Manufacturing Wage Inequality in the Appalachian Region

A Report for the Appalachian Regional Commission

By

James K. Galbraith and Pedro Conceição

Lyndon B. Johnson School of Public Affairs
The University of Texas at Austin
Austin, Texas 78713-6450

May 2001

Executive Summary

This study examines inequality in manufacturing pay between manufacturing plants within states and counties of the United States, and compares the trends in such inequality with those in Appalachia over the 1963 to 1992 period. The statistical measure of inequality used in this study, the Theil statistic, is applied to measure the dispersion of average payrolls across manufacturing plants *within* and *between each state* and *county*.

This approach allows one to determine whether the patterns of inequality differ across states and regions, and to assess whether the experience of the Appalachian Region, taken as a whole, differs from that of the rest of the nation.

While a measure of inequality across manufacturing facilities cannot substitute for broader measures of inequality in household income distribution, and other measures of well-being, it does have two advantages. First, it provides an accurate measurement at the county level that permits the grouping of counties—using this harmonized measure of inequality—at the state and regional level. Second, the trends in the dispersion of pay across manufacturing plants are likely to track broader measures of inequality that cannot be directly measured at this level of geography.

For comparative purposes, the analysis of manufacturing pay inequality proceeds from the national, to the state and county level, and thence to the regional level. At the national level, the analysis indicates that there was a reduction in pay inequality within states from 1963 to 1967, followed by a steady increase at an annual average rate of 2 percent from 1967 to 1992. Over this time, increasing inequality in manufacturing pay overall has been largely driven by a *widening average pay gap within states, rather than an increasing gap between rich and poor states*. This does not mean that there has been a *convergence* of average wages across states, only that between-state inequality has not worsened markedly, while inequalities within-states have grown.

The measure of inequality within states can be broken into two components: the contribution to pay inequality of the dispersion *between* counties, and the contribution of dispersion *within* counties. This analysis shows that *changes within counties account for the largest share of the widening pay gap*, increasing from 69 percent in 1963 to 76 percent in 1992 of total measured inequality. Thus, the importance of differences across states has diminished, while the increase in manufacturing pay inequality is largely due to the growing dispersion of wages within states and counties.

In order to analyze the relative effect of changes in inequality in the Appalachian Region, the inequality measure is divided into five components. The first component measures inequality between Appalachian counties; while the second component measures dispersion of wages within Appalachian counties. The third and fourth components measure inequality between and within non-Appalachian counties. The final component measures dispersion between the group of Appalachian counties and the group of non-Appalachian counties.

In general most of the Appalachian states exhibit a similar overall pattern of a decrease in inequality from 1963 to 1967, and then an increase in inequality from 1972 to 1992. The exceptions are the states of Kentucky, Tennessee and Virginia, which exhibited a halt to increasing inequality after 1987, and a decrease through 1992.

Differences *between Appalachian and non-Appalachian counties* are small for every state, which indicates that there are not sharp differences in average wages. In Alabama and Mississippi the differences narrowed, but most states showed stable differences. The differences between Appalachian and non-Appalachian counties increased in Kentucky, Maryland, Pennsylvania and Virginia from 1987 to 1992.

Inequality *between and within non-Appalachian counties* accounted for the largest share of wage dispersion within most Appalachian states, except in West Virginia, Alabama, Tennessee and Pennsylvania where the Appalachian portion of the state makes up all or a significant share of employment and wages.

In all 13 Appalachian states inequality *within Appalachian counties is higher than inequality between Appalachian counties*. In other words average wage levels are similar across Appalachian counties. Furthermore, inequality *within* Appalachian counties is more *volatile over time* than inequality between counties. Six states exhibit an increase in inequality *within Appalachian counties* from 1967 to 1992 (Alabama, Georgia, Mississippi, Ohio, Pennsylvania, and South Carolina). Four states show increases in inequality *within Appalachian counties* from 1967 to 1987, and then declines from 1987 to 1992 (Kentucky, Maryland, Tennessee, and Virginia). Three states exhibit fluctuations in inequality *within Appalachian counties*, but over the whole period registered higher levels of inequality in 1992 than in 1963 (New York, North Carolina, and West Virginia).

A comparison of the most unequal and most equal counties in the Appalachian and non-Appalachian counties in the region seems to indicate that Appalachian counties are neither substantially more nor substantially less unequal than non-Appalachian counties.

A comparison of relative contributions to overall pay inequality in manufacturing shows, not surprising, that the contribution of differences between Appalachian and non-Appalachian counties is extremely small. Perhaps more interesting is the finding that there are indications of a widening gap between Appalachian and non-Appalachian counties beginning in 1982. Overall, however, the regional trend toward widening inequality of pay within counties mirrors the national pattern.

Finally, the study reflects on the role of macroeconomic factors on the evolution of inequality across geographic regions. The study presents evidence that the national economic situation, particularly trends in national unemployment, conditions local changes in wages and wage inequality. Thus, during prosperous times such as 1963-67, the preponderance of counties experienced rising wages and falling inequality. In tougher times, such as 1977-82, falling wages and rising inequality predominate. In essence, periods of exceptional prosperity are necessary to reduce inequality in American manufacturing pay. Conversely, steep recessions tend to worsen pay inequality.

1. Introduction

This report presents and describes a new measure of inequality in manufacturing pay, suitable for assessing geographic patterns in the change of inequality over time in the United States. The measure is a between-groups component of Theil's T statistic, calculated between manufacturing plants within counties of the United States. Data are drawn from the non-public tabulations of the *Longitudinal Research Database* of the Bureau of the Census[1]. We consider only data for the census years of 1963, 1967, 1972, 1977, 1982, 1987 and 1992. The data set thus provides a time-series measure of inequality for each covered county, with the potential for extension as the Economic Census for 1997 becomes available. Due to limitations on disclosure, data are available for 1,486 counties out of 3,150. In addition, the data set includes an aggregated index of inequality for those counties within each state that could not be individually disclosed.

Until recently, measures of inequality in the United States have been almost exclusively available at the national or state level only. This is due to a wide reliance on sample surveys and the need for very large samples if inequality is to be measured separately over a large number of distinct spatial units. There has also been a tendency in the economic literature on the evolution of inequality to suppose that the forces affecting inequality—whether technological change, trade, or such macroeconomic factors as the rate of unemployment—apply with equal force to all regions. In effect, this amounts to the implicit assumption that a single uniform national labor market exists.

Such an assumption may or may not be correct. A national measure of inequality can mask different patterns of inequality itself, and of changes in inequality over time in different regions of the country. In a study of inequality across American states, Bernard and Jensen (1998) argue that patterns of inequality do differ. In particular, these authors find that education premia differ across states, and that these differences have persisted over time. Topel (1994) also disaggregates the dynamics of US inequality, again finding that differences exist across Census regions.

Our results indicate that national conditions seem to determine the local dynamics of inequality, which suggests that, at least at the level of analysis at which our data allows us to draw inferences, the assumption of a national labor market is valid. Still, this report takes the measurement of inequality down to a much smaller fundamental unit of geographic analysis than States or Census regions: the county. The sacrifice involved, as compared with prior analysis, lies in the fact that one cannot obtain measures of individual or household income at the county level; the unit of observation in the Economic Census is the manufacturing plant or other business facility. Thus, the measure of dispersion within each county is a measure of inequality across the average wages of manufacturing plants[2].

At this level, we find that national conditions seem to be the most important determinants of changes in inequality. However, we still found significant differences in

---

[1] The Longitudinal Research Database is described in Appendix 3.

[2] This average does not consider any control for industry, urban/rural differences or any other possibly relevant characteristics of firms. The average wage lumps all establishments of each county together and includes both production and non-production workers.

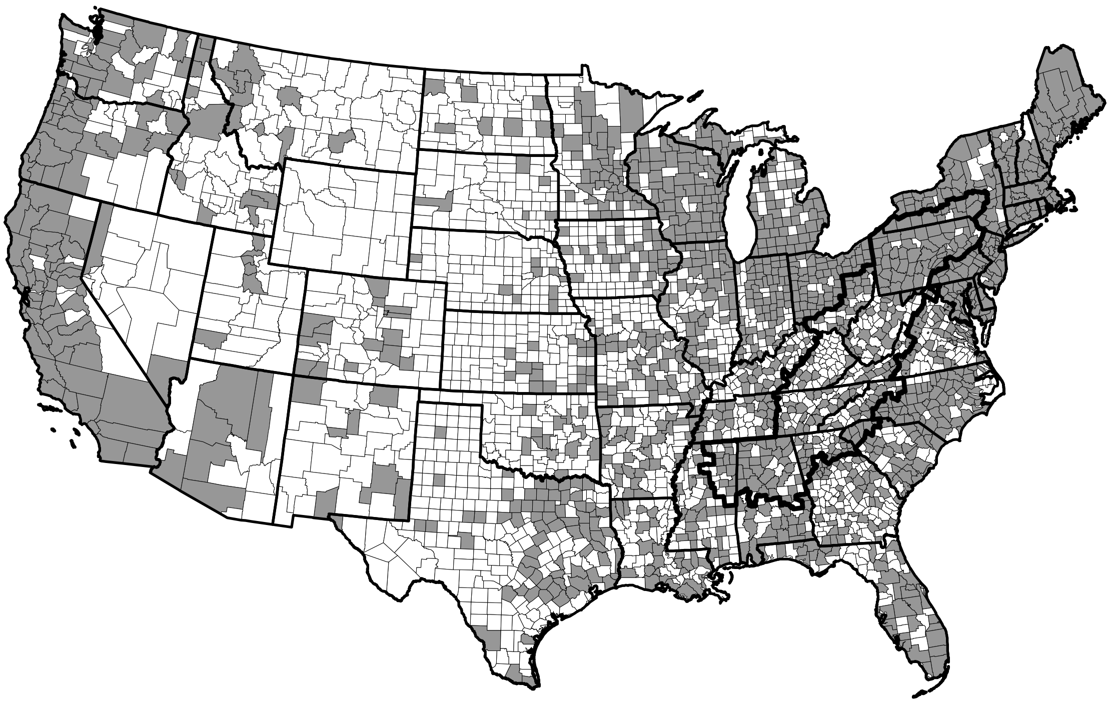
the dynamics of average wages in ARC and non-ARC counties from the 1980s onwards. These suggest that in this period wage growth in Appalachia was slower than elsewhere, on average.

Our approach necessarily excludes many forms of income that are relevant for a larger analysis of social welfare. It also excludes all those sources of pay that lie outside of manufacturing facilities. Indeed, it is fair to state that a measure of inequality across manufacturing facilities provides at most a narrowly focused and limited frame of reference from which to address the question of economic inequality broadly speaking. It cannot substitute—and we do not suggest that it can—for research into household income and well-being using other sources of data.

Nevertheless, the approach has two basic advantages. The first lies in achieving accurate measurement of this aspect of inequality in very fine geographic detail. And the second lies in the mathematical properties of the Theil index itself, which permit us to group counties together and so to achieve harmonized measures of inequality at the state and regional levels. Not least, this permits us to assess whether the experience of the Appalachian region, taken as a whole, differs from that of the rest of the country. We also believe that trends in the dispersion of pay across manufacturing plants are likely to indicate the direction of trends in broader measures of inequality that we do not observe. If pay in the garment shops is declining relative to pay in the mines, for instance, we think it highly likely – though we cannot prove the point from the evidence before us -- that other forms of inequality are increasing as well.

Figure 1 presents a map showing the coverage of our data set—those counties for which an inter-plant measure of inequality could be disclosed. The number of plants in each county ranges from 19,964 in New York County in 1967 to 10 establishments in 1987 and 1992 in Hancock County (Georgia). In 1992, the county with the most establishments (18,439) was Los Angeles County.

Figure 1. Coverage of Between-Plant Pay Dispersion measures, 1963-1992, U.S. Counties. Shaded counties were used in the data set. Appalachian Region counties are shown within the thick border.



In what follows, we develop the information in this rich new source of inequality measures in hierarchic fashion, beginning at the national and state levels, and working our way toward an analysis of inequality at the county level.

2. Manufacturing Pay Dispersion at the National Level

This section describes the dynamics of dispersion in our measure of inter-plant pay dispersion at the national level. We first present a picture of the evolution of this dimension of economic inequality in the country as a whole, including a decomposition into the contributions of the between state and within state components of national inequality.

Figure 2 shows the Theil measure of pay inequality across manufacturing plants at the national level. There is a reduction from 1963 to 1967, following which inequality steadily increased from 1967 to 1992, at an annual average rate of 2%, in a pattern that resembles a "check mark."

Figure 2. Manufacturing pay dispersion aggregated to the national level: relative importance of inequality within and between states.

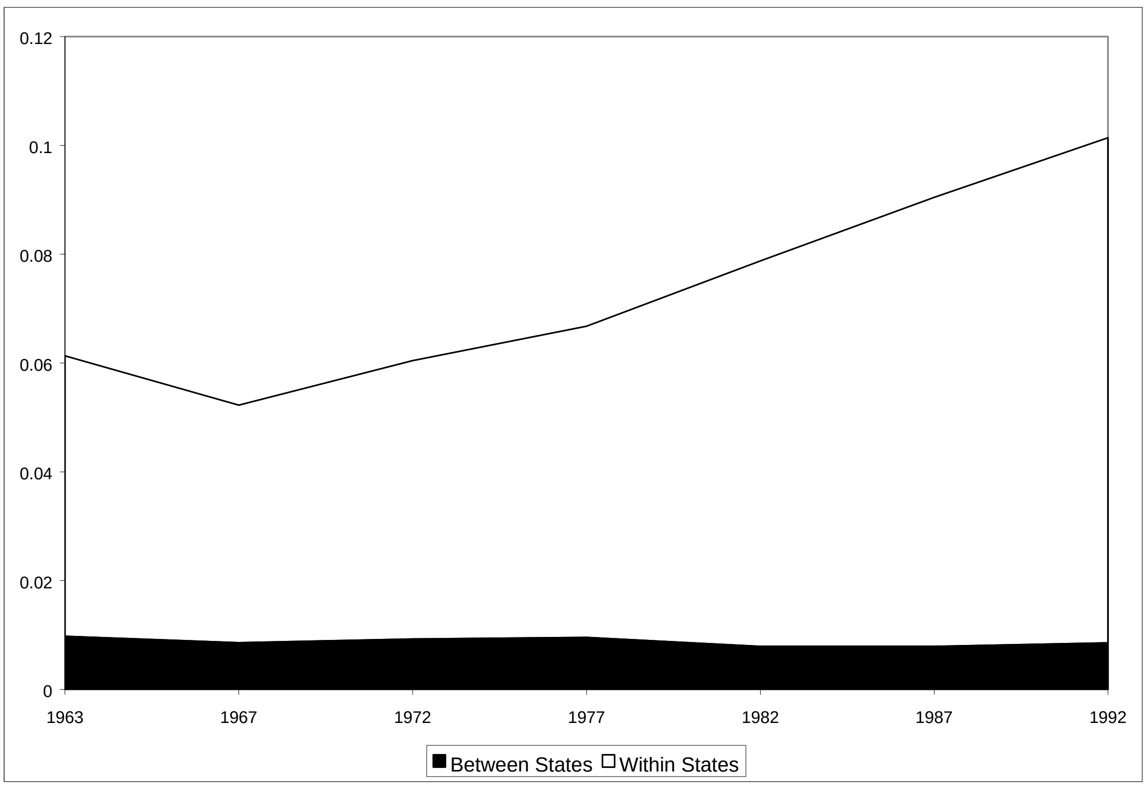


Figure 2 also shows the contribution to US inequality of the dispersion in wages due to differences *across* states (area shaded in black) and the contribution that is accounted for by differences *within* states (white area). The contribution of cross-state inequality has remained stable throughout the period. Since overall inequality almost doubled from 1963 to 1992, the contribution of the between states component has decreased almost by half. This leads us to an opening inference. Clearly, the dynamics of rising national pay inequality have been driven by the within-states contribution. They have not been driven by an increasing average gap between rich states and poor.

The within states component of inequality can be further decomposed into two sub-components: the dispersion of wages across counties and the dispersion within counties (Figure 3). Again, changes within the smaller geographic units predominate. It is a changing within-counties component that is driving the increase in manufacturing pay dispersion.

Figure 3. Contributions of Inequality Across Counties and Within Counties.

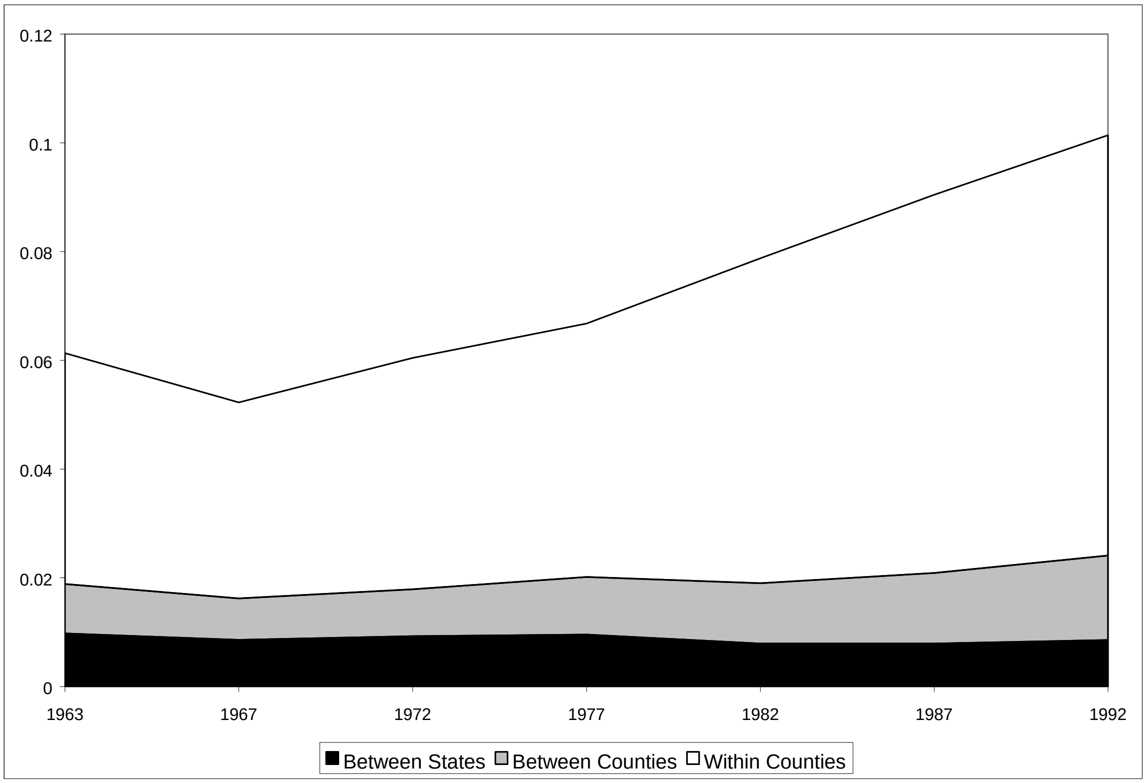


Table 1 shows the relative contribution of each of the components depicted in Figure 3 to US inequality. As the charts indicate, the largest contribution to US inequality comes from the within counties component. The share of the within counties component in total inequality increased from 69% in 1963 to 76% in 1992. Note that this increase in the share of the within counties component was made “at the expense” of the between states contribution, which has seen its share decrease by half, from 16% in 1963 to 8% in 1992. The share of the between counties component remained relatively stable throughout the period at around 15%.

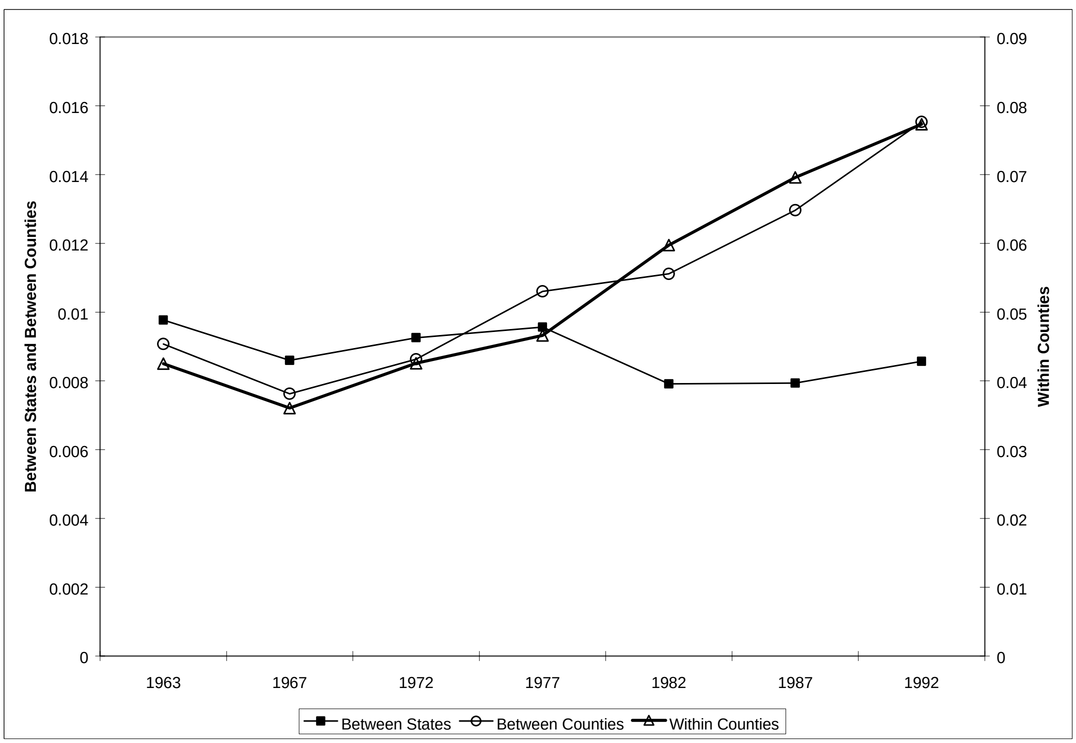
Table 1. Contributions of the Between States, Between Counties and Within Counties Components

	1963	1967	1972	1977	1982	1987	1992
Between States	16%	16%	15%	14%	10%	9%	8%
Between Counties	15%	15%	14%	16%	14%	14%	15%
Within Counties	69%	69%	70%	70%	76%	77%	76%

These results indicate that differences in average wages across states in the US have become less important, with a steady decrease of the between states contribution to US inequality since 1967. The largest drop in this contribution occurred between 1977 and 1982. Note, however, that this decrease in the *share* of the between state inequality does not mean that there has been convergence in average wages across states. As we saw above, the level of inequality between states has remained constant. What the results suggest, instead, is that the *increase* in inequality in the US since 1967 is accounted for mainly by a growing dispersion of wages within and across counties, within states.

Yet another way to compare these three components of inequality is to contrast their dynamics. Figure 4 plots a line for the Theil index of each of the three components. Given the differences in magnitude, two scales are used, with the left-hand scale corresponding to the between states and to the between counties Theil, which are about one-third smaller than the within counties Theil (which corresponds to the scale on the right-hand side).

Figure 4. The Dynamics of Pay Inequalities: Between States, Between Counties and Within Counties



The dynamics of the between states component have remained essentially stable. There are two intervals of slight decreases (1963 to 1967 and 1977 to 1987) with the between states Theil increasing modestly in the remaining time periods. The dynamics of the between-states Theil seems to be independent of the evolution of the two other components.

On the other hand, the dynamics of the between counties component and of the within counties component are essentially the same. For both components, there is a decrease in the Theil from 1963 to 1967, with the Theil increasing steadily from 1963 to 1977. It is clear that a positive change in one component is associated with a positive change in the other component, and vice-versa.

This section has therefore shown that after a decrease in wage inequality in the US from 1963 to 1967, wage dispersion in US manufacturing steadily increased through the early 1990s. The importance of differences across states has diminished; virtually all of the increase in US inequality is due to growing dispersion of wages within states.

3. Manufacturing Pay Inequalities in the Appalachian Region

About four hundred counties in thirteen US states define the Appalachian region. This section describes the dynamics of manufacturing wage dispersion in each of the thirteen states with counties that are members of the Appalachian Regional Commission (ARC). The share of population in ARC counties varies from state to state. Taking 1990 as a reference, Table 2 shows that the population in the ARC goes from 100% in West Virginia (where all counties are in the ARC) to 5% in Maryland.

Table 2. Share of State Population in ARC Counties

	1990
West Virginia	100%
Alabama	63%
Pennsylvania	49%
Tennessee	44%
Kentucky	28%
South Carolina	25%
Georgia	24%
Mississippi	20%
North Carolina	20%
Ohio	13%
Virginia	10%
New York	6%
Maryland	5%

Source: ARC and Bureau of the Census.

The shares of manufacturing employment in the ARC (Table 3) are similar, but not exactly the same, as the shares of population. For most states, shares of manufacturing employment in the ARC are higher than the corresponding population shares. The exceptions occur for Pennsylvania, Kentucky and Ohio.

Table 3. Share of State Manufacturing Employment in Plants Located in ARC Counties

	1963	1967	1972	1977	1982	1987	1992
West Virginia	100%	100%	100%	100%	100%	100%	100%
Alabama	70%	71%	70%	68%	66%	67%	67%
Tennessee	52%	52%	52%	51%	51%	49%	49%
Pennsylvania	46%	46%	47%	47%	45%	42%	44%
South Carolina	40%	39%	38%	37%	37%	37%	38%
Mississippi	30%	29%	31%	31%	33%	35%	34%
Georgia	25%	26%	28%	29%	28%	30%	32%
North Carolina	25%	24%	23%	23%	24%	23%	22%
Kentucky	13%	13%	15%	17%	18%	18%	18%
Virginia	14%	16%	15%	15%	14%	14%	14%
New York	7%	8%	7%	8%	8%	8%	9%
Maryland	9%	10%	10%	9%	9%	8%	8%
Ohio	8%	7%	7%	8%	8%	8%	8%

The shares of manufacturing wages in the ARC (Table 4) are nearly identical to the employment shares[3]. Both the employment and the wage shares provide a way to see the importance of the ARC region in each state. However, the wage shares are particularly relevant in our analysis because wage shares are used to weight the Theil index components for each county when computing cross-county measures of inequality.

Table 4. Share of State Manufacturing Wages in Plants Located in ARC Counties

	1963	1967	1972	1977	1982	1987	1992
West Virginia	100%	100%	100%	100%	100%	100%	100%
Alabama	75%	74%	73%	70%	67%	68%	68%
Tennessee	53%	53%	51%	51%	51%	48%	47%
Pennsylvania	46%	45%	46%	47%	45%	40%	41%
South Carolina	38%	38%	38%	37%	36%	36%	38%
Mississippi	26%	25%	28%	28%	29%	32%	33%
Georgia	22%	23%	26%	27%	25%	27%	29%
North Carolina	26%	25%	23%	24%	24%	23%	21%
Kentucky	10%	11%	13%	15%	15%	14%	14%
Virginia	13%	15%	13%	13%	12%	11%	11%
New York	7%	7%	7%	8%	9%	8%	8%
Ohio	7%	6%	6%	7%	8%	6%	7%
Maryland	9%	10%	10%	9%	9%	7%	7%

For each state with county-members of the ARC, we divide the evolution of inequality into five components. The first component measures inequality *across* ARC counties; the second component measures dispersion in wages *within* ARC counties. These two components are weighted by the ARC wage shares as they enter into the overall

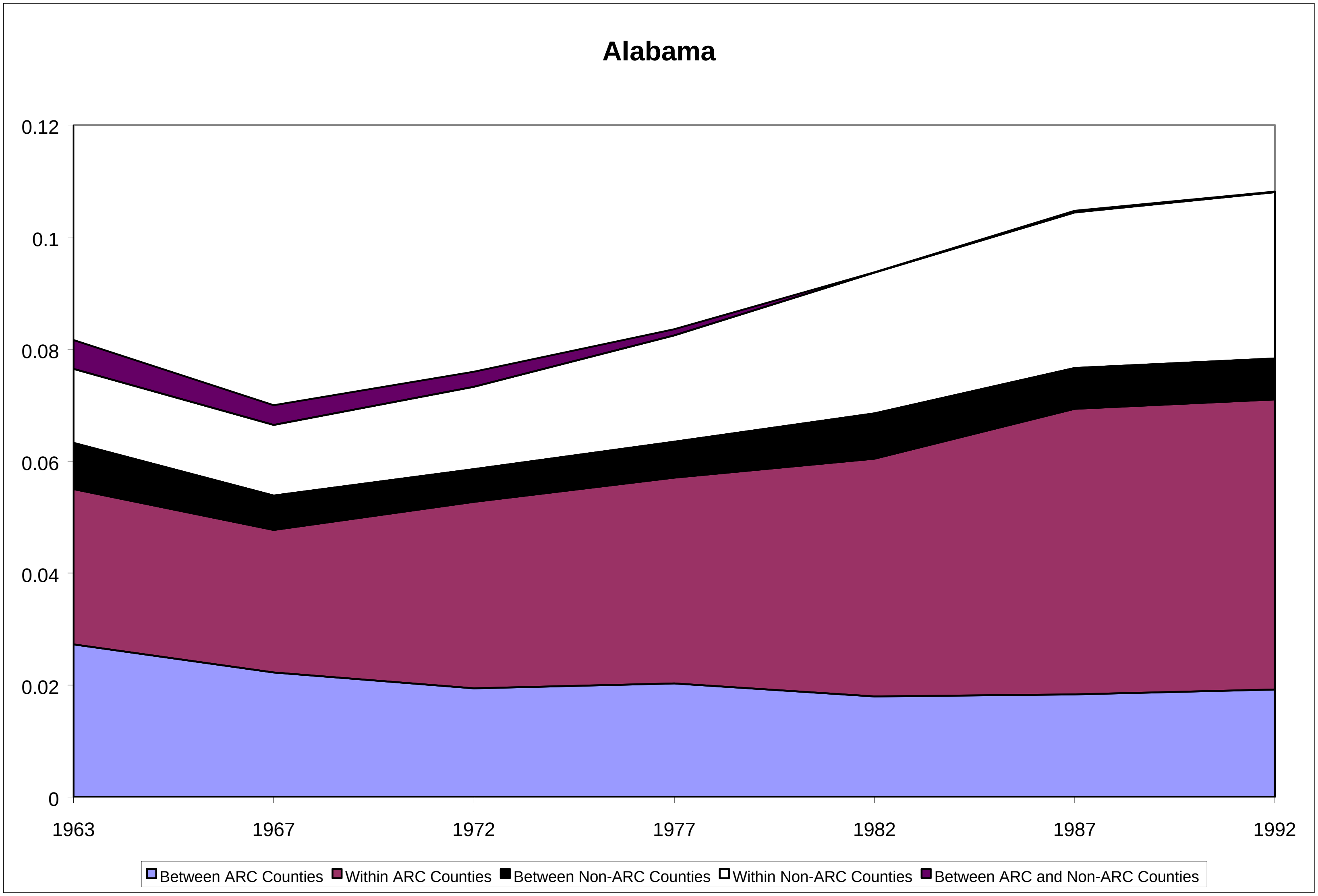
[3] Data from the County Business Pattern for 1996 suggests that the share of employment in plants located in ARC remained virtually the same as 1992.

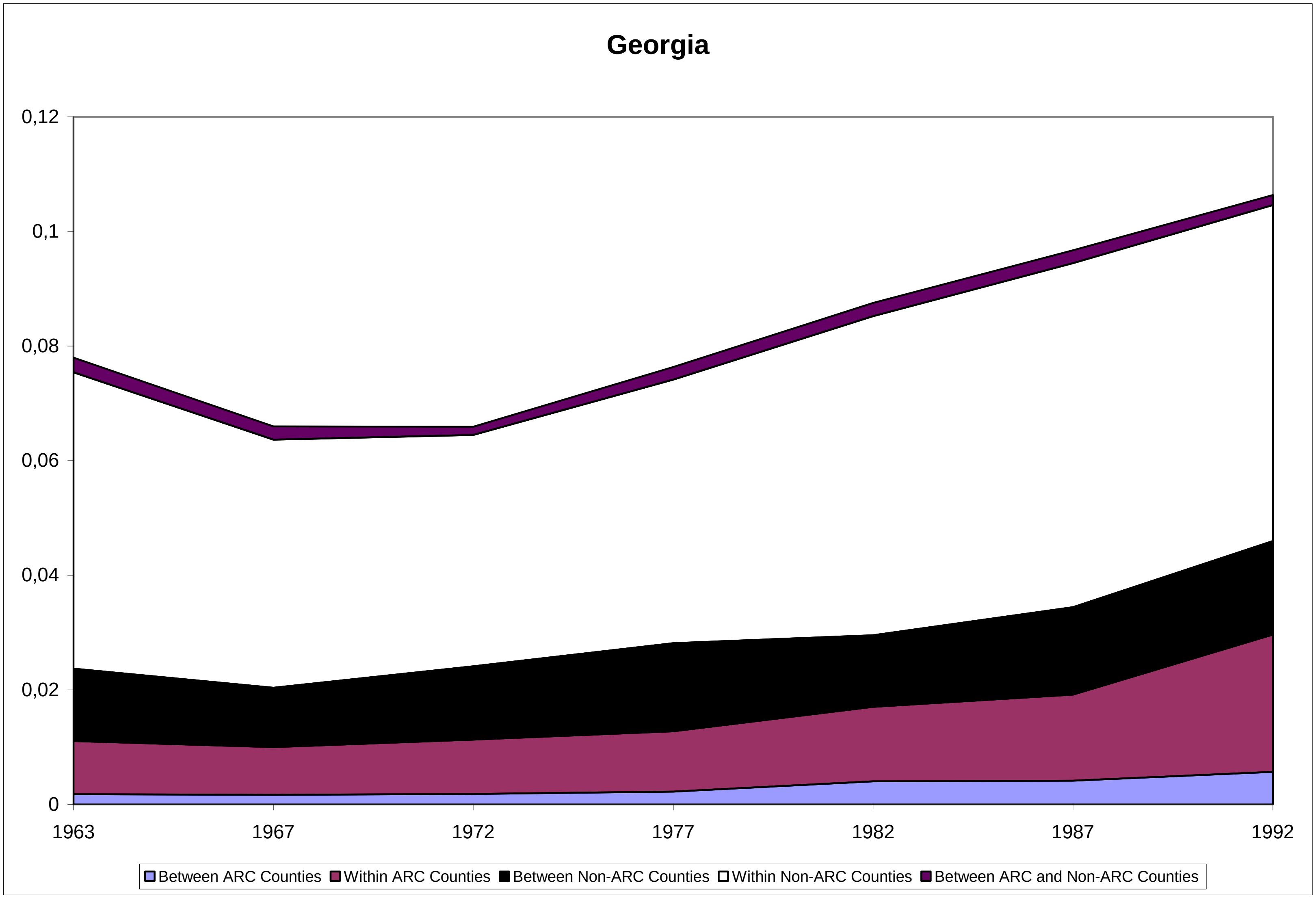
state inequality measure. Two other components account for inequality between and within non-ARC counties in the state. The fifth and final component measures the dispersion *between* the group of ARC and the group of non-ARC counties.

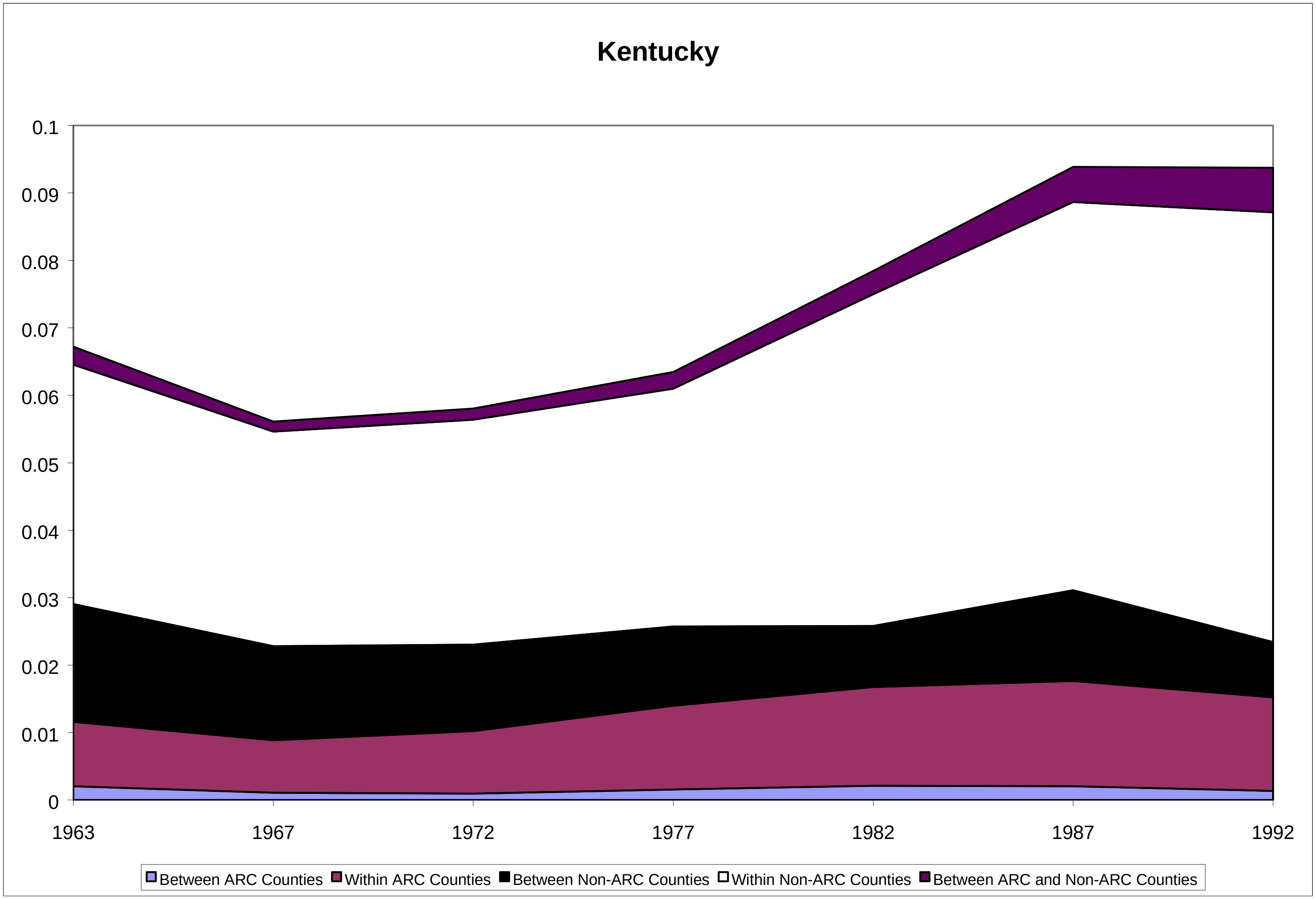
The following thirteen charts, one for each state, show the contributions of each component to the state level of wage dispersion. Most states exhibit the same pattern in overall inequality: a decrease in inequality from 1963 to 1967 and a steady increase from 1972 to 1992. In Kentucky, Mississippi, Tennessee and Virginia the trend of increasing inequality stops in 1987, with a decrease in inequality from 1987 to 1992.

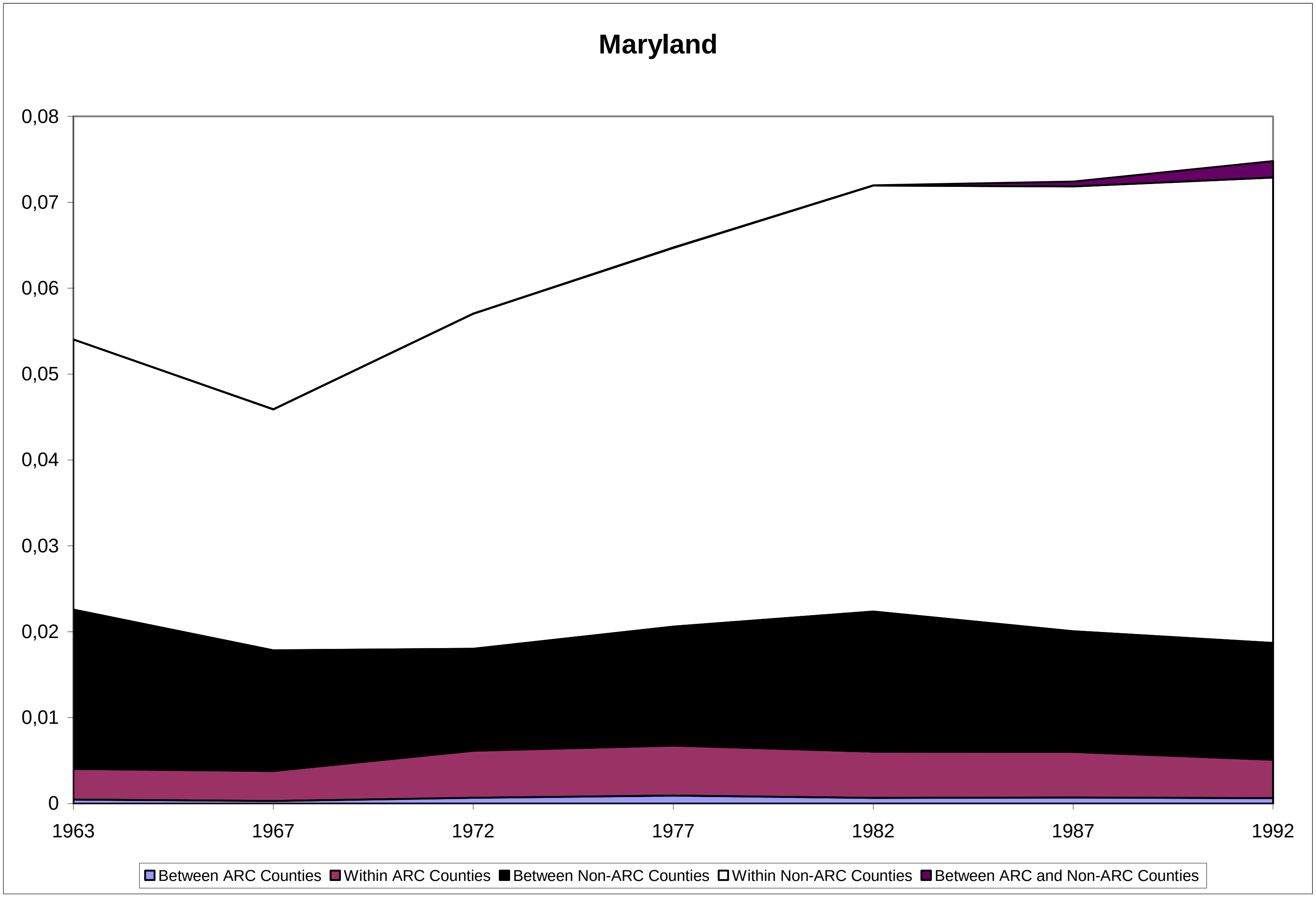
The component of inequality associated with differences between the ARC and the non-ARC counties is small for every state, which indicates that there is not a sharp difference in average wages in manufacturing for plants that are in ARC counties as compared with those that are in non-ARC counties. In Alabama and in Mississippi the between ARC and non-ARC counties component decreased over the period under analysis, but for most states this component has either remained stable or has witnessed a slight increase. The increase in this component has occurred primarily in the last period, from 1987 to 1992, especially in Kentucky, Maryland, Pennsylvania and Virginia.

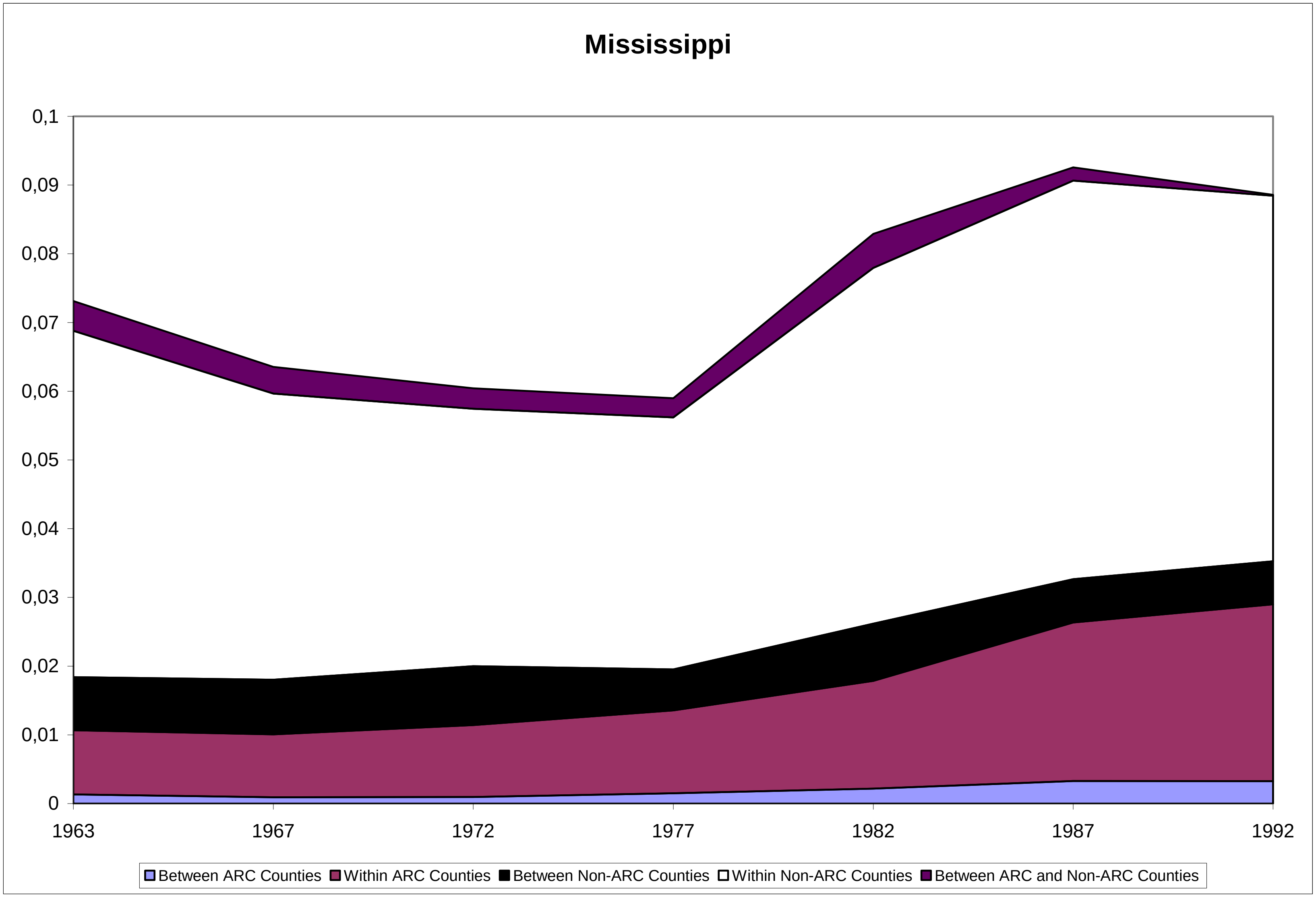
The remaining four components of the Theil index represented in the charts are associated with inequality between counties and inequality within counties for plants located in ARC and in non-ARC counties. The two components of inequality associated with non-ARC counties (between-counties and within-counties) dominate the level of state inequality in most cases. The obvious exceptions are those states for which employment and wage shares of ARC counties are large. These exceptions include West Virginia, Alabama, Tennessee and Pennsylvania.


Alabama
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

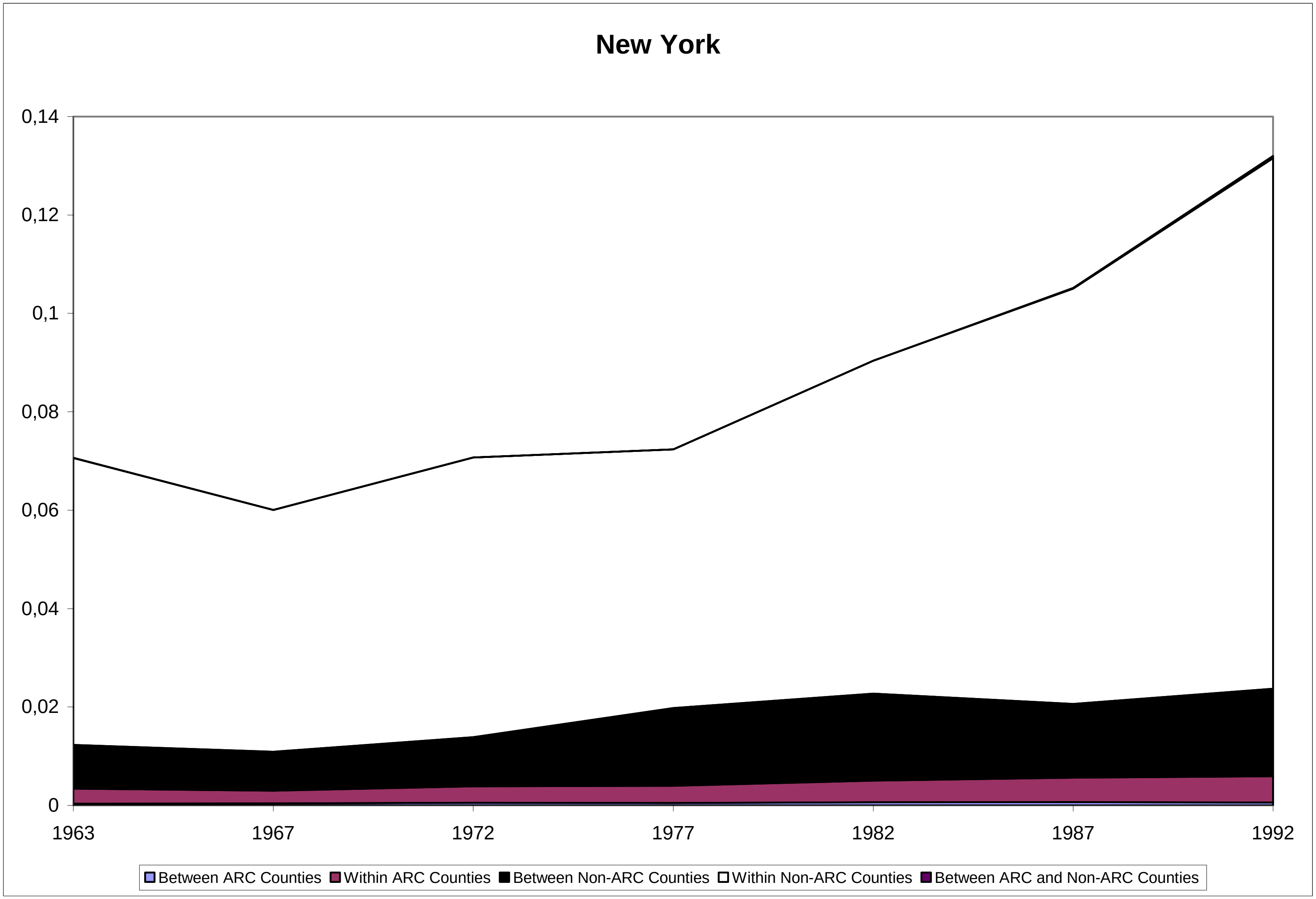

Georgia
0,12
0,1
0,08
0,06
0,04
0,02
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

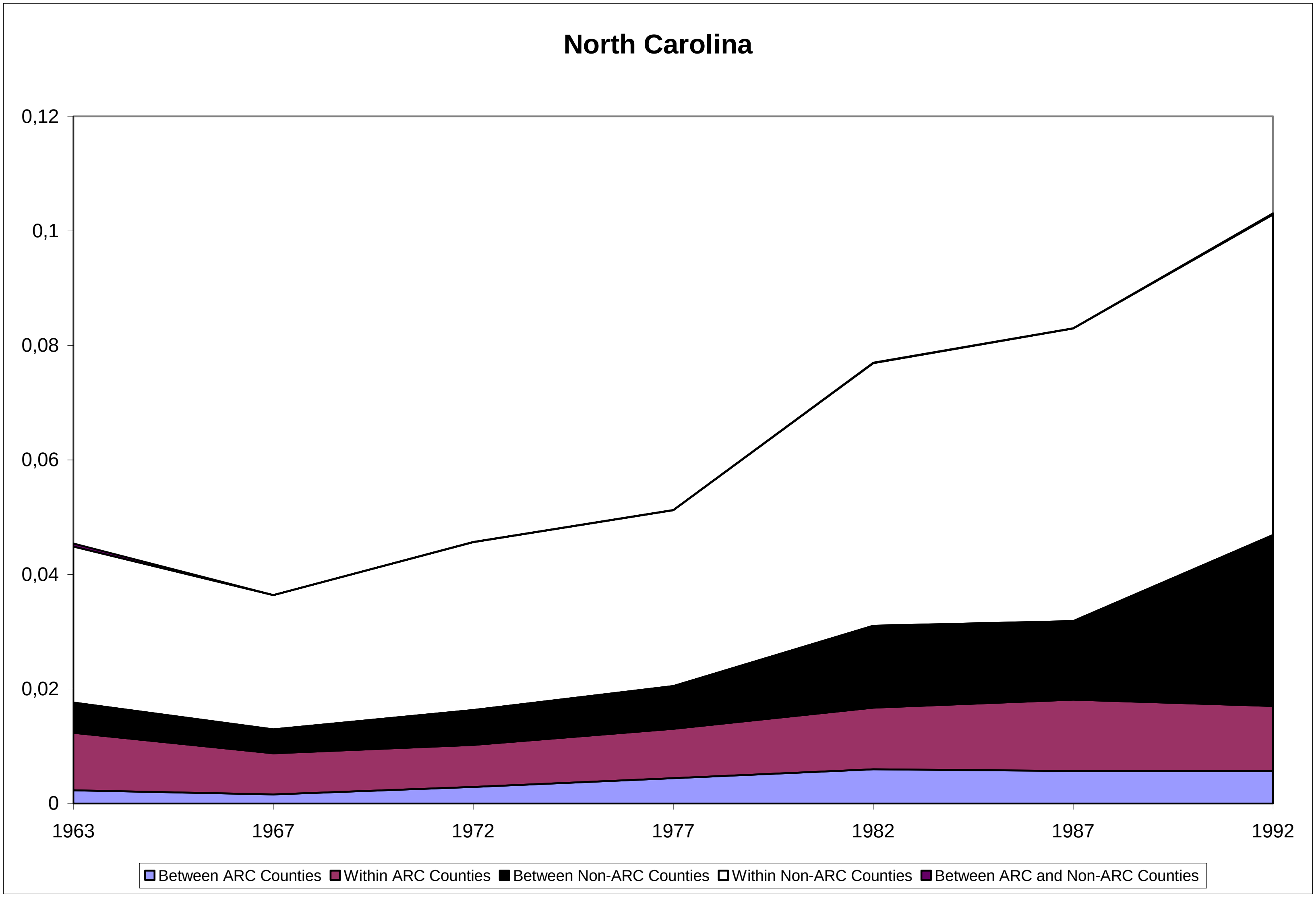

Kentucky
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

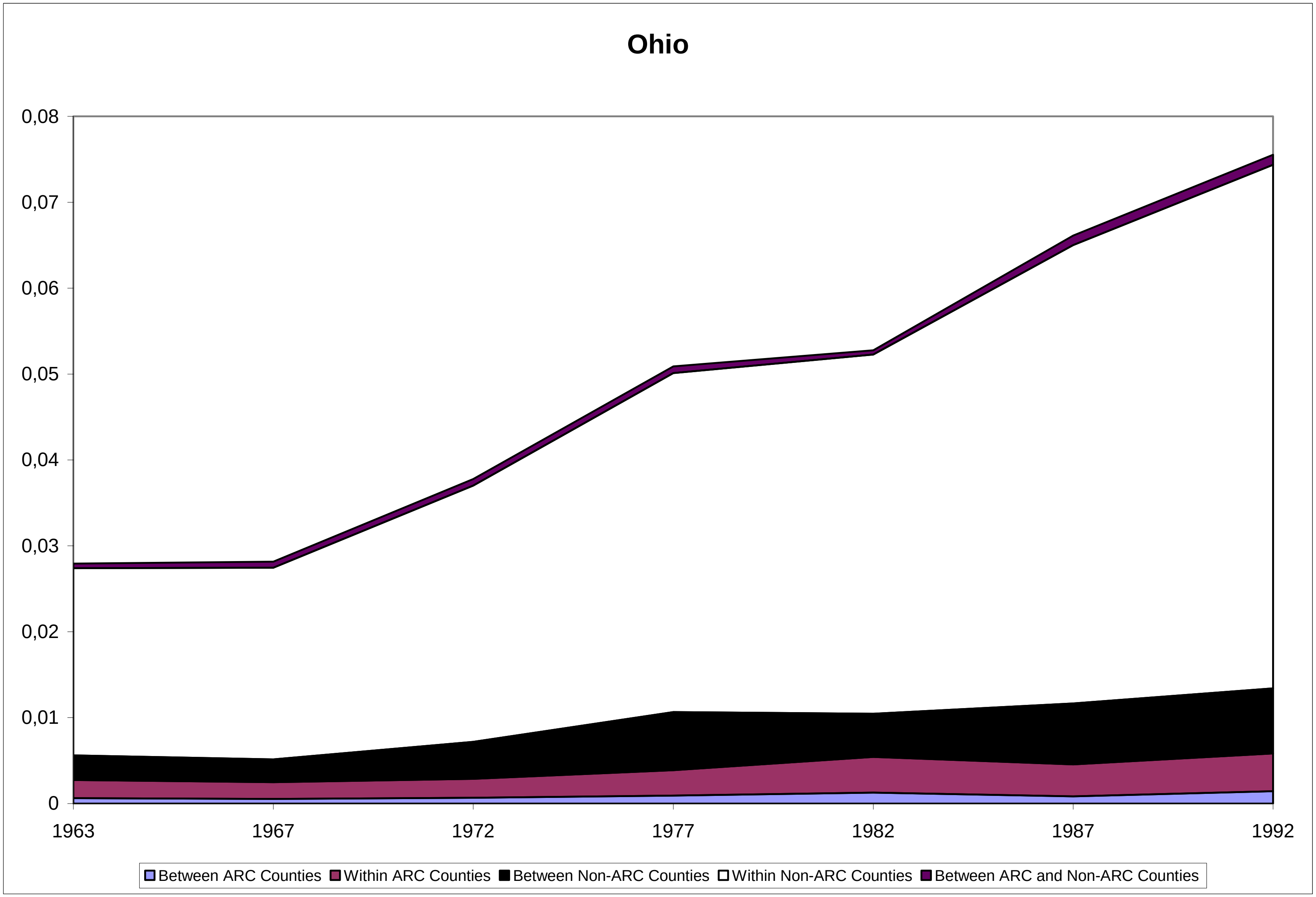

Maryland
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

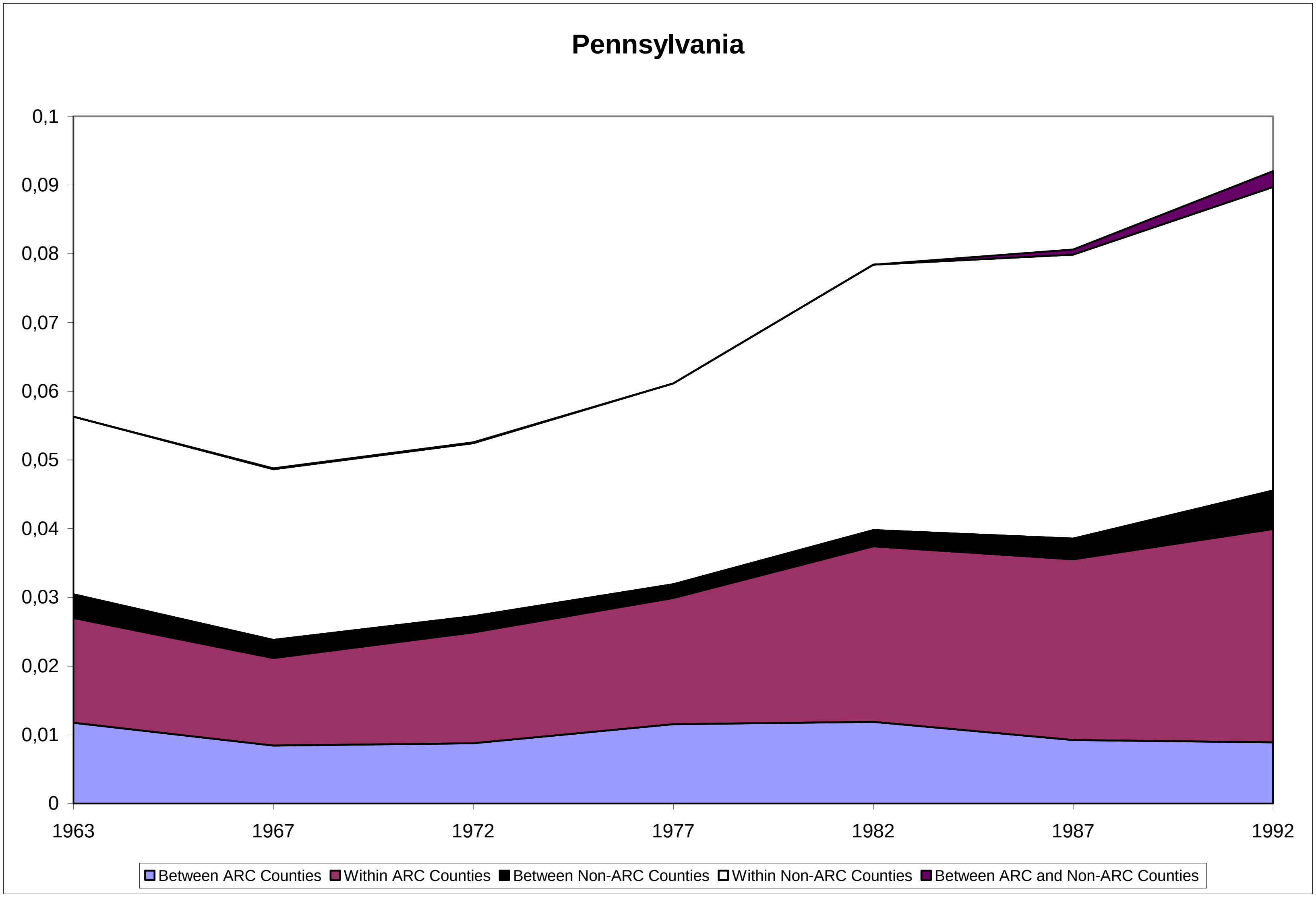


Mississippi
0,1
0,09
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

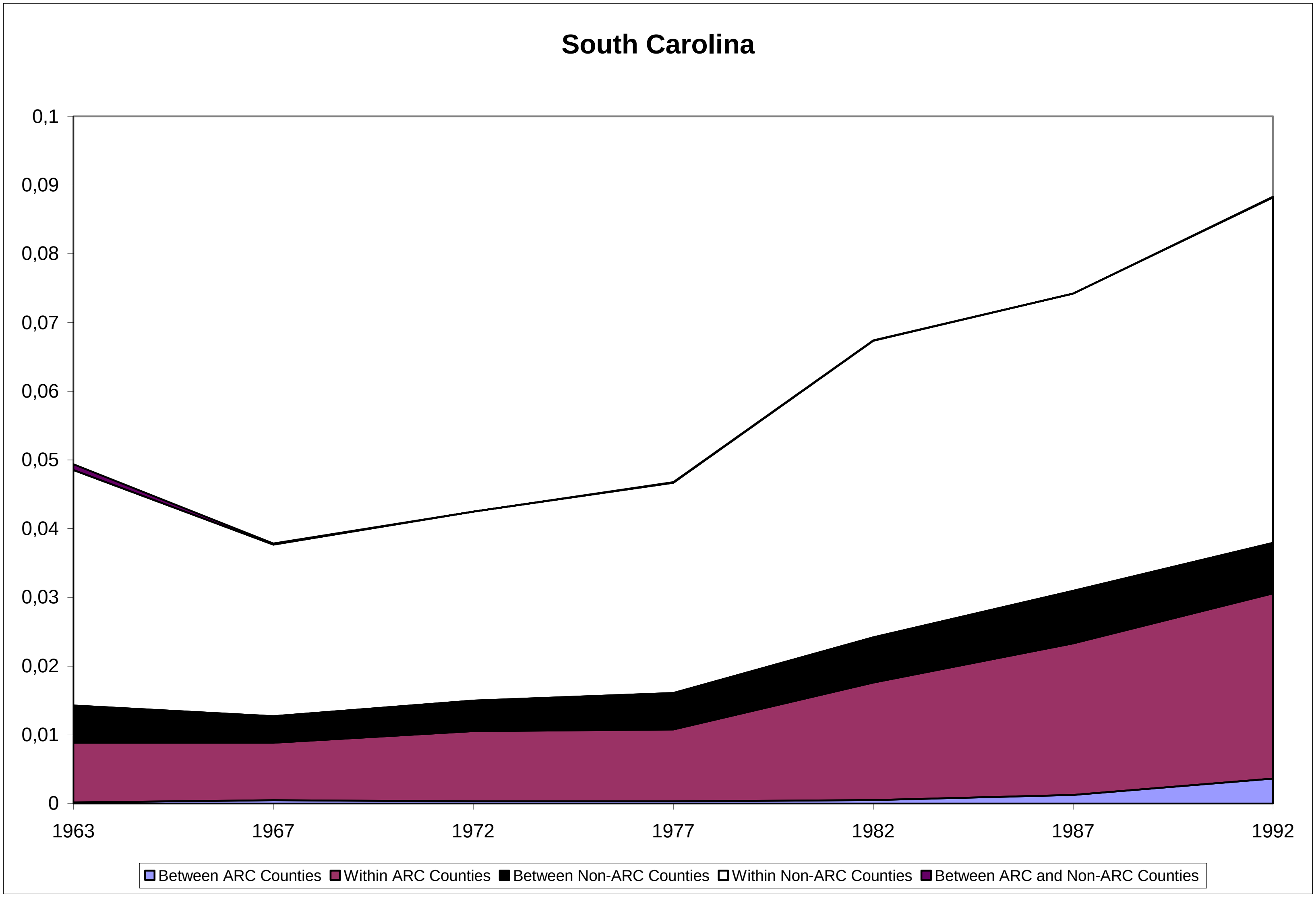

New York
0,14
0,12
0,1
0,08
0,06
0,04
0,02
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

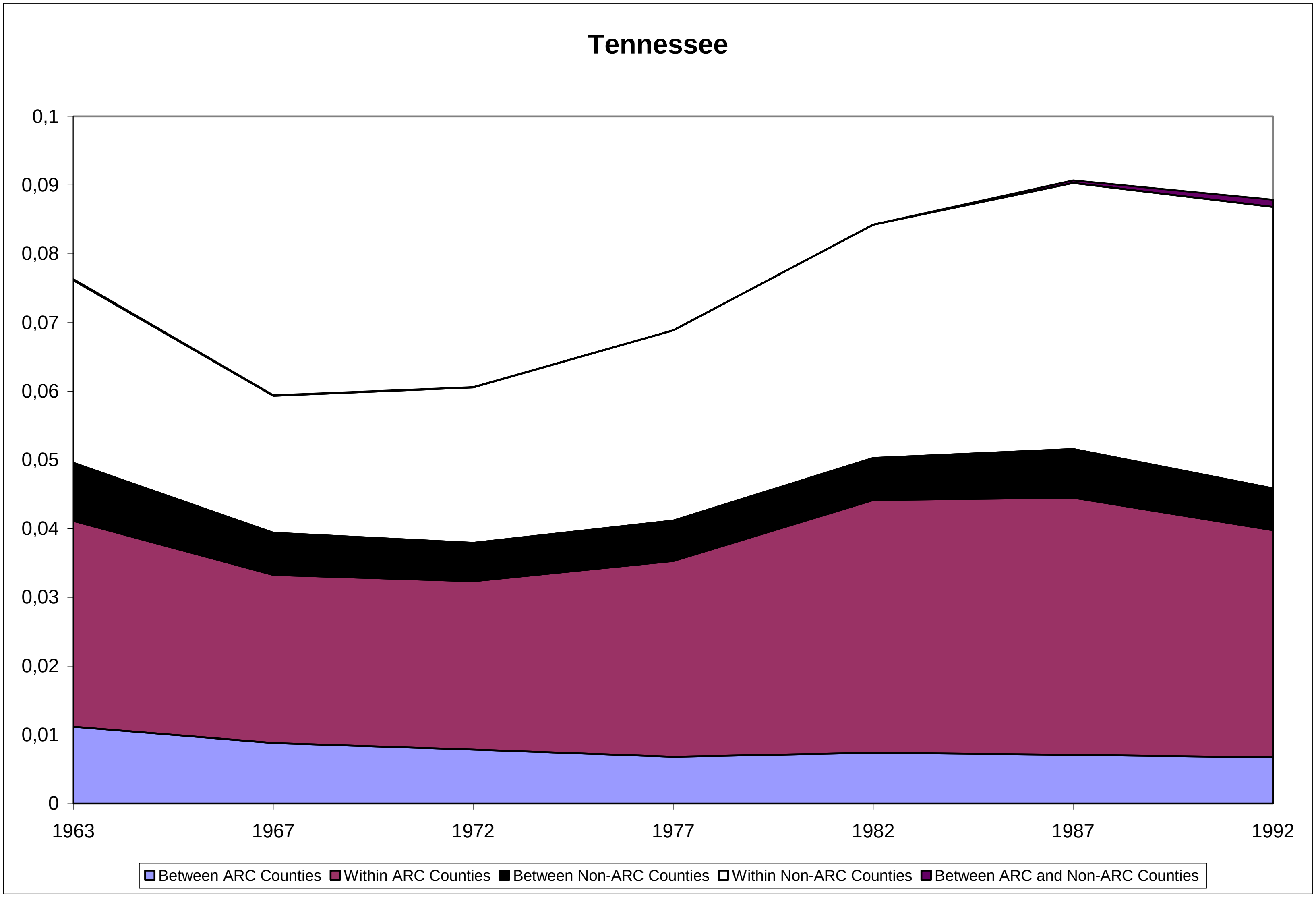

North Carolina
0,12
0,1
0,08
0,06
0,04
0,02
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

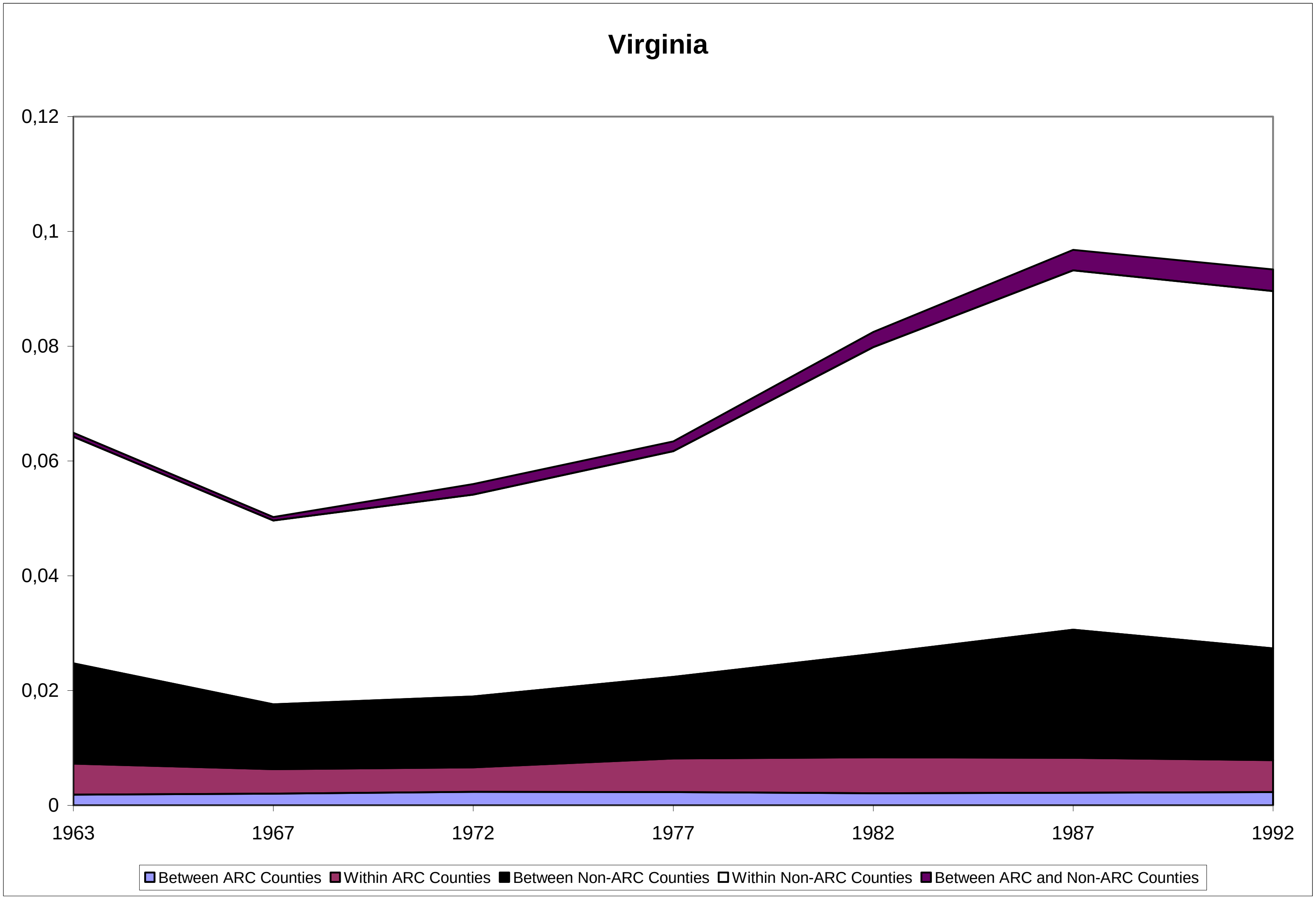

Ohio
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties



Pennsylvania
0,1
0,09
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties


South Carolina
0,1
0,09
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties


Tennessee
0,1
0,09
0,08
0,07
0,06
0,05
0,04
0,03
0,02
0,01
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties


Virginia
0,12
0,1
0,08
0,06
0,04
0,02
0
1963
1967
1972
1977
1982
1987
1992
Between ARC Counties
Within ARC Counties
Between Non-ARC Counties
Within Non-ARC Counties
Between ARC and Non-ARC Counties

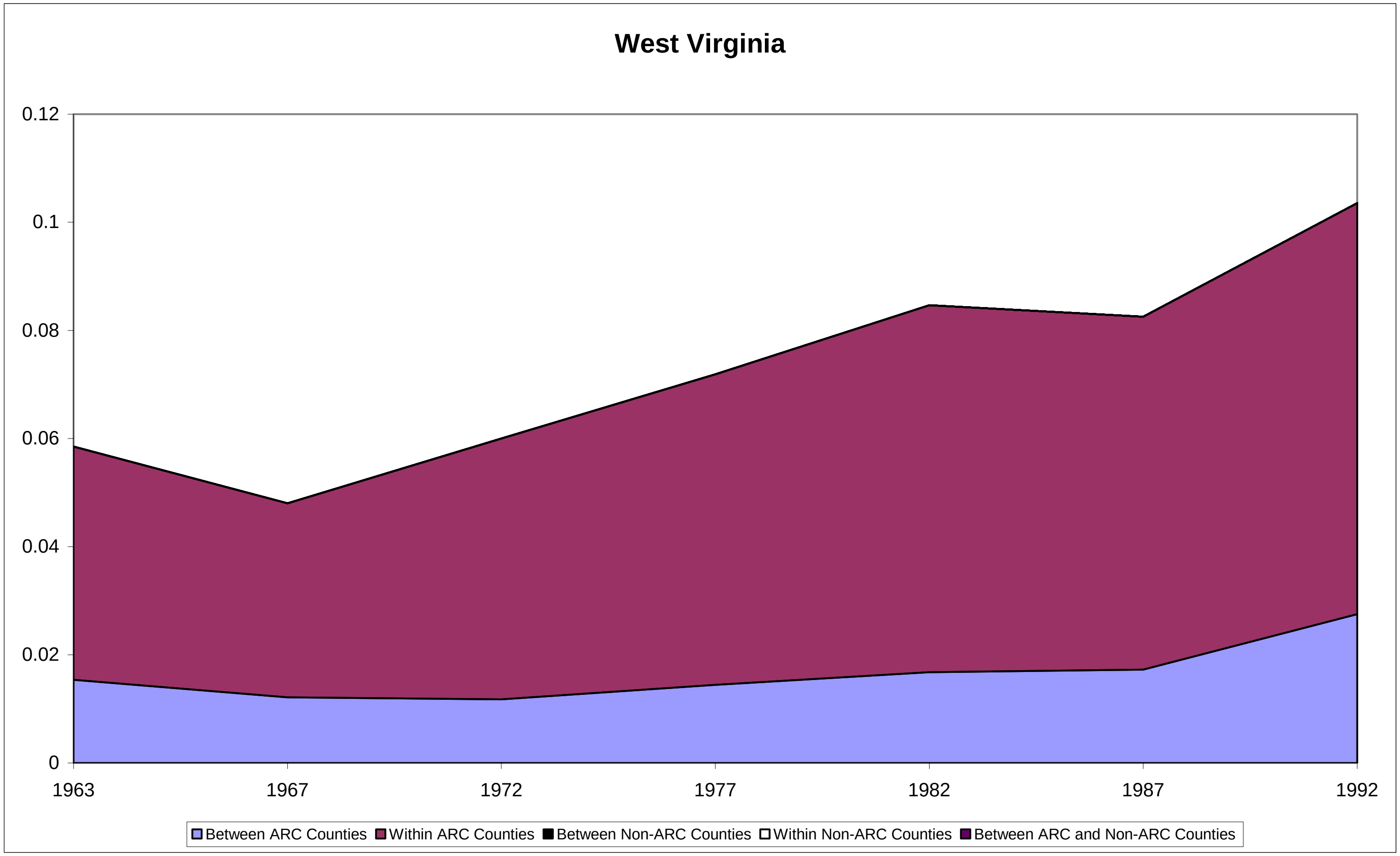


Figure 5 - Inequality in States with ARC Counties

The major determinant of the importance of the ARC-related components to state inequality is the wage share of plants in ARC counties. Figure 6 shows the relationship of the contribution of ARC counties' inequality to the statewide inequality, on the one hand, and the share of wages of plants in ARC counties, on the other. The chart shows the relationship for 1992, which indicates that the two variables are almost perfectly correlated. The relationship is the same for any other year.

The only state that is almost out of place is New York, where the share of wages in ARC counties is much larger than the contribution of ARC related counties to the state's inequality. This indicates that inequality in New York is largely driven by what happens outside the counties included in the ARC region – not a surprising finding, to be sure.

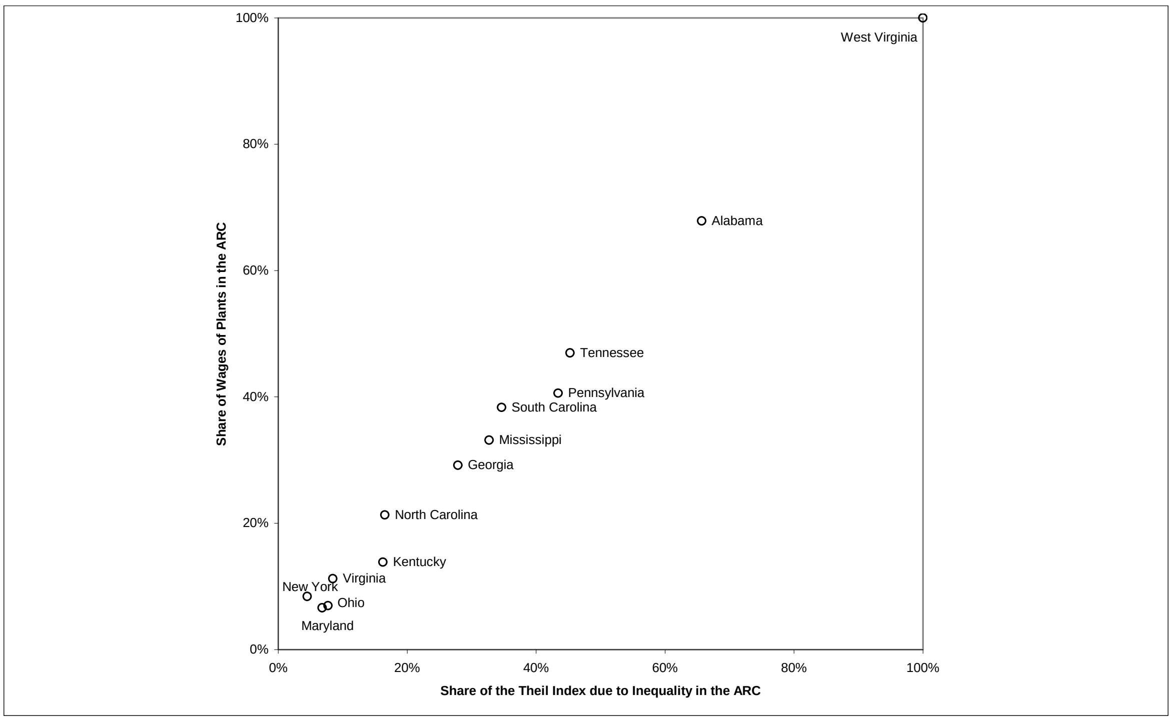
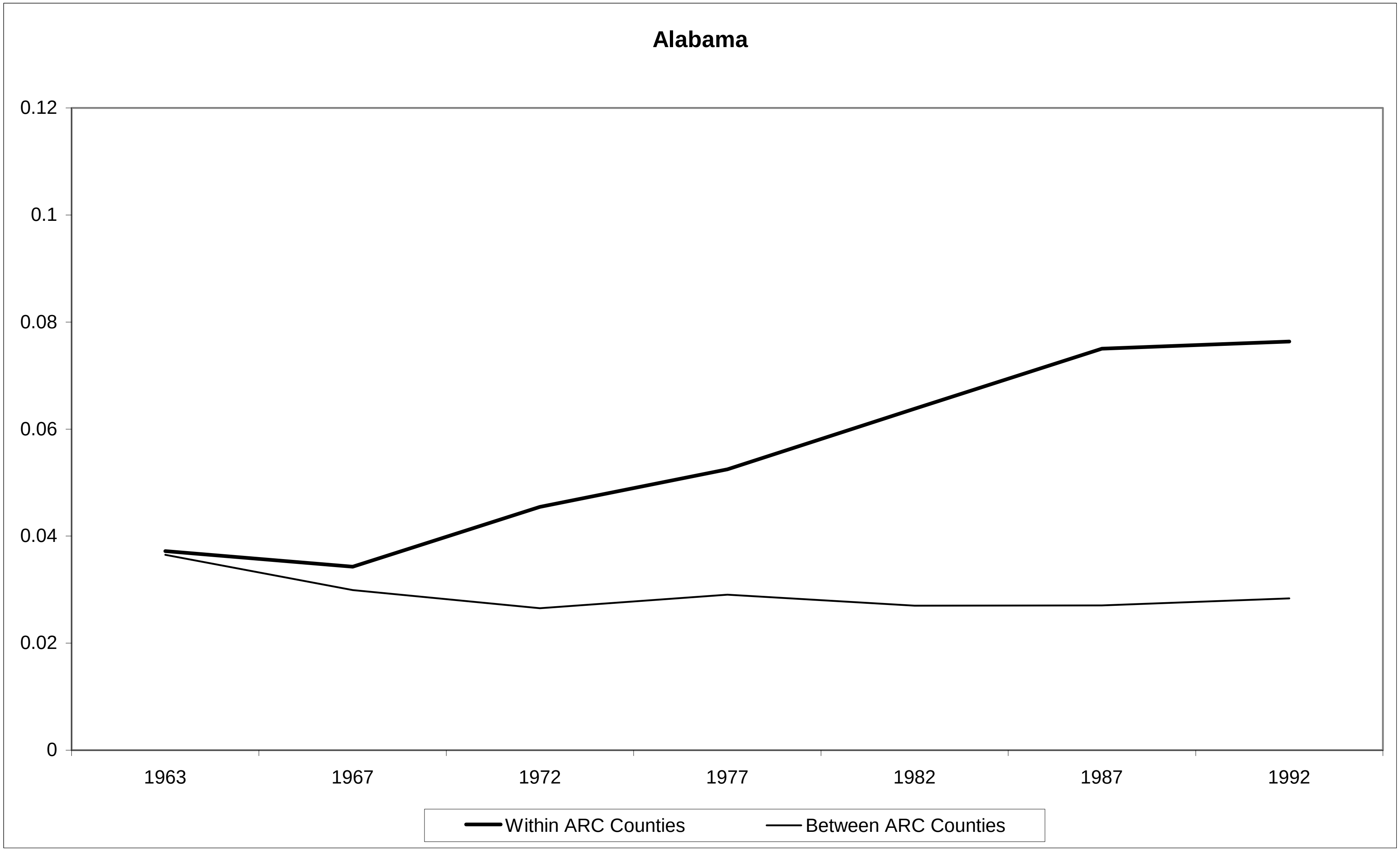
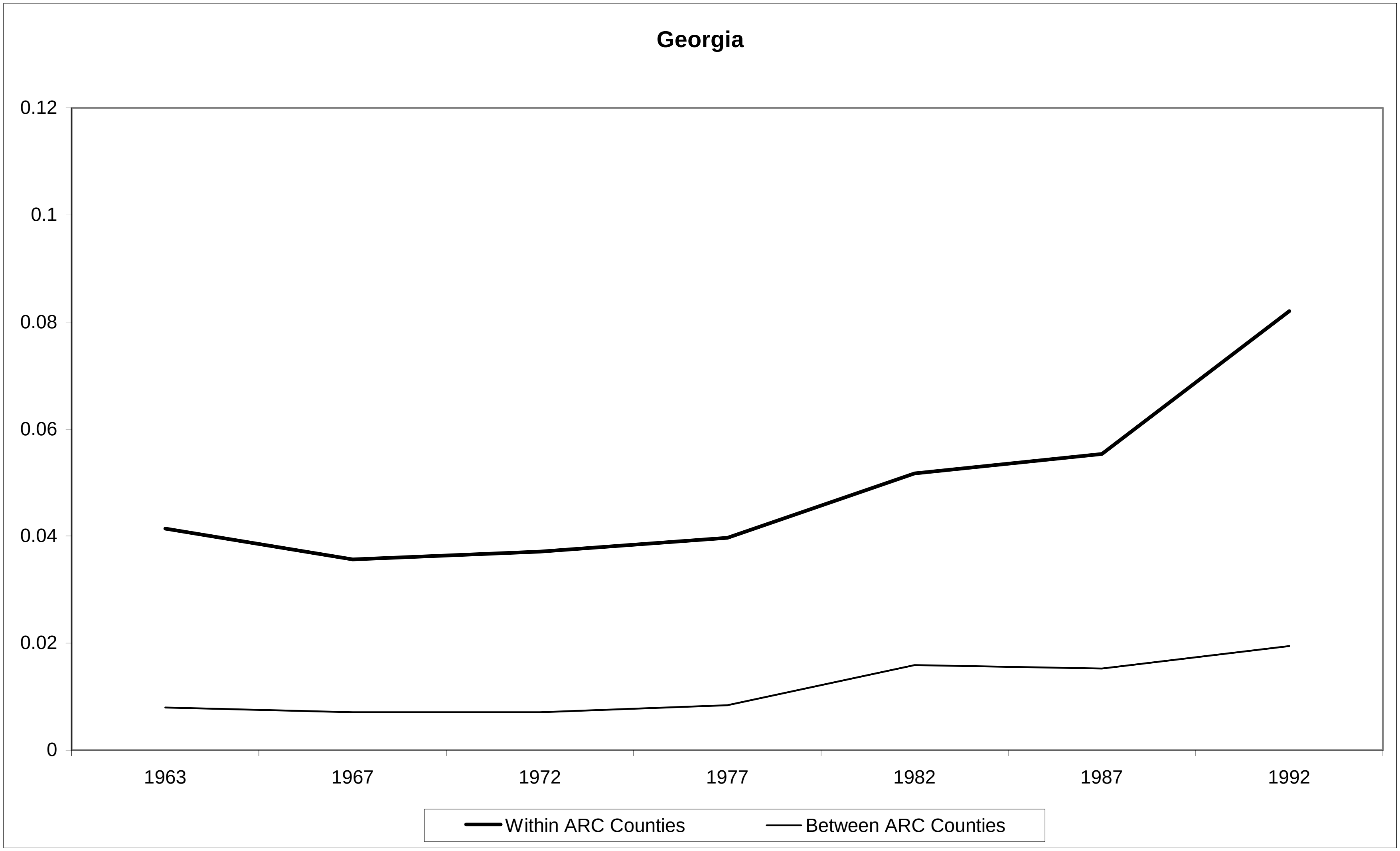
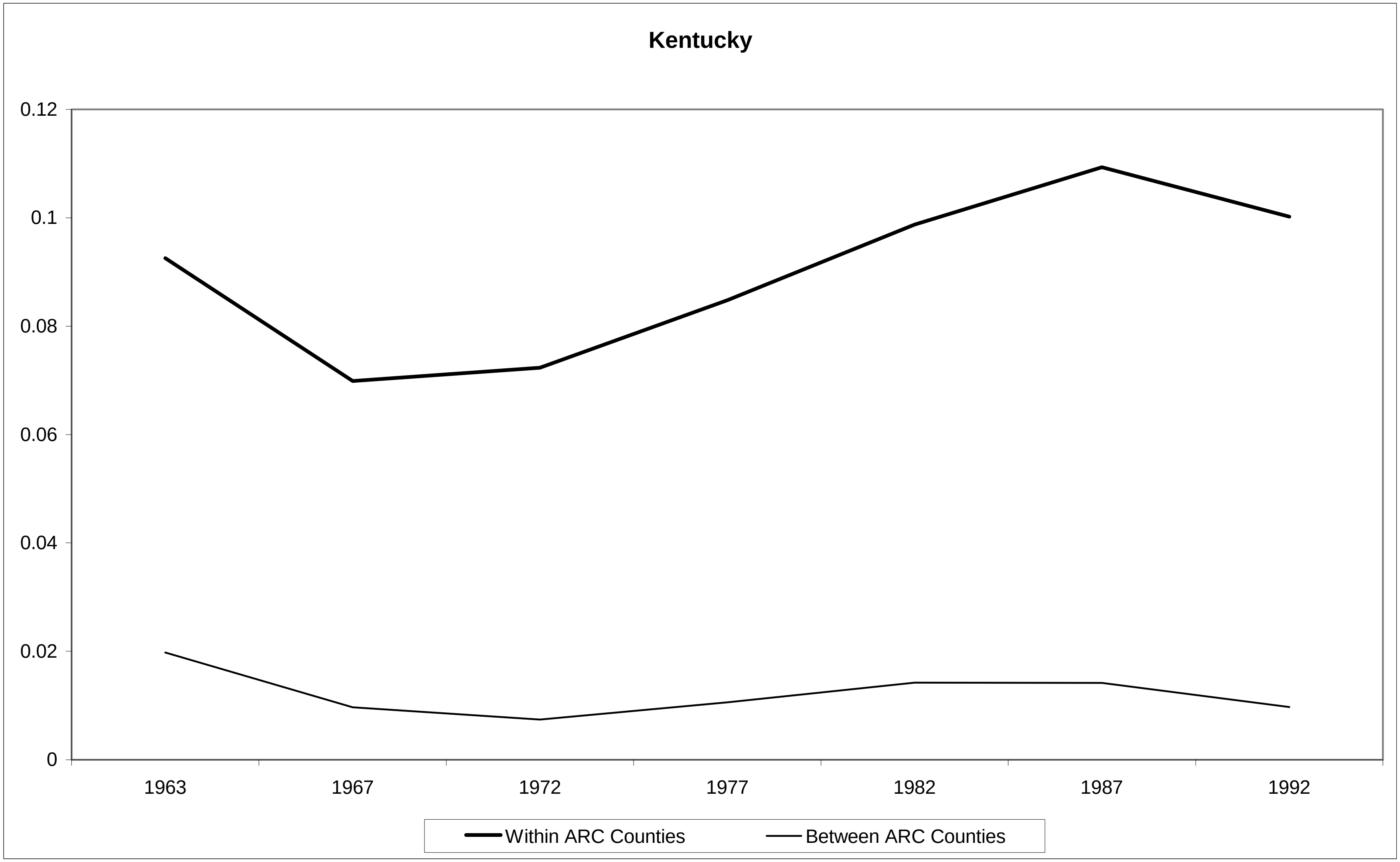


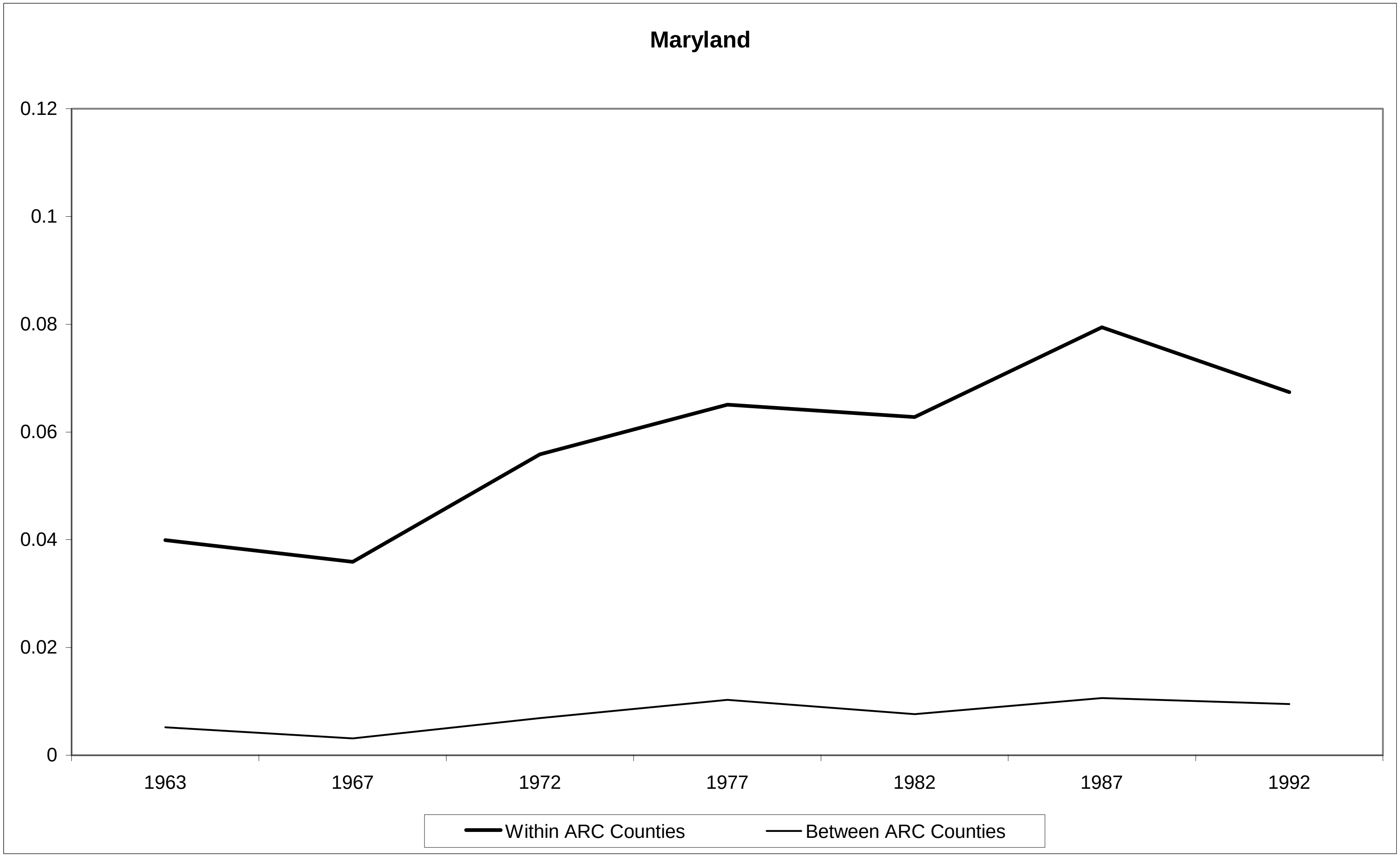
Figure 6- Relationship between Wage Share of ARC counties and their Contribution to Inequality, 1992.

It is also of interest to look at the dynamics of the "pure inequality" components, that is, at the evolution the between- and within-counties components of inequality in the ARC region. Figure 7 shows a sequence of thirteen charts, each showing, in the same scale for all states, two lines, one representing inequality within ARC counties (the thicker line) and the other inequality between ARC counties.

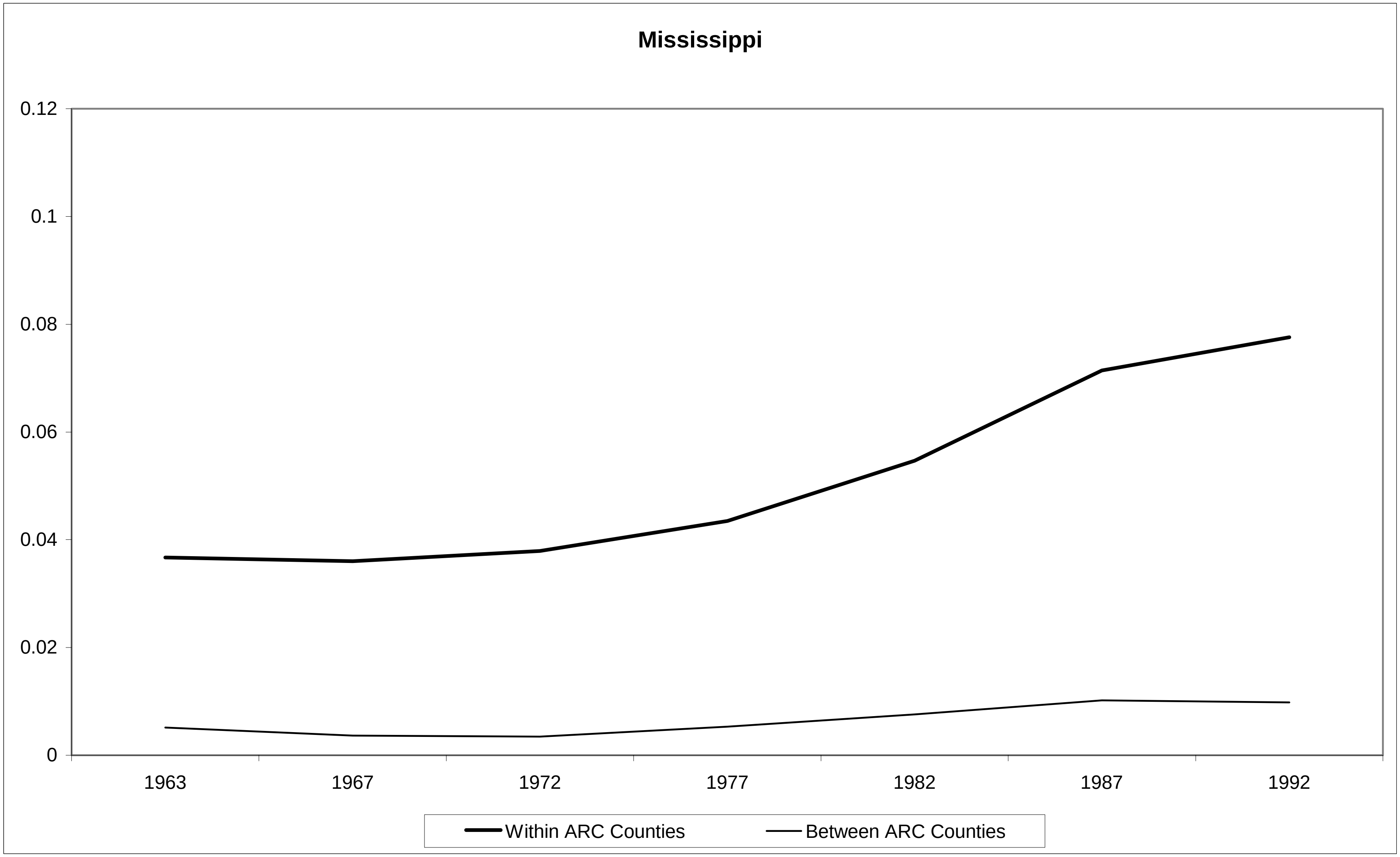

Alabama
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

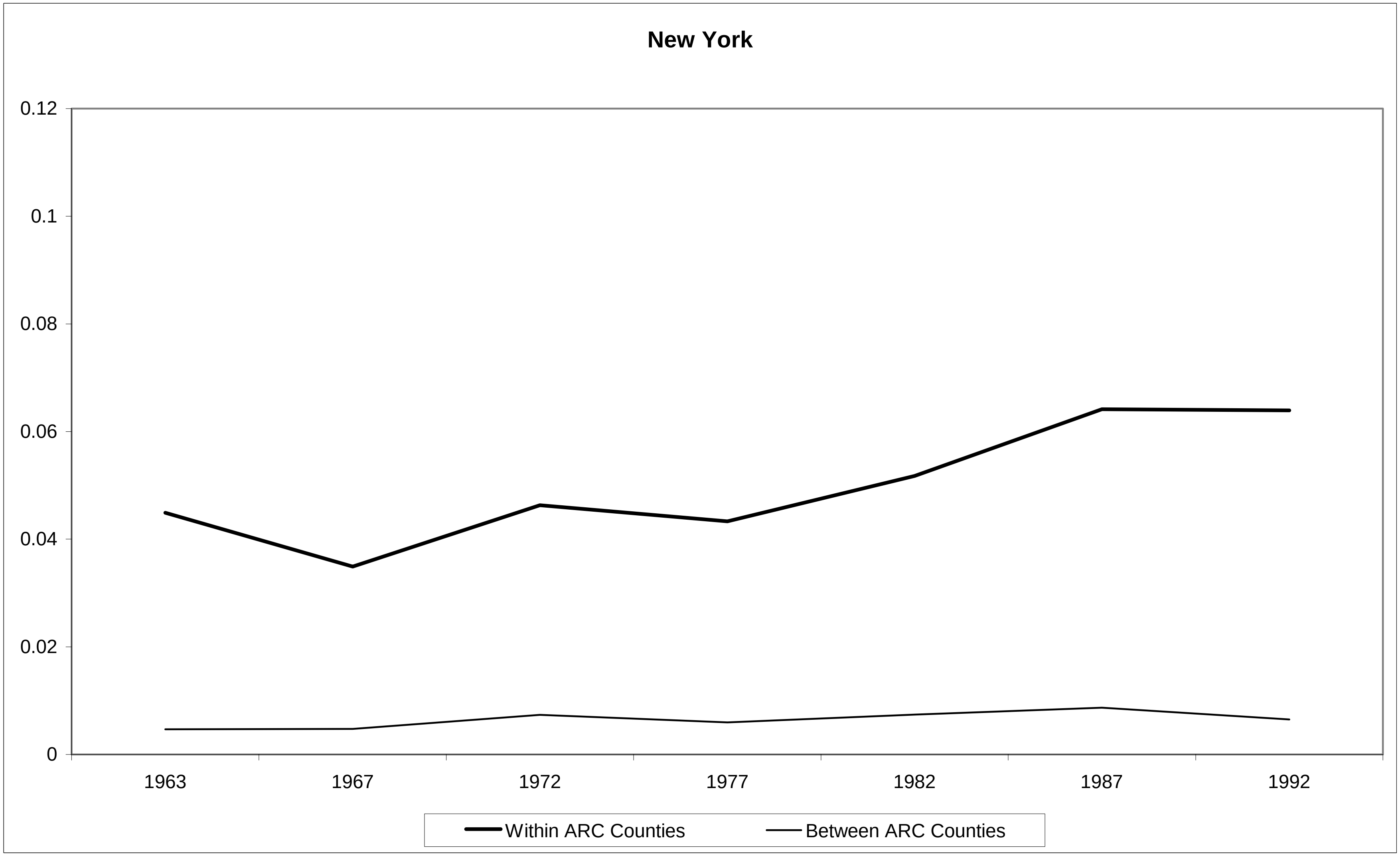

Georgia
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

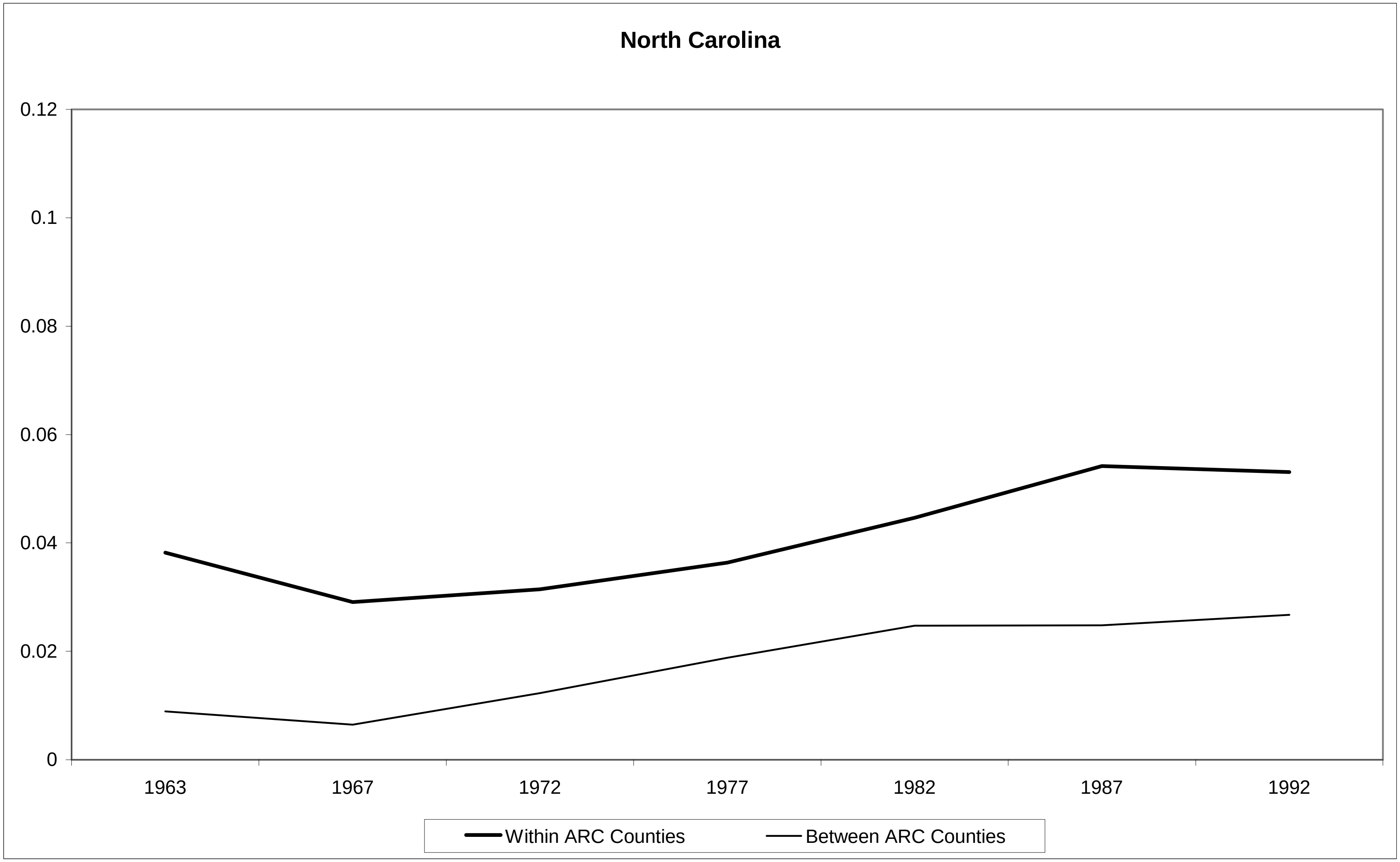

Kentucky
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

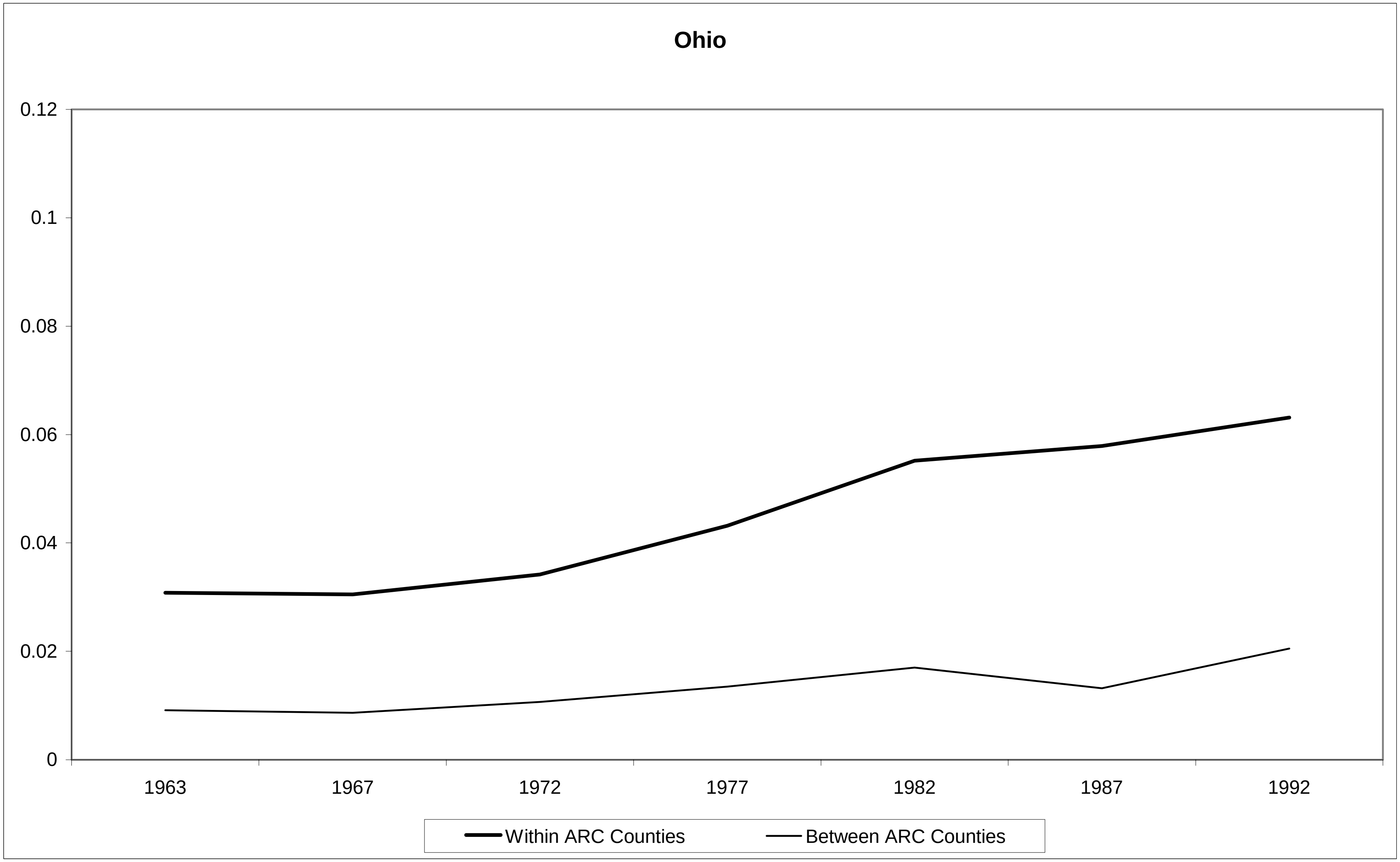


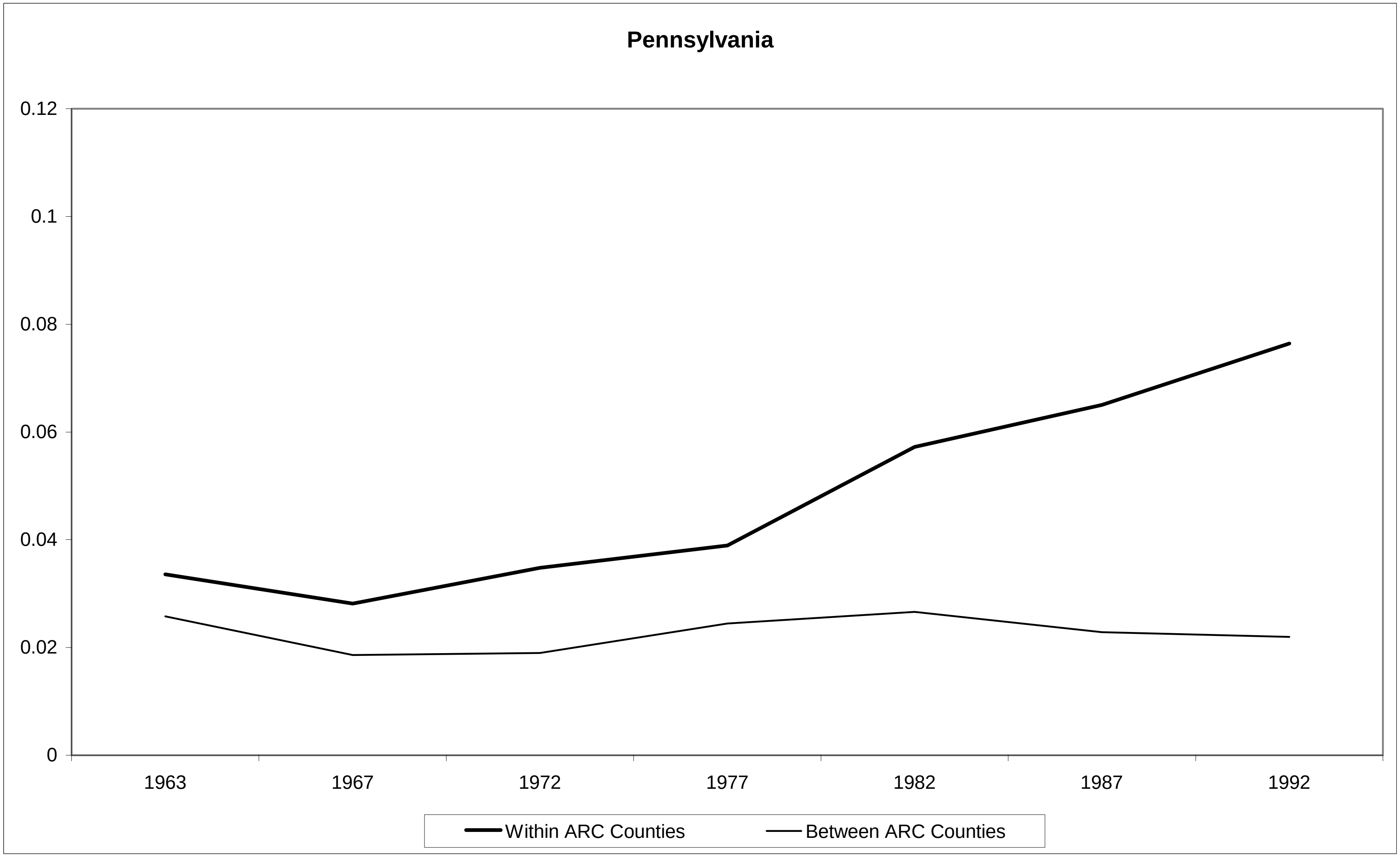
Maryland
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

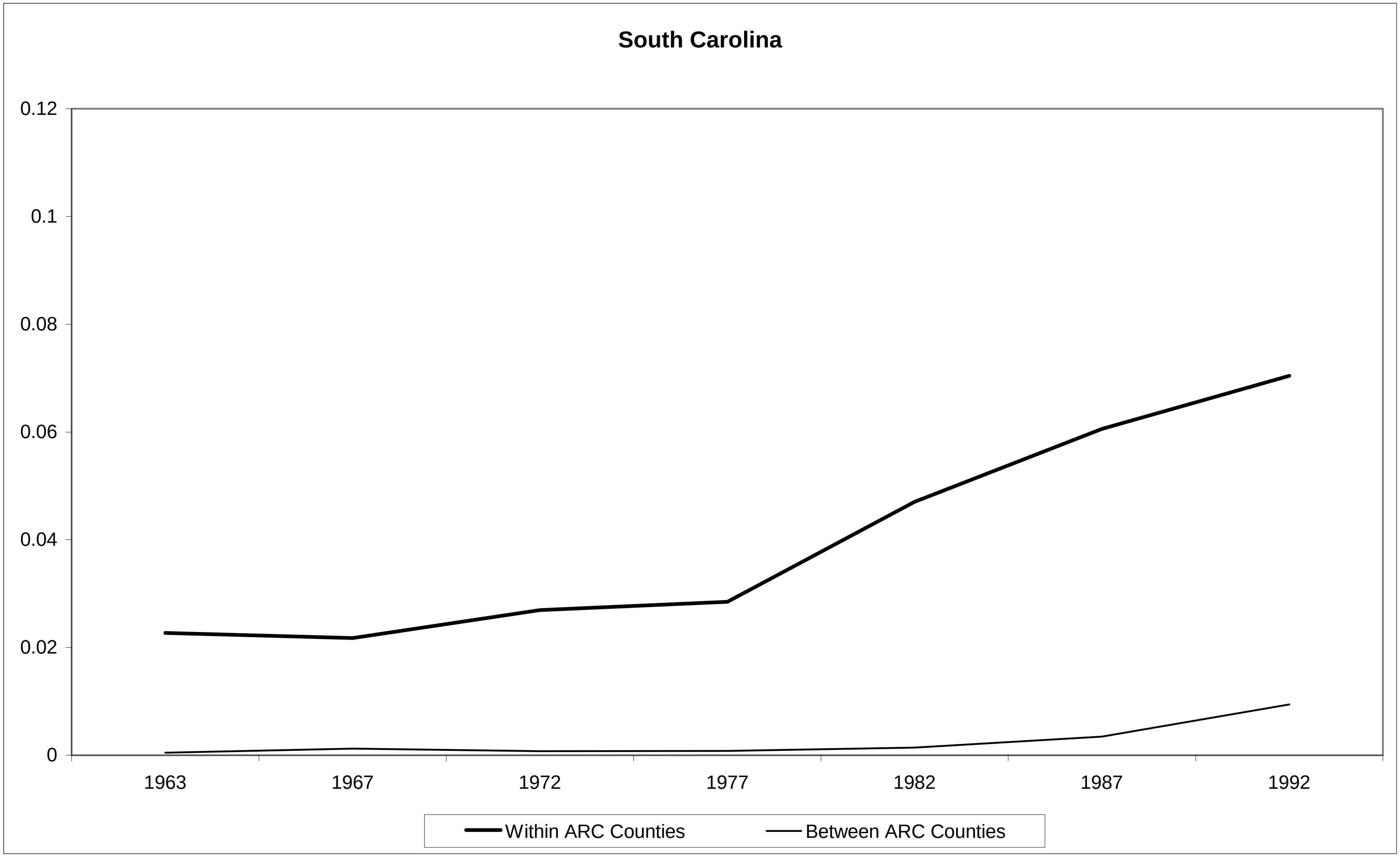

Mississippi
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

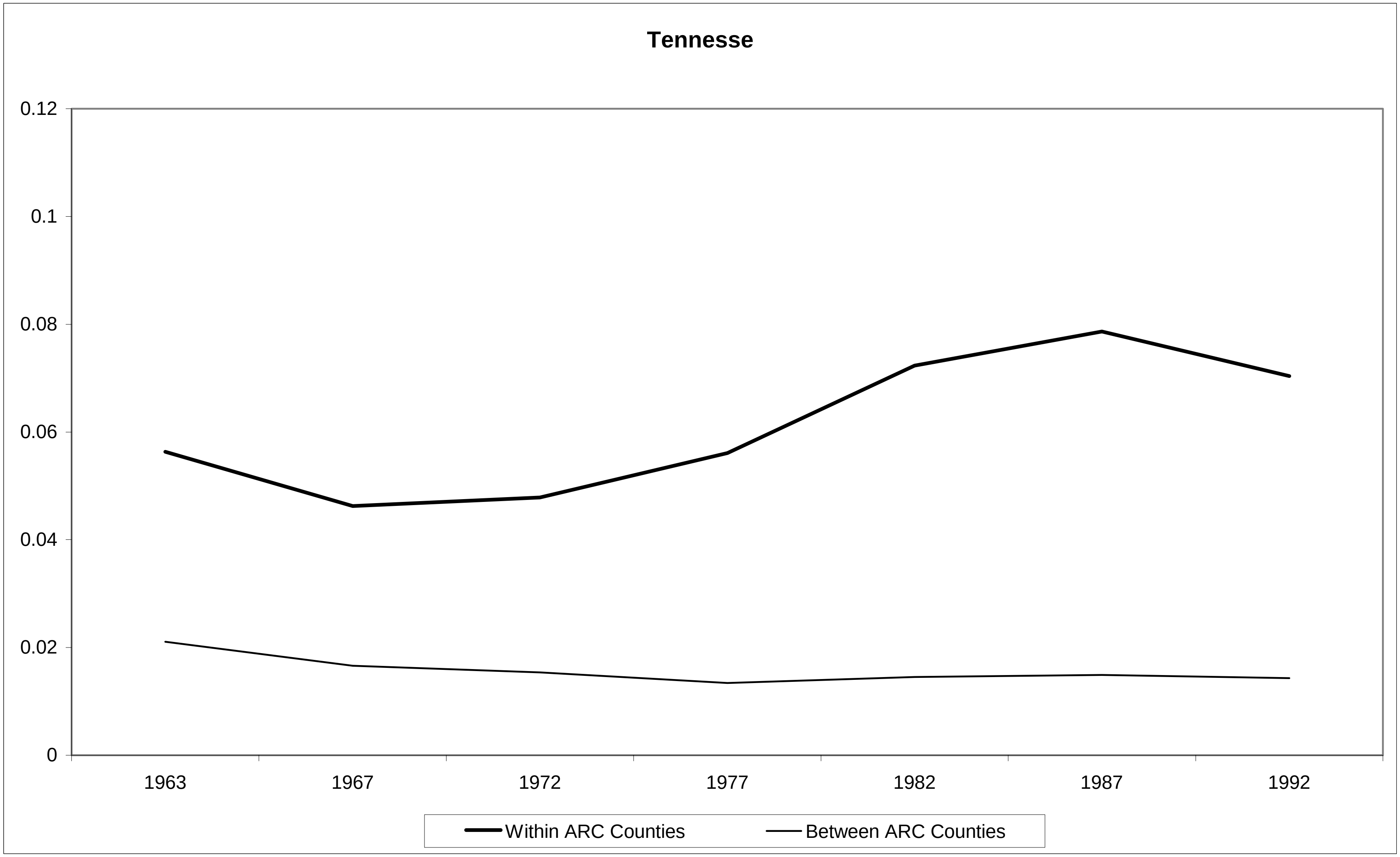

New York
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

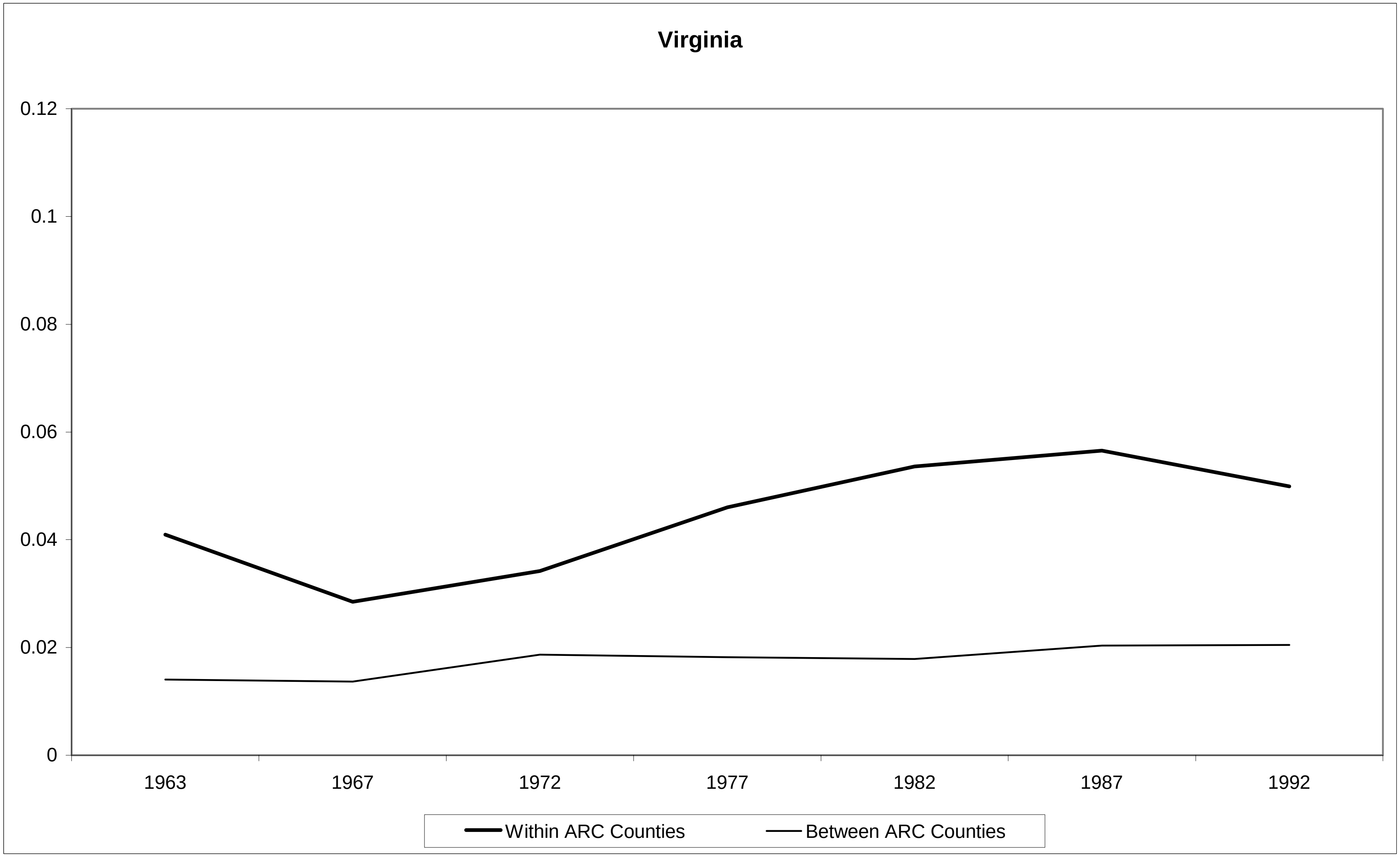

North Carolina
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties


Ohio
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties


Pennsylvania
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties


South Carolina
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties


Tennesse
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties


Virginia
0.12
0.1
0.08
0.06
0.04
0.02
0
1963
1967
1972
1977
1982
1987
1992
Within ARC Counties
Between ARC Counties

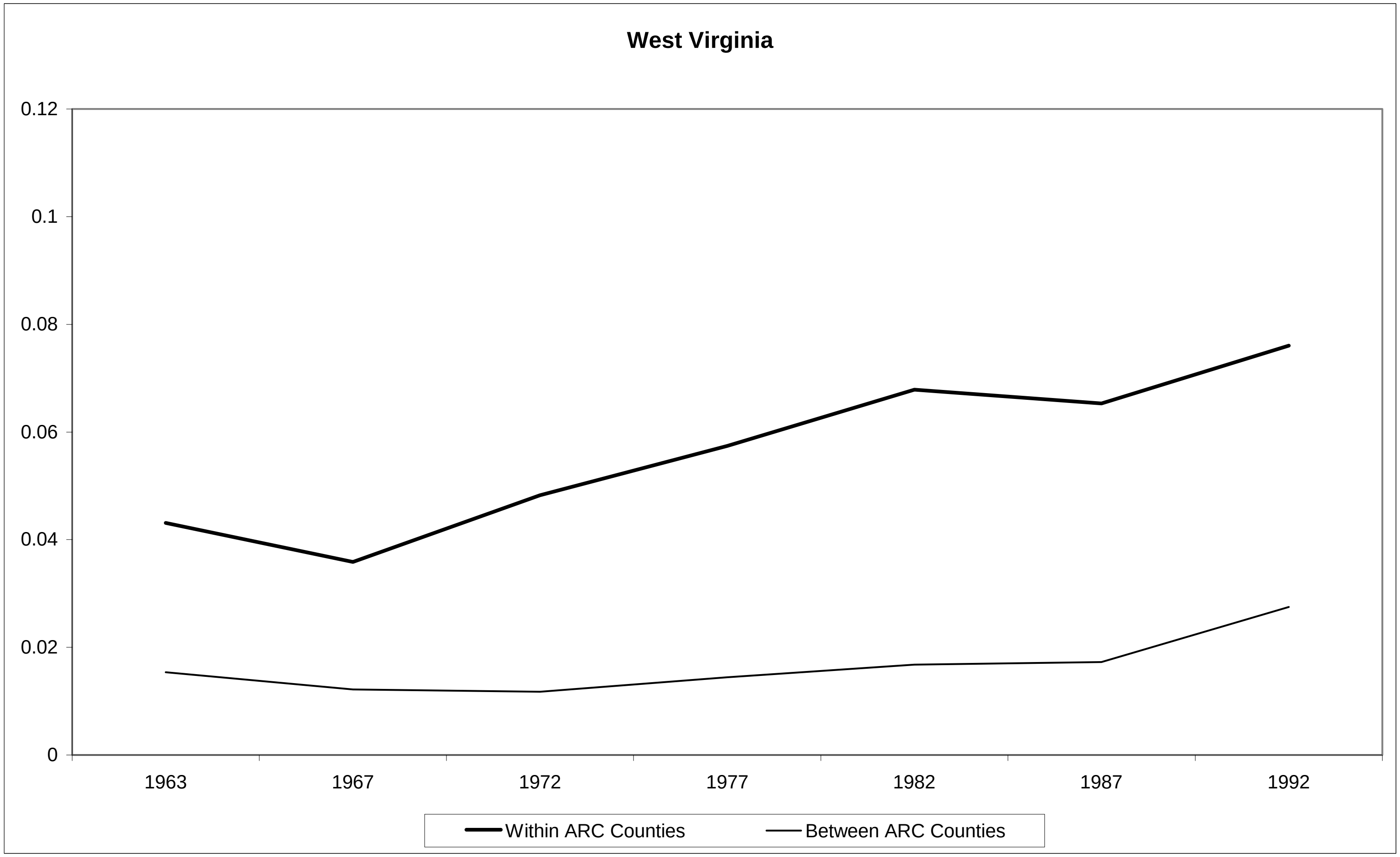


Figure 7- Within ARC Counties and Between ARC Counties Components of State Inequality

The dynamics of the two components vary from state to state, but there are some common features. First, inequality within counties is always higher than inequality between counties: compared to differences across plants within counties, the average levels of wages across counties are relatively similar. Secondly, inequality within counties (across plants) is more volatile and is subject to larger changes than inequality between counties. Beyond these general features, clearly common to all states, the dynamics of the two components differ from state to state. For a number of states there is a co-evolution of the two components throughout the period under analysis, but for other states the two components diverge, either since the beginning of the period or from some later point in time. We look at the dynamics of the two components in each state below.

In Alabama the levels of the Theil index for the within and between counties components were the same in 1963, a singular feature unique to this state. While the between-counties component remained stable, which means that there was not any divergence across the average wage levels across counties, inequality within counties increased steadily. In 1992 inequality within counties is about two times as large as inequality between counties. It must also be noted that the level of inequality across counties in Alabama, while stable, is large when compared with the other states. The Theil index measuring inequality across counties in Alabama never drops below 0.03.

In Georgia the dynamics of the two components is very similar, with the "distance" between the two lines remaining the same from 1963 to 1987. In 1992 inequality within counties increases sharply, with inequality between counties increasing as well from its 1987 value of 0.015 to about 0.02.

The dynamics of the two components also track each other in Kentucky, but in this state the level of inequality within counties is much larger than inequality between counties. In fact, inequality within counties in Kentucky is the largest of all states, reaching a Theil of 0.1 in 1987. The evolution of the two components is close to each other in both Maryland and in Mississippi, but in different ways. While in Maryland the two components have a rather volatile behavior, in Mississippi the within counties component increases monotonically, with the between counties component increasing as well or remaining virtually stable. In New York, again, the dynamics of the within and of the between counties components is similar to each other, with the dispersion in average wages across counties remaining small throughout the period.

In North Carolina and in Ohio the dynamics of the two components are the same until 1982. From 1967 to 1982 the difference between the measure of inequality within counties and the measure between counties remains rather stable, with both components increasing. From 1982 to 1987 the within counties component continues to increase in both states, but the between counties component shows a slight decrease. In North Carolina the trends are reversed from 1987 to 1992, with inequality within counties decreasing slightly and inequality between counties showing a small increase. In Ohio, both components recover an upward movement.

Pennsylvania, like Alabama, shows divergent behavior of the two components, with inequality within counties increasing steadily from 1967 onwards, while inequality between counties oscillates around the 0.02 level. From 1982 on, inequality between counties actually decreases, while the within counties component continues the upward trend. In 1992 the difference between the two lines is more than five times larger than the difference in 1963.

The dispersion across counties in South Carolina is virtually nonexistent until 1982. The increase in the within county inequality is also modest up to 1977, but it picks up then, more than doubling its Theil level from 1977 to 1992. The between counties component also starts to increase in 1982, doubling from 1982 to 1987 and then doubling again from 1987 to 1992.

In Tennessee and in Virginia the within counties component increases up to 1987 and decreases from1987 to 1992. The between counties component in Tennessee decreases steadily throughout the period, while in Virginia it remains stable. For West Virginia the two lines represent the level of overall inequality in the state, since all counties belong to the ARC. The increase in inequality within counties starts, as with most states, in 1967, but stops in 1982, resuming again in 1987. From 1987 to 1992 the between counties component, which had remained virtually stable, starts to increase as well.

We next consider the issue of inequality in ARC counties at a more detailed level. As we have already seen, the major portion of inequality is due to the dispersion of wages within counties, regardless of whether we are considering ARC or non-ARC counties. Is there systematically higher inequality within ARC counties than within non-ARC counties? To explore this question, we list in Table 5 the thirty counties with the highest level of the Theil index in 1967 and in 1992. For most states, as for the US as whole, 1967 corresponds to the year of lowest inequality and 1992 to the peak year.

There is not an overwhelming presence of ARC counties in the two lists. In fact, ARC counties are almost absent from the 1967 list, with only five making the top thirty of inequality. In 1992 the number of ARC counties increases to eight, but the results of the two lists suggest that there is no reason to conclude that ARC counties are particularly unequal. The distribution of the most unequal counties among states reveals that there is no concentration of very unequal counties in a single state, or even a handful of states.

The table indicates that there have been major changes in the structure of the geographic distribution of inequality across counties, since few that were most unequal in 1967 list make the list in 1992. An important exception is New York County, which is in the top ten in 1967 and is the most unequal US county in 1992, with a level of the Theil index far higher than any other county.

Table 5. The Most Unequal Counties in States with ARC Counties in 1967 and in 1992[4]

	1967			1992		
	STATE	COUNTY	THEIL	STATE	COUNTY	THEIL
1	Maryland	Somerset	0.17	New York	New York	0.26
2	Virginia	Northumberland	0.12	South Carolina	Saluda	0.19
3	Maryland	Talbot	0.11	**Alabama**	**Shelby**	**0.19**
4	Georgia	NonARC Counties	0.11	**Georgia**	**Whitfield**	**0.18**
5	North Carolina	Lenoir	0.10	New York	Westchester	0.18
6	**Tennessee**	**ARC Counties**	**0.09**	Kentucky	Muhlenberg	0.17
7	**Kentucky**	**ARC Counties**	**0.09**	North Carolina	Columbus	0.16
8	New York	New York	0.09	Mississippi	Franklin	0.16
9	Alabama	Marengo	0.09	Virginia	Fairfax	0.15
10	**Georgia**	**Haralson**	**0.09**	Alabama	Russell	0.15
11	Kentucky	NonARC Counties	0.09	Alabama	Marengo	0.15
12	Mississippi	NonARC Counties	0.09	Georgia	Washington	0.14
13	Virginia	Westmoreland	0.09	Virginia	Prince Edward	0.14
14	**Tennessee**	**McMinn**	**0.08**	Alabama	NonARC Counties	0.14
15	Maryland	Dorchester	0.08	Virginia	Amherst	0.14
16	North Carolina	Columbus	0.08	Mississippi	Marion	0.14
17	Virginia	Arlington	0.08	Alabama	Clarke	0.14
18	Alabama	Escambia	0.08	Tennessee	NonARC Counties	0.13
19	Alabama	Russell	0.08	**Kentucky**	**ARC Counties**	**0.13**
20	North Carolina	Pamlico	0.08	**Alabama**	**Limestone**	**0.13**
21	Alabama	NonARC Counties	0.07	Georgia	Thomas	0.13
22	Georgia	Glynn	0.07	**Mississippi**	**Monroe**	**0.13**
23	**New York**	**Schoharie**	**0.07**	Georgia	Greene	0.13
24	Virginia	Hampton	0.07	**Georgia**	**Barrow**	**0.13**
25	Georgia	Lowndes	0.07	South Carolina	Dorchester	0.13
26	Kentucky	Boone	0.07	**Virginia**	**Pulaski**	**0.13**
27	Virginia	Lancaster	0.07	Alabama	Dallas	0.13
28	South Carolina	Clarendon	0.07	**Georgia**	**Carroll**	**0.13**
29	Alabama	Dallas	0.07	Georgia	Ben Hill	0.12
30	Kentucky	Christian	0.07	Kentucky	Warren	0.12

Note: ARC counties are in bold.

If ARC counties are not particularly unequal, could it be that they are especially equal? Table 6 shows the least unequal counties in the US again for 1967 and for 1992. Again, ARC counties, while more abundant here than in the most unequal list, do not appear to be particularly more equal than the non-ARC counties. However, there is now a concentration of least unequal counties in two states, since Georgia and Virginia have, together, about half of the most equal counties in the US.

[4] The listings of "ARC counties" and of "Non-ARC counties" correspond to groupings of either ARC or Non-ARC counties in each state due to disclosure concerns.

Table 6. The Most Equal Counties in States with ARC Counties in 1967 and in 1992

	1967			1992		
	STATE	COUNTY	THEIL	STATE	COUNTY	THEIL
1	**Virginia**	**Floyd**	**0.003**	**New York**	**Schoharie**	**0.002**
2	North Carolina	Cabarrus	0.005	Georgia	Hancock	0.005
3	Mississippi	Scott	0.005	Virginia	Essex	0.012
4	**North Carolina**	**Macon**	**0.005**	Tennessee	Marshall	0.013
5	Virginia	Louisa	0.006	Georgia	Randolph	0.014
6	Tennessee	Carroll	0.006	**North Carolina**	**Rutherford**	**0.014**
7	**Alabama**	**Bibb**	**0.006**	**Georgia**	**Walker**	**0.014**
8	**North Carolina**	**Cherokee**	**0.007**	Virginia	Newport News	0.018
9	Virginia	Page	0.007	Virginia	King and Queen	0.018
10	Georgia	Tattnall	0.007	**Ohio**	**Gallia**	**0.020**
11	**Georgia**	**Gilmer**	**0.007**	North Carolina	Chowan	0.020
12	Virginia	Brunswick	0.007	**Kentucky**	**Adair**	**0.020**
13	**New York**	**Schuyler**	**0.007**	**North Carolina**	**Ashe**	**0.020**
14	**North Carolina**	**Rutherford**	**0.007**	**Kentucky**	**Casey**	**0.020**
15	Georgia	Hancock	0.007	**North Carolina**	**McDowell**	**0.020**
16	Georgia	Greene	0.008	Mississippi	Montgomery	0.021
17	Georgia	Baldwin	0.008	North Carolina	Anson	0.021
18	Virginia	Prince Edward	0.008	Georgia	Johnson	0.021
19	Alabama	Conecuh	0.008	Ohio	Champaign	0.021
20	**Mississippi**	**Itawamba**	**0.008**	**Georgia**	**Hart**	**0.021**
21	**Alabama**	**Lamar**	**0.008**	Virginia	Harrisonburg	0.021
22	Kentucky	Grayson	0.009	Virginia	Caroline	0.021
23	Georgia	Grady	0.009	**Ohio**	**Hocking**	**0.021**
24	**South Carolina**	**Pickens**	**0.009**	Georgia	Henry	0.022
25	Alabama	Geneva	0.010	Ohio	Madison	0.022
26	Tennessee	Dickson	0.010	Virginia	Page	0.022
27	**New York**	**Delaware**	**0.010**	Tennessee	Tipton	0.023
28	Virginia	New Kent	0.010	Ohio	Putnam	0.023
29	Ohio	Trumbull	0.010	**Georgia**	**Gordon**	**0.023**
30	**Pennsylvania**	**Beaver**	**0.010**	Maryland	Kent	0.023

Note: ARC counties are in bold.

In Table 9, shown in an appendix, the five most and least unequal counties in each state, again for 1967 and for 1992, are shown. There is no systematic relationship between being an ARC county and the level of inequality.

4. Contribution of the Appalachian Region to US Inequality

In this section we look at the contribution of the ARC region as whole to the level of US inequality. Dividing all counties in the US into those that are members of the Appalachian Regional Commission and those that are not, we can create an intermediate sub-national level of aggregation. Thus, following the same strategy as in section 2, we can divide US inequality into six components.

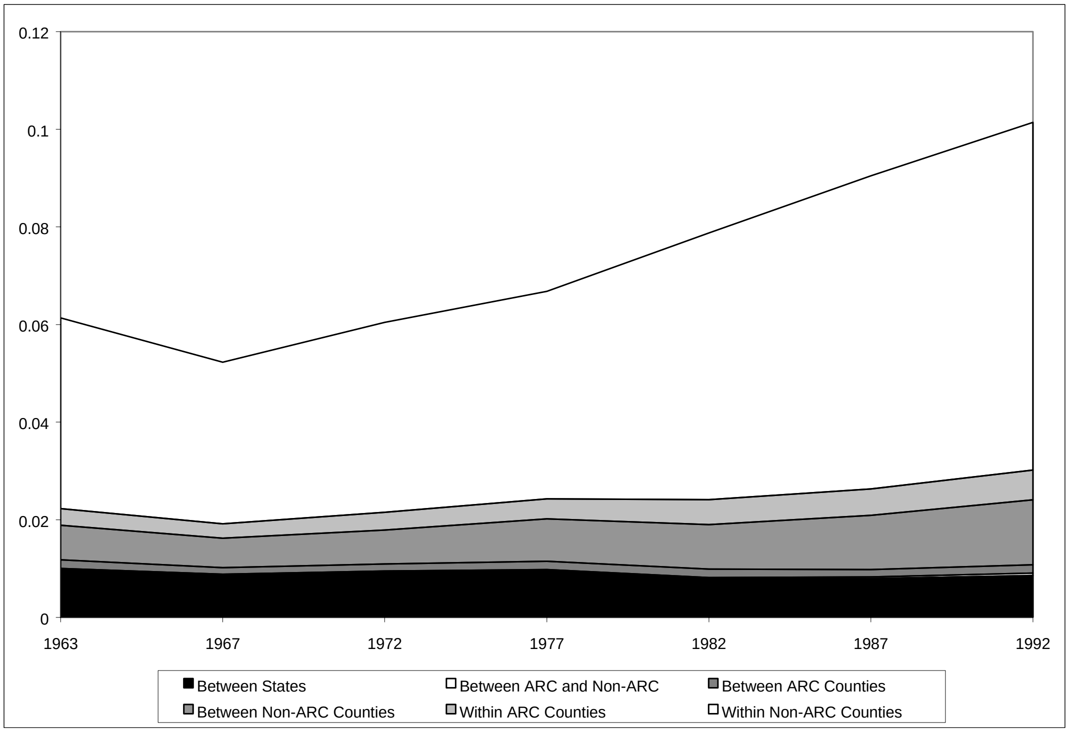
The first component is associated with the differences across the ARC and the non-ARC counties, when counties in each division are considered as a whole. This component, between ARC and non-ARC, measures the contribution of differences in average wages across the two divisions. Each of the two divisions can then be decomposed into between counties and within counties components. The sixth and final component is the contribution to US inequality of differences across states.

Figure 8 shows the results of this decomposition. Naturally, the between states component remains the same as in section 2, but the within states component is now divided into five contributions. The share with which each component contributes to the US Theil is given in Table 7, complementing the information in the chart.

The contribution of the between ARC counties remains small, decreasing from almost 3% in 1967 to slightly more than half that percentage in 1992. The contribution of the within ARC counties is equally small, contributing between 5.5% and 6.5% throughout the period, with the highest contribution, of 6.5% having been reached in 1982.

Clearly, the increase in US inequality is accounted for by the non-ARC counties, which is not surprising, given the structure of the Theil index in which the contributions of the groups are weighted by each group’s wage share. Given that the ARC’s wage share is relatively small, the small contribution of the ARC counties to US inequality measured by the Theil index is not surprising.

Figure 8. Contribution of Inequality within and between ARC and Non-ARC counties to the US



Theil

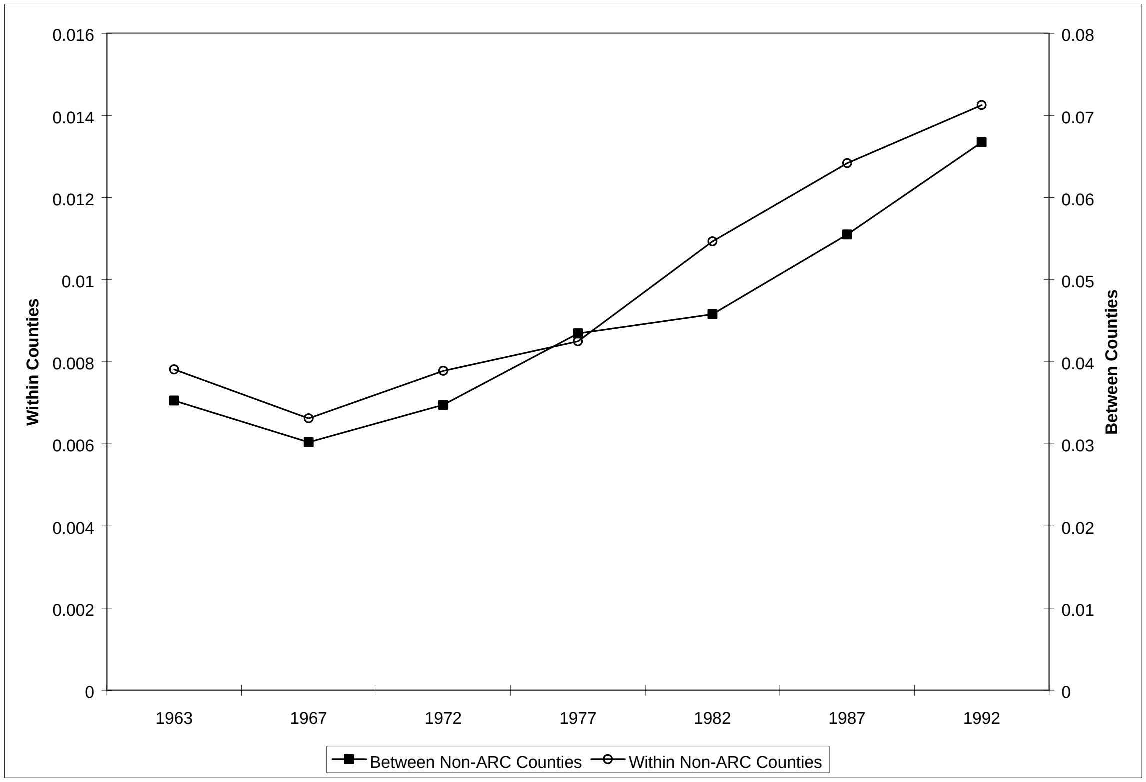
Table 7. Share of the Contribution to US Inequality of Each of the Five Components

	1963	1967	1972	1977	1982	1987	1992
Between ARC and Non-ARC	0.39%	0.41%	0.38%	0.36%	0.30%	0.39%	0.48%
Between ARC Counties	**2.90%**	**2.63%**	**2.39%**	**2.50%**	**2.18%**	**1.67%**	**1.67%**
Between Non-ARC Counties	11.50%	11.55%	11.50%	13.01%	11.62%	12.26%	13.16%
Within ARC Counties	**5.59%**	**5.62%**	**6.06%**	**6.18%**	**6.46%**	**5.95%**	**5.98%**
Within Non-ARC Counties	63.69%	63.34%	64.35%	63.62%	69.40%	70.96%	70.25%
Between States	15.93%	16.45%	15.31%	14.33%	10.05%	8.77%	8.45%

Is it worthwhile to consider the dynamics of the non-ARC and of the ARC components of US inequality? The difference in the levels of these components is such that a comparison of the contributions is almost meaningless. Still, we can ask: Are there differences in the evolution of inequality components between and within both the ARC and the non-ARC counties?

Figure 9 shows that the dynamics of the within and between components of the non-ARC counties are highly similar. Both components decrease from 1963 to 1967, with a steady increase from 1967 to 1992, in the usual “check mark” pattern of overall US inequality.

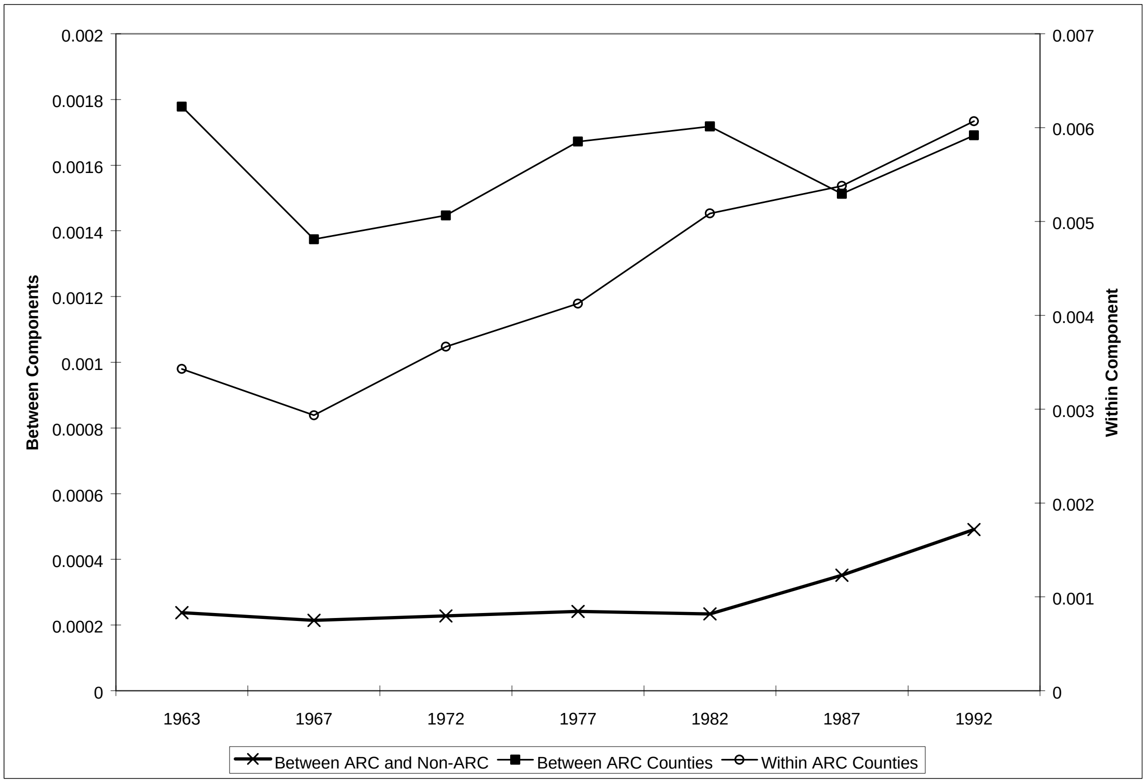
Figure 9. Contributions to Inequality: Between Non-ARC Counties and Within Non-ARC Counties



If we consider the same two components (across counties and within counties) for the ARC counties (Figure 10) the dynamics of the two series are again similar, but they are not as close as the dynamics of the non-ARC components. For the ARC counties components there is a decrease from 1963 to 1967, with both increasing from 1967 to 1982. However, from 1982 to 1987 there was a decrease in the component associated with inequality across ARC counties, while the within ARC counties component continued to increase. From 1987 to 1992 both components increased.

Jensen (1998) found that there has been a persistence in the gap in manufacturing between Appalachia and the rest of the country from 1967 to 1992, with Appalachia showing consistently lower wage and lower productivity than the national average. Our results suggest, though, that in the 1980s, while the difference has remained small, it increased considerably in relative terms. Jensen also found a high reliance in Appalachia of branch plants, as opposed to single plants. Additionally, Jensen concluded that branch plants pay higher wages than single unit plants, and that this premium has increased over time. This effect may offer a possible explanation for the dynamics of within county inequality in Appalachia.

Figure 10- Dynamics of the Between ARC Counties, of the Within ARC Counties and of the Between ARC and Non-ARC Counties Components of US Inequality



Perhaps the most interesting finding in Figure 10 is the increase in the component of US inequality measuring the dispersion between the ARC and the non-ARC groups of counties. This component remains small throughout the period and is constant up to 1982. However, from 1982 to 1992 the gap between the ARC and non-ARC counties starts to increase. This indicates that there has been a widening gap between the ARC and the non-ARC counties considered as whole beginning in the early 1980s, possibly due to slower wage growth in Appalachia than in the non-ARC counties as a whole.

5. Dynamics of Pay Inequality: A Representation in Maps

In this section, we present highly summarized information on the evolution of inequality at the county level. This information is most readily conveyed in map form. To produce the maps in Figures 11A through 11G, we first the measured inequality levels for 1992 in gray-scale so that equal numbers of counties would fall into each bin, ranging from white for low-inequality counties to black for high-inequality counties. We then froze the bin boundaries at their 1992 values, and asked the question: to what degree did the distribution of counties across bins shift in earlier years? Thus, as the maps progress over time, a movement toward the black indicates a rising proportion of counties that would be considered high-inequality in 1992; a shift toward white indicates a falling proportion of such counties. Figures 11A through 11G present a sequence of maps for the ARC counties.

Figure 11A- Inequality within ARC Counties in 1963

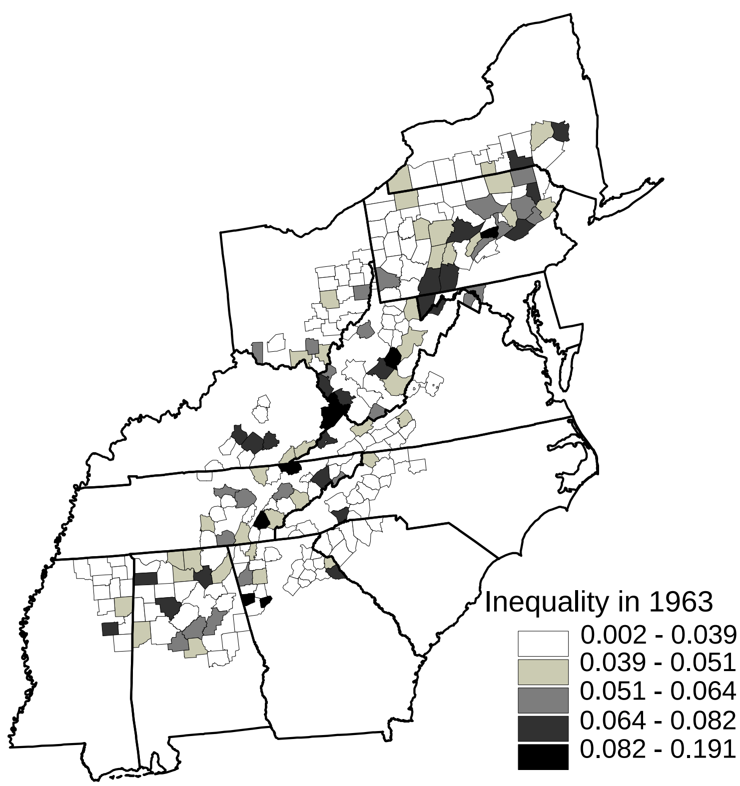


Figure 11B- Inequality within ARC Counties in 1967

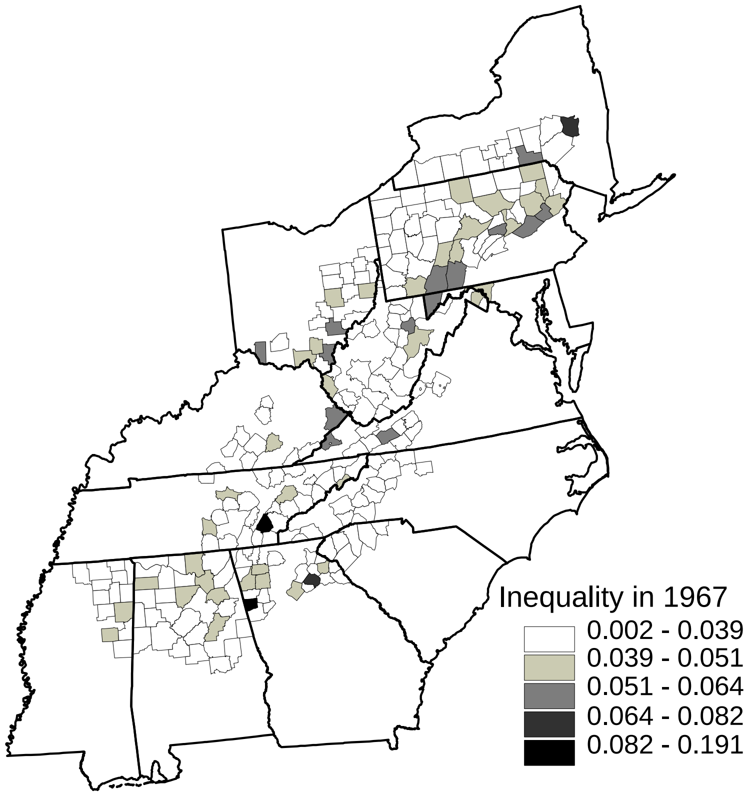


Figure 11C- Inequality within ARC Counties in 1972

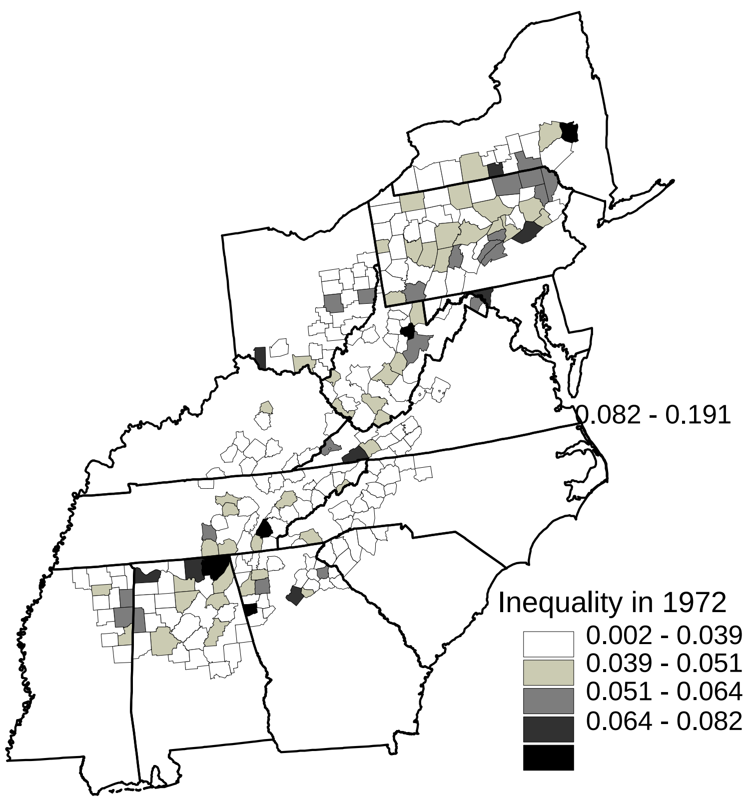


Figure 11D- Inequality within ARC Counties in 1977

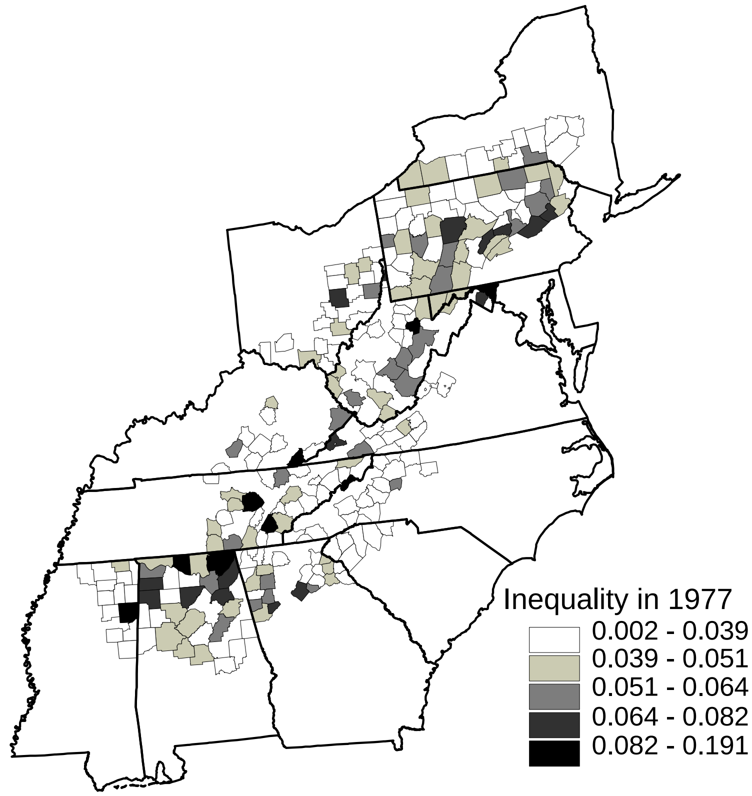


Figure 11E- Inequality within ARC Counties in 1982

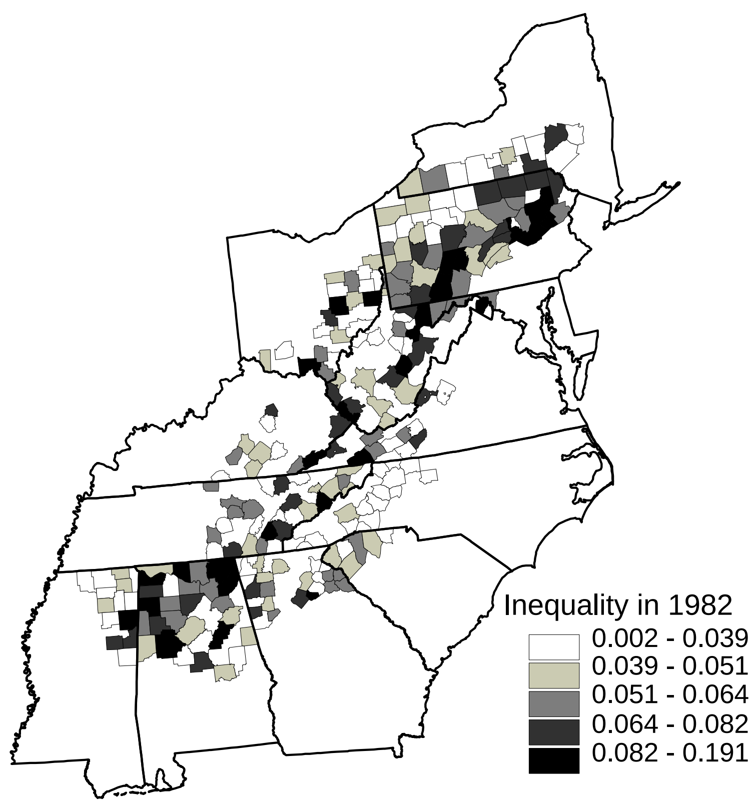


Figure 11F- Inequality within ARC Counties in 1987

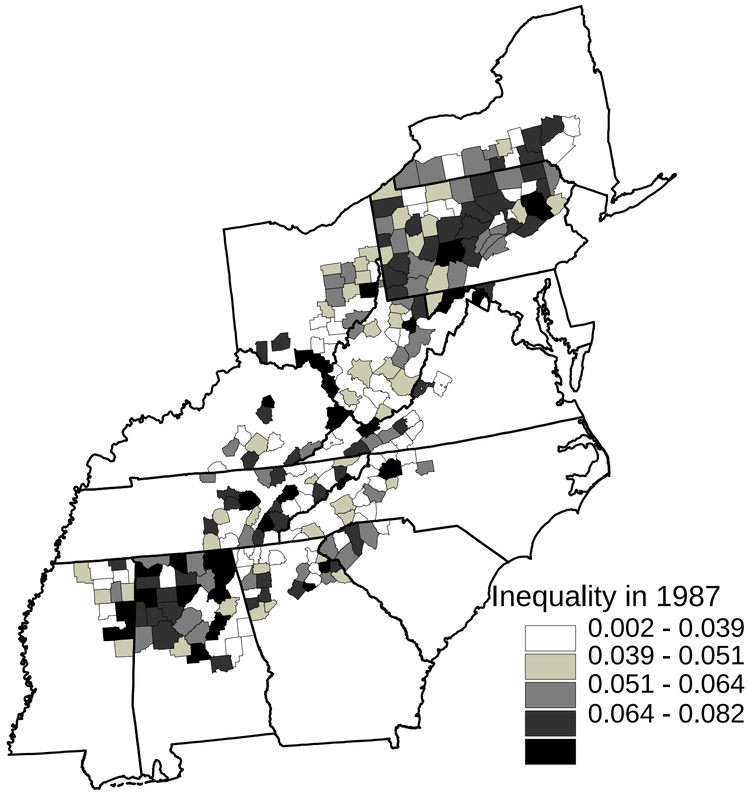
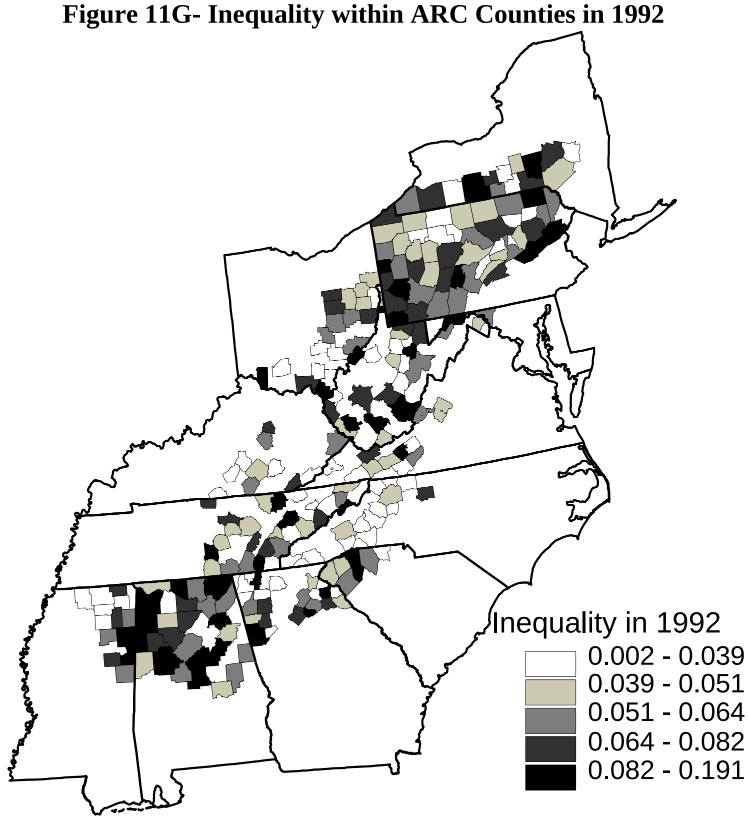


Figure 11G- Inequality within ARC Counties in 1992



While some counties manage to escape, the large increase in within-county inequality in general is a nationwide phenomenon of the 1980s. The national pattern is, essentially, the local pattern in this instance.

6. Macroeconomic Factors and Inequality Dynamics

We now turn briefly to consider the role of macroeconomic factors in the evolution of inequality across geographic regions.

We already know that there is a macroeconomic pattern: this is the "check-mark" pattern of rising inequality at the national level in the years following 1967. The fact that this pattern is replicated in so many instances throughout this study at the regional, state and within-state levels tells us that we need not seek too far for regionally-specific sources of the movement of inequality; the affair was a national phenomenon.

Clearly, a reasonable inference from this data might be that the national rate of unemployment – a measure of national economic performance that declined from 1963 to 1967 and then increased dramatically until the early 1990s – played an important role in driving up the dispersion of pay rates among those who remained employed in the manufacturing sector. The dramatic increase in pay inequality nearly everywhere in the early 1980s, when unemployment reached 10 percent, lends further weight to this supposition. However, with only six distinct time-series data points, a statistical test of this proposition seemed unlikely to be highly persuasive. Table 8 summarizes unemployment rates over this time.

Table 8. Unemployment rates overall and in manufacturing

	Unemployment Rates	
	US	Manufacturing Only
1963-1967	4.6	4.3
1967-1972	4.6	4.7
1972-1977	6.6	6.9
1977-1982	7.2	7.8
1982-1987	7.9	8.6
1987-1992	6.1	6.2

Can unemployment at the local level explain the change in inequality at the local level? The answer to this question is clearly, no. There is in fact no correlation whatever between changes in unemployment at the county level and changes in the county measure of pay dispersions. The force of the effect of national economic conditions, in other words, appears to be firmly from national causes to local consequences, at least in this data. This can result from the way in which local perceptions are tied to national economic conditions. For example, a context of recession or slower growth at the national level can influence the sentiment of economic actors and affect their decisions and their interactions; the reaction to national conditions will be played out at changes at the local level.

We also examined the relationship between average wage changes in each county and the change in inequality in that county.

In our view, a correct characterization of this relationship is that changes in local unemployment and local wages are *conditioned* by national economic conditions. Figure

12 illustrates this point, showing the change in inequality and the change in average wages for each of the counties in our data set over each time interval.

The figure illustrates the absence of local cross-section correlation in any time period: the counties form a shapeless ball in the space of inequality changes and wage changes. But the position of the shapeless ball in the space as a whole differs from period to period. Most notably, in 1963-1967 the center of gravity of the ball lies in the upper-left quadrant, indicating that a preponderance of counties experience rising wages and falling inequality during this prosperous time. In the 1977-1982 period, in contrast, the ball moves squarely to the lower right quadrant: falling average wages and rising inequality predominate in this period—even though the scale of one effect does not predict the scale of the other for any particular county. In the other periods, the force of national economic performance, either positive or negative, does not appear sufficient to impart a systematic pattern to the relationship between wage change and inequality change.

Figure 12. Changes in Average Wages and Changes in Inequality by County 1963-1992

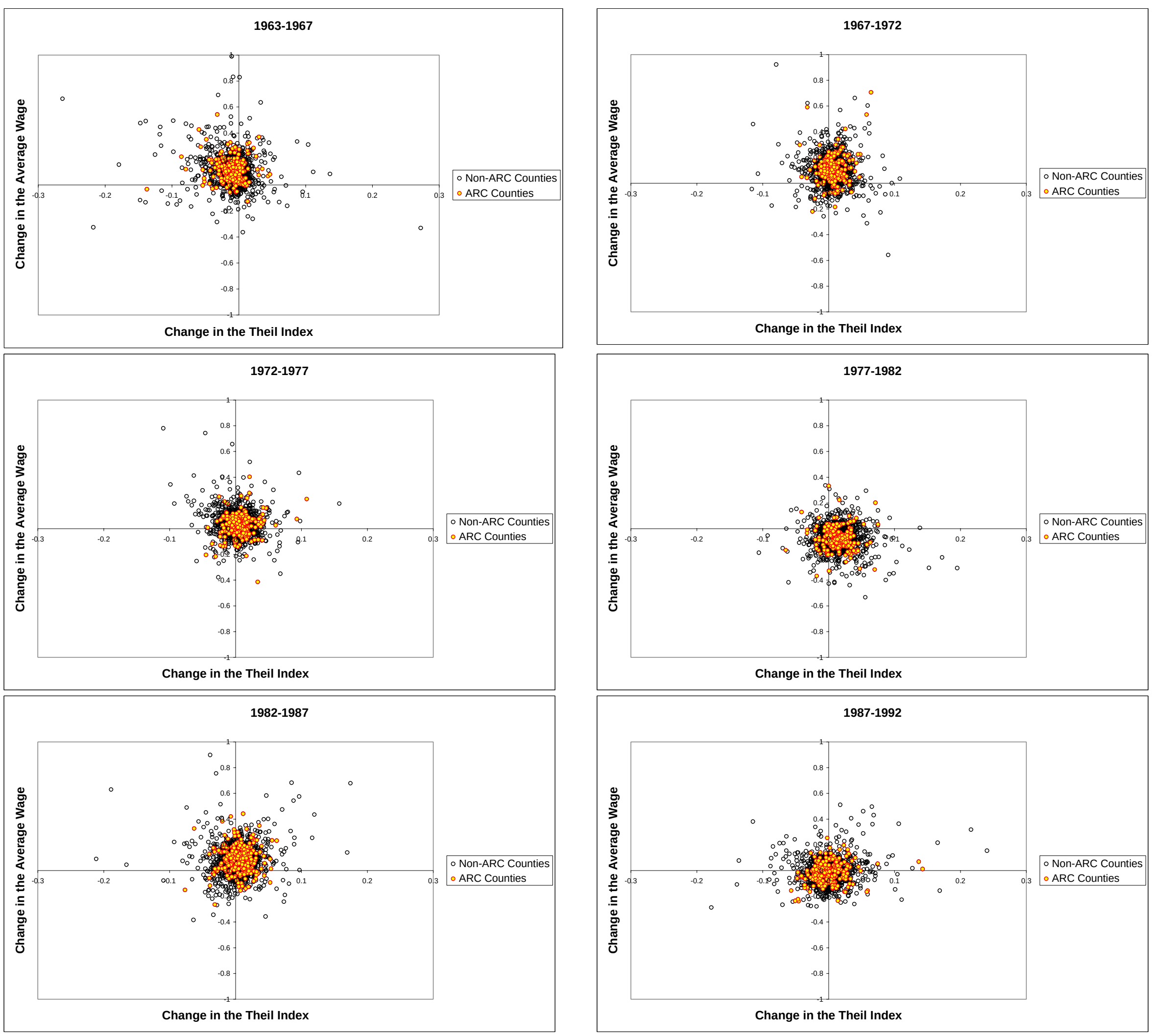


Table 9, finally, summarizes this information. In the strongly prosperous and strongly recessionary periods of the mid-1960s and early 1980s, respectively, there is a broadly negative relationship between wage growth and inequality change. In the remaining periods, the relationship is more evenly balanced; we suggest that there is not a statistically significant difference between the proportion of counties for whom the relationship is positive and those for whom it is negative.

Table 9. Share of Counties (points in the graphs of

Figure 3) that are either in quadrants where the relationship between wage changes and inequality changes is positive or negative

	Positive	Negative
1967	32%	68%
1972	63%	37%
1977	53%	47%
1982	28%	72%
1987	58%	42%
1992	53%	47%

In short, periods of exceptional prosperity—full employment over a sustained time—are required to markedly reduce inequality in American manufacturing pay. Periods of crisis—a steep recession, even if relatively short—are sufficient to alter the profile of inequality in the country for the worse, and for a generation. This is true inside and outside the Appalachian region, though we also find evidence that suggest that in the absence of great and sustained national prosperity, there is a tendency for a low-income region like Appalachia to slip behind the rest of the country on average. This conclusion results from the comparison of the difference between ARC and non-ARC counties in the 1980s, when the gap between the averages in the two groups of counties widened. We would expect, of course, if this view is correct, that some narrowing of differentials would have occurred in the late 1990s. Unfortunately, while this trend may have started and prove detectable in the 1997 Census, it will take until the 2002 Census becomes available before it can be tested definitively by the methods developed here. By that time, moreover, it may be that a slowdown or even a recession will have undone any progress achieved.

A final inference thus seems in order. Our evidence gives no reason to suppose that rising inequality within and between the counties of Appalachia and elsewhere has served any positive social or economic purpose. In particular, it is not associated with rising productivity or increasing living standards. And so, if inequality in American society is itself a problem—as we believe it is—it may be time to address that problem by concerted efforts of policy. Such policies would have to aim to reduce the pay gaps we observe prominently in plant-based geographic data: not inequalities across states, regions, and large distances, but inequalities that are socially conspicuous precisely because they are increasing in small neighborhoods. What measures can achieve this aim? In our view, stronger social minimums—a rising minimum wage—would be a constructive first step. Support for stronger labor organizations, to recapture some of the solidarity that characterized American communities, comparatively speaking, in the late 1960s, would be a second step and, for manufacturing in particular, perhaps an even more important measure.

References

Bernard, A. B., Bradford Jensen, J. (1998). *Understanding Increasing and Decreasing Wage Inequality*, NBER Working Paper No. W6571.

Conceição, P., Galbraith, J. K. (2000), "Constructing Long and Dense Time-Series of Inequality Using the Theil Index," *Eastern Economic Journal*, **26**(1): 61-74.

Jensen, Bradford J. (1998). *Birth and Death of Manufacturing Plants and Restructuring of Appalachia's Industrial Economy, 1963-1992*, Washington DC: Appalachian Regional Commission.

McGuckin and Pascoe (1988) "The Longitudinal Research Database (LRD): Status and Research Possibilities." *Survey of Current Business*, Vol. 68, No. 11, pp. 31-37.

Shannon, C. E. A Mathematical Theory of Communication, *Bell System Technical Journal*, vol. 27, 1948, 379-423.

Theil, H. *Economics and Information Theory*. Chicago: Rand McNally and Company, 1967.

Topel, R. H. (1994). "Regional Labor Markets and the Determinants of Wage Inequality," *American Economic Review*, 84(2), 17-22.

Appendix 1. Notes on the Theil Index[*]

Henri Theil (1967) first noted the possibility of using Claude Shannon's (1948) information theory to produce measures of income inequality. Shannon's theory was motivated by the need to measure the value of information. Shannon argued that the more unexpected an event is, the higher the yield of information it would produce. To formalize this idea, Shannon proposed to measure the information content of an event as a decreasing function of the probability of its occurrence. Adding some axiomatic principles, most importantly that independent events should yield information corresponding to the sum of the individual events' information, Shannon chose the logarithm of the inverse of the probability as the way to translate probabilities into information. The logarithm allows the decomposition of the multiplicative probabilities into additive information content.

If we have a set of *n* events, one of which we are certain is going to occur, and each with a probability x_i of occurring, then $\sum_{i=1}^{n} x_i = 1$ and the expected information content is given by Shannon's measure:

[1] $$H = \sum_{i=1}^{n} x_i \log \frac{1}{x_i}$$

The information content is zero when one of the events has probability 1; we draw no information from the occurrence of an event we are sure is going to happen. The information content is maximum when $x_i = \frac{1}{n}, i = 1, \ldots, n$; in this case $H = \log n$. In other words, maximum information is derived from the occurrence of one event in a context of maximum uncertainty. To borrow from thermodynamics, maximum information is derived from a state of maximum disorder, or maximum entropy. This is the reason why entropy is used as a synonym of expected information.

Theil was attracted to information theory because it might lead toward a general partitioning theory. Beyond dividing certainty (probability 1) into various uncertain probabilities, information theory presented an opportunity to devise measures for the way in which some set is divided into subsets. Theil considered it natural to apply information theory to the partitioning of overall income throughout the taxpayers of a country. If we were to apply Shannon's measure directly to individual shares of income, we would have a measure of equality (recall that the maximum of Shannon's measure occurs when all the shares are equal). Therefore, Theil proposed to subtract Shannon's measure from log *n*, leading to his well-known measure of *in*equality:

[2] $$T = \frac{1}{n} \sum_{i=1}^{n} r_i . \log r_i$$

[*] Adapted from Conceição and Galbraith (2000).

Where r_i is the ratio between individual income (y_i) and average income (μ_Y):

$r_i = \frac{y_i}{\mu_Y}$, $\mu_Y = \frac{\sum_{i=1}^{n} y_i}{n}$. The value of the Theil index (T index) is a monotonically increasing measure of inequality in the distribution of income, bounded by $T \in [0, \log n]$.

Theil argues that the fact that T does not have an upper bound but depends always on population size is desirable. Consider a society with only two individuals in which one earns all the income. In this case, $T = \log 2$. Next, consider another society in which all the income is again concentrated in one person, but the overall population is now one thousand. In this case, $T = \log 1000$, a much higher value as desired in a much more unequal society.

Consider now a different situation: if the division of income in this larger society were in the same proportion as in the first (half of the population having all the income), then we would have again $T = \log 2$ for the larger society, as is to be expected. In general, Theil showed that $T = \log \frac{1}{\theta}$, in which θ is the proportion of the population having all the income (1/2 in our last example). This is independent of the size of the population.

This feature of the Theil index is important in the context of the report because the value of the Theil index depends on the size of the population. Since we are considering counties with fairly different number of plants, the comparison of the Theil values should be made having in mind that the values of the Theil index that we are reporting are not relative, in the sense that the upper bound differs according to the number of plants being considered. We chose not to normalize the Theil values to include in the inequality measure the added information associated with the size of the population, in the context described in the previous two paragraphs.

Theil's measure has all of the desirable properties of an inequality measure: it is symmetric (invariance under permutations of individuals), replication invariant (independent of population replications), mean independent (invariant under scalar multiplication of income), and satisfies the Pigou-Dalton property (inequality increases as a result of a regressive transfer). It is also Lorenz-consistent, meaning that it agrees with the quasi-ordering that can be derived from comparing Lorenz curves.

An important characteristic of entropy-based indexes such as the Theil index is that they are decomposable. If individuals are grouped in a mutually exclusive, completely exhaustive way, overall inequality can be separated into a between-group component and a within-group component. If we consider that the population is divided into m groups, $g_1, g_2, \ldots, g_m$, each with n_j individuals, $j=1, \ldots, m$, then the decomposition takes the self-similar form of a fractal:

[3] $$\begin{cases} T = \sum_{j=1}^{m} p_j R_j \log R_j + \sum_{j=1}^{m} p_j R_j T_j \\ \\ T_j = \frac{1}{n_j} \sum_{i \in g_j} r_i \log r_i \end{cases}$$

The population proportion in each group is represented by $p_j = \frac{n_j}{n}$ and the ratio of average group income to overall average income by $R_j = \frac{\mu_j}{\mu_Y}$.

There are several reasons why it may be of interest to have a decomposable measure of inequality. One might be interested in analyzing the functional distribution of income according to some criterion that divides the overall population into groups. Examples are race, gender (both of which were explored by Theil in 1967), education level, economic sector, age, to name a few. Another reason might be associated with geography (different regions, like, say, states or countries, which were explored also by Theil in 1967). Another possibility is study differences in urban vs. rural populations. Yet another reason may be related to the differentiation of sources of income.

A further important motivation, again recognized by Theil himself, is associated with data. Data on income is often reported in income brackets, which do not give information on what is the distribution of income *within* the income bracket. Theil explored how the decomposition properties of the *T* index might help in devising measures of inequality not based on percentiles.

Appendix 2. Extremes of Equality and Inequality at the County Level in the US by State

Table 9- Most Unequal (Left) and Most Equal (Right) Counties in the US by State

Most Unequal 1967		Most Unequal 1992		Most Equal 1967		Most Equal 1992	
Alabama				**Alabama**			
Marengo	0.09	**Shelby**	**0.19**	**Bibb**	**0.006**	Geneva	0.025
Escambia	0.08	Russell	0.15	Conecuh	0.008	Pike	0.036
Russell	0.08	Marengo	0.15	**Lamar**	**0.008**	**Winston**	**0.042**
NonARC Counties	0.07	NonARC Counties	0.14	Geneva	0.010	**Pickens**	**0.042**
Dallas	0.07	Clarke	0.14	**Elmore**	**0.012**	Coffee	0.045
Georgia				**Georgia**			
NonARC Counties	0.11	**Whitfield**	**0.18**	Tattnall	0.007	Hancock	0.005
Haralson	**0.09**	Washington	0.14	**Gilmer**	**0.007**	Randolph	0.014
Glynn	0.07	Thomas	0.13	Hancock	0.007	**Walker**	**0.014**
Lowndes	0.07	Greene	0.13	Greene	0.008	Johnson	0.021
Jackson	**0.06**	**Barrow**	**0.13**	Baldwin	0.008	**Hart**	**0.021**
Kentucky				**Kentucky**			
ARC Counties	**0.09**	Muhlenberg	0.17	Grayson	0.009	**Adair**	**0.020**
NonARC Counties	0.09	**ARC Counties**	**0.13**	**Adair**	**0.011**	**Casey**	**0.020**
Boone	0.07	Warren	0.12	**Monroe**	**0.012**	**Laurel**	**0.027**
Christian	0.07	Franklin	0.12	**Wayne**	**0.013**	**Monroe**	**0.028**
Pike	**0.06**	NonARC Counties	0.11	Simpson	0.013	Nelson	0.029
Maryland				**Maryland**			
Somerset	0.17	**Allegany**	**0.09**	Baltimore	0.014	Kent	0.023
Talbot	0.11	Baltimore	0.08	Howard	0.022	Carroll	0.034
Dorchester	0.08	Montgomery	0.08	**Allegany**	**0.026**	Dorchester	0.036
Charles	0.06	Prince George's	0.08	Anne Arundel	0.026	**Garrett**	**0.036**
Cecil	0.06	Harford	0.07	Carroll	0.027	Talbot	0.037
Mississippi				**Mississippi**			
NonARC Counties	0.09	Franklin	0.16	Scott	0.005	Montgomery	0.021
ARC Counties	**0.06**	Marion	0.14	**Itawamba**	**0.008**	Panola	0.025
Jones	0.06	**Monroe**	**0.13**	Leflore	0.013	Leflore	0.028
Forrest	0.06	Pike	0.11	Washington	0.013	**Prentiss**	**0.036**
Harrison	0.05	Copiah	0.10	**Noxubee**	**0.014**	Grenada	0.037
New York				**New York**			
New York	0.09	New York	0.26	**Schuyler**	**0.007**	**Schoharie**	**0.002**
Schoharie	**0.07**	Westchester	0.18	**Delaware**	**0.010**	Wyoming	0.025
Suffolk	0.07	Richmond	0.11	**Tompkins**	**0.013**	Putnam	0.029
Bronx	0.07	Kings	0.11	**Cortland**	**0.018**	Lewis	0.034
Saratoga	0.06	Queens	0.11	Genesee	0.018	Franklin	0.035
North Carolina				**North Carolina**			
Lenoir	0.10	Columbus	0.16	Cabarrus	0.005	**Rutherford**	**0.014**
Columbus	0.08	Craven	0.12	**Macon**	**0.005**	Chowan	0.020
Pamlico	0.08	Cumberland	0.11	**Cherokee**	**0.007**	**Ashe**	**0.020**
ARC Counties	**0.07**	NonARC Counties	0.11	**Rutherford**	**0.007**	**McDowell**	**0.020**
Beaufort	0.06	Pamlico	0.11	**Ashe**	**0.011**	Anson	0.021
Ohio				**Ohio**			
Clermont	**0.06**	Richland	0.10	Trumbull	0.010	**Gallia**	**0.020**
Athens	**0.06**	**Clermont**	**0.10**	Shelby	0.011	Champaign	0.021
Gallia	**0.05**	Summit	0.09	NonARC Counties	0.012	**Hocking**	**0.021**
Preble	0.05	Cuyahoga	0.09	Defiance	0.013	Madison	0.022
Hancock	0.05	**Lawrence**	**0.09**	**Holmes**	**0.014**	Putnam	0.023
Pennsylvania				**Pennsylvania**			
Schuylkill	**0.06**	**Allegheny**	**0.12**	**Beaver**	**0.010**	**Elk**	**0.028**
Bedford	**0.06**	**Susquehanna**	**0.10**	**Elk**	**0.011**	**Mifflin**	**0.034**
Carbon	**0.06**	Lehigh	0.10	**Sullivan**	**0.017**	**Sullivan**	**0.034**
Philadelphia	0.06	**Greene**	**0.10**	**Juniata**	**0.018**	**Union**	**0.035**
Snyder	**0.06**	Northampton	0.10	**Mercer**	**0.018**	**Mc Kean**	**0.037**
South Carolina				**South Carolina**			
Clarendon	0.07	Saluda	0.19	**Pickens**	**0.009**	Chesterfield	0.025
Charleston	0.07	Dorchester	0.13	Horry	0.012	Laurens	0.028
Darlington	0.06	Darlington	0.12	Newberry	0.014	**Cherokee**	**0.038**
Dorchester	0.06	Bamberg	0.11	Allendale	0.014	Sumter	0.038
Florence	0.06	NonARC Counties	0.10	**Cherokee**	**0.015**	**Oconee**	**0.043**
Tennessee				**Tennessee**			
ARC Counties	**0.09**	NonARC Counties	0.13	Carroll	0.006	Marshall	0.013
McMinn	**0.08**	**Campbell**	**0.10**	Dickson	0.010	Tipton	0.023
NonARC Counties	0.07	**ARC Counties**	**0.10**	**Warren**	**0.010**	Lauderdale	0.025
Madison	0.05	Carroll	0.10	**White**	**0.010**	Dyer	0.025
Putnam	**0.05**	**Unicoi**	**0.10**	**Campbell**	**0.014**	Madison	0.038
Virginia				**Virginia**			
Northumberland	0.12	Fairfax	0.15	**Floyd**	**0.003**	Essex	0.012
Westmoreland	0.09	Prince Edward	0.14	Louisa	0.006	Newport News	0.018
Arlington	0.08	Amherst	0.14	Page	0.007	King and Queen	0.018
Hampton	0.07	Pulaski	0.13	Brunswick	0.007	Harrisonburg	0.021
Lancaster	0.07	Arlington	0.11	Prince Edward	0.008	Caroline	0.021
West Virginia				**West Virginia**			
Barbour	**0.06**	**Barbour**	**0.11**	**Marion**	**0.012**	**Ritchie**	**0.031**
ARC Counties	**0.06**	**Logan**	**0.11**	**Lewis**	**0.012**	**Cabell**	**0.036**
Berkeley	**0.05**	**ARC Counties**	**0.10**	**Logan**	**0.013**	**Harrison**	**0.038**
Randolph	**0.05**	**Ohio**	**0.09**	**Kanawha**	**0.014**	**Webster**	**0.039**
Wayne	**0.04**	**Wood**	**0.09**	**Mineral**	**0.014**	**Marion**	**0.040**

Note: ARC counties are in bold.

Appendix 3. Data Sources and Description

The data used in this paper are from the Longitudinal Research Database (LRD). The LRD is composed of the Census of Manufactures from each year it was collected starting in 1963 (63, 67, 72, 77, 82, 87, and 92). The data items available on the LRD include employment, wages, output, capital, material inputs, among other. The plants on the LRD are assigned unique identifiers that allow the plants to be tracked over time, creating a panel data set. For more information on the LRD, see McGuckin and Pascoe (1988).

<u>LRD Coverage by Two-digit Major Industry Groups based on the Standard Industrial Classification of 1987</u>

SIC 20	Food and kindred products
SIC 21	Tobacco products
SIC 22	Textile mill products
SIC 23	Apparel and other finished fabric products
SIC 24	Lumber and wood products
SIC 25	Furniture and fixtures
SIC 26	Paper and allied products
SIC 27	Printing, publishing and allied industries
SIC 28	Chemicals and allied products
SIC 29	Petroleum refining and related industries
SIC 30	Rubber and miscellaneous plastic products
SIC 31	Leather and leather products
SIC 32	Stone, clay, glass and concrete products
SIC 33	Primary metal industries
SIC 34	Fabricated metal products, except machinery and transportation equipment
SIC 35	Industrial and commercial machinery and computer equipment
SIC 36	Electronic and other electrical equipment & components
SIC 37	Transportation equipment
SIC 38	Measuring, analyzing and controlling instruments; photographic, medical and optical goods; watches and clocks
SIC 39	Miscellaneous manufacturing industries