

STATE AND REGIONAL PER CAPITA PERSONAL INCOMES AND INCOME INEQUALITIES

*George H. Carter**

Introduction

Per capita personal income differences between states and regions have been observed by many researchers including Coughlin and Mandelbaum (1988, 1989), Amos (1988, 1989, 1990, 1991), Rowley, Redman, and Angle (1991), and Garnick (1990). Amos contends that although the differences in per capita income among states declined between 1932 and 1978, a noticeable increase occurred from 1978 to 1985. Similarly, Garnick supports the view of a widening difference in regional per capita income in the 1980s as compared with the previous five decades.

This article complements the findings of these studies by examining state and regional per capita personal income inequalities. Annual per capita personal income data by state and region for 1979, 1988, and projections to 2000 obtained from Johnson, Kort, and Friedenbergl (1990) are examined for convergence/divergence in regional and state economic trends. A measure based on the concept of entropy of information (the Theil index) and simple linear regression are used to compare state and regional annual data during the three periods.

Inequality Between States

Theil Index

The model is a version of the Theil index of inequality derived from the notion of entropy in information theory. This measure is

$$(1) L = \sum_{i=1}^n p_i \log(p_i/y_i).$$

* Associate professor and chair of economics and international business at the University of Southern Mississippi. The comments of two anonymous referees and Dr. Edward Nissan are gratefully acknowledged.

where:

- p_i = The population share of state i ; and
 y_i = The income share of state i

such that $\sum_i p_i = 1$, $\sum_i y_i = 1$. If $p_i = y_i$ for all i , then $p_i/y_i = 1$ and $\log(p_i/y_i) = \log 1 = 0$. Therefore, the measure L attains a minimum of zero when each state has identically proportionate shares of income and population, a condition of complete income equality among the states. L is usually between zero and one, but it may exceed one. Though L is always non-negative, some of the components will be negative, which indicates that a ratio p_i/y_i is less than one and therefore its logarithm is negative. In this case, for a given population share p_i , state i receives a larger share of income y_i .

As Theil and Friedman (1973) explain, the measure of equation (1) is the expected information of the message. The message transforms the shares of income (y_i) of the states as prior probabilities into their shares of population (p_i) as posterior probabilities weighted by p_i . Theil and Friedman explain that an alternative to equation (1) is an interchange of roles between p_i and y_i resulting in the income shares y_i as weights, which makes the measure more dependent on income than on population. Population is preferable as the weight, however, because it is more precisely measured than is income.

The interstate inequalities (or total income inequalities), L_s , from equation (1) for 1979, 1988, and 2000 are, respectively, 0.313, 0.524, and 0.424. L rising from 0.313 to 0.524 provides further support for the conclusion in the literature that there was a trend toward increasing income inequality from 1979 to 1988. L falls from 0.524 in 1988 to 0.424 in 2000 which anticipates an income inequality decrease.

Regression

The analysis above is based on aggregation of states' data into a single measure of inequality. An analysis that takes into account differences among the states is regression of per capita personal income for 1988 on per capita personal income during the former period 1979 and then regression of the projected data for year 2000 on corresponding data of the former period 1988.

If it is assumed that the distribution of Y and X are two measurements on the same entity at two time periods, each separately normally distributed, then together they are jointly normally distributed. It follows that the conditional mean is

$$E[Y|X] = \mu_Y + \beta(X - \mu_X)$$

where:

$$E[Y|X] = \text{The conditional expectation of } Y \text{ given } X; \text{ and}$$

$$\beta = \rho\sigma_Y/\sigma_X.$$

If the estimators \bar{X} , \bar{Y} , S_X , S_Y , r , and b are substituted for the parameters μ_X , μ_Y , σ_X , σ_Y , ρ , and β in the above expressions, the result is the linear regression

$$Y' = \bar{Y} + b(X - \bar{X})$$

where:

$$Y' = \text{An estimator of } E[Y|X].$$

This equation is written alternatively as

$$(2) Y'_i - \bar{Y} = b(X_i - \bar{X})$$

where:

- Y'_i = The expectation from regression;
- \bar{Y} = The average in a final period;
- X_i = The observed; and
- \bar{X} = The average in an original period.

An upward (downward) divergence takes place if $b > 1$ and $X_i > \bar{X}$ ($X_i < \bar{X}$) because deviations of Y'_i from their mean \bar{Y} exceed the deviations of X_i from their mean \bar{X} . Performance above the mean during the initial period is enhanced further during the final period for a variable such as per capita personal income. There is a systematic tendency when $b > 1$ and $X_i > \bar{X}$ for those states in the higher income groups to receive on average proportionately higher increases than those states in the lower income groups. By a reversal of arguments, a downward (upward) divergence occurs when $b < -1$ and $X_i > \bar{X}$ ($X_i < \bar{X}$). The change $(Y_i - Y'_i)$, which is a residual difference, estimates a relative shortfall (overflow) when negative (positive) for a particular state.

The regression slopes are 1.12 for the regression of 1988 on 1979 and 1.02 for the regression of 2000 on 1988. The respective t-values are 8.77 and 60.15. Thus, $b > 1$ and is statistically significant (p-value close to zero) for the regressions, indicating divergence. The divergence of per capita income with a slope of $b = 1.12$ for the 1980s provides a result similar to that observed by many who report a rise in state per capita income inequality for this period. Because $b > 1$, the divergence is upward for states with higher than average per capita incomes in the 1970s ($X_i > \bar{X}$), and downward for states with lower than average per capita personal incomes in the 1970s ($X_i < \bar{X}$).

The slope for the regression of 2000 on 1988 is a statistically significant $b = 1.02$. Technically, $b > 1$. The relationship, however, is essentially one of neither divergence nor convergence in state per capita incomes because 1.02 is so close to 1.00. Although this conclusion seems to differ from the Theil index conclusions that $L = 0.524$ in 1988 and $L = 0.424$ in 2000, the two measures of inequality are compatible as developed in Appendix A. L indicates a lessening of inequality, while the regression indicates a lessening in the divergence.

The correlation coefficients of 0.78 and 0.99 are significant when tested for significance by:

$$t = \frac{(n - 2)^{1/2} r}{(1 - r^2)^{1/2}}$$

where:

- n = The number of observations; and
- r = The sample correlation coefficient.

These correlation coefficients depict the degree of relative ranking permanence by the leaders among the states. Permanence of leadership is strong because the r values are positive and fairly high for both regressions. Thus, throughout the historic periods under consideration, as well as the projected period, those states with high per capita incomes maintained their leadership. Table 1 reports the ratios (p/y_i) for the three periods as well as the regression residuals ($Y_i - Y_i'$) for 1988 on 1979 and 2000 on 1988. The residuals ($Y_i - Y_i'$) are tested with $(n - 2)$ degrees of freedom by:

$$t = \frac{(Y_i - Y_i')}{[S_y^2(1 - r^2)]^{1/2}}$$

A one-sided test is chosen for overflow [$H_a: (Y_i - Y_i') > 0$] or shortfall [$H_a: (Y_i - Y_i') < 0$]. Significance at the 10 percent level is indicated by an asterisk in Table 1, columns (4) and (5).

Inequality Between Regions

An attraction of the measure L is its decomposability, making it suited for estimating the contribution to income inequality of subsets of the population. This paper divides the United States into eight regions as classified by the Bureau of Economic Analysis (BEA). Appendix B identifies the states in each of the BEA regions.

Table 1—State Population Share to Income Share and State Per Capita Overflows or Shortfalls of Personal Income Expectation^a

State (1)	1979 p_i/y_i (2)	1988 p_i/y_i (3)	2000 p_i/y_i (4)	$Y_i - Y_i'$ 1988 on 1979 (5)	$Y_i - Y_i'$ 2000 on 1988 (6)
Alabama	1.28	1.28	1.26	118	-22
Alaska	0.72	0.86	0.92	-2755*	-602*
Arizona	1.09	1.10	1.11	44	-92
Arkansas	1.30	1.35	1.30	-220	-156
California	0.86	0.88	0.89	-64	256
Colorado	0.96	0.92	1.01	795	-1129*
Connecticut	0.84	0.72	0.75	3088*	-133
Delaware	0.98	0.93	0.98	988	-512*
District of Columbia	0.80	0.77	0.78	862	558*
Florida	1.04	0.99	0.99	782	137
Georgia	1.19	1.08	1.08	1284	39
Hawaii	0.95	0.99	1.01	-230	-253
Idaho	1.16	1.31	1.26	-1151	121
Illinois	0.90	0.94	0.95	-387	-30
Indiana	1.04	1.20	1.10	-579	105
Iowa	0.99	1.13	1.11	-1312	88
Kansas	0.97	1.05	1.03	-697	334
Kentucky	1.22	1.29	1.01	-354	11
Louisiana	1.18	1.34	1.32	-1185	-138
Maine	1.23	1.09	1.10	1517	-101
Maryland	0.93	0.84	0.87	1820*	-135
Massachusetts	0.96	0.79	0.80	3143*	330
Michigan	0.94	1.00	1.00	-490	55
Minnesota	0.98	0.99	0.99	93	111
Mississippi	1.40	1.48	1.45	-373	-225
Missouri	1.05	1.07	1.05	-3	178
Montana	1.11	1.28	1.23	-1385	170
Nebraska	1.02	1.12	1.07	-871	453*
Nevada	0.86	0.94	0.97	-994	-262
New Hampshire	1.04	0.86	0.89	2898*	-153

Table 1—State Population Share to Income Share and State Per Capita Overflows or Shortfalls of Personal Income Expectation (cont.)

State (1)	1979 p_i/y_i (2)	1988 p_i/y_i (3)	2000 p_i/y_i (4)	$Y_i - Y_i^i$ 1988 on 1979 (5)	$Y_i - Y_i^i$ 2000 on 1988 (6)
New Jersey	0.88	0.75	0.77	3005*	35
New Mexico	1.21	1.32	1.29	-726	-32
New York	0.94	0.86	0.86	1623	315
North Carolina	1.24	1.15	1.14	949	22
North Dakota	1.08	1.29	1.23	-1820*	263
Ohio	1.01	1.06	1.06	-414	46
Oklahoma	1.08	1.24	1.19	-1362	279
Oregon	0.98	1.11	1.11	-1260	-32
Pennsylvania	1.00	1.02	1.01	76	132
Rhode Island	1.07	0.98	0.99	1371	-2
South Carolina	1.31	1.28	1.25	422	-21
South Dakota	1.12	1.29	1.25	-1377	138
Tennessee	1.22	1.19	1.17	470	84
Texas	1.02	1.13	1.11	-1006	153
Utah	1.23	1.35	1.33	-825	-124
Vermont	1.16	1.08	1.08	1093	-121
Virginia	1.04	0.93	0.94	1640	118
Washington	0.92	1.00	1.00	-929	72
West Virginia	1.25	1.41	1.41	-1031	-400*
Wisconsin	1.00	1.06	1.06	-572	86
Wyoming	0.89	1.21	1.31	-3714*	-16

*If $(Y_i - Y) > 0$ then personal income exceeds expectation (overflow) and vice versa. *Significant at $\alpha = .10$ level. Source: Computations from equation (2)

Partition the United States into disjoint regions R_1, \dots, R_g . Let P_g and Y_g be, respectively, the total population and total income shares of R_g , $g = 1, \dots, g$; then the inequality among regions by extension of equation (1), according to Theil (1989), is:

$$(3) L_R = \sum_{g=1}^8 P_g \log(P_g/Y_g),$$

and the inequality among the states in region R_g is:

$$(4) L_g = \sum_{i \in R_g} (p_i/P_g) \log \left[\frac{(p_i/P_g)}{(y_i/Y_g)} \right].$$

An identity is established such that:

$$(5) L = L_R + \bar{L}$$

where:

$$\bar{L} = \sum_{g=1}^8 P_g L_g.$$

This way, the total inequality L among the states is composed of the sum of interregional inequality L_R and the weighted average of intraregion inequality \bar{L} , the weights being P_1, \dots, P_8 .

Table 2 shows regional per capita personal income as well as the ratio of the share of the population to the share of income in a region (P_g/Y_g). ($P_g/Y_g < 1$ indicates that the size of the share of the population is smaller than that of income, meaning that the region is better off economically than if ($P_g/Y_g > 1$). For 1979 the ratios of the New England, Mideast, Great Lakes, and Far West regions are less than one, while the other four regions have ratios greater than one; the Southeast region occupied the last position. In 1988 the Great Lakes region parted company with the top four and joined the other five, with the Southwest instead of the Southeast in the lowest ranked position. Because the Great Lakes region was once dominant in industrial production, these findings support the prevailing consensus of a recent rise in income inequality due to manufacturing structural change (Leigh, 1994). For the year 2000 the same three regions occupy the top positions. Schematically the relationships are:

1979: [FW, ME, GL, NE/PL, RM, SW, SE]

1988: [ME, FW, NE/GL, RM, PL, SE, SW]

2000: [ME, FW, NE/PL, GL, RM, SE, SW].

Table 3 displays total income inequality L which is calculated from equation (1) and interregional inequality L_R which is calculated from equation (3). L_R increased from 0.212 in 1979 to 0.301 in 1988, with an

Table 2—Summary of Regional Per Capita Personal Income and Population

(1)	Per Capita			Population Share/ Income Share		
	Total Personal Income					
	1979	1988	2000	1979	1988	2000
	(2)	(3)	(4)	(5)	(6)	(7)
Coastal						
New England	11,989	16,205	18,156	0.96	0.82	0.84
Mideast	12,256	15,239	17,477	0.94	0.87	0.97
Southeast	9,816	11,614	13,691	1.18	1.14	1.12
Far West	13,198	14,575	16,622	0.87	0.91	0.92
Interior						
Great Lakes	11,999	13,042	15,098	0.96	1.02	1.02
Plains	11,412	12,362	14,562	1.01	1.07	1.06
Southwest	11,019	11,527	13,633	1.05	1.15	1.13
Rocky Mountain	11,071	11,531	13,577	1.04	1.11	1.13
All States	11,551	13,245	15,345			

Source: *Survey of Current Business*, 70 (1990)

anticipated decrease to 0.244 in 2000. Column (4) of Table 3 reports the average within-region inequality which is calculated using equation (5) where the total inequality in column (2) for each year must be the sum of the two terms in columns (3) and (4) (i.e., by subtraction). Interregional income inequality as a proportion of total income inequality among the states, shown in column (5), decreased from 68 percent in 1979 to 57 percent in 1988 and is expected to remain constant through 2000. The average within-region income inequality, therefore, rose from 32 percent to 43 percent. Regional income inequality became less dominant than the average within-region inequality in 1988 in comparison with 1979. Regional income inequality, while increasing, has decreased in its contribution to total inequality in recent years. The growing disparity among regions is of concern, but the more rapidly growing average disparity within the regions is of greater concern. The stable proportion in column (5) from 1988 to 2000 shows that the mix of income inequality between regions and within regions is not expected to change in the near future.

To test the null hypothesis of equality of the measure L_R across regions, analysis of variance is performed for all three periods. The F statistic is significantly large for every period under consideration; it is of interest to determine how the regions differ in their measure of inequality. A multiple comparison procedure by the method of least-significant-difference therefore is employed. The grouping of regions in

Table 3—Income Inequality By State and Region In the United States

Year (1)	L State Inequality (2)	L_R Regional Inequality (3)	Average Within Regions (4)	Regional as a Proportion of State (5)
1979	0.313	0.212	0.101	0.68
1988	0.524	0.301	0.223	0.57
2000	0.424	0.244	0.180	0.58

Column (2) is calculated from equation (1), and column (3) is calculated from equation (3). The calculations are based on state data of per capita personal income and state population. Alaska, Hawaii, and District of Columbia are included

ascending order of inequality resulting from the multiple comparisons is shown below. Any two regions that are not included inside the same parentheses are considered to be different. The groupings provide insights for the change in the rank order of regions.

1979: [(FW, ME, GL, NE), (ME, GL, NE, PL, RM, SW), (RM, SW, SE)],

1988: [(ME, FW, NE), (FW, NE, GL, RM, PL), (GL, RM, PL, SE, SW)],

2000: [(ME, FW, NE), (FW, NE, PL, GL, RM), (PL, GL, RM, SE, SW)].

Concluding Remarks

This paper analyzes interstate change and regional change in per capita personal income inequality using a model based on the entropy concept in information. The years covered are 1979, 1988, and 2000. We find that income inequality among the states increased considerably between 1979 and 1988 and is expected to decline a bit by the year 2000. Regional inequality, while rising, decreased as a percent of total inequality from 68 percent in 1979 to 57 percent in 1988 and is expected to remain constant, with 58 percent predicted by the year 2000. A general assessment is that during recent decades there has been a trend toward income divergence for states and regions. The most remarkable evidence is that average per capita personal incomes for states within regions are becoming more diverse.

References

1. Amos, Orley M., "Unbalanced Regional Growth and Regional Income Inequality in the Latter Stages of Development," *Regional Science and Urban Economics*, 18 (1988), pp. 549-566.
2. Amos, Orley M., "An Inquiry Into the Causes of Increasing Regional Income Inequality in the United States," *The Review of Regional Studies*, 19 (1989), pp. 1-12.
3. Amos, Orley M., "Growth Pole Cycles: A Synthesis of Growth Pole and Long Wave Theories," *The Review of Regional Studies*, 20 (1990), pp. 37-48.
4. Amos, Orley M., "Divergence of Per Capita Real Gross State Product By Sector—1963 to 1986," *The Review of Regional Studies*, 21 (1991), pp. 221-234.
5. Coughlin, Cletus C., and Thomas B. Mandelbaum, "Why Have State Per Capita Incomes Diverged Recently," *Federal Reserve Bank of St. Louis Review*, 70 (1988), pp. 24-36.
6. Coughlin, Cletus C., and Thomas B. Mandelbaum, "Have Federal Spending and Taxation Contributed to the Divergence of State Per Capita Incomes in the 1980s?" *Federal Reserve Bank of St. Louis Review*, 71 (1989), pp. 29-42.
7. Garnick, Daniel H., "Accounting for Regional Differences in Per Capita Personal Income Growth," *Survey of Current Business*, 70 (1990), pp. 29-40.
8. Johnson, Kenneth P., John R. Kort, and Howard L. Friedenberg, "Regional and State Projections of Income, Employment, and Population to the Year 2000," *Survey of Current Business*, 70 (1990), pp. 33-54.
9. Leigh, Nancy G., *Stemming Middle-Class Decline: The Challenge to Economic Development Planning* (New Brunswick, NJ: Center for Urban Policy Research, 1994).
10. Rowley, Thomas D., John M. Redman, and John Angle, *The Rapid Rise in State Per Capita Income Inequality in the 1980s* (Washington, DC: United States Department of Agriculture, Data Service Center, 1991).
11. Theil, Henri, "The Development of International Inequality, 1960-1985," *Journal of Econometrics*, 42 (1989), pp. 145-155.
12. Theil, Henri, and Yoram Friedman, "Regional Per Capita Incomes and Income Inequalities: Point Estimates and Their Standard Errors," *Journal of the American Statistical Association*, 68 (1973), pp. 531-535.

Appendix A—Comparison of L and b

The apparent contradiction between L dropping from 0.524 in 1988 to 0.424 in 2000 and $b = 1.02$ results because L is a static measurement of inequality during a particular year. The driving term in L is (p_i/y_i) which can be rearranged into $[(\text{national per capita income})/(\text{state } i \text{ per capita income})]$:

$$\begin{aligned} p_i/y_i &= \frac{[(\text{state } i \text{ population})/(\text{national population})]}{[(\text{state } i \text{ income})/(\text{national income})]} \\ &= \frac{[(\text{national income})/(\text{national population})]}{[(\text{state } i \text{ income})/(\text{state } i \text{ population})]} \\ &= \frac{(\text{national per capita income})}{(\text{state } i \text{ per capita income})} \end{aligned}$$

Thus, the states' per capita incomes are generally closer to the national per capita income in 2000 than they were in 1988. On the other hand, the regression of per capita income in 2000 on per capita income in 1988 is a dynamic concept relating state i per capita income to the mean of all states' per capita incomes. The mean of all states' per capita incomes is a different measure from national per capita income because all states are weighted equally in the mean. (In 1988 the average of the states' per capita incomes was \$12,619, while the national per capita income was \$13,259.)

L is a static or snapshot measure of the variation in the states' per capita incomes from the national per capita income. The measure is particularly useful because L ranges from zero when all states' per capita incomes equal the national per capita income (perfect income equality) to numbers without bound when all income is held by one state. Mathematically, L can be rearranged to show that it is the weighted logarithm of the geometric mean, the weights being p_i , as follows:

$$\begin{aligned} 10^L &= \prod (p_i/y_i)^{p_i} \\ L &= \log[\prod (p_i/y_i)^{p_i}] \\ &= \sum \log(p_i/y_i)^{p_i} \\ &= \sum (p_i) \log(p_i/y_i) \end{aligned}$$

In contrast to L, b is a dynamic measure between two time periods of the variation in states' per capita incomes from their arithmetic means. The means in the regression model are the means of state per capita incomes. With all states being equally weighted, these means are

not national per capita income because low population states receive equal weights with high population states.

In summary, L and b are different measures of income inequality measuring two different concepts. The static measurement in 1988 was $L = 0.524$. In 2000, $L = 0.424$ which indicates that the state per capita incomes were generally closer to the national per capita income. On the other hand, the dynamic measure of divergence indicated by $b = 1.12$ for the period 1988 on 1979 dropped to essentially no divergence from the mean ($b = 1.02$) for the period 2000 on 1988.

Appendix B—Bureau of Economic Analysis Regions

Regions of the United States according to the Bureau of Economic Analysis are:

Coastal

1. New England

Maine, New Hampshire, Vermont, Massachusetts, Connecticut, Rhode Island

2. Mideast

New York, Pennsylvania, New Jersey, Delaware, Maryland, District of Columbia

3. Southeast

Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana

4. Far West

Nevada, California, Oregon, Washington, Alaska, Hawaii

Interior

5. Great Lakes

Ohio, Michigan, Indiana, Wisconsin, Illinois

6. Plains

Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas

7. Southwest

Oklahoma, Texas, Arizona, New Mexico

8. Rocky Mountain

Colorado, Wyoming, Montana, Idaho, Utah

PEDAGOGY

The editors of *Regional Science Perspectives* wish to encourage our readers to submit their suggestions/techniques for improving the teaching and learning environment in courses devoted to regional science. Papers submitted may describe new data sources appropriate for class exercises, applied techniques designed to illustrate facets of our discipline, new software utilized in the classroom environment, or, simply, new perspectives on presenting the core concepts of our courses. While not undergoing the usual blind review process, the editors and editorial board will try to maintain a level of utility and clarity as well as a lack of redundancy in publishing submissions. We hope this section of *RSP* becomes a regular feature for the exchange of pedagogic techniques designed to stimulate nascent scholars in our discipline.