Evaluation of State Industry Employment Projections to 2000

Report prepared for the
Projections Workgroup

Employment and Training Administration
U.S. Department of Labor

Harvey A. Goldstein, Ph.D.*

The University of North Carolina at Chapel Hill
Center for Urban and Regional Studies
Hickerson House, CB# 3410
Chapel Hill, NC 27599-3410

EMAIL: hgold@email.unc.edu

December 2004

*Professor of City and Regional Planning.

The author gratefully acknowledges the research assistance of Chris Harder and John Guerra, and the comments and suggestions of Larry Less.

Evaluation of State Industry Employment Projections to 2000

Harvey A. Goldstein

1. Introduction

The periodic evaluation of states' employment projections is an activity that can yield important benefits to producers of projections as well as to various client users of projections. First, evaluations of the accuracy of employment projections can provide diagnostic information and feedback to the states for improving subsequent rounds of projections. By knowing which projections were highly accurate and which had relatively large errors, analysts will have more diagnostic information, for example, deciding to use alternative projection techniques, or use supplementary information in the review and adjustment process. Second, evaluation of employment projections is an activity that increases the public accountability of the projections process and of the providers of labor market information. It signals to client users and the public that the accuracy of the employment projections do matter, and providers of employment projections are committed to learning from experience in order to improve their products.

This study reports on the results of the statewide projections developed to the target year of 2000. Although there are a number of relevant dimensions and criteria for evaluating the states' employment projections process, including timeliness and usefulness, this study focuses on only the *accuracy* of the statewide industry employment projections. Here, accuracy is measured by the percentage difference between the projected industry employment in 2000, and the actual year 2000 industry employment. More details on measurement and methods are provided in section 2 below. Neither sub-state industry employment projections nor occupational employment projections were included in the scope of this study, not because they are not considered important and valuable, but because of the considerably greater data and requirements and costs.

All 50 states plus the District of Columbia were invited to participate in the evaluation. Letters were sent to all the LMI directors asking for participation and cooperation in providing projections data as well as detailed 2000 actual employment.
About one-half of the states agreed to participate. In the end we could not include several of these states because of missing data or because projections were developed at only highly aggregated industry categories (major industry divisions).[1]

This evaluation closely follows the approaches and measures used in two previous multi-state evaluations of states industry employment projections in order to be able to make historical comparisons. These previous evaluations were conducted for the 1976-1982 round of projections – published in the October 1987 issue of the *Monthly Labor Review* – and the 1985-1995 round of projections written in a June 1998 report for the ALMIS

[1] States included in the evaluation are: Alaska, California, Colorado, Illinois, Indiana, Kentucky, Maine, Missouri, Montana, New Jersey, New Mexico, North Dakota, Ohio, Oklahoma, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas.

Long-Term Industry Employment Projections and Census Tools Consortium. Although the states participating in these respective evaluations are not necessarily the same, there are some general observations to be made about employment projection accuracy over the last eighteen years.

2. Measures and Methods

In this section we describe the measures, methods, and procedures used in the evaluation. More details, including mathematical formulas, can be found in the Technical Appendix.

The accuracy of a specific industry employment projection can be measured in a number of ways. We use two measures in this evaluation. The first of these is the percent projection error. This is calculated as the difference between the projected level of employment and actual employment in 2000, divided by the actual level of employment in 2000. Thus the minimal percent error is 0.0, which would occur when the projected employment is equal to the actual employment. There is no maximum value. The second measure used is qualitative: was the *direction* of employment change between the base year and the projection year correctly predicted? The possible values for this measure are “yes” and “no” (we treat the specific case of no employment change as positive). The four possible outcomes on this measure for a particular industry projection are: (1) the projected change in employment was 0.0 or greater and the actual change was 0.0 or greater; (2) the projected change was 0.0 or greater, but the actual change was negative; (3) both the projected change and actual changes were negative; and (4) the projected change was negative but the actual change was 0.0 or greater.

When aggregating industry sectors into groups, such as all 3-digit manufacturing industries, we calculate two different measures of the average projection error for the group. The first is the mean absolute percentage error, or MAPE, which is the average of the individual percent projection errors (absolute values) for that particular group of industries. The second measure is the weighted MAPE (or WMAPE), which takes into account the difference in sizes among the industries in the group, giving larger weights to larger industries when calculating the “average” percent projection error. The weights used here are the actual 2000 employment level for the particular industry. To help explain the difference between the MAPE and WMAPE, let’s assume there are two hypothetical industry projections in a state. The projection error for the first industry equal to 10.0 percent (and with employment equal to 12,000) while the projection error for the second industry is equal to 20.0 percent (and with employment equal to 4,000). The MAPE is easily calculated as 15.0 percent, which is the average between 20 percent and 10 percent. The WMAPE, however, is 12.5 percent because the first industry is weighted three times as much as the smaller second industry. Both of these criteria are valuable. We use the MAPE when we treat the magnitude of the projection errors the same regardless of the size of each industry, while we emphasize the WMAPE when we wish do not want to place as much weight on the projection errors for industries with low employment compared to industries with high employment.

The MAPES and WMAPES were calculated for a number of different types of groups of state industry sectors in order to easily compare the magnitude of the average projection error for different categories. Thus, we can compare the average projection error by state, by SIC level, by major industry division, by category of employment size, and by category employment growth rates. These results are displayed in Tables 1-13. We also have developed regression models to identify the factors that are most important in explaining the variation in magnitude of percent projection error, holding other factors constant. These results are discussed and shown below in Section 4 and Tables 14 and 15.

For the 2000 round of projections, states used different base years. Thus the length of projection period varied among the states.[2] Given that we know from previous evaluations that the size of the error tends to increase with longer projection periods, it may be unfair to compare projection errors across states. Also, there are differences in the number of staff and resources available for projections among the states. For these reasons, we do not identify or rank the states by name within this report. Each state will be informed of the "code" used for their state in separate communication.

3. Projection Accuracy – Descriptive Results

3.1. Results Overall

The weighted MAPE for all projections for all nineteen states is 12.6 percent (see Table 1). The overall MAPE, on the other hand is considerably larger – 86.0 percent. This discrepancy reflects the fact that a few very small industries had very large projection errors (in some cases over 1,000 percent). We note that the WMAPE compares quite closely with the overall WMAPE from the evaluation of state projections from the 1985-1995 projection period of 11.8 percent.

Of all the 3-digit industry projections, the direction of change in employment from base year to projection year was predicted correctly for 75.6 percent of the cases (see Table 2). This compares quite favorably with 71.0 percent for the 1985-1995 round of projections. If the industry actually grew, the likelihood of correctly projecting growth was an impressive 0.89 (out of 1.0). On the other hand, if the actual employment declined, the likelihood of correctly predicting decline falls off to 0.55. This is an indication of a well-known optimistic bias in economic forecasting.

3.2. Projection Accuracy by SIC Level

The results in Table 1 show that both the WMAPE and MAPE monotonically increase with higher SIC level of disaggregation. Projections of total nonagricultural employment had a weighted MAPE of only 5.9 percent (8.7 percent for the 1985-1995 period). The WMAPE for all 3-digit SIC industry projections was highest at 19.7 percent (nearly identical to the 19.6 percent for the 1985-1995 round). The WMAPE for 2-digit

[2] Base years used: 1987 (California, Kentucky); 1988 (Illinois, Indiana, Maine, New Mexico, North Dakota, Tennessee); 1990 (Missouri, Oklahoma, South Carolina); 1991 (Ohio, Pennsylvania, South Dakota); 1993 (Texas); 1994 (Montana); 1995 (Alaska, Colorado).

industries (17.5 percent) was not much larger than for the 3-digit industries, indicating that the small loss of accuracy is outweighed by the benefits of using more disaggregated industry sectors for the development of occupational industry employment projections. The very large MAPE of 3-digit industries (100.6 percent) again reflects the effects of having a small number of very small industries with percent errors sometimes in excess of 1000 percent. It should be noted that the differences in magnitude of projection error between SIC levels (based upon the MAPEs) are not statistically significant. This is largely because of such large internal variation in the size of projection errors from a small number having extreme values. The differences among the WMAPEs would be, no doubt, statistically significant.

Table 1. Projection Error By SIC Level

		MAPE	WMAPE
Total Non-Ag	Mean	7.20	5.89
	N	15	
	Std. Deviation	4.40	2.32
1-Digit SIC	Mean	15.67	9.44
	N	136	
	Std. Deviation	24.34	10.00
2-Digit SIC	Mean	53.07	17.49
	N	1199	
	Std. Deviation	602.89	46.34
3-Digit SIC	Mean	100.60	19.65
	N	3444	
	Std. Deviation	1519.49	34.84
Total	Mean	86.01	12.55
	N	4794	
	Std. Deviation	1322.87	29.70

F statistic = 0.540, significant @ 0.655

Table 2. Frequency of Types of Projection Error, 3-Digit Industries

	Frequency	*Percent*
Projected Empl. > Base, Actual Empl. > Base	2,196	66.0
Projected Empl. > Base, Actual Empl. < Base	548	16.5
Projected Empl. < Base, Actual Empl. > Base	262	7.9
Projected Empl. < Base, Actual Empl. < Base	321	9.6
Total	3,327	100.0

3.3. Projection Accuracy by Industry Size

The variation in projection accuracy by industry employment size for 2- and 3-digit industries is shown in Tables 3 and 4. The average error (MAPE) for state industries with less than 1,000 employment is between eight and ten times larger than the next to worst size category! So size makes a huge difference in accuracy at the smallest levels, but the differences in accuracy are quite small once one reaches the 3,000-9,999 employment size category and above. When we look more closely at the smallest size category, we can see the difference between the MAPE and WMAPE is very large (335 percent versus 56 percent for 2-digit industries; 252 percent versus 55 percent for 3-digit industries), indicating that the "culprits" are the very smallest industries, i.e., well under 1,000 employees).

Table 3. Projection Error By Employment Size Category, 2-Digit Industries

		MAPE	WMAPE
Less than 1,000	Mean	335.53	56.03
	N	124	
	Std. Deviation	1855.05	439.20
1,000 - 2,999	Mean	32.21	32.46
	N	134	
	Std. Deviation	31.20	29.85
3,000 - 9,999	Mean	20.75	19.42
	N	222	
	Std. Deviation	23.81	21.83
10,000 - 24,999	Mean	18.29	18.48
	N	230	
	Std. Deviation	20.44	21.21
25,000 +	Mean	18.19	17.24
	N	489	
	Std. Deviation	39.75	46.22
Total	Mean	53.07	17.49
	N	1199	
	Std. Deviation	602.89	46.34

F statistic = 7.77, significant at .000

3.4. Projection Accuracy by Industry Employment Growth Rate

Tables 5 and 6 shows the projection accuracy among four categories of employment growth rate between the base year and projection year for 2- and 3-digit industries. The projection accuracy was clearly highest among industries which grew, but at a moderate rate. The next most accurate employment growth rate category was the set of state industries that declined but at a moderate rate, closely followed by industries which grew at a high rate (20 percent or greater). The accuracy of projections for industries that declined at a high rate ($\leq$ -20.0 percent) falls off drastically. These results indicate that industries that suffer large declines are often unanticipated. This is not an unexpected result given the predominant use of projection models that rely heavily on historical growth trends. The results are consistent between the set of 2-digit and 3-digit industries. Also, the "U" shaped relationship between projection error and employment growth rate is similar to the results from the evaluation of the 1985-1995 round of projections.

Table 4 Projection Error By Employment Size, 3-Digit Industries

		MAPE	WMAPE
Less than 1,000	Mean	252.32	55.34
	N	1132	
	Std. Deviation	2644.23	207.23
1,000 - 2,999	Mean	36.60	35.95
	N	689	
	Std. Deviation	49.48	50.53
3,000 - 9,999	Mean	24.28	23.71
	N	869	
	Std. Deviation	22.14	21.63
10,000 - 24,999	Mean	19.32	18.97
	N	447	
	Std. Deviation	22.04	27.27
25,000 +	Mean	19.20	17.09
	N	307	
	Std. Deviation	29.95	26.01
Total	Mean	100.60	19.65
	N	3444	
	Std. Deviation	1519.49	34.84

F statistic = 5.643, significant @ 0.001

Table 5. Projection Error By Growth Rate, 2-Digit Industries

		MAPE	WMAPE
-20% or less	Mean	351.70	75.54
	N	117	
	Std. Deviation	1895.06	107.11
-19.9% to 0%	Mean	18.39	14.27
	N	158	
	Std. Deviation	16.09	11.41
0% to 19.9%	Mean	8.66	9.02
	N	287	
	Std. Deviation	7.80	7.25
20% or geater	Mean	23.32	20.54
	N	436	
	Std. Deviation	34.09	60.68
Total	Mean	56.82	18.14
	N	998	
	Std. Deviation	655.75	50.75

F statistic = 9.353, significant @ .000

Table 6. Projection Error By Growth Rate, 3-Digit Industries

		MAPE	WMAPE
-20% or less	Mean	529.43	79.62
	N	494	
	Std. Deviation	3985.21	142.83
-19.9% to 0%	Mean	22.01	17.76
	N	400	
	Std. Deviation	19.80	14.15
0% to 19.9%	Mean	11.85	10.11
	N	587	
	Std. Deviation	13.25	8.11
20% or greater	Mean	30.69	19.11
	N	1325	
	Std. Deviation	22.78	16.31
Total	Mean	113.31	19.04
	N	2806	
	Std. Deviation	1681.89	36.54

F statistic = 12.40, significant @ 0.000

3.5 Projection Accuracy by Industry Type

Traditionally, the sectors of mining, construction, and durable goods manufacturing tend to have the highest projection errors because of their relatively high volatility caused by susceptibility to external events that in themselves are difficult to predict. The results for the 1990-2000 round of projections display this ordinality, but overall the differences in accuracy among major industry divisions are not statistically significant because of very large variation in magnitude of the projection errors *within* categories (see Tables 7 and 8). For both 2- and 3-digit industries, mining and durable goods manufacturing had the highest WMAPES, and government, retail trade, and wholesale trade had the lowest WMAPES. Yet such sectors as transportation services, communications and public utilities and services (for 2-digit industries) also had relatively high WMAPES. The large differences between the MAPE and the WMAPE and the large variation in magnitude of projection errors within categories indicate that industry volatility alone is not responsible for extremely large errors; rather it is the interaction of volatility and small size that leads to high percent errors.

3.6. Projection Accuracy by State

Differences in the WMAPES among states for 2-digit industries were not statistically significant (see Table 9). Indeed, the range was only between 37.4 percent and 9.2 percent, and the state with the 37.4 percent WMAPE could be considered an outlier since the next highest WMAPE was only 22.4 percent. Significantly larger variation in projection accuracy as measured by the MAPE is again indicative of the problem of small and volatile industries being concentrated among the smaller states.

When we examine the differences in projection accuracy among states for 3-digit industries (Table 10), there is a larger range of WMAPES (52.9 percent to 14.8 percent) but still a relatively tight clustering around the overall WMAPE (all states) of 19.7 percent. Of the thirteen states that developed projections at the 3-digit level, only two states had WMAPES larger than 30 percent. Large differences in both magnitude and ordinality between the MAPE and the WMAPE in a number of states again show that the state industries with the highest projection errors are highly concentrated in a small number of very small industries.

Table 7. Projection Error By Industry Sector, 2-Digit Industries

		MAPE	WMAPE
Mining	Mean	204.82%	25.85%
	N	55	
	Std. Deviation	1099.95%	18.30%
Construction	Mean	20.98%	18.97%
	N	54	
	Std. Deviation	14.61%	11.72%
Durable Goods Manufacturing	Mean	24.51%	22.87%
	N	183	
	Std. Deviation	32.35%	29.92%
Non-Durable Goods Manufacturing	Mean	84.82%	17.47%
	N	167	
	Std. Deviation	462.12%	24.82%
Transportation	Mean	181.25%	18.99%
	N	117	
	Std. Deviation	1685.07%	66.67%
Communications & Public Utilities	Mean	20.02%	20.81%
	N	38	
	Std. Deviation	11.47%	9.07%
Wholesale	Mean	8.14%	11.36%
	N	38	
	Std. Deviation	6.63%	6.97%
Retail	Mean	12.23%	9.47%
	N	142	
	Std. Deviation	10.17%	7.67%
FIRE	Mean	20.16%	17.06%
	N	120	
	Std. Deviation	17.59%	19.01%
Services	Mean	23.12%	21.50%
	N	240	
	Std. Deviation	53.72%	71.56%
Government	Mean	13.37%	0.29%
	N	45	
	Std. Deviation	10.19%	13.44%
Total	Mean	53.07%	17.49%
	N	1199	
	Std. Deviation	602.89%	46.34%

F statistic = 1.193, significant at 0.291

Table 8. Projection Error By Industry Sector, 3-Digit Industries

		MAPE	WMAPE
Mining	Mean	211.79	33.09
	N	89	
	Std. Deviation	947.91	43.86
Construction	Mean	28.56	23.43
	N	154	
	Std. Deviation	54.04	15.89
Durable Goods Manufacturing	Mean	131.01	30.98
	N	724	
	Std. Deviation	1338.80	62.04
Non-Durable Goods Manufacturing	Mean	258.88	24.03
	N	587	
	Std. Deviation	3322.66	63.93
Transportation	Mean	56.27	21.63
	N	196	
	Std. Deviation	185.15	47.61
Communications & Public Utilities	Mean	40.24	27.82
	N	101	
	Std. Deviation	40.55	17.36
Wholesale	Mean	18.38	15.42
	N	201	
	Std. Deviation	18.71	14.96
Retail	Mean	38.64	12.21
	N	421	
	Std. Deviation	166.82	14.86
FIRE	Mean	42.92	19.79
	N	271	
	Std. Deviation	471.86	35.51
Services	Mean	29.54	17.81
	N	690	
	Std. Deviation	39.35	16.44
Government	Mean	11.97	6.17
	N	10	
	Std. Deviation	12.32	4.40
Total	Mean	100.60	19.65
	N	3444	
	Std. Deviation	1519.49	34.84

F statistic = 1.072, significant @ 0.380

3.6. Projection Accuracy by State

Differences in the WMAPES among states for 2-digit industries were not statistically significant (see Table 9). Indeed, the range was only between 37.4 percent and 9.2 percent, and the state with the 37.4 percent WMAPE could be considered an outlier since the next highest WMAPE was only 22.4 percent. Significantly larger variation in projection accuracy as measured by the MAPE is again indicative of the problem of small and volatile industries being concentrated among the smaller states.

When we examine the differences in projection accuracy among states for 3-digit industries (Table 10), there is a larger range of WMAPES (52.9 percent to 14.8 percent) but still a relatively tight clustering around the overall WMAPE (all states) of 19.7 percent. Of the thirteen states that developed projections at the 3-digit level, only two states had WMAPES larger than 30 percent. Large differences in both magnitude and ordinality between the MAPE and the WMAPE in a number of states again show that the state industries with the highest projection errors are highly concentrated in a small number of very small industries.

Table 9. Projection Error By State, 2-Digit Industries

		MAPE	WMAPE
A	Mean	99.23	9.20
	N	68	
	Std. Deviation	646.24	31.55
B	Mean	45.17	37.35
	N	40	
	Std. Deviation	104.61	105.97
C	Mean	286.28	16.20
	N	68	
	Std. Deviation	2211.06	67.09
D	Mean	18.07	10.96
	N	71	
	Std. Deviation	29.37	10.15
E	Mean	31.07	22.36
	N	68	
	Std. Deviation	64.14	51.97
F	Mean	30.32	20.28
	N	69	
	Std. Deviation	30.40	17.35
G	Mean	150.81	17.47
	N	66	
	Std. Deviation	1006.36	22.42
H	Mean	17.44	11.88
	N	69	
	Std. Deviation	20.37	10.63
I	Mean	16.14	10.02
	N	67	
	Std. Deviation	15.55	10.87
J	Mean	28.56	20.17
	N	48	
	Std. Deviation	34.26	24.11
K	Mean	48.01	21.51
	N	26	
	Std. Deviation	99.95	26.64
L	Mean	45.11	17.84
	N	67	
	Std. Deviation	116.97	21.67
M	Mean	14.97	10.05
	N	67	
	Std. Deviation	14.80	9.78
N	Mean	21.87	15.78
	N	69	
	Std. Deviation	18.84	13.54
O	Mean	15.86	12.59
	N	67	
	Std. Deviation	11.52	7.96
P	Mean	59.18	15.60
	N	65	
	Std. Deviation	316.78	18.48
Q	Mean	26.62	15.46
	N	65	
	Std. Deviation	36.73	16.43
R	Mean	29.17	18.00
	N	69	
	Std. Deviation	47.79	18.95
S	Mean	15.43	12.24
	N	70	
	Std. Deviation	13.20	8.15
Total	Mean	53.07	17.49
	N	1199	
	Std. Deviation	602.89	46.34

F statistic = .806, significant @ 0.695

Table 10. Projection Error By State, 3-Digit Industries

		MAPE	WMAPE
A	Mean	47.18	52.86
	N	61	
	Std. Deviation	71.38	87.47
B	Mean	44.20	15.70
	N	341	
	Std. Deviation	140.24	17.62
C	Mean	132.55	14.83
	N	317	
	Std. Deviation	1695.47	31.84
D	Mean	56.66	29.29
	N	311	
	Std. Deviation	161.73	50.05
E	Mean	619.83	34.42
	N	279	
	Std. Deviation	4979.71	158.66
F	Mean	46.58	18.26
	N	335	
	Std. Deviation	111.48	24.69
G	Mean	26.37	15.50
	N	229	
	Std. Deviation	63.04	18.18
H	Mean	67.31	25.34
	N	322	
	Std. Deviation	265.37	27.02
I	Mean	28.05	20.08
	N	286	
	Std. Deviation	26.25	15.69
J	Mean	58.56	24.13
	N	274	
	Std. Deviation	207.47	44.12
K	Mean	56.87	26.72
	N	287	
	Std. Deviation	154.91	31.21
L	Mean	37.59	25.42
	N	52	
	Std. Deviation	49.06	20.90
M	Mean	29.42	14.85
	N	350	
	Std. Deviation	55.74	12.37
Total	Mean	100.60	19.65
	N	3444	
	Std. Deviation	1519.49	34.84

F statistic = 3.071, significant @ 0.000

3.7. Projection Accuracy by Industry Size and by Employment Growth Rate (Two-Way)

With results that show that both industry size and industry employment growth rate are related to the magnitude of projection error, it is valuable to analyze whether differences remain across industry size categories when we control for employment growth rate, and vice versa. To accomplish this we develop two-way tables that cross-classify categories of size and growth rate (Tables 11 and 12), and examine the differences in projection error among the cells. The results show that while the MAPEs are higher for small industries (less than 1000 employment) *controlling for growth rate*, employment growth rate is the more important factor. That is, small industries that did *not* precipitiously decline over the projection period did not have unusually large projection errors. And the large industries that suffered steep declines in employment tended to have very large projection errors. But the large majority of state industries that did suffer large percentage declines in employment happened to be small industries (less than 1000 actual 2000 employment). Our findings, then, indicate that it is primarily the phenomenon of large employment declines that lead to unusually high projection errors, but these cases tend to be concentrated disproportionately in among the smallest industries.

Table 11. Projection Error (MAPE) By Size and Growth Rate, 3-Digit Industries

Employment Size Category		-20% or less	-19.9% to 0%	0% to 19.9%	20% or greater	Total
Less than 1,000	Mean	835.7%	30.9%	16.9%	41.9%	272.3%
	N	296	121	148	433	998
	Std. Deviation	5128.8%	26.2%	20.0%	26.3%	2813.8%
1,000 - 2,999	Mean	85.5%	21.9%	11.1%	30.9%	34.7%
	N	95	98	111	254	558
	Std. Deviation	60.7%	16.7%	11.3%	20.1%	38.3%
3,000 - 9,999	Mean	54.5%	15.8%	10.4%	25.3%	24.0%
	N	77	111	140	337	665
	Std. Deviation	34.9%	12.7%	9.5%	18.7%	22.8%
10,000-24,999	Mean	57.9%	15.4%	8.6%	20.5%	18.5%
	N	21	46	104	169	340
	Std. Deviation	62.1%	13.9%	8.2%	14.8%	22.6%
25,000 +	Mean	131.5%	19.1%	10.4%	20.2%	19.0%
	N	5	24	84	132	245
	Std. Deviation	188.5%	14.0%	7.7%	16.8%	32.5%
Total	Mean	529.4%	22.0%	11.8%	30.7%	113.3%
	N	494	400	587	1,325	2,806
	Std. Deviation	3985.2%	19.8%	13.2%	22.8%	1681.9%

Table 12. Projection Error (MAPE) By Industry Type and Growth Rate, 3 - Digit Industries

Employment Size Category		-20% or less	-19.9% to 0%	0% to 19.9%	20% or greater	Total
Mining	Mean	489.7%	21.7%	14.0%	48.9%	228.8%
	N	34	8	16	21	79
	Std. Deviation	1504.4%	16.8%	13.1%	28.1%	1005.0%
Construction	Mean	148.0%	22.8%	12.8%	22.2%	30.0%
	N	9	4	18	93	124
	Std. Deviation	183.4%	8.2%	12.7%	13.1%	58.7%
Durable Goods Manufacturing	Mean	552.0%	20.3%	12.7%	35.4%	146.4%
	N	135	99	136	217	587
	Std. Deviation	3061.8%	18.7%	13.8%	22.6%	1481.0%
Non-Durable Goods Manufacturing	Mean	866.8%	15.3%	12.3%	38.7%	305.1%
	N	158	103	81	132	474
	Std. Deviation	6378.5%	15.3%	9.9%	25.0%	3696.3%
Transportation	Mean	234.4%	31.2%	13.9%	3623.3%	57.3%
	N	21	16	33	96	166
	Std. Deviation	535.2%	24.3%	12.9%	25.6%	199.6%
Communications and Utilities	Mean	78.6%	19.3%	26.5%	41.0%	40.0%
	N	16	22	14	28	80
	Std. Deviation	47.2%	12.7%	36.1%	27.1%	37.2%
Wholesale	Mean	48.4%	20.9%	6.5%	19.4%	16.9%
	N	11	27	59	65	162
	Std. Deviation	25.7%	14.9%	5.3%	17.2%	17.8%
Retail	Mean	164.1%	28.7%	8.4%	27.4%	40.6%
	N	43	49	76	173	341
	Std. Deviation	505.8%	23.0%	7.1%	18.5%	184.7%
FIRE	Mean	535.3%	27.8%	10.3%	38.4%	83.7%
	N	24	26	42	139	231
	Std. Deviation	1537.3%	28.6%	10.1%	26.7%	510.6%
Services	Mean	127.3%	30.8%	13.7%	24.0%	30.5%
	N	43	40	108	361	552
	Std. Deviation	92.9%	21.7%	15.4%	19.5%	42.4%
Government	Mean	-	14.8%	4.2%	-	12.0%
	N	-	6	4	-	10
	Std. Deviation	-	13.8%	2.1%	-	12.3%
Total	Mean	529.43%	22.01%	11.85%	30.7%	113.3%
	N	494	400	587	1,325	2,806
	Std. Deviation	3985.21%	19.80%	13.25%	22.8%	1681.9%

F statistic = 12.697, significant @.000

3.8. Direction of Error and Industry Type

To assess whether there is an association between ability to correctly predict direction of employment change and industry type, we have aggregated the major industry divisions into four aggregate categories: infrastructure (construction, transportation services, communications and utilities), production (mining, durable and nondurable manufacturing), services (FIRE, services, government), and trade (retail and wholesale).

The correct direction of employment change was highest for services (79.3 percent correct direction among all 3-digit industries, all states combined), while correctly predicting direction of change for production-oriented industries was lowest (61.8 percent). Infrastructure and trade were in between.

Table 13 Frequency of Type of Error By Aggregated Industry Type

Industry Type	**Type of Error**				
	Proj Empl. > Base, Act Empl. > Base	**Proj Empl. > Base, Act Empl. < Base**	**Proj Empl. < Base, Act Empl. > Base**	**Proj Empl. < Base, Act Empl. < Base**	***Total***
Infrastructure	250	65	32	23	370
	(67.6)*	(17.6)	(8.6)	(6.2)	(100.0)
Production	456	289	147	248	1,140
	(40.0)	(25.4)	(12.9)	(21.8)	(100.0)
Services	608	118	46	21	793
	(76.7)	(14.9)	(5.8)	(2.6)	(100.0)
Trade	332	101	41	29	503
	(66.0)	(20.1)	(8.2)	(5.8)	(100.0)
Total	1,646	573	266	321	2,806
	(58.7)	(20.4)	(9.5)	(11.4)	(100.0)

*(Numbers in parentheses are percent)

A more insightful measure, however, is the conditional probability of correctly predicting direction of change, given either actual growth or decline of employment, since we already know that industries that decline in employment have larger percent projection errors. The results here show that when there was actual employment growth experienced by an industry sector, the likelihood of correctly predicting growth was 0.89 for infrastructure, 0.76 for production, 0.93 for services, and 0.89 for trade. But when there was actual decline of industry employment, the likelihood of correctly predicting decline was low across the board, and *lowest for services*: 0.26 for infrastructure, 0.46. for production, 0.15 for services, and 0.36 for trade. The rather dismal record in correctly predicting decline in the services is tied to the fact that the services are hailed as the growth sector of the economy, and are most immune to employment decline. In general this is true about the services, but it does not obviate the fact that analysts are often not able to predict decline when it does occur. The relatively high conditional probability of correctly predicting decline in the production sectors is likely tied to downward secular trends in the historical time-series, at least in manufacturing. But the general tendency of not being able to correctly predict employment decline still holds across all industry types.

4. Explaining Variation in Projection Accuracy

In this section we estimate the how much of the variation in magnitude of projection error, among all 2- and then 3-digit SIC industries and across all states, that is accounted for by each specific industry or state-level factor that we think is an important explanatory variable. These factors include the industry size, its growth rate, the major industry division, certain characteristics of the overall state economy, and length of projection period. By using a multiple regression model, we can isolate the separate effect of each of these factors on the magnitude of projection error, by statistically holding all other factors constant. We estimate separate models for 2-digit and 3-digit industries. However, all state industries that had 2000 actual employment of less than 50 were not included in the estimation of the models because we felt these were outliers (and indeed accounted for the large majority of projection errors over 500 percent.

The results for 2-digit industries are shown in Table 14. The overall explanatory power of the model is relatively low, with an R^2 of 0.21. The F-statistic = 9.9. The statistically significant variables (@0.05) are all three growth rate dummy variables, the state's location quotient for nondurable goods manufacturing, and the length of the projection period. The growth rate dummy variables included in the model are the three highest categories (the lowest growth rate variable is the default) and all three of the coefficients for the dummies are strongly negative (meaning negatively associated with higher projection errors). These variables turn out to be the most important in the model.

The estimate of the coefficient for the variable measuring the state economy's concentration in nondurable goods manufacturing is also negative. This is not so readily interpretable, as we might expect states that have a higher concentration of manufacturing industries to display greater volatility in their employment, and hence higher projection error. Yet, analysts in these states did a better job in projecting employment across the full range of industries. This tendency may be due to state analysts in such states being less likely to be biased toward optimism on employment growth, and/or that economic trends in these states have been more consistent over the recent historical, albeit slower or negative in many industries.

The coefficient estimate for length of projection period, which varied among the 19 states (but did not vary within state), was predictably positive. That is the longer the length of the projection period, everything else equal, the higher the average projection error for all industries in the state.

Interestingly, neither size of industry categories nor major industry sector categories were statistically significant after controlling for other factors. Neither were the state-level variables measuring the concentration of high-tech industry employment or the concentration of jobs in managerial or professional occupations.

The results of the regression model for explaining the variation in magnitude of projection error among 3-digit industries from the 19 states were not as good in terms of explanatory power. The R^2 was only 0.114 with an F-statistic = 14.7. The statistically significant variables include (in order of magnitude of standardized coefficient): the three growth rate category (dummy variables), three of the industry size categories (dummies), and the FIRE major industry division (dummy). No state-level variables, nor length of projection period, are significant in this model.

The estimation results from both models suggest the following: (1) there is a lot of unexplained variation in the magnitude of the projection errors. Much of this comes from a number of industry projections with extremely large errors. Even after controlling for growth rate and industry size, the degree of fit is not good given these "outliers". But there are also some factors not accounted for in the model, such as projection model type, and information about the projection adjustment process that was followed; (2) the most important single factor accounting for variation in projection error included in the models is the industry's employment growth rate over the projection period. If an industry were in the lowest growth rate category (percent decline of 20.0 percent or more), its absolute percent projection error was significantly higher. Industry size was also important for 3-digit industries, after controlling for employment growth rate, but not for 2-digit industries; (3) in general, state-level economic conditions or industry composition were not significant factors in explaining projection error variation. Length of projection period, which varied among the states, was significant for 2-digit industries, but not for 3-digit.

5. Conclusions and Lessons

The general *pattern* of errors for industry employment projections to the year 2000 developed by the 19 participating states is quite similar to the results of the previous evaluation for state projections to 1995 (Goldstein, 1998). Level of SIC disaggregation, industry employment growth rates and to a lesser extent, industry size, are the attributes of state industries that are most closely associated with magnitude of projection error. The more disaggregated the industry sector, the higher the average projection error. And if the industry has declined significantly over the projection period, then that predicts a significantly higher projection error. In general, state-level differences – either in the industry composition of state's economies or macroeconomic performance – are not significant factors in explaining the differences in projection errors. The one exception is the length of projection period, which did vary among the states. As expected, the longer the projection period, the higher the average error.

There are, however, some significant differences in the *magnitude* of projection errors in this latest round compared to the earlier one. Most prominent is the increase in the number of industry projections with *very* large projection errors, e.g., 500 percent or greater.

A closer examination of the cases of very large projection errors reveals that the large majority of these occurred when the "bottom" fell out of the state industry between the base year and the projection year. That is, an industry that had, say, a base year employment of 300, but by 2000 had disappeared entirely or had dropped below 50.

In selected cases this may have been the result of the collapse of the dot com sector of the economy in the year 2000. And it is possible that a few of these cases might be the result of data error or reclassification of employment among SIC codes that was not diagnosed prior to the analysis. Yet there were enough of these cases of very large projection errors spread around a variety of SICs to make us believe that there is a systematic problem rather than idiosyncratic data error. We suspect that the historical employment series in these cases suggested either employment remaining relatively constant, or else a modest decline, when in fact there was a precipitous decline. For these cases, projection models that are based upon linear time-series analysis are not likely to “pick-up” a sudden change in the time trend.

Table 14 Regression Model Results for 2-Digit Industries

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.458	.210	.189	.45548683

Coefficients

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	.581	.246		2.358	.019
	UE	1.387E-03	.021	.003	.066	.947
	STTEMPGR	.405	.238	.069	1.699	.090
	STGROW	.188	.392	.034	.479	.632
	YEARS	2.986E-02	.010	.155	3.068	.002
	GRCATD2	-.561	.062	-.398	-9.084	.000
	GRCATD3	-.657	.059	-.584	-11.230	.000
	GRCATD4	-.546	.057	-.537	-9.591	.000
	10-24K	-4.658E-02	.046	-.036	-1.024	.306
	3-10K	-4.831E-02	.052	-.038	-.930	.353
	1-3K	9.507E-02	.061	.076	1.555	.120
	mining	5.115E-02	.084	.022	.609	.543
	construction	5.544E-02	.082	.023	.678	.498
	nondurable goods	6.468E-02	.057	.044	1.138	.255
	transportation	2.520E-02	.063	.015	.401	.689
	communic/utilities	5.879E-02	.093	.021	.633	.527
	wholesale	-1.787E-02	.094	-.006	-.190	.849
	retail	-7.554E-03	.060	-.005	-.127	.899
	FIRE	5.667E-02	.062	.034	.917	.359
	services	8.735E-02	.053	.069	1.648	.100
	govt	-2.337E-02	.092	-.008	-.254	.799
	LQNDURMF	-.288	.085	-.166	-3.383	.001
	LQDURMFG	4.587E-02	.054	.033	.846	.398
	% managerial, professional jobs	-.475	.742	-.027	-.641	.522
	%high tech employment	.123	1.404	.004	.087	.930

The key to reducing the number of cases of extremely high projection errors is devoting more effort and thought to the projection review and adjustment process. In particular enlisting the assistance of industry experts in the state may provide information of movements taking place within industry sectors which has not yet affected employment but which may serve as "leading" indicators of changes to come.

Table 15 Regression Model Results for 3-Digit Industries

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.337	.114	.105	1.12389

Coefficients

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.344	.388		3.462	.001
	UE	9.065E-03	.022	.008	.408	.684
	STGROW	.146	.330	.008	.443	.657
	STTEMPGR	.193	.258	.014	.747	.455
	YEARS	1.762E-02	.020	.029	.881	.379
	GRCATD2	-.911	.081	-.266	-11.197	.000
	GRCATD3	-1.005	.076	-.347	-13.282	.000
	GRCATD4	-.861	.068	-.362	-12.680	.000
	25k+	-.180	.095	-.044	-1.895	.058
	10-24.9k	-.246	.083	-.069	-2.956	.003
	3-9.9k	-.279	.064	-.102	-4.328	.000
	1-2.9k	-.244	.063	-.084	-3.898	.000
	mining	-.128	.156	-.016	-.824	.410
	construction	-6.682E-02	.115	-.012	-.583	.560
	nondurable	-.121	.074	-.036	-1.645	.100
	transportation	7.798E-02	.101	.016	.773	.440
	comm/utilities	-7.167E-02	.138	-.010	-.519	.604
	wholesale	-.123	.102	-.025	-1.208	.227
	retail	-.113	.077	-.032	-1.457	.145
	fire	.189	.090	.044	2.103	.036
	services	-7.010E-02	.070	-.024	-1.001	.317
	govt	-4.648E-02	.364	-.002	-.128	.898
	LQ non durable manf	.277	.173	.046	1.598	.110
	LQ durable manf	-2.536E-02	.154	-.004	-.165	.869
	% prof, managerial, technical jobs	-1.797	1.634	-.038	-1.100	.272
	% high tech jobs	-2.284	2.841	-.023	-.804	.421

REFERENCES

Goldstein, Harvey and Andrius Nemickas. *Evaluation of States Industry Employment Projections to 1995*. Final Report for the ALMIS Long-Term Industry Employment Projections and Census Tools Consortium. June 1998.

Goldstein, Harvey A. and Alvin M. Cruze. “An Evaluation of State Projections of Industry and Occupational Employment,” *Monthly Labor Review*, 110:10 (1987).

Less, Larry. *Ohio Industry Employment Projections for 1995: An Evaluation*. Final report to the ALMIS Long-Term Industry projections and Census Tools Consortium, Employment and Training Administration, U.S. Department of Labor. October 1997.

Progressive Policy Institute, *The State New Economy Index, Benchmarking Economic Transformation in the States*. July,1999. (www.neweconomyindex.org)

Technical Appendix
Variable Definitions and Measures

Base year employment: The level of employment in the last year (annual average) for which annual time series were available for developing year 2000 projections. The base year varies somewhat among the states.

Absolute percent error: Calculated as the absolute value of the difference between projected and actual employment in 2000, divided by the actual 2000 employment.

> $APE = ABS (P_{2000} – A_{2000}) / A_{2000}$ where P is projected, A is actual, and ABS is absolute value.

Mean Absolute Percent Error (MAPE): The mean, or average, of the absolute percent errors for a set of industry employment projections, i..

> $MAPE = [SUM (APE_i)] / N$, where N = the number of industry employment projections in the set.

Weighted MAPE (WMAPE): the weighted mean, or average of the absolute percent errors for a set of industry employment projections, i..

> $WMAPE = [SUM (APEi \times A_i)] / SUM (A_i)$

Size Category (Industry): State industries were divided into five size categories based upon actual 2000 employment

SIZCAT1	Less than 1,000
SIZCAT2	1,000-2,999
SIZCAT3	3,000-9,999
SIZCAT4	10,000-24,999
SIZCAT5	25,000 or more

Growth Rate Category (Industry): Industry employment growth rate category. Based upon the percent employment growth rate between the base year and projected year. Source: State Labor Market Information organizations.

GRCAT1	Less than –20.0 %
GRCAT2	-20.0 % - -0.1 %
GRCAT3	0.0 % - 19.9 %
GRCAT4	20.0 % or more

Combined (Aggregate) Industry Groups: Combinations of major industry divisions

Infrastructure	Construction, transportation services, communications and utilities
Production	Mining, manufacturing.

Services	FIRE, services, government
Trade	Wholesale trade, retail trade

Industry Sectors: Dummy variable codes for major industry divisions (1987 SIC)

INDSECD1	Mining
INDSECD2	Construction
INDSECD3	Nondurable goods manufacturing
INDSECD4	Durable goods manufacturing
INDSECD5	Transportation, communications, and utilities
INDSECD6	Wholesale trade
INDSECD7	Retail trade
INDSECD8	FIRE
INDSECD9	Services
INDSECD10	Government

LQDurableMfg: The calculated location quotient in each state for durable goods am,nufacturing activity in each state. Calculated using 2000 annual wage and salary employment. This is a measure of the relative specialization, or concentration of durable goods manufacturing in each state.

LQNondurableMfg: The calculated location quotient in each state for nondurable goods manufacturing activity. Calculated using annual wage and salary employment for 2000. This is a measure of the relative specialization of nondurable goods manufacturing in each state

STGROW: Average annual state employment (total nonagricultural wage and salary) growth rate, 1990-2000. Source: BLS.

STTEMP: Average annual state employment growth rate (total nonagricultural wage and salary) from 1984 to the base year. Source: BLS.

UE: Average annual state unemployment rate, 1990-2000. Source: BLS.

YEARS: Length of projection period, in years. Varies among states.

% HIGH TECH: Percent of state employment in high tech sectors, 1997. Progressive Policy Institute, The State New Economy Index, July 1999. (www.neweconomyindex.org).

% MANAGERIAL, PROF: Percent of total state employment in managerial, professional and technical occupations, 1997. Source: Progressive Policy Institute, The The State New Economy Index , July 1999 (www.neweconomyindex.org).