RICH VERSUS POOR?

AN ECONOMETRIC ANALYSIS OF RELATIVE INCOME EXTREMES

Philip N. Jefferson
Swarthmore College

and

Frederic L. Pryor
Swarthmore College

A standard feature of both populist and Marxist political rhetoric about the class struggle and long-term trends in income distribution is that the rich are getting richer at the expense of the poor. Mainline neoclassical economists pay little attention to this issue, not just because they have no theory to explain why this might happen but also because they believe the problem to be poorly specified. Both of these strands of thought are found, for instance, in a recent essay by Martin Feldstein [1998]. He argues that if we take seriously the Pareto welfare criterion, then we should not worry about income inequality *per se* but, instead, we should focus our attention on poverty. If the rich are getting richer while the poor are not made worse off, this represents an increase in overall welfare.

Although the theoretical differences between these two approaches should be obvious, the two groups also look at the data on income in different ways, with the Marxist/populists focusing on relative income and the neoclassical economists focusing on absolute income. The implications of these two approaches can be seen in a dramatic fashion by looking at the income of families in the 10th and the 95th percentiles. Measured in real dollars, from 1947 to 1997, the two incomes levels are positively and significantly related to each other. Measured as ratios of the median income (hereafter called relative income), the two income levels are negatively and significantly related to each other.

A key question addressed in this essay is whether there is any causal relation between the relative income levels of the families at the extremes of the income distribution. This is not exactly the same question as the determinants of the income distribution in general because changes in the ratio between these two relative income levels do not necessarily determine the direction of change of overall measures of income inequality.[1]

The line of argument in this essay is straightforward. We present econometric evidence suggesting that when the different causal factors acting on income at the

---

Frederic L. Pryor: Department of Economics, Swarthmore College, Swarthmore, PA 19081.
E-mail: fpryor1@swarthmore.edu

two ends of the income distribution are held constant, the extremes of relative income are still significantly and inversely related to each other. This result seems to counter a hidden assumption in the neoclassical view that no causal relationship exists. A crucial statistical question is, of course, whether this relationship is spurious. We use standard time-series methods to explore this issue and obtain results suggesting that it is not. Furthermore, changes in the relative income of the rich are related to changes in the relative income of the poor, but that the reverse situation does not hold. We examine and discard several explanations for these asymmetric results, which remain a puzzle to be solved in the future. We conclude with a discussion of some broader issues.

DESCRIPTIVE EVIDENCE ABOUT THE RELATIONSHIP OF RELATIVE INCOMES OF THE RICH AND POOR

Definition of Income

It is customary to analyze income distribution problems in terms of either some overall measure of income inequality or the shares of total income received by different groups. The first approach does not distinguish between different income trends at the tails of the income distribution. It also does not allow one to isolate the impact of causal factors influencing only one tail of the income distribution. The second approach raises problems in dealing with the very poor and the very rich. More specifically, considerable underground economy income is not included in the measurement of income of the very poor who, ostensibly, are not employed. Among the top income receivers, one important source of income is excluded, namely capital gains (both realized and unrealized). Other data bases such as IRS statistics include such capital gains income, but they have other biases that make them less suitable for the exercise we wish to carry out. Appendix A provides the data sources used in our study.

We have, therefore, focused on the dollar income of those in the 10th and 95th percentiles, calculating both as ratios of the median income.[2] To provide perspective on our results, we carry out the same statistical exercises using the 20th and 80th percentiles, presenting the qualitative results in the text and the quantitative results in Appendix B. For income at these various percentiles, we employ the family income data of the Census Bureau and estimate, using their data, average income only at the 10th percentile. These relative income ratios allow us to examine more closely the impact of variables such as the percentage of families headed by single mothers, which might influence relative income ratios of those in the 10th percentile, but not at the 95th percentile.

Specification Issues

It is possible in the specification of determinants of income inequality, to conceive of all causal factors working within a relatively short time period. For instance, Bishop, Formby, and Sakano [1994] analyze income shares only in terms of contemporaneous variables and note (their footnote 9) that none of their results hold if lags are introduced. This approach contrasts with that of Blank and Blinder [1986], who analyze the distribution of income in a given period in terms of various determinants and an equilibrium income distribution that takes some time to achieve. Assuming a simple geometric distributed lag, their final specification is: $Y_t = a + bY_{t-1} + cX_t + e_t$, where Y is the income share measure, X is a vector of determinants, and e is a random variable. In this regard, our initial approach (Table 1) is similar.

We also assume that different causal forces influence relative income levels at the low and high ends of the income distribution. For instance, the percentage of single-parent families should have little influence on the relative income level of those in the 80th or 95th percentile. Moreover, we try to select variables that are linked with the relative income level by some discernable mechanism.[3]

To explain relative income levels of those in the 10th percentile of the income distribution, we have selected four variables: the lagged value of the relative income level; the percentage of one-parent families; governmental (federal, state, and local) monetary welfare expenditures as a percentage of aggregate wages and salaries; and the weakness of the labor market as measured by the percentage of prime age men (those between 25 and 50) who are jobless. This choice of explanatory variables warrants three special comments:

- In contrast to other econometric studies of income distribution, we have chosen joblessness rather than unemployment, as a proxy for labor market weakness. This is because the former includes discouraged workers and others who are without employment and not counted as unemployed under the Census definition. The joblessness rate fluctuates much less than the unemployment rate and seems to more accurately reflect long-term trends in labor market conditions.
- Other demographic indicators might be added as additional explanatory variables. We believe, however, that they have only secondary effects.[4]
- The variables we have selected are particularly important for those with the lowest income and the statistical results should be weaker if, for instance, we carry out the same statistical calculations for those with incomes at the 20th percentile. As indicated below in more detail, this is exactly what we find.

To explain relative income levels in the 95th percentile, the choice of variables is not so easy, except, of course, for the lagged value of the dependent variable. A promising causal variable is the share of women working. Men have a greater likelihood of marrying women whose level of education is similar to their own. Highly educated women, in turn, earn more than those with less education. Since employment rates of less educated prime age women did not change greatly between 1964 and 1994, while such employment rates among women with more education increased [Pryor and Schaffer, 1999, 8), family incomes of highly educated (and well-paid) men should rise faster than family incomes of those with less-educated and less well-paid men. This conjecture receives empirical support from the detailed analyses of *Current Population Survey* data by Burtless [1998] and Karoly and Burtless [1995].[5] Other possible explanatory variables come to mind, but they seem of secondary importance.[6]

TABLE 1
Two Stage Least Squares Regressions Explaining Relative Income Levels, 1948 - 1997
Contemporaneous Values of Independent Variables (except lagged dependent variable)

Dependent variables:	**Relative income level at 10th percentile**		**Relative income level at 95th percentile**
Regressors:			
Constant	+0.493[a] (0.100)	Constant	+1.995[a] (0.585)
Dependent variable lagged one year	+0.148 (0.162)	Dependent variable lagged one year	+0.506[a] (0.111)
Year (1947 = 1)	+0.0026[a] (0.0005)	Year (1947 = 1)	+0.0101[a] (0.0049)
Relative income at 95th percentile	−0.082[a] (0.025)	Relative income at 10th percentile	−2.349[a] (0.705)
Prime age male joblessness rate	−0.088 (0.094)	Share of women in labor force (>19)	−0.675 (1.078)
Ratio of governmental money transfers to total compensation of employees	+0.689[a] (0.237)		
Percentage of female-headed families	−0.901[a] (0.300)		
Adjusted-R^2	0.8657		0.9585
Durban-Watson	1.746		2.051

Standard errors in parentheses.
a. Statistically significant at the 5 percent level.
Sources of data are presented in Appendix A.

To test the rich-at-the-expense-of-the-poor hypothesis, we also include the relative income level of those in the 95th percentile to help explain relative income of those in the 10th percentile. The reverse procedure is used to help explain relative income of those in the 95th percentile.

Results of a Simple Simultaneous Equation Model

The variations of the relative incomes at the extremes of the income distribution are quite different. Between 1947 and 1997, the relative income at the 10th percentile varied only between 25.9 and 33.5 percent of the median income; while the relative income at the 95th percentile varied between 231.5 and 309.5 percent of the median income. Nevertheless, their coefficients of variation are more similar: 6.5 percent for the former, 8.6 percent for the latter.[7]

The most appropriate first step for exploring the problem is estimating a system of simultaneous equations in which the dependent variables of the two equations are the relative income ratios at the 10th and 95th percentiles. Table 1 presents the results of such a calculation using two-stage least-squares. In both regressions, the sign of

TABLE 2
Augmented Dickey-Fuller (ADF) Tests

	No trend	Trend
Relative income level at the 10th percentile	−1.603 (2)	−1.673 (2)
Prime age male joblessness rate	−1.787 (0)	−2.267 (0)
Ratio of governmental money transfers to total compensation of employees	−0.843 (2)	−2.875 (1)
Percentage of female-headed families	−0.200 (3)	−2.251 (3)
Relative income level at the 95th percentile	+0.182 (0)	−2.082 (3)
Share of women in the labor force (only those 20 and over taken into account)	−1.692 (1)	−0.275 (1)

The number in parentheses are the optimal lag lengths: k − max = 4. The 5 and 10 percent critical values are: −2.914 and −2.598 (no trend) and −3.501 and −3.179 (trend). Sources of data are presented in Appendix A.

the variable reflecting the relative income level at the opposite end of the income distribution is negative and statistically significant. In short, the relative (to the median) income *levels* of the rich and poor move in opposite directions. For the 20th and 80th percentile income ratios (the regressions are shown in Table B-1 in the Appendices), the same inverse relationship is found, although for the 20th percentile income ratio regression, the coefficient for the 80th percentile income ratio is not quite statistically significant at the 0.05 level.

On a descriptive level, therefore, the populist and Marxist political rhetoric about the rich getting richer at the expense of the poor appears to be correct. Conversely, the neoclassical approach appears to be wrong. These calculations using contemporaneous variables do not, however, indicate whether the inverse relation between the relative income levels at the 10th and 95th levels is causal, is influenced by the same set of unspecified causal variables, or is spurious. It is necessary, therefore, to examine this problem in greater depth.

TESTING FOR SPURIOUS CORRELATION AND COINTEGRATION

Statistical Problems Arising Because of Unit Roots

It is possible that the relative income levels of the rich and poor are inversely correlated because both wander away from a determinate trend in the opposite direction. To explore this possibility, which Mocan [1999] found when examining income *shares*, we start by examining a key stochastic property of the individual data series, namely whether they have unit roots. Table 2 shows the results of tests of the null hypothesis of the existence of a unit root in all of the variables over the full sample using an augmented Dickey-Fuller (ADF) test.[8]

At conventional significance levels, we cannot reject the hypothesis that all of the variables have unit roots. The correlations shown in Table 1, therefore, may be spurious [Granger and Newbold, 1974]. As a result, further examination of the inverse relationship between relative incomes at the extremes of the income distribution is in

TABLE 3
DOLS Cointegrating Vectors and Tests for Cointegration

Dependent variables	Relative income level at 10th percentile		Relative income level at 95th percentile
Regressors:			
Constant	+0.6123[a] (0.0457)	Constant	+4.3371[a] (0.6476)
Year (1947 = 1)	+0.0043[a] (0.0005)	Year (1947 =1)	+0.0279[a] (0.0091)
Relative income at 95th percentiles	−0.0983[a] (0.0267)	Relative income at 10th percentile	−4.0006[a] (0.6241)
Prime age male joblessness rate	−0.0867 (0.1445)	Share of women in labor force (>19)	−3.3043 (2.1966)
Percentage of female headed families	−1.5829[a] (0.3725)		
Ratio of governmental money transfers tototal compensation of employees	+1.0285[a] (0.3382)		
DOLS Lags	+2 to −2		+2 to −2
ADF test	−6.0461[a] (3)		−5.2563[b] (0)

Standard errors in parentheses for the point estimates. DOLS Lags are the leads and lags of the first differences of the right-hand side variables used in the estimation of the cointegrating vector. For the ADF test, the numbers in parentheses are the optimal lag lengths: k−max = 4. The 5 and 10 percent critical values for the ADF tests are: −5.4914 and −5.1010.
a. Statistically significant at the 5 percent level.
b. Statistically significant at the 10 percent level.
Sources of data are presented in Appendix A.

order. Although the individual variables are not stationary, it is possible that some linear combination of the variables is stationary, that is, that they are cointegrated. In such a case, the inverse relation between the relative income of the rich and the poor would not be spurious.

Testing for Cointegration

In Table 3, we report the test for whether linear combinations of the variables are stationary using Stock and Watson's [1993] dynamic estimator for the cointegrating vector (DOLS). This technique corrects for simultaneity bias in small samples by including leads and lags of the differences of the right-hand side variables in the cointegrating regression. Also, DOLS is asymptotically equivalent to Johansen's [1991] full information maximum likelihood procedure.

The ADF test statistics indicate that there is a cointegrating relationship for each of the relative income level measures. For relative income in the 10th percentile, the null hypothesis of no cointegration is rejected at the 5 percent level. For relative income in the 95th percentile, the null hypothesis of no cointegration is rejected at the 10 percent level. An interpretation of the cointegrating vector is that it represents the

long-run equilibrium levels relationship among the variables. It is worth noting that when the 20th and 80th, instead of the 10th and 95th, relative income ratios are used, such cointegration is not found. This suggests that the causal variables that we believe influence the distribution extremes are not completely applicable for percentiles in the income distribution that are closer to the median.

The chief significance of the results presented in Table 3 can be simply stated: The inverse relationship between the *levels* of relative income at the extremes of the income distribution in the post-World War II years in the 20th century does not appear to be spurious. Something important is happening between the relative income ratios of the rich and the poor, making them move in opposite directions.

EXPLORATION OF CHANGES IN RELATIVE INCOME RATIOS

It is also worthwhile to look at the causal factors underlying *changes* in these relative income ratios. Of course, relative income levels are not expected to be in equilibrium year by year but, to the extent they are not, relative income may partially adjust to close the gap from equilibrium. This consideration suggests the following error correction specification:

$$\Delta X_t = \alpha_0 + \alpha_1 \Delta X_{t-1} + \alpha_2 \Delta Z_{t-1} + \beta_1 \upsilon_{t-1} + \epsilon_t , \qquad (1)$$

where X = the relative income variable to be explained
Z = a vector of explanatory variables
υ = error term from the corresponding equation in Table 3 (called "equilibrium error" in Table 4)
ϵ = error term for this equation.

In equation (1), the change in relative income responds to lagged changes in its determinants as well as deviations from the long-run equilibrium. The estimated value of the error correction coefficient, β_1, should be negative as the system at time period t moves to close the gap between the actual value of relative income at time period $t-1$ and its long-term equilibrium value.

Table 4 reports the results of estimating equation (1) for the relative income ratios at the extremes of the income distribution. The results suggest that relative incomes at the tails of the income distribution respond to deviations from long-run equilibrium level in the expected manner. For example, if relative income for the 95th percentile was above its equilibrium level by one percentage point in the previous year, then it is expected to fall by 0.81 percentage points in the current year, everything else held constant.[9] Similar results occur when the model is calculated using relative income ratios at the 20th and 80th percentiles.

If we look at annual changes in the relative incomes of rich and poor, rather than their annual levels, the results in Table 4 yield an asymmetry: at the upper tail of the income distribution, the expected tradeoff that annual changes in the relative income of the poor forecast annual changes in the relative income of the rich in the opposite direction is apparent. At the lower tail of the income distribution, by way of contrast,

TABLE 4

Error Correction Model for Changes in Relative Income Levels, 1948 - 1997

Dependent variables: **Regressors**	**Relative income level at 10th percentile**		**Relative income level at 95th percentile**
Constant	+0.0004	Constant	+0.0031
	(0.0014)		(0.0142)
Dependent variable	+0.2300	Dependent variable	+0.0281
	(0.1570)		(0.1439)
Relative income at 95th percentile	+0.0387	Relative income at 10th percentile	−1.9028[a]
	(0.0220)		(0.8600)
Equilibrium error, 10th percentile	−1.3653[a]	Equilibrium error, 95th percentile	−0.8060[a]
	(0.4024)		(0.1973)
Prime age male joblessness rate	−0.2683[a]	Share of women in labor force (>19)	+2.9373
	(0.1212)		(3.1652)
Ratio of governmental money transfers to total compensation of employees	+0.5175		
	(0.3558)		
Percentage of female headed families	−0.5780		
	(0.4611)		
Adjusted-R^2	0.2263		0.3517
Durban-Watson test	2.375		1.9853

All of the regressors are lagged one period and all variables, except the equilibrium error term, are in first differences. A more exact description of the equations is presented in the text. Because we use first differences, one observation had to be dropped; four observations were also dropped due to the DOLS leads and lags in the first stage equation. (a) Statistically significant at the 5 percent level. Sources of data are presented in Appendix A.

the tradeoff is not supported empirically. It is important to note that the inverse relation between relative income *levels* at the extremes of the income distribution can be obtained if the inverse relationship of *changes* in these relative incomes shown in Table 4 is statistically significant for only the wealthy, and not the poor. Using relative incomes at the 20th and 80th percentile, Table B-4 shows no statistically significant relationship for either relative income change variable.

We know of no economic theory that explains such an asymmetry. It is certainly not a feature of the traditional class struggle approach. In that regard, it should also be noted that many of those at the 10th percentile are not engaged in the formal sector, so that their only impact on workers in this sector is to depress wages. The Marxist/populist approach suffers from other difficulties as well.[10] The asymmetry might also be statistical and due to our omission of a variable positively related to relative income at the 95th percentile which explains part of the movement of relative incomes in the 10th percentile, although we have not been able to come up with any such variable. Other explanations can be offered, but they are also not convincing.[11] In brief, although the inverse relation between high and low incomes is not spurious, the underlying economic mechanism of change in these relative income relations remains a puzzle.

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Neoclassical economists can point out that even if a negative relation exists between relative incomes of the rich and the poor, it does not negate their argument that policy analysis should be concerned with absolute, not relative, income. Moreover, although such a relationship may exist, it is irrelevant for questions of equity, which only concern absolute income.

This raises a much larger issue, namely what is the impact of income inequality on other variables of economic, political, or social concern? For instance, for a number of reasons that have received attention in the literature, a widening gap in relative income can lead to lower economic growth [Aghion, Caroli, and García-Peñalosa, 1999]. Income inequalities may raise social tensions and lead to socially wasteful expenses for police and other types of guard labor [Bowles, Gordon, and Weisskoff, 1990, 194], or to the conditions fostering hate groups [Jefferson and Pryor, 1999].

These broad and crucially important issues lie outside the scope of this essay, which has focused on the technical issue of whether the inverse relationship between relative incomes of the rich and the poor is spurious and whether a causal mechanism exists that links these two tails of the income distribution. Three key tasks lie ahead for future research: (a) To clarify the mechanisms underlying the inverse relation between relative incomes at the high and low ends of the income distribution; (b) To isolate various economic, political, and social effects that stem from the widening gulf in relative incomes between the rich and the poor; and (c) to determine what policy measures, if any, might be effective in alleviating problems arising from such a widening gap in relative income.

APPENDIX A
SOURCES OF DATA

The data on median family income at the 20th, 80th, and 95th percentiles, and female headed families come from the Census Bureau web site: http://www.census.gov/hhes/income/histinc . For family income at the 10th percentile we used data from various issues of the *Current Population Reports*, series P-60, that give the percentage of families with incomes below a certain income level, interpolating by assuming a log normal distribution of income.

Data on joblessness among prime age males (25 through 50) for 1964 through 1994 come from Current Population Survey, as calculated for each year in Pryor and Schaffer [1999, Chapter 1]. This series is highly correlated to a series of employed men divided by total population of men from 20 through 64 minus an estimate of male students, male members of the armed forces, and adult male prisoners (to derive the non-institutionalized population). The latter series was used to extrapolate from 1947 through 1964 and 1995 through 1997.

Data on government money transfers and total compensation of employees come from Department of Commerce, Bureau of Economic Analysis [1998], Tables 1.14, 3.15 and 3.16, supplemented by data from *Survey of Current Business*, August 1998. The transfers include those for welfare and social services, disability, and unemployment on both the federal, state, and local levels.

Data on women over 19 years as a share of the total labor force over 19 years come from various editions of the Council of Economic Advisers' *Economic Report of the President*, (annual).

APPENDIX B: RESULTS OF REGRESSION EXPERIMENTS WITH RELATIVE INCOME RATIOS AT THE 20TH AND 80TH INCOME PERCENTILES

TABLE B-1
Two Stage Least Squares Regressions Explaining Relative Income Levels, 1948 - 1997
Contemporaneous values of independent variables (except lagged dependent variable)

Dependent variables: / Regressors:	Relative income level at 20th percentile		Relative income level at 80th percentile
Constant	+0.503[a] (0.160)	Constant	+0.938[a] (0.305)
Dependent variable lagged one year	+0.455 (0.147)	Dependent variable lagged one year	+0.627[a] (0.099)
Year (1947 = 1)	+0.00007 (0.00042)	Year (1947 = 1)	−0.00091 (0.00181)
Relative income level at 80th percentile	−0.155[b] (0.080)	Relative income at 20th percentile	−1.019[a] (0.352)
Prime age male joblessness rate	-0.152 (0.115)	Share of women in labor force (>19)	−0.531 (0.427)

TABLE B-1 (Cont.)
Two Stage Least Squares Regressions Explaining Relative Income Levels, 1948 - 1997
Contemporaneous values of independent variables (except lagged dependent variable)

Dependent variables: / Regressors:	Relative income level at 20th percentile	Relative income level at 80th percentile
Ratio of governmental money transfers to total compensation of employees	+0.167 (0.214)	
Percentage of female headed families	−0.165 (0.333)	
Adjusted-R^2	0.7974	0.9573
Durban-Watson	1.616	1.909

Standard errors in parentheses.
a. Statistically significant at the 5 percent level.
b. Statistically significant at the 10 percent level.
Sources of data are presented in Appendix A.

TABLE B-2
Augmented Dickey-Fuller (ADF) Tests for New Variables in Regressions

	No trend	Trend
Relative income at the 20th percentile	−0.589 (2)	−1.774 (2)
Relative income at the 80th percentile	+0.189 (0)	−2.433 (1)

The number in parentheses are the optimal lag lengths: k − max = 4. The 5 and 10 percent critical values are: −2.914 and −2.598 (no trend) and −3.501 and −3.179 (trend). Sources of data are presented in Appendix A.

TABLE B-3
DOLS Cointegrating Vectors and Tests for Cointegration

Dependent variables / Regressors:	Relative income level at 20th percentile		Relative income level at 80th percentile
Constant	+0.6884[a] (0.1773)	Constant	+2.4014[a] (0.3992)
Year (1947 = 1)	+0.0012 (0.0008)	Year (1947 =1)	+0.0030 (0.0041)
Relative income at 80th percentile	−0.0879 (0.1376)	Relative income at 20th percentile	−1.8835[a] (0.4520)

TABLE B-3 (Cont.)
DOLS Cointegrating Vectors and Tests for Cointegration

Dependent variables / Regressors:	Relative income level at 20th percentile		Relative income level at 80th percentile
Prime age male joblessness rate	−0.1327 (0.2480)	Share of women in labor force (>19)	+0.2404 (1.0039)
Ratio of governmental money transfers to total compensation of employees	+1.1500 (0.6160)		
Percentage of female headed families	−0.9173 (0.6334)		
DOLS Lags	+2 to−2		+2 to−2
ADF test	−4.6651 (0)		−3.4172 (0)

Standard errors in parentheses for the point estimates. DOLS Lags are the leads and lags of the first differences of the right hand side variables used in the estimation of the cointegrating vector. For the ADF test, the numbers in parentheses are the optimal lag lengths: k−max = 4. The 5 and 10 percent critical values for the ADF tests are: −5.4914 and −5.1010. (a) Statistically significant at the 5 percent level. Sources of data are presented in Appendix A.

TABLE B-4
Error Correction Model for Changes in Relative Income Levels, 1948 - 1997

Dependent variables: / Regressors:	Relative income level at 20th percentile		Relative income level at 80th percentile
Constant	−0.0003 (0.0016)	Constant	+0.0066 (0.0058)
Dependent variable	+0.0481 (0.1710)	Dependent variable	+0.0384 (0.1431)
Relative income at 80th percentile	+0.0579 (0.0777)	Relative income at 20th percentile	−0.3570 (0.3041)
Prime age male joblessness rate	−0.1487 (0.1504)	Share of women in labor force (>19)	−0.4134 (1.3182)
Ratio of governmental money transfers to total compensation of employees	+0.1210 (0.4479)		
Equilibrium error, 20th percentile	−0.5523[a] (0.2646)	Equilibrium error, 80th percentile	−0.3583[a] (0.1580)
Percentage of female headed families	−0.6514 (0.5556)		
Adjusted-R^2	0.0276		0.0790
Durban-Watson test	2.094		2.003

All of the regressors are lagged one period and all variables, except the equilibrium error term, are in first differences. A more exact description of the equations is presented in the text. Because we use first differences, one observation had to be dropped; four observations were also dropped due to the DOLS leads and lags in the first stage equation. (a) Statistically significant at the 5 percent level. Standard errors in parentheses. Sources of data are presented in Appendix A. Since Table B-3 shows no cointegration, this type of calculation is not completely appropriate, but it is included to provide a basis of comparison for the other tables.

APPENDIX C: RESULTS OF REGRESSION EXPERIMENTS INCLUDING AVERAGE FAMILY SIZE AND GOVERNMENTAL TRANSFERS IN BOTH EQUATIONS

TABLE C-1
Two-Stage Least-Squares Regression Explaining Relative Income Levels, 1948 - 1997
Contemporaneous values of independent variables (except lagged dependent variable)

Dependent variables: / Regressors:	Relative income level at 10th percentile		Relative income level at 95th percentile
Constant	+0.609[a] (0.163)	Constant	+2.222[a] (1.056)
Dependent variable lagged one year	+0.129 (0.163)	Dependent variable lagged one year	+0.422[a] (0.114)
Year (1947 = 1)	+0.0030[a] (0.0006)	Year (1947 = 1)	+0.0119[a] (0.0060)
Ratio of governmental money transfers to total compensation of employees	+0.670[a] (0.238)	Ratio of gov't money transfers to total comp of employees	+2.619[a] (1.217)
Relative income level at 95th percentile	−0.081[a] (0.025)	Relative income at 10th percentile	−3.365[a] (0.889)
Average family size	−0.024 (0.027)	Average family size	+0.081 (0.166)
Prime age male joblessness rate	−0.088 (0.094)	Share of women in labor force (>19)	−1.103 (1.587)
Percentage of female-headed families	−1.188[a] (0.443)		
Adjusted-R^2	0.8649		0.9625
Durban-Watson	1.729		2.196

Standard errors in parentheses. (a) Statistically significant at the 5 percent level. Sources of data are presented in Appendix A.

TABLE C-2
Augmented Dickey-Fuller (ADF) Tests for New Variables in Regression

	No trend	Trend
Average family size	−0.843 (3)	−3.273 (3)

The numbers in parentheses are the optimal lag lengths: k− max = 4. The 5 and 10 percent critical values are: −2.914 and −2.598 (no trend) and −3.501 and−3.179 (trend). Sources of data are presented in Appendix A.

TABLE C-3
DOLS Cointegrating Vectors and Tests for Cointegration

Dependent variables / Regressors:	Relative income level at 10th percentile		Relative income level at 95th percentile
Constant	+1.2813[a] (0.4204)	Constant	+4.1847[a] (1.5815)
Prime age male joblessness rate	−0.1678 (0.1456)	Share of women in labor force (>19)	−2.5522 (2.9996)
Year (1947 = 1)	+0.0068[a] (0.0020)	Year (1947 =1)	+0.0255[a] (0.0110)
Ratio of governmental money transfers to total compensation of employees	+1.5808[a] (0.5175)	Ratio of gov't money transfers to total comp of employees	+2.2000 (2.7470)
Relative income at 95th percentile	−0.1007[a] (0.0271)	Relative income at 10th percentile	−7.7424[a] (2.7968)
Average family size	−0.1406 (0.0876)	Average family size	+0.2617 (0.4112)
Percentage of female-headed families	−3.5472[a] (1.5476)		
DOLS Lags	+2 to −2		+2 to −2
ADF test	−8.8164[a] (1)		−5.5103[a] (0)

Standard errors in parentheses for the point estimates. DOLS Lags is the leads and lags of the first differences of the right hand side variables used in the estimation of the cointegrating vector. For the ADF test, the numbers in parentheses are the optimal lag lengths: k−max = 4. The 5 and 10 percent critical values for the ADF tests are: −5.4914 and −5.1010. 0 (a) Statistically significant at the 5 percent level. Sources of data are presented in Appendix A.

TABLE C-4
Error Correction Model for Changes in Relative Income Levels, 1948 - 1997

Dependent variables: / Regressors:	Relative income level at 10th percentile		Relative income level at 95th percentile
Constant	+0.0006 (0.0015)	Constant	+0.0015 (0.0146)
Dependent variable	+0.1684 (0.1700)	Dependent variable	+0.1686 (0.1562)
Prime age male joblessness rate	−0.2700[a] (0.1300)	Share of women in labor force (>19)	+4.5874 (3.4804)
Relative income at 95th percentile	+0.0349 (0.0238)	Relative income at 10th percentile	−1.8964[a] (0.8731)
Equilibrium error, 10th percentile	−1.3776[a] (0.5484)	Equilibrium error, 95th percentile	−0.9688[a] (0.2334)
Ratio of governmental money transfers to total compensation of employees	+0.4968 (0.3804)	Ratio of gov't money transfers to total comp of employees	−3.5951 (1.9911)

TABLE C-4 (Cont.)
Error Correction Model for Changes in Relative Income Levels, 1948 - 1997

Dependent variables: / Regressors:	Relative income level at 10th percentile		Relative income level at 95th percentile
Average family size	−0.0029 (0.0564)	Average family size	+0.6278 (0.3451)
Percentage of female headed families	−0.6235 (0.4983)		
Adjusted-R^2	0.1172		0.3682
Durban-Watson test	2.345		2.011

All of the regressors are lagged one period and all variables, except the equilibrium error term, are in first differences. A more exact description of the equations is presented in the text. Because we use first differences, one observation had to be dropped; four observations were also dropped due to the DOLS leads and lags in the first stage equation. (a) Statistically significant at the 5 percent level. Standard errors in parentheses. Sources of data are presented in Appendix A.

NOTES

The authors thank Samuel Bowles, Thomas Dee, Philip Everson, the editor of this *Journal*, Zora Pryor, and two anonymous referees for extremely helpful comments on earlier versions of this paper. Jefferson also gratefully acknowledges the support of a Swarthmore College Faculty Research Grant and a Eugene M. Lang Faculty Fellowship.

1. For instance, the ratio of income of the poor to the rich might fall, and yet incomes of families between these income extremes might cluster more closely to the median so that the overall income distribution is more equal according to some measures of inequality. Nevertheless, Raj and Slottje [1994] show graphically, using nine different measures of income inequality, that the various measures of income inequality move downward and together in an irregular pattern from 1947 to 1967 - 69, at which point they move more smoothly upward and together. Therefore, the theoretical possibility that we have raised does not seem likely. Raj and Slottje focus their attention on econometrically documenting the structural break that occurred in the late 1960s; this paper focuses more on the underlying causal factors.
2. We have chosen to use family money income before taxes because this series is available from the Bureau of the Census' *Current Population Reports*, Series P-60 for the longest period (and also found on the Census Bureau's website). This definition of income omits capital gains, non-monetary income (such as imputed rents of owner occupied housing or non-monetary fringe benefits such as health insurance), and the impact of taxes. Studies using shorter time periods have employed a broader definition of income and, in addition, have focused on adjusted income per person, where each person in a family is considered to have an income equal to the total family income divided by the adjusted number of people in the family. While this is a better measure of income and welfare, the various measures of income inequality derived from such measures move closely together with the family income series.
   The Census Bureau presents data on money income of families at various quintiles, the median, and the 95th percentile as well as income data of the percentage of families with incomes less than various absolute amounts. From the latter we can interpolate, assuming a log normal distribution of income, to calculate the money income at the 10th percentile as well.
3. By way of contrast, some studies of income inequality include both a variable for unemployment and inflation. Although the linkage between labor market weakness and relative money income in the lowest part of the income distribution is clear, its linkage with the top is not. Furthermore, once labor

market variables are included in the regressions, the theoretical linkage between inflation and income inequality is not apparent. Certainly annual price increases and the higher interest rates accompanying them should have little impact on the money income of the poor; for the rich, their interest income may be higher, but this is a very minor part of their total income. For these reasons we have not included an inflation variable.

4. In particular, average family size comes to mind because it has two important impacts on the family income of the poor. First, a larger family size might indicate the presence of two parents. Second, for single parent families it might indicate that the family receives more welfare income. The first effect is taken care of with our variable indicating the percentage of single parent families; the latter is taken care of with our welfare expenditures variable. In Appendix C, we report the results of statistical experiments that include this demographic variable and find that the overall results strengthen slightly the conclusions reported in the text.

5. Some serious conceptual issues arise in determining whether the income of wives equalizes total family income or not, and in some types of comparisons, for instance, those made by Cancian and Reed [1999], the income of wives appears to make the distribution of income more equal. In our regressions, it turns out that the female participation rate has no significant impact on the relative income ratios that we calculate.

   On another matter, it might be argued that we should include some indicator of the share of property income or the profit rate, since these might be causally related to the income of the rich, but not the poor. The use of such a variable, however, raises a serious simultaneity problem in the model. More specifically, property income is not only a component of the income of the wealthy, but the spending of a marginal dollar by the wealthy raises rents on property held by another segment of those with high incomes. Moreover, experiments with this variable at a very early stage of this research showed that it added very little explanatory power to the regressions in Table 1 and, as a result, was dropped from further consideration.

6. A referee suggested that we include governmental transfers as a ratio of total compensation of workers and employees, presumably because such transfers come at the expense of the wealthy and serve as a work disincentive. We experimented with this variable in Appendix C but it has the wrong sign and is not statistically significant.

7. For the 20th percentile, the variation is from 45.7 to 53.0 percent of the median income, with a coefficient of variation of 4.1 percent. For the 80th percentile, the variation is from 150.8 to 180.7 percent and the coefficient of variation is 5.4 percent.

8. Blough [1992] and Cochrane [1991] discuss some of the limitations of unit root tests. MacKinnon's [1991] Table 1, is used to generate the critical values for all of the unit root tests reported. The lag lengths indicated resulted from the use of the selection procedure suggested by Campbell and Perron [1991].

9. For relative income in the 10th percentile, the 95 percent confidence interval for the error correction coefficient is $(-2.154, -0.577)$. Thus, the hypothesis that if relative income for the 10th percentile was above its equilibrium level by one percentage point in the previous year, it is expected (other things held equal) to fall by one percentage point or less in the current year, cannot be rejected.

10. For instance, they do not take into account income mobility of individual families. Various empirical studies, such as Gottschalk and Danziger [1998], show some year-to-year movement of many families from one income quintile to another.

11. A microeconomic analogy might be suggested. Changes in the relative income at the 95th percentile are like the situation where, if wages change and the company maintains the same price, changes in high incomes (managerial bonuses and profits) and low incomes (wages) move in opposite directions. If high incomes are, in addition, influenced by shocks, for instance, a change in price or in labor productivity, while low incomes are not, then low incomes may not be inversely related to high incomes. In brief, high income receivers absorb the major price and productivity shocks of the system so that the income of the poor is not greatly affected. Results from some preliminary tests of this approach did not seem promising and it was abandoned.

REFERENCES

Aghion, P., Caroli, E. and García-Peñalosa. C. Inequality and Economic Growth: The Perspective of the New Growth Theories. *Journal of Economic Literature* , December 1999, 1615-61.

Bishop, J. A., Formby, J. P. and Sakano, R. Estimating Changes in the Distribution of Income in the United States. *Journal of Income Distribution* No. 1, 1994, 79-105.

Blank, R. M. and Blinder, A. S. Macroeconomics, Income Distribution, and Poverty, in *Fighting Poverty: What Works and What Doesn't*, edited by S.H. Danziger and D. H. Weinberg. Cambridge, Mass: Harvard, 1986.

Blough, S. The Relationship between Power and Level for Generic Unit Root Tests in Finite Samples. *Journal of Applied Econometrics*, July-September 1992, 295 -308,

Bowles, S., Gordon, D. M. and Weisskoff, T. E. *After the Waste Land: A Democratic Economics for the Year 2000*. Armonk: M.E. Sharpe, 1990.

Burtless, G. Technological Change and International Trade: How Well Do They Explain the Rise in U.S. Income Inequality, in *The Inequality Paradox: Growth of Income Disparity*, NPA Publication 228, edited by J. A. Auerbach and R.S. Belous. Washington, D.C.: National Policy Association, 1998, 66-91.

Campbell, J. Y. and Perron, P. Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots in *NBER Macroeconomics Annual 1991*, edited by O. J. Blanchard and S. Fisher. Cambridge, Massachusetts:MIT Press, 1991, 141-201.

Cancian, M. and Reed, D. The Impact of Wives' Earnings on Income Inequality: Issues and Estimates. *Demography*, May 1999, 173- 4.

Cochrane, J. H. A Critique of the Application of Unit Root Tests. *Journal of Economic Dynamics and Control,* April 1991, 285-84.

Council of Economic Advisers. *Economic Report of the President*. Washington, D.C.: G.P.O, Annual.

Federal Reserve Bank of Kansas City, ed. *Income Inequality: Issues and Policy Options*. Kansas City, 1998.

Feldstein, M. Overview, in Federal Reserve Bank of Kansas City, 1998, 357-67.

Granger, C. W. J. and Newbold, P. Spurious Regressions in Econometrics. *Journal of Econometrics*, July 1974, 111-20.

Gottschalk, P. and Danziger, S. Family Income Mobility - How Much Is There, and Has It Changed? in *The Inequality Paradox: Growth of Income Disparity*, NPA Publication 228, edited by J. A. Auerbach and R.S. Belous. Washington, D.C.: National Policy Association, 1998, 92-112.

Jefferson, P. N. and F. L. Pryor. On the Geography of Hate. *Economics Letters* No. 65, 1999, 389-95.

Johansen, S. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica,* December 1991, 1551-80.

Karoly, L. A. and Burtless, G. Demographic Change, Rising Earnings Inequality, and the Distribution of Personal Well-Being, 1959 - 1989. *Demography,* August 1995, 379-405.

MacKinnon, J. G. Critical Values for Cointegration Tests, in *Long-Term Economic Relationships: Readings in Cointegration*, edited by R. F. Engel and C. W. J. Granger. New York: Oxford University Press, 1991. 267 - 76.

Mocan, H. Naci. Structural Unemployment, Cyclical Unemployment, and Income Inequality. *Review of Economics and Statistics*, February 1999, 122- 34.

Pryor, F. L. and Schaffer, D. L. *Who's Not Working and Why: Employment, Cognitive Skills, Wages, and the Changing U.S. Labor Market.* New York: Cambridge University Press, 1999.

Raj, B. and Slottje, D. The Trend Behavior of Alternative Income Inequality Measures in the United States From 1947 - 1990 and the Structural Break. *Journal of Business and Economic Statistics* No. 4, 1994, 479-87.

Stock, J. H. and Watson, M.W. A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems, *Econometrica*. July 1993, 783-820..

U.S. Bureau of the Census. *Current Population Reports.* Washington D.C.:GPO, Annual.

U. S. Department of Commerce, Bureau of Economic Analysis. *National Income and Product Accounts of the United States, 1929-94.* Washington, D.C.: GPO, 1998.