The Rich, The Poor, and The Changing Gap:
An Investigation of the Determinants of Income Inequality from 1996-2002

Thomas Clark

The College of New Jersey

April 2004

I. Introduction

The gap between the countries richest and poorest is widening again. Wage earners in the lowest 10% have suffered a decrease in their wages in the last 4 years, while average income in the top 10% has remained unchanged. (Ip, 2004) Income inequality is an important issue for public policy makers. The latest round of tax cuts passed amid accusations that the cuts helped only the wealthiest Americans. Others characterized recent changes in welfare benefits as harmful to the poor. While this paper will make no recommendations on whether specific public policy is correct or not, this paper will investigate the determinants of the changing gap in income inequality of American households.

It is the aim of this paper to take further the research done by the Economic Policy Institute. Bernstein et al. (2003) offered a good assessment of trends in the income levels across the 50 states, comparing the top 20 percent to the bottom 20 percent, but failed to offer a detailed statistical analysis of the causes of the inequality. What follows is an investigation of the causes, drawn from extensive past research, using the income inequality measurement employed by Bernstein and regressed over the years 1996-2002.

II. Background

During the period of economic prosperity after World War II, most Americans were able to benefit from the great opportunity the country faced.

America entered the 1950's as the single most powerful country in the world. At that time, Kuznets (1955) theorized in his paper, that with continued economic development, a more equalized income distribution would follow. Kuznets himself admits in his paper that his theory was 5% empirical, 95% theoretical. This was the case because a serious lack of empirical data prevented him from doing any kind of serious statistical analysis.

Since 1955, a series of studies have been conducted since to test Kuznet's claim. It took some 25 years for the empirical data to catch up, but the 1970's proved to be the turning point for income inequality in the U.S. Between the 1950's and 1970's the income gap between the richest and poorest moved in accordance with Kuznet's projections, while after 1970 the gap began to expand, even though economic growth continued. (Levy and Murnane, 1992).

While we may argue that growth was anemic during the 1970's and as a consequence inequality worsened, such assertions became untenable by the 1980's. During the 1980's a huge boom of economic prosperity was followed by the fastest growing inequality ever (Levy and Murnane, 1992). Studies have found during the decade, the richest in the country prospered while the economic situation of the poor failed to improve. A series of studies (Burtless, 1990; Bound and Johnson, 1992; Katz and Murphy, 1992) all concluded that the changing labor market was the reason for the income gap separation.

The 1990's saw one of the biggest economic expansions the United States ever experienced. If Kuznets was right, the gap between the richest and the poorest should have narrowed, or at the very least stay the same, with both groups moving at

the same rate of change. Bernstein et. al. (2002) compared the top 20 percent of income earned by families as compared to the bottom 20 percent. They found that across the board, each quintile experienced rising income levels. But, the highest gains registered for individuals in the top 20 percent, with only slight gains for those at the bottom 20 percent. When a country prospers as the United States did in the 90's, there is an ethical argument that everyone should be gaining from it equally. Bernstein et. al. found that towards the end of the decade (1998 through 2000), the gap between the top quintile and bottom quintile began to shrink. To address what they felt was an anomaly, they referenced the limited nature of their data source, the Current Population Survey, which was unable to capture income that might be deferred to later years like stock options.

Bernstein et. al. found that in 44 out of 50 states, the gap between the top 20 percent and bottom 20 percent widened over the last two decades. Out of those 44 states, statistics showed that in 5 states average incomes for the top 20 percent rose while average income for the bottom 20 percent fell. In 39 states income for both the top 20 percent and the bottom 20 percent rose, however incomes for the top 20 percent rose faster. Of the other 6 states, only 4 states had the incomes of the richest and poorest growing at the same rate, 1 state had no change, and only 1 state showed a decrease in the gap between the rich and poor decrease because of growth only in the bottom quintile. By the late 1990's, the average income of the bottom 20 percent was $14,620, while the average for the top 20 percent was $145,990. For the sake of the study, the size of the families did not matter when calculating the percentages.

Much research has been done across the last 3 decades, encompassing different methods of measurement, different techniques, and different goals. A series of studies (Bishop, 1992; Partridge, 1996; Topel, 1994; Gottschalk, 1997; Lerman, 1996; Kim, 2003; Bernstein, 2003) all show that greater female participation in the work force has led to an overall lowering of household inequality. The inflow of female workers could possibly be due to hardships felt at home, so more females have gone to work. In addition, adding a second income in a household with one low-wage worker can reduce or eliminate the effect of low hourly earnings on family income. Bernstein et. al. (2002) found that more households were headed by females than in the past. Lerman theorized that due to divorce and the failing of the family structure, inequality will continue to rise as more women begin to work and live on their own. Female-headed households generally increase inequality. To the extent that female-headed households also raise labor force participation of women, the effect must be smaller than the two-income effect described above.

Other studies (Partridge, 1996; Topel, 1994; Gottschalk, 1997; Davies, 1992; Bernstein, 2002) have found a link between immigration and income inequality. The relationship stems from the inflow of low-wage workers, who then compete with the low-wage workers who are already here. This not only creates a larger number of unskilled workers at the bottom of the income spectrum, but the over supply of unskilled workers could possibly bring down the wages for other unskilled workers. Davies (1992) considers different types of immigration, low wage or high wage, the latter he refers to as "brain drain." An influx of low-wage workers tends

to increase inequality, but an influx of "brain drain" can have the effects of both raising and lowering inequality. The effect depends on the nature of the host country and its level of international trade and globalization.

Related to issues of immigration are the theories around concentrations of people in cities. Partridge (1996) and Long (1977) find that inequality is higher when there is a bigger share of people living in cities. Long's research found that the bigger the city size, the larger the inequality. Partridge's panel-data study concluded that the metropolitan shares of population were significant determinants of inequality.

There is no consensus on the effects of unionization and the manufacturing sector on inequality. (Partridge, 1996; Topel, 1994; Kim, 2003; Bernstein, 2002) Topel, Bernstein, and Kim all found that the decrease in unionization has led to the increase in income inequality. They argue that decreases in unionization reduce real wages because of decreases in bargaining power. As technological changes reduce manufacturing employment, the service sector adds jobs that require higher levels of training and education. Only Partridge, in his panel data analysis, found that the decline in unionization had little effect on the increase in inequality.

Autor et. al. (2004) has found that over the last two decades, there has been a substantial increase in the demand for skilled labor that can complete non-routine tasks. The reason is that computer capital substitutes for workers who complete a limited and easily described set of activities (i.e. routine tasks). While on the other hand, computer capital compliments problem solving and complex communication activities (non-routine tasks). They also found that there has been an increase in the

demand for the number of college-educated workers as compared to those with high school educations. The results of their study find that the increase in demand for nonregular computer skilled labor is one of the main factors behind the increased demand for the college educated work force. The percentage of college-educated labor has been shown in a study by Partridge (1996) to increase income inequality.

III. Empirical Model and Data

Analysis of the determinants of income inequality will be conducted using panel data for the 50 United States and for the years 1996-97 and 1999-2002. In the past, panel data has been used, but only incorporating information from the census done every decade. Using individual years allows the research to capture changing trends that affect earnings in the short term. The fixed effect model used in this study overcomes a biased result that can result when a common intercept would be used. The fixed-effects model assumes that differences across states are captured by differences in the constant term.

The dependent variable for this study is an income inequality measure that was calculated for this paper and not available elsewhere. As mentioned before, the variable is a ratio of average income of households, of the highest quintile to the lowest quintile, for each state (income).[1] This information is not published on an

[1] Braun (1988) notes limitations of the GINI coefficient as a measure of income inequality. The GINI coefficient is based on the Lorenz curve. Both of these methods of inequality measurement are indices that will rank the states in some ordinal fashion of least equal to most equal. The GINI coefficient can sometimes hide or exaggerate

annual basis from the Bureau of Labor Statistics (BLS). Any annual data is collected using the Current Population Survey that is published in the March supplement by BLS. In the past, typically 60,000 people were surveyed, asking a variety of questions. In recent years, this number has been increased. BLS does not publish annual data by state for income, because they believe that an annual sample of size of 500 to 1000 per state is too small. While the sample size is relatively small, we contend that the length of the time-series for each state is sufficient to produce reliable estimates using a panel estimation procedure.

In order to calculate the ratio, the raw data sets from the March supplement were obtained from BLS. Two variables were drawn out from the data. The first was **HTOTVAL,** which is defined as the total household income amount for the year before. Due to the lag in time, the data published for 2003 was used for the ratio in 2002. 2002's data was used for 2001's ratio, etc. It incorporates all different types of income, including wages, dividends, interest, rental, and public assistance. The second variable used was **GMSTCEN.** This variable is the 1960 census code for the state that each household resides in. To calculate the average incomes, the data was sorted by state, and then ranked. Each ranked group was

---

relationships of independent variables. For example, the GINI coefficient will have a greater change when the middle of the income spectrum changes. This is not helpful when testing income differences between the richest and the poorest. This assessment is also backed up by a studies done in the 1970's (Soltow, 1971; Reynolds and Smolensky 1977). Braun also discusses a method developed by Nelson (1984) which compares the income of the lowest 5 percent to the highest 5 percent. Comparing the average incomes of the top 20 percent to the bottom 20 percent allows us to directly assess how the top and bottom incomes of the country are changing.

separated into quintiles, and then average income was then taken from each quintile. The mean inequality ratio across all observations was 12.127. The number can be defined by saying, that on average for every $1 earned by a household in the lowest quintile, a household in the highest quintile earned $12.18.

The majority of independent variables were obtained from the Statistical Abstract of The United States. Female labor force participation rate (female) is the percentage of the female civilian population that is currently employed or is classified as unemployed. This includes those who are seeking employment and those that are available for employment. The reason for the year 1998 being excluded from the study stems from the use of this variable. For reasons that are not clear, the participation rate was not calculated for 1998 and has not been published by BLS. The percentage of the population with high school diplomas (high) and with a bachelor's degree or better (bach) includes only those who are 25 years or older. The variable used for immigration (immig) was calculated from dividing the number of immigrants intended for each state, provided by the abstract and dividing it by the population estimate for each state, also provided by the abstract. The number of immigrants calculated are those that have registered with all relevant federal agencies and are coming in from foreign countries. It does not include an estimate of illegal immigrants or inter state migration. This resulted in a percentage of the population that is new legalized immigration each year. Percentage of the work force that has manufacturing jobs (manuf), those that are covered by unions (union), and the unemployment rate (unemp), were also obtained from the Abstract.

The two remaining variables were calculated using similar methods as used to find the income inequality measure. The first remaining variable is the percentage of the population of each state that lives in an urban area (urban). The variable **GMSTCEN** was again used to sort by state. The second variable **GMMETSTA** is a coded variable asked of each respondent that categorizes them into either living in a Metropolitan Statistical Area (MSA), a non MSA, or not identifiable. The percentage in the data represents those respondents that answered that they lived in a MSA.

The remaining variable is a calculation of the percentage of households that are headed by unwed females that contain family (unwed). According to CPS, a family is defined as a group of two persons or more residing together and related by birth, marriage, or adoption. The variable **H_TYPE** codes each household into different categories, including those headed by married male and females, single non family males or females, and unwed family males or females. The respondents were sorted by state and the percentages were calculated.

Based on the discussion in the background, the inequality measure (income) was regressed against female participation rate (female), new immigration rate (immig), percentage of manufacturing jobs (manuf), percentage of the population with high school diplomas (high), or the percentage of the population with bachelors degrees or better (bach), percentage of the population living in an urban area (urban), percentage of population being covered by a union (union), the unemployment rate (unemp) and the percentage of households head by an unwed female family member (unwed). An increase in female, manuf, high, bach, and

union are expected to lessen the inequality gap, while an increase in immig, urban, unemp, and unwed are expected to widen the gap. The equation is as follows

(1) $$Income_{it} = \beta_1 Female_{it} + \beta_2 Manuf_{it} + \beta_3 Immig_{it} + \beta_4 High_{it} + \beta_5 Bach_{it} + \beta_6 Urban_{it} + \beta_7 Union_{it} + \beta_8 Unemp_{it} + \beta_9 Unwed_{it} + u_{it}$$

where i indexes states, t indexes time and u_{it} = the transitory error term that varies across states and time-periods.

IV. Results

The summation of important statistics for the variables can be found in Table 1. In the year 2002, Tennessee (15.75), New York (15.61), and Louisiana (14.72) had the highest ratio, while Nevada (9.39), Utah (9.23), and Idaho(9.05) had the lowest. Regressions results can be found in Table 2. A collinearity problem arose between the variables high and bach, thus two different regressions were run instead. Results for the remaining variables changed only slightly for each of the regressions. The significant variables found were female labor force participation rate, percentage of new immigration, and percentage of population living in a metropolitan statistical area. For both regressions, a 1 percentage point increase in female labor force participation rate, decreases the income that a household in the highest quintile makes by 19 cents for every $1 that is made by someone in the lowest quintile. This is consistent with findings by Bishop, 1992; Partridge, 1996; Topel, 1994; Gottschalk, 1997; Lerman, 1996; Kim, 2003; and Bernstein, 2003. One can theorize that household income for the lowest quintile is increasing because many homes are becoming a two-income household, and the female income is

supplementing male incomes. This argument is also supported by the insignificant results for the percentage of the households headed by unwed females. As described above, supplementing a male income, overpowers any effect on inequality that might have been affected from unwed female led households.

Higher levels of immigration raise income inequality. An increase in the level of immigration (as percentage of population), by one tenth of one percentage point increased the inequality ratio by about 37 cents for regression 1 and 39 cents for regressions 2. This was consistent with work done by Partridge, (1996); Topel, (1994); Gottschalk, (1997); Davies, (1992); Bernstein, (2003). Immigrants coming into the United States are mostly made up of either two groups. The first are relatively unskilled low-wage workers, who will now compete with low-wage workers that are already here, thus continuing to lower overall household income in the lowest quintile. The other group is highly educated immigrants who are coming to the United States to seek high-paid positions that further increase the income of the highest quintile.

A one-percentage point increase in a states population living in an urban area decreased the inequality ratio by about 6 cents for both regressions. This was consistent with Kuznet's theory that increased urban populations, means continued economic growth, therefore lower inequality. But the results found in the paper are different than those found by Partridge (1996). He found that increased urban population caused greater inequality. Those results were found using the GINI coefficient as the measure of inequality and a study was done of the 48 lower states

in 1960, 1970, 1980, and 1990. Unfortunately, Partridge did not offer any explanation for the empirical results.

The average unemployment level proved to be an insignificant cause of income inequality. These results are also duplicated by Partridge (1996). The results show that with either an increase or decrease in the employment level, the jobs gained or lost are from across the spectrum of jobs, not just high or low paying jobs. This proves that the solution job creation will not be successful in lowering inequality.

Educational attainment measured as either percentage of the population with a bachelor's degree or percentage of the population with a high school degree had no significant effect on income inequality. This is inconsistent with the hypothesis discussed earlier stemming from the research done by Autor (2004) and Partridge (1996). The percentage of the population with a high school degree or college degree, were being used as a proxy for basic skills. This suggests that low-skilled and high-skilled workers and the firms that employ them, migrate across U.S. states, so that at a point in time the return to skill is the same across states, even if it is growing over time. Another possible explanation for the lack of significance for the percentage of high school graduates is that all high school educations are not created equal. The possibility exists that those households that are in the bottom 20 percent for income, tend to live in areas where public education problems exist. The education received in these areas, amounts to little or no difference than if they had no education at all.

The level of the population covered by unions had very insignificant results on income inequality. This result is similar to what Partridge (1996) found in his panel data study of the 48 lower states. Results in the past had found that a decrease in unionization led to a decrease in income inequality. (Topel, 1994; Kim, 2003; Bernstein, 2002) The explanation for this might be that unionization does not cover those who earn the lowest 20 percent income. A decline in unionization quite possibly has affected the incomes of those in the middle class, but not enough for that to affect those in the lowest 20 percent. The percentage of manufacturing jobs on income inequality was also found to be insignificant. The decline in manufacturing jobs forcing the labor force to either become service related or move down the scale in income to lesser paid jobs had little effect on the income inequality ratio.

V. Conclusion

An empirical panel data study was done of the determinants of household income inequality for the 50 U.S. states from the years 1996-2002. In recent years, as economic growth continues, income inequality has increased. A ratio of household income of the top quintile compared to the bottom quintile was constructed to measure the inequality. The results have left more questions than answers. This paper has found that increased female labor force participation rate has a decreasing effect on household income inequality. So does more of a

population living in an urban area. An increased percentage of new immigration in a state increases household income inequality.

Percentage of the population with a high school diploma, percentage of the population with bachelors degrees or better, percentage of manufacturing jobs, percent of population covered by a union, the unemployment level, and the percentage of households headed by an unwed mother had no effect on the level of household income inequality.

Work Cited

Al-Samarrie, Ahmad and Herman P Miller, 1967, "State Differentials in Income Concentration," *American Economic Review*, 57: 59-72

Autor, David H., Frank Levy, and Richard J. Murnane, 2004, "The Skill Content of Recent Technological Change: An Emperical Exploration," *The Quarterly Journal of Economics*, 18: 1279-1333

Bernstein, Jared, Heather Boushey, Elizabeth McNichol and Robert Zahradnik, 2002, "Pulling Apart: A State-by-State Analysis of Income Trends," *Economic Policy Institute*

Bound, John and George Johnson, 1992, "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *American Economic Review*, 82: 371-392

Bishop, John A, John P Formby and Paul D Thistle, 1992, "Explaining Interstate Variation in Income Inequality," *Review of Economics and Statistics*, 74: 553-557

Braun, Denny, 1988, "Multiple measurements of U.S. Income Inequality," *Review of Economics and Statistics*, 70: 398-405

Bureau of Labor Statistics, 2003, "Local Area Unemployment Statistics," US GPO, Washington DC. (http://www.bls.gov/lau/home.htm).

Burtless, Gary, 1990, "A Future of Lousy Jobs?" Washington, DC: Brookings Institution, 1-30

Davies, James B and Ian Wooton, 1992, "Income Inequality and International Migration," *The Economic Journal*, 102:789-802

Gottschalk, Peter and Timothy M Smeeding, 1997, "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature*, 35: 633-687

Ip, Greg, 2004, "The Gap in Wages is Growing Again For U.S. Workers," *Wall Street Journal*, January 23

Kim, Jongsung, 2003, "U.S. Earnings Inequality in the 1990's," *Journal of American Academy of Business*, Cambridge, 2: 371-378

Kuznets, Simon, 1955, "Economic Growth and Income Inequality," *American Economic Review*, 45: 1-28

Lerman, Robert I., 1996, "The Impact of the Changing U.S. Family Structure on Child Poverty and Income Inequality," *Economica*, 63: 119-139

Levy, Frank and Richard J. Murnane, 1992, "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations," *Journal of Economic Literature*, 30: 1333-1381

Long, James E, David W Rasmussen, and Charles T Haworth, 1977, "Income Inequality and City Size," *Review of Economics and Statistics*, 59: 244-246

Nelson, Joel I., 1984, "Income Inequality: The American States," *Social Science Quarterly*, 65: 854-860

Partridge, Mark D, Dan S Rickman and William Levernier, 1996, "Trend in U.S. Income Inequality: Evidence from a Panel of States," *Quarterly Review of Economics and Finance*, 36: 17-37

Topel, Robert H., 1994, "Regional Labor Markets and the Determinants of Wage Inequality," *American Economic Review*, 84: 17-22

U. S. Bureau of the Census, various years, *Statistical Abstract of the United States,* US GPO, Washington DC. (http://www.census.gov/statab/www/)

Table 1. Means and Standard Deviations

Variable	Mean	Standard Deviation	Minimum	Maximum
Income	12.127	1.888	7.689	18.927
Female	61.27	4.298	47.700	71.200
Manuf	.135	.051	.027	.239
Immig	.00202	.00162	.000216	.008293
High	84.51	4.318	72.70	92.80
Bach	24.44	4.539	12.70	38.70
Urban	.680	.214	.211	1.00
Union	14.11	5.711	3.80	28.20
Unemp	4.63	1.153	2.20	8.10
Unwed	.118	.025	.057	.200

$Income_{it}$ = Ratio of average household income of lowest and highest quintile for state i in year t.
$Female_{it}$: Female labor force participation rate for state i in year t.
$Manuf_{it}$: Percentage of the labor force that is a manufacturing job for state i in year t.
$Immig_{it}$: Percentage of new foreign immigration for state i in year t.
$High_{it}$: Percentage of the population with high school diploma for state i in year t.
$Bach_{it}$: Percentage of the population with bachelor's degree or better for state i in year t.
$Urban_{it}$: Percentage of the population living in a MSA for state i in year t.
$Union_{it}$: Percentage of the labor force covered by a union for state i in year t.
$Unemp._{it}$: Unemployment rate for state i in year t.
$Unwed_{it}$: Percentage of households headed by unwed family females for state i in year t.

Table 2. Fixed-effects regression results for household income inequality of lowest to highest quintile

	Fixed Effects	Fixed Effects
Variable	(1)	(2)
Female	-0.196***	-0.199***
	(0.0643)	(0.0643)
Manuf	-5.759	-4.279
	(6.885)	7.141
Immig	365.04**	386.72***
	(150.4)	149.9
High		0.0731
		0.0480
Bach	0.0562	
	(0.0429)	
Urban	-6.056*	-6.206*
	(3.689)	(3.689)
Union	0.0138	0.0063
	(0.0701)	(0.0690)
Unemp	0.00525	0.02322
	(0.0976)	(0.1020)
Unwed	-1.523	-1.869
	(5.363)	(5.350)
R-Squared	0.7497	0.7503
Observations	n = 300	n = 300

Dependent variable: $Income_{it}$ = Ratio of average household income of lowest and highest quintile for state i in year t.
Standard errors in parentheses.
*** = significant at 0.01.
** = significant at 0.05.
* = significant at 0.10.
All cross-section estimates are suppressed.
$Female_{it}$: Female labor force participation rate for state i in year t.
$Manuf_{it}$: Percentage of the labor force that is a manufacturing job for state i in year t.
$Immig_{it}$: Percentage of new foreign immigration for state i in year t.
$High_{it}$: Percentage of the population with high school diploma for state i in year t.
$Bach_{it}$: Percentage of the population with bachelor's degree or better for state i in year t.
$Urban_{it}$: Percentage of the population living in a MSA for state i in year t.
$Union_{it}$: Percentage of the labor force covered by a union for state i in year t.
$Unemp._{it}$: Unemployment rate for state i in year t.
$Unwed_{it}$: Percentage of households headed by unwed family females for state i in year t.

Figure 1- Income Inequality of Top Quintile to Bottom Quintile for the U.S.