An evaluation of State projections of industry, occupational employment

Analysis of the first projections by States using BLS occupational employment data identifies a number of causes of projection errors, and offers suggestions for improving the projections procedures

HARVEY A. GOLDSTEIN AND ALVIN M. CRUZE

State Employment Security Agencies develop and publish statewide and substate industry and occupational employment projections to help meet the information needs of planners and administrators in vocational education, Job Training Partnership Act programs, educational counseling, private sector training programs, and government economic development agencies. Almost all States now use the Occupational Employment Statistics (OES) program of the Bureau of Labor Statistics for the development of their projections. The methodological core of the Bureau program is the industry-occupational (or staffing pattern) matrix produced for each State from the results of the OES survey and other supplementary data.

Because data from the OES survey first became available in 1976, the State agencies had their first opportunity to develop projections using the OES results for the 1976–82 projection round. This article summarizes the results of an evaluation of the accuracy of those projections for 20 States.[1] Based on the evaluation results, we provide some recommendations to improve subsequent rounds of statewide projections.

Evaluation methodology

The basic approach of the analysis was to calculate the projection error by industry and occupation for each State in the sample by comparing the projected 1982 employment levels developed by the respective State agency and the actual 1982 employment levels directly calculated by BLS from State reports. The particular error measure used for each industry or occupation is the adjusted absolute percent error. The average error for various aggregates of industries or occupations is the weighted adjusted mean absolute error.[2] Projection errors were calculated for industries and occupations at all levels of detail. The focus, however, was on 3-digit Standard Industrial Classification (SIC) industry sectors and the most detailed occupational categories.[3]

The evaluation was complicated because many of the 1982 industry employment projections were based on the 1967 SIC coding system, while the actual 1982 industry employment estimates were based on 1972 SIC codes. So that the projected and actual employment data would be comparable, the 1982 industry employment projections were converted to the 1972 SIC code basis using conversion factors calculated from first-quarter 1975, dual-coded data for each State from the Bureau's ES-202 program. But because these conversion factors were more than 10 years old, some error unrelated to the projection error was introduced into the transformed 1982 industry employment projections. To minimize the effect of this spurious error in the evaluation but still retain as many industry sectors as possible to avoid biasing the sample, we deleted all observations for which the difference between the dual-coded employment levels was greater than 15 percent.

Harvey A. Goldstein is an associate professor of planning at the University of North Carolina, Chapel Hill. Alvin M. Cruze is a senior economist at the Research Triangle Institute, Research Triangle Park, NC.

To keep the evaluation manageable, other rules were used to reduce the number of observations involved. For industries, a minimum size cutoff of 500 employees in both the base and projection years was used. The final number of 3-digit SIC industries in the 20-State sample was 3,010; the number of 2-digit industries was 1,120. Occupations with fewer than 50 employees in both the base and projection years were deleted. Also, occupations for which there had been definitional changes between the two years were deleted for reasons of noncomparability. For the remaining observations, a stratified sample of occupations was drawn in each State. Each State sample included one subsample of occupations that were common to all of the States. On average, there were about 120 occupations from each State in the evaluation.[4]

In addition to the procedures and calculations described above, other methods were used for several specific aspects of the evaluation. These are described below, with the respective results.

Accuracy of industry projections

We attempted to explain variation in the magnitude of the projection error among all the industry observations in the sample, rather than focusing on the error magnitude itself. In other words, we wanted to see if there was a pattern to the projection errors that could be explained by different attributes of the industries themselves, by different projection techniques used, or by the economic conditions or other characteristics of the States during the projection period. The results of this approach should serve as a guide to identifying problem industries or occupations in future projection rounds and directing efforts to reduce projection errors for these industries and occupations.

The results indicated, first, that the more detailed the industry category, the larger the error, an intuitively reasonable result. (See table 1.) On average, sampling and reporting errors in the data and nonsystematic events (such as large establishment openings or closings, or strikes) will have larger proportional effects on projection errors at a more disaggregated industry level because of the smaller number of establishments. The projection error by employment size of the industry, with industry detail held constant, showed a similar pattern.

Projection errors varied significantly among major industry divisions. Mining and durable goods manufacturing, which tend to be the most volatile sectors of the economy, had the largest average errors. Wholesale trade, retail trade, and services had the lowest errors.

It had been expected that there would be significant differences in average projection error among the 20 States in the sample. This proved to be the case, but there were no obvious attributes of State economic performance, size, or location that accounted for the differences. No linear relationship was found between average projection error and a State's total employment, census region, total employment growth rate, percent of employment in manufacturing industries, or annual average unemployment rate during the projection period.

The differences in employment growth rates by industry explained by far the largest portion of the variation in projection error. Four industry growth rate categories for the period 1976–82 were formed: (1) −15.0 percent or under; (2) −14.9 percent to −0.1 percent; (3) 0.0 percent to 14.9 percent; and (4) 15.0 percent or over. It is clear from table 1 that if industry employment declined by over 15 percent during the projection period, the error, on average, was about twice the average projection error for all 3-digit SIC industries. However, if an industry experienced modest growth (0.0 percent to 14.9 percent) during the projection period, the projection error was about one-half the average error for all 3-digit industries. If an industry experienced either modest decline or high growth in employment, the projection error tended to be close to the overall average projection error.

There are several complementary interpretations of this result. The first is that the simple time-series regression models or shift-share techniques used extensively by the State agencies in the 1976–82 projection round implicitly assume that the historical employment growth trend will continue into the future. For the majority of industries, the historical data used tended to be for the 1960–76 period, a span characterized by modest but steady employment

Table 1. Industry employment projection error by selected characteristics, 20-State sample
[Error in percent]

Characteristic	Sample size	Mean absolute percent error	Standard deviation[1]	Weighted mean absolute percent error	Standard deviation[2]
Industry level					
Total, all industries	20	6.9	4.3	7.3	4.7
1-digit SIC	157	11.8	11.7	10.6	9.6
2-digit SIC	1,120	16.7	14.5	15.2	13.2
3-digit SIC	3,010	22.6	20.7	19.2	17.8
Industry sector					
Mining	35	32.0	22.0	66.8	24.1
Construction	139	23.5	20.3	20.5	15.6
Durable goods manufacturing	611	30.6	23.3	27.6	20.5
Nondurable goods manufacturing	540	23.4	20.1	20.6	15.7
Transportation	123	23.3	21.5	16.3	15.6
Communications and utilities	100	18.7	21.3	15.7	15.7
Wholesale trade	306	16.9	16.6	14.5	11.6
Retail trade	532	18.4	17.6	14.9	14.3
Finance, insurance, and real estate	208	20.8	19.2	16.8	15.5
Services	416	19.5	20.7	15.3	15.1
Growth rate					
−15.0 percent or less	550	45.7	24.7	39.1	19.4
−14.9 percent to −0.1 percent	591	20.2	12.1	18.5	10.6
0.0 percent to 14.9 percent	641	11.2	9.7	9.3	8.0
15.0 percent or more	1,228	19.3	18.5	19.2	18.6

[1] The standard deviation around the unweighted group mean.

[2] Standard deviation around the weighted group mean.

Table 2. Type of projection error, 3-digit SIC industries, 20-State sample

Type of error	Sample size	Percent distribution	Weighted mean absolute percent error
Total	3,010	100.0	19.2
Predicted 1982 employment > 1976 base year employment; actual 1982 employment > 1976 base year employment	1,778	59.1	16.3
Predicted 1982 employment > 1976 base year employment; actual 1982 employment < 1976 base year employment	956	31.8	29.1
Predicted 1982 employment < 1976 base year employment; actual 1982 employment > 1976 base year employment	91	3.0	21.2
Predicted 1982 employment < 1976 base year employment; actual 1982 employment < 1976 base year employment	185	6.1	16.0

growth in most industry sectors in most States. The employment effects of the structural changes in the U.S. economy, concentrated in the manufacturing industries, had already begun but were not yet large enough to show up in the time-series data as shifts in long-term trends.

The second interpretation is that there may be a systematic, optimistic bias in the projections process—specifically, an unwillingness on the part of analysts to project employment declines. There may indeed be a sincere fear of creating a self-fulfilling prophecy, because economic growth is less likely to occur where markets and overall local economic activity are seen to be stagnant or declining. The results of our evaluation do not confirm this interpretation but they do clearly show the overwhelming tendency for the State agencies to have predicted increases rather than decreases in industry employment for the 1976–82 period. (See table 2.) Employment had been projected to grow in 90.9 percent of the cases but actually did so in only 62.1 percent. Put another way, if employment in an industry sector actually declined, the chances that the decline had been predicted were less than 1 in 6.

Effect of the 1980–82 recessionary period. A third possible interpretation of the relation between industry employment growth rates and projection errors is that the target year of the projections, 1982, was the trough of the deepest national recession since the 1930's. One might then conclude that, except for the unfortunate timing of the 1980 and 1981–82 recessions, the overall projection errors (and particularly the errors for those industries most affected by the recessions) would have been much lower. Moreover, BLS and the State agencies acknowledge that they do not attempt to take into account cyclical fluctuations when making long-term (5- to 10-year) employment projections, but only attempt to project secular trends. For these reasons, we attempted to separate that portion of the projection errors that could be attributed to the recession alone from other sources of error.

A multiple regression model was developed to estimate the effects of the recessionary period on industry employment projection error. The model was fitted to cross-sectional data in which State-level industries were the units of observation. The sample of industries consisted of all 2-digit SIC industries for which monthly CES employment data were available in six sample States.[5] These States were selected, in part, for geographical representation, diversity of State industrial structure, and variation in the statewide severity of the 1980–82 recessionary period. The dependent variable was the projection error for the given industry. The independent variables were the cyclical severity (CS_i) experienced by the State industry during the 1979–82 period; and several control variables, including State industry growth rate category (GROCAT1, GROCAT2, and GROCAT4 as dummy variables), level of employment of the State industry (SIZE), and total State employment (STSIZE).[6] CS_i was measured as the percentage change in industry employment from peak to trough in the 1979–82 period after the trend (linear) component had been removed from the monthly, seasonally adjusted time series. The peak and trough were dated uniquely for each State industry.

The results of the estimated model (in reduced form) are presented below. t-ratios are indicted in parentheses.

Variable	*Parameter estimates* (a_i)
CS	−0.39
	(−7.7)
GROCAT1	24.53
	(7.7)
GROCAT2	6.82
	(2.6)
GROCAT4	−18.46
	(−8.6)
SIZE	−18.21
	(−4.7)
STSIZE	−4.10
	(−3.2)
R^2	0.74
Sample size (N)	183
F-statistic	83.9

The parameter estimates for CS_i indicate that, on average, for every full percentage-point decrease in industry employment due to recessionary conditions alone, the percent projection error increased by 0.39 points.

The parameter estimates then were used to simulate a counterfactual scenario of "no recession" for the full sample of industries and for each subsample by employment growth rate category. These results are shown in table 3. They indicate that both the absolute and relative effect of the recession years on the projection error varied considerably, depending on the growth rate of the industry. The lower the growth rate, the larger the effect of the recessionary period on the projection error. The percentage decline in the percent projection error with "no recession" gets larger with increasing growth rates, except for the highest growth rate category. In the last case, recession conditions actually had the effect of lowering the projection error—that is, had there been no recession, the underprediction in high growth rate industries would have been even larger.

From these results, we infer that while recessionary conditions during the latter part of the projection period had a significant positive effect on the magnitude of the projection errors, they were not the most important factor. Indeed, the evidence from tables 1 and 3 lends support to the hypothesis that forces leading to changes in the long-term employment growth trends of many State industries in the late 1970's were more important in explaining the variation in industry employment projection errors. These structural, rather than cyclical, forces included changes in the international division of labor, the terms of international trade, technological change, rapid movements of capital among U.S. regions, and regional demographic shifts. The industries most affected by these structural changes in the national and international economies were more likely to be those with high rates of employment decline or growth. Because the "turning points" in the long-term employment trends occurred near the end of the historical time series, no statistically based projection models—shift-share, single-regression, or even fully specified econometric models—would have been able to project accurately 1982 employment in those industries affected by structural change. The implications of this plausible interpretation of the results for improving State and area projections are discussed below.

Occupational projections examined

In the OES program, projections of occupational employment are developed by multiplying projections of industry employment by staffing pattern estimates entered into an industry-occupation matrix. This method could lead to two major types of errors in projecting occupational employment: (1) errors in projecting industry employment totals, and (2) errors in projecting the distribution of employment by occupation within an industry—that is, errors in projecting staffing patterns to the target year.

Table 3. Estimated effects of the 1980 and 1981–82 recessions on percent projection error, by 1976–82 industry employment growth rate, 6-State sample

Growth rate category	$\overline{CS}$[1] (1)	$\overline{PCERR}$[2] (2)	PCERR*[3] (3)	Reduction in PCERR (2) - (3) (4)	Percent reduction in PCERR (5)
All industries	-12.2	5.2	0.4	4.8	92.3
−15.0 percent or less	-25.1	38.3	28.5	9.8	25.6
−14.9 percent to −0.1 percent	-19.1	17.3	9.9	7.4	42.8
0.0 percent to 14.9 percent	-9.6	6.8	3.1	3.7	54.4
15.0 percent or more	-5.7	-12.9	-15.1	-2.2	17.1

[1] Average percent decline in industry employment (peak to trough) due to recession.
[2] The average actual percent projection error (not absolute value).
[3] The simulated, "no recession" scenario projection error.

Table 4. Occupational employment projection error, 15-State sample
[In percent]

State	Mean absolute percent error	Standard deviation	Weighted mean absolute percent error	Standard deviation
A	25.3	23.2	13.7	13.6
B	27.1	27.9	14.9	16.5
C	23.9	20.6	16.1	14.3
D	27.6	23.6	16.5	13.6
E	30.7	32.0	16.6	16.9
F	27.4	27.4	17.4	13.8
G	23.7	21.9	17.5	15.3
H	29.6	26.2	18.1	15.2
I	23.5	19.4	18.4	16.1
J	33.8	30.3	19.6	17.5
K	30.9	31.0	19.8	16.4
L	28.0	24.6	19.8	16.2
M	26.3	21.8	20.7	15.9
N	31.7	25.5	20.8	16.5
O	34.3	29.4	22.8	21.1

NOTE: See footnotes to table 1 for definitions of the types of errors.

To evaluate the 1976–82 projections, we first examine the total occupational employment projection error, with particular emphasis on identifying factors that may be associated with systematic variation in the projection errors. Second, the total error is decomposed into (1) errors in projecting industry employment, and (2) errors in projecting staffing patterns within industries. Third, the effects of sampling error in the OES survey on occupational employment projection errors are analyzed. And fourth, the effects of industry and regional aggregation in the OES staffing pattern matrix on projection errors are evaluated.

Total occupational error. Adjusted absolute percentage errors in occupational employment projections for each of 15 sample States are presented in table 4. (Because data for Colorado, the District of Columbia, Kentucky, Missouri, and Oregon were not available, those jurisdictions are exluded from this portion of the analysis.) The weighted average projection error across the State sample is 18.6 percent, while the unweighted average error is 28.8 percent. On an individual State basis, the weighted average errors range from a low of 13.7 percent to a high of 22.8 percent. The unweighted averages range from 23.5 percent to 34.3 percent. In general, there is a high degree of correlation between the two measures. The product moment correlation coefficient is 0.59, while the rank correlation coefficient is 0.53. Both of these correlation coefficients are significantly different from zero at the 95-percent confidence level.

As indicated by the relative magnitudes of the percentage errors and their associated standard deviations, there are no statistically significant differences between these measures across the 15 States in our sample. For this reason, no formal tests of the statistical significance of these differences were made.

The next step in the evaluation was to identify factors that may be associated with systematic differences in the projection errors. In analyzing the relationships between occupational employment projection error and employment level, we formed four size categories of occupational employment: under 1,000, 1,000 to 1,999, 2,000 to 4,999, and 5,000 and over. As shown in table 5, there is a definite inverse relationship between the magnitude of the projection error and the size of occupational employment. The weighted projection error ranged from a high of 37.6 percent for occupations with fewer than 1,000 workers to a low of 16.4 percent for those with employment greater than 5,000. In fact, the results for our 15 sample States indicate that the projection error is a monotonically decreasing function of the size of employment. In addition, the variation in projection error decreased with size of employment.

In contrast to these findings, we noted a U-shaped relationship between projection error and occupational growth rate. As indicated in table 5, occupations with an employment decline greater than 15 percent over the 6-year projection period had the highest mean error—43.4 percent. At the other end of the distribution, occupations with a growth rate in excess of 15 percent had an average projection error of 19.7 percent. The lowest error, 9.2 percent, occurred for those occupations that grew less than 15 percent.

These results indicate that projections for occupations that exhibited significant turning points or changes in growth rates are more likely to be in error, a finding that is consistent with that reported in the evaluation of the accuracy of industry employment projections.

As in the analysis of industry employment projection errors, it is useful to examine an alternative measure of projection error—the extent to which the predicted direction of occupational employment change is the same as the actual direction. Overall, the direction of change was predicted correctly in only 61.8 percent of the cases. (See table 6.) Of these, a large majority (94.4 percent) were instances of correctly predicting increases in occupational employment. Of the cases in which the direction of change was incorrectly predicted, 97.5 percent were predictions of positive change when actual employment declined between 1976 and 1982. Expressed in another way, 95.6 percent of the sample occupations were predicted to have an increase in employment over the 6-year period, while only 59.2 percent actually did so.

Table 5. Occupational employment projection error by selected characteristics, 15-State sample

[Error in percent]

Characteristic	Sample size	Mean absolute percent error	Standard deviation	Weighted mean absolute percent error	Standard deviation
Occupation size					
Fewer than 1,000 workers	490	36.7	31.8	37.6	32.2
1,000 to 1,999 workers	384	32.8	28.4	30.1	24.8
2,000 to 4,999 workers	382	27.0	24.5	25.4	20.9
5,000 workers or more	534	19.9	17.0	16.4	13.6
Growth rate					
−15.0 percent or less	416	57.5	29.6	43.4	18.8
−14.9 percent to −0.1 percent	313	21.1	9.0	19.6	7.6
0.0 percent to 14.9 percent	307	10.1	7.0	9.2	6.1
15.0 percent or more	754	23.8	22.3	19.7	18.5

Decomposition of occupational projection error. The difference between actual and projected occupational employment may be decomposed into two components: the portion due to changes in staffing patterns and the portion due to errors in projecting industry employment. (See the appendix for a mathematical proof of this observation.) The second component can be readily calculated by multiplying the 1982 staffing patterns by errors in projections of industry employment. This component can then be subtracted from the total projection error to provide the portion of the total error due to changes in staffing patterns. These two sources of error can then be averaged across selected industry or occupational groups to identify and analyze patterns of sources of occupational projection error.

As shown in table 7, total projection error for our 1,790 sample occupations was 440,105, or an average of 246 per occupation. The industry component of this error was −185,299, while the occupational component was 625,404. In other words, although total occupational employment was overprojected, the component due to industry employment projections resulted in an underprojection of actual 1982 totals. The absolute value of the occupational component was approximately 3.4 times greater than the absolute value of the industry component, indicating that changes in staffing patterns over the 6-year period were a greater source of error in the occupational employment projections than were errors in projecting industry employment.

However, it should be noted that for the 1982 projection round, none of the States developed projections of staffing patterns. Instead, 1976 State-level staffing patterns were assumed to remain unchanged over the 1976–82 period. The effects of this assumption are vividly illustrated by this decomposition analysis. For later projection rounds, States are constructing projections of their staffing patterns, using change factors developed and estimated by BLS for projecting the national staffing pattern matrix.

By definition, the total projection error will be positive if the direction of error is greater than zero and negative if the direction of error is less than zero. According to the error decomposition, situations in which the direction of error is greater than zero arise more from changes in staffing patterns (average staffing pattern error component = 1,137) than from errors in projecting industry employment (average industry error component = 272). Occupations with a projection error less than zero (that is, actual 1982 employment was greater than the predicted value) were characterized by more equal industry and staffing pattern error components.

In other words, situations in which predicted 1982 employment exceeded actual values were due more to changes in staffing patterns than to errors in projecting industry employment.

OES sampling error. The OES staffing pattern matrices used to develop projections of occupational employment are based on surveys of a sample of establishments in each of the relevant industry sectors. The effects of survey sampling error on projection errors were measured by determining whether the projected values of occupational employment fell within statistically acceptable confidence limits around the actual values. The confidence limits were calculated from parameters of the OES sample survey design.

As indicated in the *OES Survey Manual*,[7] the sample design for the OES survey calls for a complete census of all establishments with more than 100 employees in an industry sector and a sample of the remaining establishments. Given the sample design implemented in each State, the standard error of the number of workers in occupation i in industry sector j, σ_{Eij}, can be readily calculated.[8] Given this standard error, the 90- and 95-percent confidence intervals around the actual 1982 estimate of the number of workers in this occupation in the industry sector can be calculated as follows:

95-percent confidence interval: $E_{ij} \pm 1.96\ \sigma_{Eij}$

90-percent confidence interval: $E_{ij} \pm 1.645\ \sigma_{Eij}$

where E_{ij} is employment in occupation i in industry j, and σ_{Eij} is the standard error of the estimate.

To undertake this analysis, the confidence intervals around the estimates of 1982 employment in individual industry-occupation cells are first computed, using results in the industry-occupation matrix benchmarked to 1982 actual industry employment totals. Projected 1982 employment totals for these cells are obtained by multiplying projected 1982 employment for relevant industry sectors by the (constant) staffing patterns from the 1976 matrix. Because this operation requires the use of an actual 1976 industry-occupation matrix, the analysis is restricted to: (1) the six southeastern States for which sufficient information was available to calculate standard errors; (2) the 59 occupations common to these States; and (3) industry employment projections for 2-digit SIC sectors. We also restricted our attention to occupations with at least 50 employees in the relevant matrix cell in 1982.

The results of the analysis are presented in table 8, in terms of the percentages of 1982 projected values that fall within 95-percent confidence intervals around actual 1982 values. To assist in interpretation, we classified these percentages according to the size of 1982 employment in the cell—50 to 99, 100 to 499, and 500 workers or more—and the year and sector in which the OES survey was conducted—1980, manufacturing; 1981, nonmanufacturing; and 1982, nonmanufacturing.

As indicated in the table, projected employment in 37.9 percent of the 2,479 industry-occupation cells falls within the 95-percent confidence intervals around the respective actual 1982 employment totals, as estimated from 1982 base year industry-occupation matrices developed from the OES surveys. This percentage is higher for the industry cells in the 1980 manufacturing survey (40.3 percent) than for the 1981 nonmanufacturing round (34.1 percent), and lower than for the 1982 nonmanufacturing round (40.0 percent). There is no consistent pattern across the six States when these percentages are broken out by size of employment in the industry-occupation cell.

These percentages do exhibit significant variations across the six States in our sample, however, with the statewide percentages of employment projections falling within the

Table 6. Type of projection error for sample occupations by employment size category, 15-State sample

Occupation size category	Total			Type of error: A[1]			B[2]			C[3]			D[4]		
	Sample size	Percent of total	Weighted mean absolute percent error	Sample size	Percent of total	Weighted mean absolute percent error	Sample size	Percent of total	Weighted mean absolute percent error	Sample size	Percent of total	Weighted mean absolute percent error	Sample size	Percent of total	Weighted mean absolute percent error
Total	1,790	100.0	18.6	1,044	58.3	16.0	667	37.3	27.1	17	0.9	14.5	62	3.5	23.8
Fewer than 1,000 workers	490	100.0	37.6	269	54.9	36.9	189	38.6	41.1	7	1.4	35.1	25	5.1	27.5
1,000 to 1,999 workers	384	100.0	30.1	206	53.6	26.8	160	41.7	39.3	5	1.3	28.8	13	3.4	22.0
2,000 to 4,999 workers	382	100.0	25.4	227	59.4	22.7	141	36.9	33.7	2	.5	19.8	12	3.1	22.0
5,000 workers or more	534	100.0	16.4	342	64.0	13.9	177	33.1	24.8	3	.6	9.9	12	2.2	24.2

[1] Predicted 1982 employment > 1976 base year employment; actual 1982 employment > 1976 base year employment.
[2] Predicted 1982 employment > 1976 base year employment; actual 1982 employment < 1976 base year employment.
[3] Predicted 1982 employment < 1976 base year employment; actual 1982 employment > 1976 base year employment.
[4] Predicted 1982 employment < 1976 base year employment; actual 1982 employment < 1976 base year employment.

Table 7. Decomposition of projection error, total and 13 selected States

State	Industry component of error		Staffing pattern component of error		Total projection error	
	Sum	Mean	Sum	Mean	Sum	Mean
Total	-185,299	-104	625,404	349	440,105	246
A	115,935	641	50,167	277	166,102	918
B	73,762	591	42,569	335	116,331	916
C	22,827	217	37,343	356	60,170	573
D	35,464	246	45,202	314	80,666	560
E	18,273	228	25,652	321	43,925	549
F	37,770	420	8,482	94	46,252	514
G	58,761	470	4,924	39	63,685	509
H	28,955	252	17,647	153	46,602	405
I	-153,027	-1,034	212,210	1,434	59,183	400
J	16,405	256	4,116	64	20,521	321
K	-392,395	-2,192	73,010	408	-319,375	-1,784
L	-311	-31	27,687	243	27,376	240
M	-41	-1	4,062	88	4,021	87

95-percent confidence interval ranging from 30.8 percent to 44.9 percent. On an individual State basis, there is no consistent pattern in these percentages across either survey rounds or size of employment in the matrix cells.

In general, these results indicate that it is extremely difficult to project employment for a given occupation in a particular industry sector with an acceptable degree of statistical precision. Factors such as small sample sizes and low response rates in the 1980–82 OES surveys result in wider confidence intervals, with a greater proportion of the projected values falling within these intervals. Conversely, recalling that the 1982 projected values were developed under the assumption of constant staffing patterns over the 6-year period, we would expect that industries undergoing rapid technological change would have a larger percentage of predicted values falling outside the confidence intervals around the 1982 estimates of actual employment. From available data, it is difficult to separate the effects of these two factors. The relative percentages for the manufacturing and nonmanufacturing rounds are, however, in the expected directions. Most likely, OES survey sampling frames are better developed and occupational titles and duties are better defined and understood in the manufacturing sector. Other things equal, each of these factors is expected to produce a higher proportion of projected values within our confidence limits in the manufacturing sector, which was indeed the case for the six States in this analysis.

Effects of aggregation

By industry. Table 9 presents a comparison of the weighted projection errors for the original, completely detailed matrix and for the 2-digit SIC level of industry aggregation. As indicated, all seven southeastern States are ranked in order of increasing weighted prediction errors calculated from the full matrix. Across the seven States, the weighted projection error increased by only 0.4 percentage points when the 2-digit industry matrix was used in place of the full matrix. Differences for individual States are also relatively small, the largest being 1.3 percentage points.

A number of factors account for these small differences. First, although the full matrices contain approximately 400 industry sectors per State, employment data are available only at the 2-digit level of detail for some of the sectors (such as government, education, and eating and drinking places). These sectors contain relatively large proportions of total employment. In fact, for the 59 common occupations across the seven southeastern States, 1976 employment in the industry sectors having only 2-digit level of detail accounted for an average of 26.9 percent of total employment. Therefore, slightly less than three-fourths of employment in these occupations can even by affected by the industry aggregations.

The second factor is that employment in the remaining 2-digit sectors may be concentrated in a single 3-digit industry. If this is the case, aggregation to the 2-digit level would not have much impact because the industry employment projections and associated staffing patterns would be dominated by the constituent 3-digit industry. This appears to be the case for the States in our sample. For all occupations, 13.2 percent of employment in 2-digit sectors with 3-digit detail is in a single 3-digit industry that accounts for over 75 percent of employment in the 2-digit sector. A total of 27.1 percent of employment is in a 3-digit industry that accounts for over 50 percent of employment in the higher-level sector.

Assuming that employment in our sample occupations follows similar patterns, between 46 percent and 59 percent of employment in the 59 common occupations could be affected by changes in the level of industry aggregation. With such distribution of industry employment across 2- and 3-digit sectors, it is not surprising that the projection errors from the 2-digit matrices are not significantly larger than those developed from the full matrices.

By region. A single regional matrix was built from staffing pattern data for the individual States and then applied to projected industry employment data for each of the seven southeastern States to develop a second set of simulated occupational projections for 1982. These simulated projections were then compared with projections developed with individual State matrices and actual 1982 occupation employment totals. Table 9 presents a comparison of the weighted projection errors for the 59 common occupations in the southeastern States that were developed from the regional matrix and from fully detailed matrices for each State.

As shown in the table, use of the regional matrix at the 2-digit industry level of detail increases the overall weighted projection error by 0.9 percentage points—from 15.8 percent to 16.7 percent. The effects on the weighted error of using the regional matrix alone are estimated at 0.5 percentage points because, as pointed out in the previous section, the 2-digit matrices yielded a weighted error of 16.2 percent. There is no obvious pattern of differences in projection errors by State, occupational employment size, or occupational employment growth rate when we examine the effects of using the regional matrix in place of the individual State matrices. In one State, the combined use of industry aggregation and the regional matrix increased the weighted average projection error by 3.9 percentage points, of which 2.6 percentage points were due to use of the regional matrix. In another State, however, use of the regional matrix alone reduced the weighted average projection error by 2.2 percentage points. In reviewing these findings, it should be noted that these results will not necessarily hold for any arbitrary selection of States to make up a "region." Both the industry structure and associated staffing patterns should be relatively similar among the States in the region to minimize the possibility of significant differences in individual State projection errors when a regional matrix is substituted for the individual State matrix.

Suggested improvements

The results of this evaluation suggest a number of improvements that can be made to the State-level industry and occupational employment projection process. These improvements can be conveniently organized into two major categories: (1) methods for OES systems design and data collection, and (2) dissemination of projection results.

Methodology. The first recommendation to improve the methodology for developing industry and occupation projections is to make the entire process more analytical and to minimize the mechanical aspects that were prevalent when the 1976–82 State projections were prepared. The greater uncertainties in the national and international economies and markets, the increasing openness of State and substate economies to worldwide developments, a more rapid rate of technological change, and the increasing diversity of economic growth and performance among State and substate areas all require a more analytical approach to developing projections. This exercise of analytical judgment would include, for example, identifying special local factors or conditions that might require adjustment of rates or ratios derived from national data and choosing the most appropriate projection models based upon the validity of their underlying economic assumptions.

While the projection process should not be mechanical, it should still be highly systematic, rather than a series of ad hoc procedures. The process can and should be made analytical and systematic at the same time by recognizing that, at each step, there are choices among alternative procedures, models, or data. Analytical judgment is exercised in choosing the most appropriate option, such that the validity and utility of the projections will be maximized within the constraints of available resources. The judgment and experience of the State Employment Security Agencies' analysts become increasingly important under this approach, and efforts to train and retain these experienced staff should be emphasized.

In facing the reality of restraints on government spending, the State agencies must make difficult choices about how they best can use the limited resources available for developing projections. For example, this may mean setting priorities among industry and occupational groups, because it would not be efficient to spend an equal amount of time developing projections for each detailed industry and occupation. In addition, choices among alternative techniques for particular elements in the projection process should take into account differences in costs. The analyst should con-

Table 8. Projections of 1982 occupational employment falling within 95-percent confidence interval around actual 1982 estimates, by size of occupational group, 6 southeastern States
[In percent]

State	Manufacturing survey (1980)				Nonmanufacturing survey (1981)				Nonmanufacturing survey (1982)				Total
	Occupational employment			Total	Occupational employment			Total	Occupational employment			Total	
	50-99	100-499	500+		50-99	100-499	500+		50-99	100-499	500+		
Total	40.3 (258)	42.2 (384)	27.9 (61)	40.3 (703)	31.6 (247)	34.3 (464)	36.4 (214)	34.1 (925)	36.9 (149)	39.2 (362)	42.4 (340)	40.0 (851)	37.9 (2,479)
A	31.8 (44)	36.5 (85)	16.7 (18)	32.7 (147)	53.3 (30)	43.4 (99)	48.4 (93)	46.9 (222)	31.0 (29)	53.4 (73)	45.6 (90)	46.4 (192)	43.0 (561)
B	36.7 (49)	43.2 (74)	20.0 (16)	38.4 (138)	25.6 (39)	24.7 (89)	23.3 (43)	24.6 (171)	33.3 (21)	32.8 (64)	37.3 (67)	34.9 (152)	32.1 (461)
C	30.0 (30)	31.4 (35)	0.0 (1)	30.3 (66)	20.8 (48)	44.4 (54)	30.0 (10)	33.0 (112)	39.1 (23)	36.0 (50)	35.5 (31)	36.5 (104)	33.7 (282)
D	48.9 (45)	51.2 (86)	38.9 (18)	49.0 (149)	34.0 (47)	27.6 (76)	36.4 (33)	31.4 (156)	50.0 (20)	30.6 (62)	42.6 (61)	38.5 (143)	39.5 (448)
E	53.9 (52)	48.2 (56)	50.0 (4)	50.9 (112)	38.1 (42)	37.3 (75)	30.0 (20)	36.5 (137)	44.8 (29)	48.3 (60)	51.1 (45)	48.5 (134)	44.9 (383)
F	34.2 (38)	35.4 (48)	40.0 (5)	35.2 (91)	24.4 (41)	29.6 (71)	13.3 (15)	26.0 (127)	25.9 (27)	30.2 (63)	39.1 (46)	32.5 (126)	30.8 (344)

NOTE: Number of observations indicated in parentheses.

Table 9. Comparison of weighted mean absolute projection errors for detailed, industry aggregated, and geographically aggregated industry-occupation matrices, 7 southeastern States
[In percent]

State	Weighted projection error		
	Detailed matrix	Industry aggregation	Geographic aggregation
All States	15.8	16.2	16.7
A	12.3	12.3	13.8
B	15.4	16.7	19.3
C	15.4	15.5	16.8
D	15.9	17.1	18.2
E	17.1	17.1	15.9
F	17.6	17.5	15.3
G	18.1	18.1	17.0

sider whether the expected gain in accuracy from using a more sophisticated technique is justified by the increased cost. The maxim here is to use the simplest, least costly technique that "works." At the same time, it is hoped that continued research on and evaluation of the projections process, such as the evaluation summarized in this article, will lead to further innovations that will improve the cost-effectiveness of the projections.

The second recommended improvement is to develop better projections of staffing patterns that in turn will lead to improved occupational employment projections. As indicated above, the absolute value of the occupational component of projection error was approximately 3.4 times greater than the industry component. This finding provides a strong indication that changes in staffing patterns over the 6-year period were a greater souce of error in the occupational employment projections than were errors in projecting industry employment.

For the 1976–82 projection round, none of the States developed projections of staffing patterns. Instead, the 1976 State-level staffing patterns were assumed to remain unchanged over the projections period. The effects of this assumption are vividly illustrated by the findings of the decomposition analysis presented above. And, as noted earlier, for later projection rounds, many States have developed or are developing projections of their staffing patterns, using factors calculated from projections of national staffing patterns prepared by BLS. This type of Federal-State cooperation should be encouraged and expanded to ensure that all States have the capability to develop meaningful projections of staffing patterns.

Dissemination of projection results. The first recommendation for improving the dissemination of projection results is to develop better documentation of the entire process. This recommendation has a number of dimensions: description of results in a clear, straightforward manner; comprehensive documentation of all assumptions underlying the analyses; simple, nontechnical description of methods, accompanied by appropriate technical appendices; and consistent presentation of tabular materials, with appropriate rules for rounding off, suppression of unreliable data, and so forth.

The second suggestion with respect to dissemination of projection results is to include, where suitable, measures of the statistical reliability of the projected values in documentation of the results. This is particularly appropriate in the case of industry employment projections developed from regression models, for which it would be relatively simple to calculate the standard errors of the projected values. General indicators of the reliability of projection results (for example, low, medium, or high) should be devised and presented in the general documentation of projections results. Additional details, including specific values of the standard errors and other statistical properties of the regression equations, can be included in more detailed technical documentation to accompany the main descriptive results.

Finally, the use of OES projection data can be extended by developing improved mechanisms for sharing BLS results among various user constituencies. This information sharing should include both the preview of preliminary projection results and dissemination of final written products. The findings from a users survey component of our study indicated that State agencies and planning staffs are increasingly turning to the OES employment projections for their individual planning needs. More widespread dissemination of both BLS and State projection results, including documentation of their reliability as discussed above, and continuing efforts to improve the quality of the entire OES program should lead to even greater use of projections estimates. In particular, BLS efforts to develop micro-matrix formats for projection results and to disseminate all OES products in these formats should be encouraged. □

—— FOOTNOTES ——

[1] The industry employment projections were evaluated for the following 20 jurisdictions: Alabama, Colorado, Delaware, District of Columbia, Florida, Georgia, Indiana, Kentucky, Maine, Maryland, Mississippi, Missouri, North Carolina, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, and West Virginia. Other portions of the analysis are limited to selected subsets of these jurisdictions because of data availability or other technical reasons.

[2] The adjusted absolute percent error, ADJAPE, for case i is calculated as follows:

$$ADJAPE_i = \frac{|PREDICTED_i - ACTUAL_i|}{0.5\ (PREDICTED_i + ACTUAL_i)} \times 100$$

The weighted adjusted mean absolute percent error, WADJAPE, is calculated as follows:

$$WADJMAPE = \frac{\sum_{i=1}^{N} ADJMAPE_i * ACTUAL_i}{\sum_{i=1}^{N} ACTUAL_i}$$

where
$$ADJMAPE = \frac{\sum_{i=1}^{N} ADJAPE_i}{N}$$

See J. Scott Armstrong, *Long Range Forecasting from Crystal Ball to Computer* (New York, John Wiley & Sons, Inc., 1978), for a detailed discussion of the merits of these and alternative measures of forecasting or projection accuracy.

[3] These are the levels of industry and occupational detail at which the State OES staffing pattern matrices yield occupational employment projections for program planning and decisionmaking.

[4] Complete details of the methods used in this evaluation are provided in Alvin M. Cruze, Harvey A. Goldstein, John E. S. Lawrence, Edward M. Bergman, and Katherine A. Desmond, *Evaluation of Industry and Occupational Employment Projections Made by State Employment Security Agencies*, RTI/2742/01-24F (Research Triangle Park, NC, Research Triangle Institute, 1985).

[5] The six States were Florida, Indiana, North Carolina, Oregon, Pennsylvania, and Texas.

[6] The full specification of the model was:

$$PCERR_i = a_0 + a_1CS_i + a_2GROCAT1_i + a_3GROCAT2_i + a_4GROCAT4_i + a_5SIZE_i + a_6TIMING_i + a_7EXPORT_i + a_8STSIZE_i + a_9STUERATE_i + a_{10}STPCMFG_i$$

where, for industry *i*:

$$PCERR_i = \frac{\text{Predicted } 1982_i - \text{Actual } 1982_i}{\text{Actual } 1982_i} \times 100$$

and:

CS_i is the percent change in industry employment from peak to trough in the 1979–82 period after removing the trend (linear) component. The peak and trough were uniquely dated for each State industry;

$GROCAT1_i$, $GROCAT2_i$, and $GROCAT4_i$ are dummy variables for industry employment growth rate between 1976 and 1982. GROCAT1 = 1 if the growth rate was ≤15.0 percent; GROCAT2 = 1 if the growth rate was between −14.9 percent and −0.1 percent; and GROCAT4 = 1 if the growth rate was ≥15.0 percent;

$SIZE_i$ is a dummy variable for size of State industry. An industry in which employment was less than 500 in the base year (1976) = 1, otherwise = 0;

$TIMING_i$ is a dummy variable that refers to whether the detrended peak of the State industry's employment was before (=1) or after (=0) the U.S. peak for total nonagricultural employment in November 1979;

$EXPORT_i$ is a dummy variable that refers to whether the State industry is primarily export-oriented (=1) or serves a State market (=0). These assignments were based on the magnitude of the location quotient computed for the State industry;

$STSIZE_i$ is a dummy variable for the size of State measured by 1976 total nonagricultural employment, =1, if >2,000,000, =0 otherwise. This is a proxy for the resources and staff available to the State agency for developing projections;

$STUERATE_i$ is a dummy variable indicating whether the State's 1982 average annual unemployment rate was above (=1) or below (=0) the U.S. average unemployment rate;

$STPCMFG_i$ is a dummy variable indicating whether the State's proportion of nonagricultural employment in manufacturing was above (=1) or below (=0) the U.S. proportion.

[7] U.S. Department of Labor, *OES Survey Manual* (Bureau of Labor Statistics, 1975).

[8] The details of this calculation are provided in chapter 5 of the *OES Survey Manual*. It should be noted that these results are restricted to industry sectors surveyed in the regular OES cycle. Sectors such as railroads, education, hospitals, private households, and Federal Government are excluded because their staffing patterns are not obtained from OES sample surveys.

APPENDIX: Error decomposition technique

The approach to decomposing the projection error can be presented in terms of the following notation, where:

- I_A is a $1 \times n$ vector of actual 1982 employment for *n* industry sectors;
- I_p is a $1 \times n$ vector of projected 1982 employment for *n* industry sectors;
- O_A is an $n \times m$ matrix of actual 1982 staffing patterns for *m* occupations in each of the *n* industry sectors (that is, the ratios of employment in each of the *m* occupations in a given industry sector divided by total employment in the industry sector); and
- O_p is an $n \times m$ matrix of projected staffing patterns for *m* occupations in each of the *n* industry sectors.

Note that when the I_A vector is multiplied by the O_A matrix, we obtain a $(1 \times n) \times (n \times m) = 1 \times m$ vector of actual employment in each of the *m* occupations. The following derivations are presented in terms of this vector. However, conclusions will hold for each of the elements (separate occupations) of the vector.

In this notation, the error in occupation projections due to errors in projecting industry employment may be represented by:

$$I_p \cdot O_A - I_A \cdot O_A$$

Similarly, occupational projection errors due to errors in projecting the staffing pattern matrix may be represented by:

$$I_A \cdot O_p - I_A \cdot O_A$$

Adding these two components and simplifying, we obtain:

$$\{I_p \cdot O_A - I_A \cdot O_A\} + \{I_A \cdot O_p - I_A \cdot O_A\} = (I_p - I_A)O_A + I_A(O_p - O_A)$$

Thus, the difference between actual and projected occupational employment may be decomposed into (1) the portion due to changes in staffing patterns, and (2) the portion due to errors in projecting industry employment.