Autocorrelation". Journal of Econometrics. 8 (1978). 1-12.

Weber, Richard E. "A Synthesis of Methods Proposed for Estimating Gross State Product". Journal of Regional Science. 19 (1979). 217- 230.

_____. "Regional Econometric Modeling and the New Jersey State Model." In M. Ray Perryman and James R. Schmidt (eds.). Regional Econometric Modeling. Boston, Mass.: Kluwer-Nijhoff, 1986.