Graphic Detail

Geographic Information Systems organize and clarify the patterns of human activities on the earth's surface and their interaction with each other. GIS data, in the form of maps, can quickly and powerfully convey relationships to policymakers and the public. This department of Cityscape *includes maps that convey important housing or community development policy issues or solutions. If you have made such a map and are willing to share it in a future issue of* Cityscape, *please contact david.e.chase@hud.gov.*

Visualizing Racial Segregation Differently: Exploring Geographic Patterns in Context

Ronald E. Wilson
U.S. Department of Housing and Urban Development

Visualizing geographical patterns of racial segregation is often done by mapping a proportion of a single racial group. The single proportion method, however, does not provide a context for understanding the social or economic conditions that interact with the pattern. This article is the second of two that examines segregation at the regional level. The previous article (Wilson, 2011) shows how to map two racial groups simultaneously to provide a comparative context for integration and regional segregation. The purpose of the analysis in this article is to move the reader beyond examining segregation with a single percentage map of one racial group without some comparative context. Not providing a comparison allows a reader to be misguided as to whether real problems exist. This article recasts the analysis of segregation to the interaction of the economic context with geographic patterns of segregation.

The data used in this analysis are from the 2009 Bureau of Economic Analysis (BEA) Local Personal Income Estimates.[1] An Income Inequality Index (III) is derived using three sources of income, which are earned wages, income maintenance, and unemployment insurance. Combined with measures of segregation, these categories can reveal changes in regional segregation patterns

---

[1] For the entire United States, 54 jurisdictions were missing from the original BEA data, all of which were in the Commonwealth of Virginia. Missing data were the result of several small independent cities—fewer than 100,000 residents—being combined with surrounding counties.

within the context of income disparities. III is multiplied with the Localized Diversity Index (LDI)[2] that was used in the analysis in the previous article to measure levels of segregation and integration within a county.

Personal income estimates are used for measuring disparities in sources of aggregate income through proportional balances. Categorized income sources represent various aspects of workforce abilities, education, job opportunities, and wage levels within a county. Net earnings represent income from wages that come from being directly employed, less contributions from other sources. Income maintenance represents supplemental government assistance toward living expenses. Unemployment insurance represents supplemental government assistance as compensation for a recent job loss. The latter two categories represent a general form of government assistance for individuals experiencing economic difficulties. Proportional measuring of all three categories situates populations with respect to the prevailing economic conditions in a county.

Proportions of income sources create a normalized context of income sources to determine whether the proportions are balanced. Net earnings will always be the dominant source of income in any county, because more adults are working than not in any local economy. Increasing proportions of income from government assistance through income maintenance or unemployment insurance, however, is indicative of negative economic conditions. As such, income disparities represent the local and regional economy through what proportion of the population is receiving how much income from what source.

In this analysis, Theil's Entropy Index[3] is used again to create the III because the mathematical properties are sensitive to disproportionality changes between two or more categories, as well as substantive meanings regarding proportionate sources of income. With net earnings always being the dominant source of population income, the proportion from government assistance will be small, and this index will be able to detect the changes between proportions. Resulting values have the same boundaries of 0 and $\ln\left(\frac{1}{\pi}\right)$ as the LDI from the previous article, allowing the two indexes to be combined. Although the application of this formula is similar to the LDI in its range of values, the upper and lower limits have different qualitative meanings. III values close to 0 represent populations whose dominant source of incomes is from employment (net earnings) and who are receiving little in government assistance. III values substantially above 0 represent populations who receive a larger share of income from government assistance in income maintenance and unemployment insurance. Each county will have an III value that can be thematically mapped to reveal patterns of where the poorest and richest areas of people are regarding the balance of income sources that populations in a county receive. Exhibit 1 shows the geographic patterns of III.

Counties in white have III values within one standard deviation greater or lesser than the mean. These counties exhibit the expected balance between income from net earnings and government assistance. Counties shaded in light gray have significant proportions of their populations receiving income from employment and receiving little income from government assistance. Counties
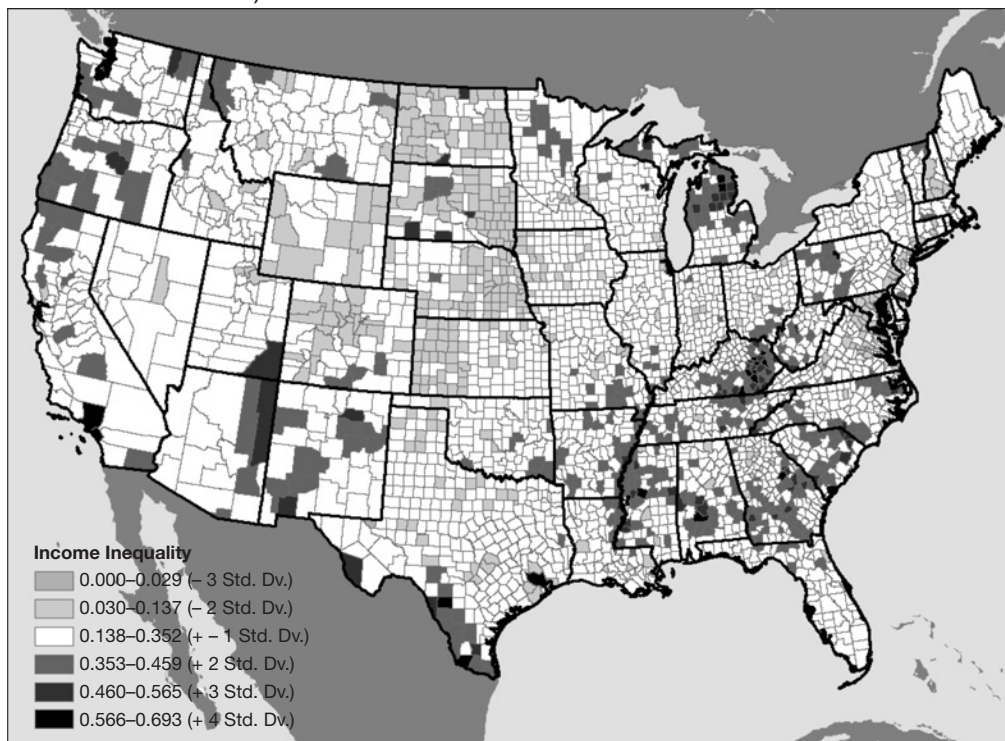
---

[2] The LDI was created using the population estimates from the U.S. Census Bureau American Community Survey 2009 5-year estimates.

[3] For the mathematical details of Theil's Entropy Index, see the appendix.

Exhibit 1

Distribution of the Income Inequality Index for the Contiguous 48 States (equal interval classification)

Income Inequality
- 0.000–0.029 (– 3 Std. Dv.)
- 0.030–0.137 (– 2 Std. Dv.)
- 0.138–0.352 (+ – 1 Std. Dv.)
- 0.353–0.459 (+ 2 Std. Dv.)
- 0.460–0.565 (+ 3 Std. Dv.)
- 0.566–0.693 (+ 4 Std. Dv.)

shaded in dark gray have significant proportions of their populations receiving income from government assistance. As the map shows, rural areas in Appalachia, the Deep South, the Northwest, and Northern Michigan often display the highest shares of income maintenance and unemployment benefits. Exhibit 2 shows that the economic conditions reflected in one county's I score are usually shared with its neighboring counties.

The thematic map in exhibit 1 indicates the clustering of income inequality, but it does not depict the extent of the geographic connection between and across those counties. Places that share similar geographies usually share similar histories in social, political, and economic conditions (Novotnỳ, 2007), one of which can be income disparity. Places are not isolated from the conditions and changes occurring in surrounding geographies. Counties in proximity to each other form interactions between themselves and establish strong economic ties that form regions. Interactions between counties in proximity to each other may exacerbate any number of local problems that affect an entire region. As such, income inequality levels are not just a reflection of a single county. These interactions require an analysis that considers the spatial relationship of income inequality

Exhibit 2

Spatial Autocorrelation of the Income Inequality Index for the Contiguous 48 States (Moran's I scatter plot)
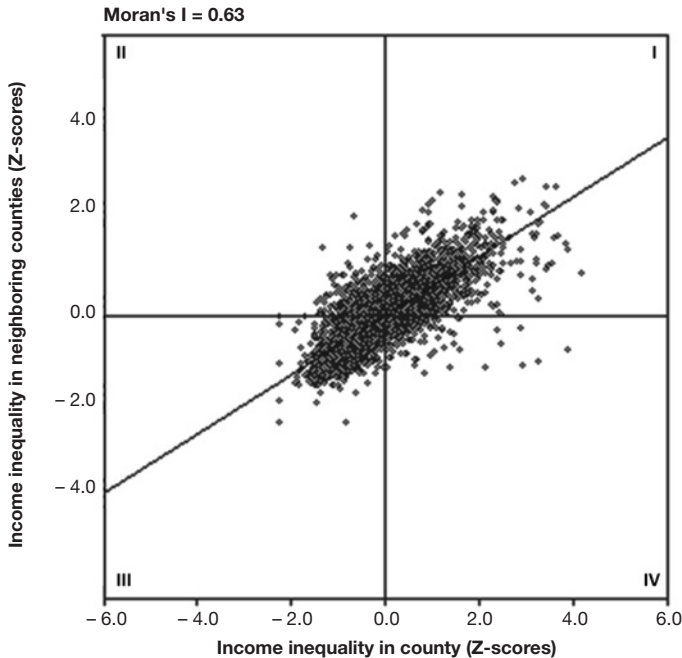
with adjacent counties.[4] To measure the strength of these spatial relationships, the Moran's I Index is used to detect significant regional spatial relationships of income inequality and racial segregation between counties.[5]

The Moran's I value of 0.63 (with a p-value of 0.001)[6] indicates the covariation patterns of III values between a county and its adjacent counties show strong regional clustering of high and low III values.[7] The data cloud in exhibit 2 depicts a strong linear pattern across quadrants I and III. Those two quadrants represent counties that are clustered together with similar values on opposite ends of the III. Counties in quadrant I are those that are surrounded by counties with similarly high

---

[4] The interactions between counties represent the First Law of Geography. For a discussion of this law and its importance, see Miller (2009).

[5] Neighboring counties were identified using the eight nearest neighbors. Nearest neighbors were used instead of adjacency to eliminate the effects of large and irregular shaped counties not capturing full regional relationships.

[6] Moran's I values closer to 1.0 indicate clustering. Values closer to -1.0 indicate dispersion.

[7] The strength of the Moran's I value is contingent on two elements. First, the steepness in the slope of the line between values in a county with the lagged values in neighboring counties. The steeper the slope, the stronger that spatial relationship is between a particular value in a county with its neighboring counties. Second is the p-value that measures how different the observed I value and a set of randomization I values are from the original data. P-values less than 0.05 indicate the observed pattern is statistically different from the random distribution of I values, which usually forms a normal distribution. The steeper the slope and the higher the p-value, the stronger the spatial relationship is between counties and neighboring counties.

values. Of the counties in quadrant I, 23 percent show statistically significant clustering of high III values with adjacent counties. Of quadrant I counties, 25 percent show statistically significant clustering of low III values with adjacent counties. This clustering supports exhibit 1 in which counties with plus or minus 2 standard deviations from the mean are clustered together in different regions of the United States. These counties share similarly high levels of income inequality through a regional connection. Regions in exhibit 1 that are two standard deviations or greater than the mean are grouped together in the most rural parts of the United States. Rural counties often have fewer economic opportunities to develop robust job markets. Regions with counties two standard deviations or below are those in which the populations are receiving little in government assistance and most income is from net earnings. These counties are located either in the heartland, where a lot of farming activity occurs, or in major urban areas that have diversified and robust economies.

Examining income inequality alone, however, does not answer questions regarding any regional relationship with segregation. Exhibit 2 shows an association with lack of economic opportunity in areas not known for robust economies, regardless of race. The question remains, does a relationship exist between the counties with high levels of income disparities and racial segregation? A new variable is created to measure the interaction between income inequality and racial diversity (*Interaction (I) = III · LDI*) where LDI and III are as defined previously. This interaction depicts a bivariate relationship that identifies counties with high and low proportions of the Black population compared with the White population. The index extremes are substantively opposite of each other regarding the positive and negative connotations. In interpreting the interaction, recall that an LDI value close to 0 signifies segregation, and a value closer to 1 signifies integration. In contrast, III values closer to 0 signify net earnings as the dominant source of income in a county; substantially higher values indicate that government supplementation of income makes up a larger proportion of income sources in the county. High values of the interaction indicate a county with a high proportion of the Black to White population (racially integrated) with a significant portion of income from government assistance greater than earned income. Low values indicate either a highly segregated county, a county with most income coming from net earnings rather than government assistance, or both.
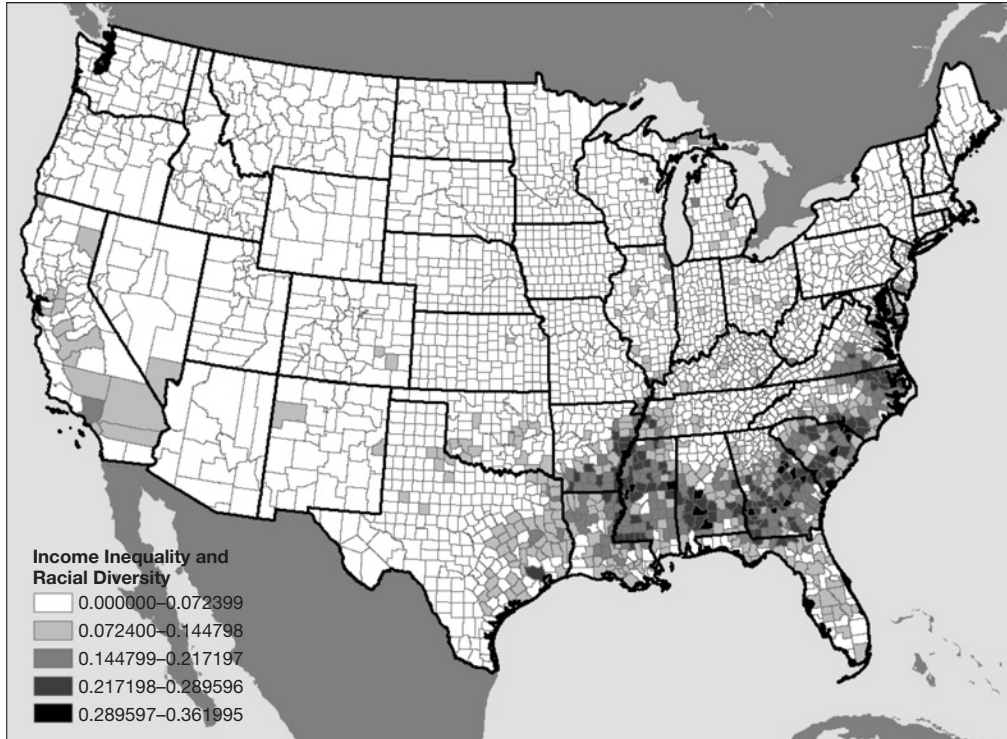
Exhibit 3 displays the map of the interaction between the two indexes. In contrast with exhibit 1, only one dominant regional pattern now exists. A belt of counties from east Texas to northeast North Carolina forms the dominant regional pattern. A lesser cluster of counties in California (mostly rural and in the interior, apart from Los Angeles County) also stands out. The latter areas have a high proportion of the Black population but far less than areas east of the Mississippi River.

As with the III, the strength of this spatial relationship needs to be measured with the Moran's I Index to detect significant regional spatial relationships of this bivariate relationship for regional income inequality and racial segregation.

Exhibit 4 displays the degree to which the interaction between the two indexes is spatially correlated. The data cloud in exhibit 4 has changed considerably and the Moran's I value (0.78 with a p-value of 0.001) is substantially closer to 1 than the III value. The covariation of interaction values between a county and its adjacent counties shows an even stronger pattern of regional clustering with the high and low LDI and III values than in exhibit 2. The tight clustering of most Z-scores near 0 is an artifact of index construction.

Exhibit 3

Distribution of the Interaction of Income Inequality and Localized Diversity Indexes
for the Contiguous 48 States (equal interval classification)

**Income Inequality and
Racial Diversity**

- 0.000000–0.072399
- 0.072400–0.144798
- 0.144799–0.217197
- 0.217198–0.289596
- 0.289597–0.361995

The data cloud still shows a strong linear association across quadrants I and III, but the patterns
have changed from those in exhibit 2. First, the slope is steeper. The spatial relationship between
counties and neighboring counties has gotten stronger. Second, the patterns in quadrants I and
III have distinctly changed. In quadrant I, the pattern is more dispersed, which signifies counties
with high values are clustering in small pockets of high values. Of these counties, 19 percent show
a statistically significant clustering of high values with adjacent counties, which is a decrease of
4 percentage points from exhibit 2. These changes reveal a strong local connection between income
inequality and racial segregation amidst a larger regional pattern. The geographic patterns from
exhibit 3 support these results because the highest proportions of income inequality and Black to
White populations are concentrated and are separated from each other—forming a loosely coupled
belt across the South. The pattern in quadrant III shows the opposite, with a more concentrated
pattern, which signifies these counties have low values and are the dominant spatial pattern. Of
these counties, 37 percent show a statistically significant clustering of low values with adjacent
counties, which is an increase of 11 percentage points from exhibit 2. These changes indicate that
a substantial number of highly segregated counties are clustered together with populations whose
income is primarily from net earnings. Exhibit 4 reveals this clustering in a pattern across the
southern states from Arkansas to North Carolina. Not only does a strong level of succinct regional
clustering exist, but the pattern forms one long belt of clusters in one part of the United States. Of

Exhibit 4

Spatial Autocorrelation of the Interaction for Counties in the Contiguous 48 States (Moran's I scatter plot)
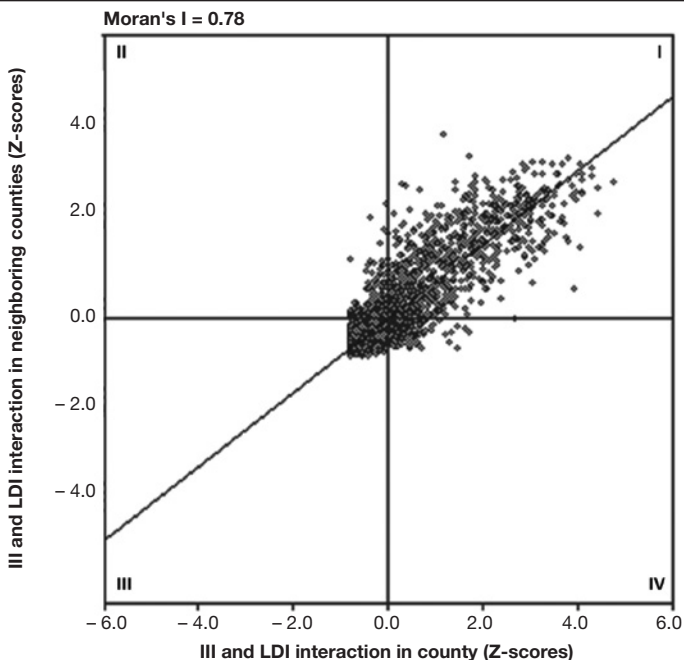
III = Income Inequality Index. LDI = Localized Diversity Index.

the southern states from Arkansas and Louisiana to Virginia, 65 percent of the area has moderate to high levels of the population receiving higher levels of income from government assistance and also higher proportions of the Black population receiving such assistance. Compared with the United States, these same counties make up 9 percent of the area. With this pattern concentrated into one part of the United States, 9 percent is a significant amount, because progress in the region is likely weighed down by a collective lack of opportunity.
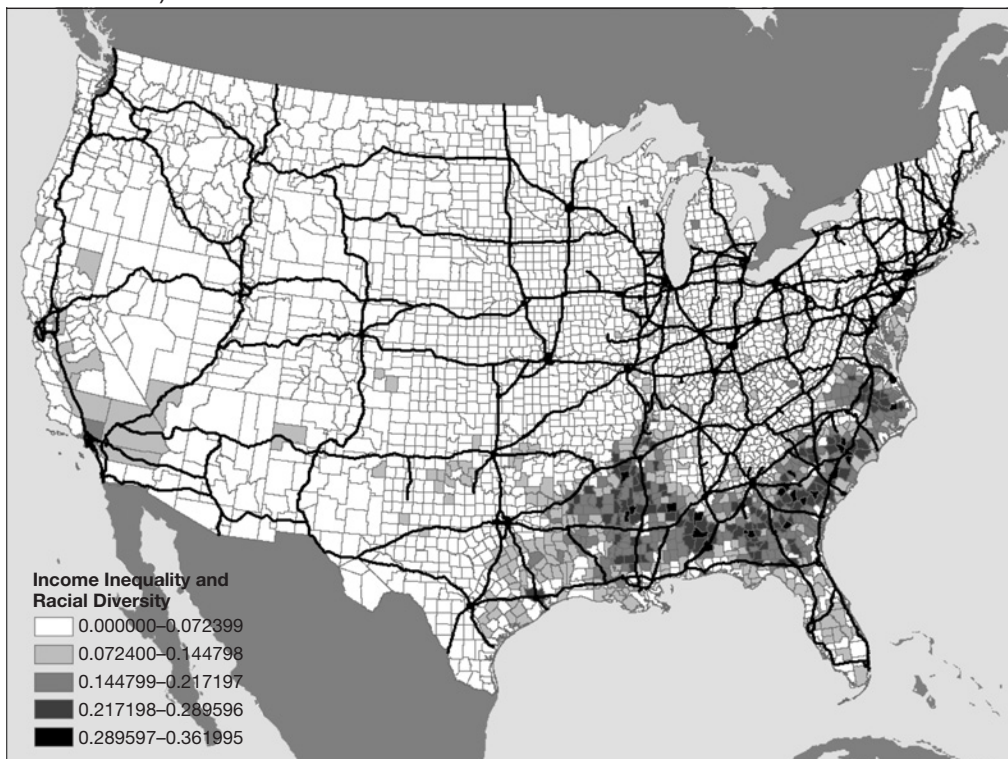
To confirm the spatial relationship that regional segregation has with income inequality, statistically significant values were removed from the scatter plot in exhibit 4 to examine the change in overall spatial relationships across the United States. When statistically significant high values are removed from the scatter plot, the Moran's I Index drops from 0.78 to 0.48. This substantial decrease demonstrates the level at which greater proportions of income inequality and Black populations affect the analysis results. These counties face specific challenges in overcoming income inequality regarding segregation that is much different from the rest of the United States. When counties with statistically significant low values were removed from the scatter plot, the Moran's I dropped to only 0.71. The slight decrease confirms that the dominant spatial pattern in the United States is of counties with low income inequality and low regional segregation and is more common than the opposite. Any challenges these counties have regarding income inequality are not associated with regional segregation.

Exhibit 5 puts regional patterns of income inequality and segregation into a final context. Overlaid with the thematic map in exhibit 3 are major U.S. highways. Highways can be lifelines for regional economic development because they provide efficient transportation of goods and services. Highways facilitate industries to locate in otherwise inaccessible areas and bring jobs and growth that work toward building robust economies. Without highways, transportation times and costs can become excessive, which leaves little incentive for economic opportunities to be created. As can be seen in exhibit 5, counties with the highest levels of income inequality and regional segregation are located in areas that are deep between these highways. This lack of access to major roadways is indicative of rural areas with little economic opportunity.

The analysis in this article reveals regional patterns of greater income inequality associated with racial segregation. The patterns are more succinct than single-race thematic map patterns and can guide place-based thinking with respect to applying policy in a geographical manner. Thematic maps using a single race or ethnicity do show patterns of segregation but do little to identify where help is needed most because the breadth of the problem is exaggerated on the map at the expense of depth. In the analysis presented here, patterns are much more succinct in identifying specific pockets of income inequality (depth) that have been identified within a larger pattern of

Exhibit 5

Distribution of the Combined Income Inequality and Localized Diversity Indexes for the Contiguous 48 States With an Overlay of Major Highways (equal interval classification)

Income Inequality and Racial Diversity

- 0.000000–0.072399
- 0.072400–0.144798
- 0.144799–0.217197
- 0.217198–0.289596
- 0.289597–0.361995

regional segregation. The most disadvantaged counties in rural areas of the South, with respect to segregation and sources of income, are highlighted in this analysis. A number of additional reasons for these patterns exist, beyond the association of rural areas with income inequality and racial segregation. But, the point of this article and the previous article is to spur analyses of race or ethnic data within a context. Other data sources should be used to conduct more combinatorial-based inquiries. Single-race thematic maps show nothing regarding why any pattern exists and leave readers unsure of which factors might be associated with segregation. Some counties in the regional belt across the South do not have high income inequality but have high proportions of the Black population. Many of those counties are urban and have different needs regarding economic policies and opportunities than rural areas. The focus on the southern belt of counties identified in this analysis does not mean other rural regions of the United States need less attention regarding economic development. Unlike other rural parts of the United States, counties in the South may need specific policies to specifically address regional segregation and expand economic development opportunities simultaneously. A third article will look at business pattern data, with the results from this analysis providing a better understanding of the current economic opportunities in this region.

Appendix

Inequality Index (Theil's Entropy Index)

Theil's Entropy Index is formally defined as

$$Inequality_j = \sum_{j=1}^{J} \pi_j \ln\left(\tfrac{1}{\pi_j}\right) \tag{1}$$

where j is a jurisdiction and π_j is the proportion of each income source (net earnings, income maintenance, and unemployment insurance) in jurisdiction j.[8] The result is I, which is a continuous value that ranges between 0 (if one income source is the only income source) and $\ln\left(\tfrac{1}{\pi}\right)$ (if all three income sources are equal in shares; this value is 3). The result is a nonlinear (logarithmic) curve that represents change effects in which each unit increment or decrement along the scale has a different rate and magnitude of change.[9] Values substantially larger than 0 indicate net earnings are the dominant source of income and are from employment. Values closer to 0 indicate net earnings are less dominant as the primary source of income and income maintenance and unemployment insurance (government assistance) make up a larger proportion of sources of income.

[8] For a review of the mechanics of this index, see Wong (2003).

[9] For additional details about how this index functions, see Wilson (2011).

Frequency Distributions

Each figure number corresponds with the respective map number in the main body of the article.
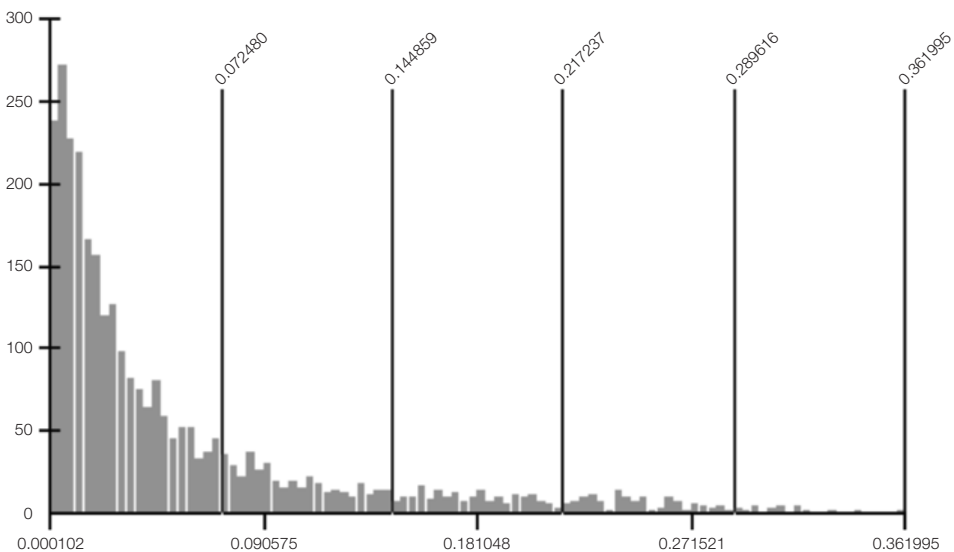
Exhibit A-1

Distribution of the Income Inequality Index for the Contiguous 48 States
(equal interval classification)
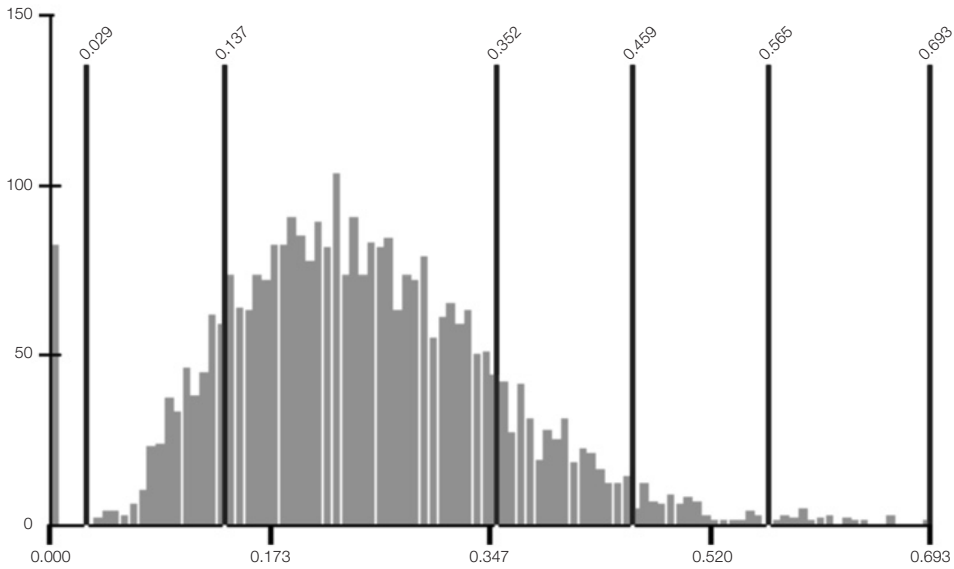
Exhibit A-3

Distribution of the Interaction Between the Income Inequality and Localized Diversity
Indexes for the Contiguous 48 States (equal interval classification)

Acknowledgments

Author

Ronald E. Wilson is a social science analyst in the Office of Policy Development and Research at the U.S. Department of Housing and Urban Development.

References

Miller, Henry J. 2004. "Tobler's First Law and Spatial Analysis," *Annals of the Association of American Geographers* 94 (2): 284–289.

Novotnỳ, Josef. 2007. "On the Measurement of Regional Inequality: Does Spatial Dimension of Income Inequality Matter?" *Annals of Regional Science* 41: 563–580.

Wilson, Ronald E. 2011. "Visualizing Racial Segregation Differently—Exploring Changing Patterns From the Effect of Underlying Geographic Distributions," *Cityscape* 13 (2): 163–174.

Wong, David W.S. 2003. "Spatial Decomposition of Segregation Indices: A Framework Toward Measuring Segregation at Multiple Levels," *Geographical Analysis* 35 (3): 179–194.