The Candidate Indicators: General Issues in Selection and Evaluation

The methodology for selecting the candidate indicators follows the principles delineated in the team's April 2008 report. These principles can be summarized as follows:

- Selection of indicator variables should be based on credibility, transparency, and acceptability to the ARC, Congress, and OMB. Moreover, we follow the current ARC practice of ranking the variables relative to the rest of the country when determining the particular category of distress, an approach is that likely to prove more defensible in the eyes of external observers.

- The component indicator variables selected for each index should be as up-to-date or produced with as little time lag as possible.

- The subsequent index constructed from the indicators variables should be accurate in terms of capturing economic distress in the ARC counties.

Based on the objectives outlined in the project proposal, our *first* priority is to consider forward-looking measures (such as population growth and educational attainment). Another priority is the inclusion of new labor market measures, such as the employment-to-population rate, as potential replacements for the unemployment rate. As a *secondary* priority, we consider other measures outlined in our April 2008 report as being potentially useful. For example, we examine how measures of building activity and expansion of bank branch offices can be tapped as potential forward-looking measures. Overall, as we explain below, we select and evaluate a total of 40 different indicators within seven specific groupings, as prospective variables that could be used to construct a new distress index.

Sources of Data and Issues in Constructing Indicators

One of the traditional critiques of the ARC distress indicators is that some of the data are typically drawn from the decennial Census, which means that these variables become rather dated as the end of a decade draws near. This major shortcoming is expected to be overcome with the planned full implementation of the American Community Survey (ACS). Officially, the Census Bureau still maintains that by late 2010, the ACS will produce *annual* data that will capture all of the elements of the decennial Census long-form at a disaggregated geographical level including county, census tract, five-digit zip code, and census block group (though there are some restrictions for small samples).[1] Because the ACS data are derived using sampling, the accuracy will be reduced for less-populated counties. Likewise, the data will not be perfectly up-to-date since a five-year moving average will be used for less-populated geographical locations—though that has the offsetting advantage of averaging out year-to-year fluctuations (as is currently the case in the housing market) that do not capture long-term distress. However, we

[1]For more details of current Census plans for the ACS, see U.S. Census Bureau (2009) and Office of Management and Budget (2009).

believe that the ACS will produce reasonably accurate estimates based on its sampling procedures. Thus, despite the shortcomings of the ACS, we view the annual availability of data for less-populous counties to be a significant improvement for calculating future ARC distress indicators.

When deriving the candidate distress indices, we examine how candidate variables measured in circa 1996/1997 are associated with current distress measured in (or as close as possible to) 2006/2007. However, as described in more detail in a latter section of this report, some of our variables are estimated due to limitations in the availability of current data. For example, the 2007 employment/population ratio and the 2007 labor-force participation rate are estimated because the denominator—population 16 years and over—is not reported annually by the Census Bureau. Instead, the Census annually produces estimates of the county population 18 and over, as well as the estimated county population between 14 and 17 years of age. For 1996/1997, data to estimate the employment/population ratio or the labor force participation rate do not exist. Therefore, we substitute their respective rates from the 2000 Census. Fortunately, such interpolation will not be necessary with the ACS because it will directly report the necessary data. Thus, we believe that when ACS data is fully available, the employment/population and labor force participation variables will be even more strongly linked to distress than the numbers being used here.

Another issue in the choice of indicators involves the time-period required to capture distress. We use ten year measures of job and population growth because they capture the persistent nature of distress that we are seeking to measure. For example, short-term changes in job or population growth may reflect a temporary, cyclical event such as a plant opening or closing, or a transitory shock to the housing market. Whether such changes are indicative of a major change in the county's underlying prosperity would depend on whether the (say) newly laid off workers find work locally, or even outside of the county through commuting. Only after some period of time has elapsed would true economic distress reveal itself through these offsetting adjustments. Indeed, as the "New Economic Geography" literature, the economic impact literature, and the military-base closing literature indicate, communities that face adverse events such as a major military-base closing, natural disaster, or even an intensive military bombing exercise, generally return to their long-term growth paths within a period of about ten years (Edmiston, 2004; Poppert and Herzog, 2003; Kilkenny and Partridge, forthcoming).

It is important to note that five-year measures of job and population growth are likely to be inadequate measures of long-term distress. This conclusion is generally confirmed by Partridge et al. (2008, Chapter 6) when they found that replacing the unemployment rate by the five-year change in population would result in a relatively small number of changes in the number of counties that would be classified as being in distress compared to the current ARC distress index. Thus, we view ten-year changes in population and job growth as good measures of distress as they are long enough to balance out short-term events, but not so long as to capture trends that are not germane.

We do have specific concerns with regard to the use of poverty rate and educational attainment in any index. These variables are not currently available on annual basis for all counties. One of

the key measures that would likely be included in any distress index is the current poverty rate. Of course, a key shortcoming of using county poverty rates is that the most accurate poverty estimates are obtained through the decennial Census. Not only is this a problem for calculating annual measures of distress, but it presents problems for our study because we would prefer to have annual measures. Specifically, much of our analysis will use data from the circa 1996/1997 and the circa 2006/2007 periods. To avoid this data-availability shortcoming, we adopted the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) as our annual estimate of poverty. One weakness is that the SAIPE does not produce direct estimates of county poverty, instead relying on estimates derived from personal income and tax data. To examine whether the SAIPE data are accurate enough for our purposes, we calculated the simple correlation coefficient between the 1999 county SAIPE poverty rate and the 1999 poverty rate derived from the 2000 Census (with the 2000 Census estimate viewed as being quite accurate). The correlation was equal to 0.94. The corresponding correlation using 1989 data equaled 0.95. We view this correlation as high enough to move forward with the SAIPE estimates in our study. Again, this would imply that using the ACS data would lead to even stronger predictors of future distress than the SAIPE poverty data used in this report—i.e., the ACS would present direct measures of poverty, not the estimates used to derive the SAIPE figures.

A final key variable is educational attainment. Unfortunately, accurate county-level annual data on educational attainment are not available between the Census decades. Thus, we are forced to rely on the 2000 decennial Census educational attainment data. Again, we expect that the annual ACS data will allow for more accurate future predictors of distress because the ACS will produce measures of educational attainment.

The Set of Indicator Variables Evaluated

We select and evaluate a total of 40 different indicators, within seven specific groupings that are described in our earlier report, as potential variables that could be used to construct a new distress index. These variable groupings and the specific variables within each group are described in the remainder of this section.

1. ***Population growth*:** This indicator, including its key component of net population migration, measures an important response for dealing with economic decline in a community; residents, in effect, vote with their feet by moving away if they believe that a place offers only a bleak economic future. As such, population growth needs to be considered carefully in any redesign of the distress index. Net population loss is a real measure of economic deterioration, whereas population growth may occur either in response to economic expansion or to features such as attractive natural amenities (as in retirement migration, for example). Over longer periods, population growth also serves as a close proxy for employment change, another key economic component that is essential to capture.

   Adjustment assistance to counties experiencing significant population loss can be motivated with two key arguments: (1) those staying behind have to deal with the very real negative fiscal and other consequences of a smaller population base; and (2) assistance could stem or even reverse the net out-migration. A third, more indirect argument is that congested or

sprawling metro areas receive a public benefit when in-migration from declining areas is held in check.

We analyze the percent population growth between 1986-1996 and 1996-2006 using population estimates from the US Census Bureau (www.census.gov/popest/counties/). As elaborated above, we use the most current data available at the time of this study and we choose ten-year periods so as to even out short-term fluctuations and to focus on longer-term trends.

In addition to actual population counts, we consider the percent change in the population 20 years of age or younger (data were also obtained from the Census Bureau). The reason for singling out this age group is that it represents, in many ways, "the future" -- both metaphorically and as the concrete workforce of a community. Furthermore, research suggests that once these individuals leave, it is difficult to attract them back to the community.

2. ***Educational Attainment:*** Education is the prime measure of human capital and serves as an underlying determinant of an individual's *current* and *future* earnings capacity. Returns to education (or spillovers) are enhanced when those with a college degree can locate near other workers who also hold college degrees. Moreover, education is highly correlated with poverty and it is associated with an individual's ability to adjust to economic change, take advantage of new opportunities, and succeed more generally in the knowledge economy. For these reasons, education is connected to current distress and it is a forward-looking measure of future distress.

   Indeed, adding high school educational attainment as an indicator may be more important than adding population growth in terms of affecting county exit/entry into/from distress (see Section 6 of our April 2008 report). We consider a full set of human capital measures, including the share of the adult population (25 + years old) in 2000 with: less than a high school diploma (nursery school through grade 12 but no diploma); only a high school degree (34.7% of the total); an associate's degree or more, including BA, MA, professional, and Ph.D. degrees (22.2%); at least four years of college with degree completion (16.5%); at least one year of college (64.7%); and the percent with less than one year of college (35.3%). These data are drawn from the US Census Bureau (www.census.gov/population/www/socdemo/educ-attn.html), and they allow us to determine whether a community lacks sufficient human capital to compete in the knowledge economy.

3. ***Income:*** Per capita market income is another important measure of economic well-being or distress, and as such, we consider it carefully as a candidate in the new index (as it is currently also used in the ARC's Distress Index). Yet, as noted in our April 2008 report, lower per-capita income can reflect a higher level of amenities, lower cost of living, and other factors that are not necessarily related to distress. For these reasons, we also evaluate other measures of income and earnings, such as *changes* in per capita income between

1986-1996 as well as 1996-2006. Income data are from the Bureau of Economic Analysis, Regional Economic Information System (www.bea.gov/regional/index.htm).

As the availability of natural amenities (including a favorable climate) tends not to change over time, we can conclude that any shifts in income observed over time are due to changes in factors other than natural amenities – such as deteriorating local economic conditions. In other words, a community with low or even negative income growth rates is truly experiencing economic difficulties, and not just attracting new residents who are willing to give up income for better amenities.

4. ***Housing or Housing Change:*** Housing construction activity, coupled with changes in property values, reflects the forward-looking economic outlook of each individual county. Housing is not only a fundamental human need but has been, until just recently, viewed as an important – if not the key -- vehicle for wealth creation. Homeownership may also be a key source of social stability and social cohesion in communities.

   Changing property values denote local on-the-ground assessments of the future direction of the community by private businesses or entrepreneurs (the market). Also, changes in new home construction can provide good forecasts of future population shifts. In fact, the University of Michigan includes building permits issued as a key indicator in its national indicator of leading economic activity. Thus, innovative measures of housing conditions can be useful additions to a forward-looking index.

   We examine building permits data (both permits issued and values of the permits) from the Census Bureau (www.census.gov/const/www/permitsindex.html). In addition, we identify an equally current data series from the same source on the number of housing units in each county, for the ten-year intervals corresponding to those used for population and employment changes. Since permits reflect intended rather than actual starts, and given the upheaval in the housing market associated with the present economic collapse and concerns about the completeness of the permitting process in some ARC states, we include only the actual number of housing units (from the Census Bureau) for the periods 1990-2000 and 2000-2007 in our analysis. In fact, the housing unit measure outperforms the building permits measure on virtually all counts. The data source is censtats.census.gov/cgi-bin/usac/usatable.pl.

   In addition, we examined changes over time in the number of bank branch offices as supplemental predictors of subsequent economic growth. Banks are key sources of funds for homebuyers and other local businesses wishing to expand. Bank branch numbers are based on official counts and as such do not suffer from potential under-reporting problems (as may occur with residential building permits). However, our analysis shows that growth in bank branch offices tends to follow population expansion, rather than the other way around. Thus, it is more of a lagging indicator and thus, we choose to rely on population growth.

5. ***Entrepreneurship and Self-Employment:*** A forward-looking measure of the local community's ability to compete is the presence of strong entrepreneurial capacity. Though "entrepreneurial capacity" cannot be directly measured from any federal data source, a plausible proxy is self-employment in the form of small business ownership. Because new business formation is motivated by a host of favorable (e.g., a new idea or business opportunity) and unfavorable (e.g., a factory lays off workers) reasons, we attempt to sort out entrepreneurship of "necessity" from entrepreneurship of "opportunity" in the region. As described (section 5.2.5) of our April 2008 report, one clue about the extent to which self-employment growth in an ARC county is a response to opportunity as opposed to necessity may be found in the returns to self-employment, measured as average earnings per self-employed worker. This distinction between different forms of entrepreneurship may be an important measure of a county's ability to adjust to new employment realities.

   We obtain data on both self-employment numbers and earnings from self-employment from the Regional Economic Information System of the U.S. Bureau of Economic Analysis, for 1996 and 2006 (www.bea.gov/regional/index.htm). In order to compare counties of various sizes, we express these numbers either as a percent of the total workforce (in the case of jobs) or as a percent of total earnings (in the case of incomes). We also examine the earnings per self-employed worker in order to sort out opportunity entrepreneurship (higher earnings) from entrepreneurship of necessity (indicated by lower earnings). After careful analysis, we conclude that self-employment, while correlated with local economic strength, does not provide sufficient additional information beyond that already contained in job changes over time to warrant inclusion of this measure in a distress index.

6. ***Labor Market Strength:*** We consider several measures of labor market conditions as candidate variables for the distress indexes. As noted in our April 2008 report, the unemployment rate fails, in general, to adequately capture labor market conditions. The *employment rate* (employment rate divided by the population that is sixteen years old and over), in conjunction with annual employment growth, are better indicators of overall labor market strength. The employment rate directly captures labor-force participation, unemployment and discouraged-worker effects. It proves more informative than the conventional unemployment rate measure. Specifically, an ARC county may have low employment rates due to a combination of high unemployment rates (i.e., those non-employed individuals who are *actively* seeking work) and low labor force participation because large numbers of non-employed individuals are not working or not actively seeking work.

   In particular, we calculate the employment to population (16 years and older) ratio for 2000 and 2007 from Census Bureau data [see discussion above on data collection issues for inter-census years], the labor force participation rate in 2000 and 2007 (workers who are in the labor force relative to all those eligible) and the 1997 unemployment rate, for comparative purposes.

7. ***Poverty Rate:*** We maintain that the poverty rate should remain one of the core variables of the ARC distress index. Our investigation considers the overlap of poverty with the other

indicators included in the complete distress index, as is described in more detail further below in this report. Data are from www.census.gov/hhes/www/poverty/poverty.html.

To summarize, along with the current variables used by the ARC, our *first* priority is to consider forward-looking measures, such as population migration and educational attainment (e.g., college and high school graduates). Another priority includes the new labor market measures, such as the employment-to-population rate, as a potential replacement for the unemployment rate. In this construction, we also explore using the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) database to derive innovative measures of employment activity and job quality. However, as this data collection effort is still not complete for all of the ARC states, we are unable to conduct the same comprehensive analysis covering all geographies as we have done with the other variables. The major shortcoming of LEHD data is that it does not go back very far in time, making historical analysis impossible for our assessment. Thus, while the LEHD is a treasure trove of data, it is not useful for the historical analysis that is necessary to derive new distress indices.

As a *secondary* priority, we consider the other measures described above. In particular, for housing, we examine how measures of building activity and expansion of bank branch offices could be used in a potential forward-looking index. Here, we determine that population growth more effectively captures local conditions and that bank offices are more of a lagging than a leading indicator (in effect, population growth drives these other measures). Likewise, we consider measures of entrepreneurial activity as another forward-looking measure.

Geographical Patterns in the 2006/2007 County Patterns of Indicator Variables

Figures 2.1 to 2.8 display maps for eight indicator variables that we are evaluating. The first set of maps (labeled with an "a") is for the U.S. and the second set of maps (labeled with a "b") is specifically for the ARC region only. The maps display the variables in standardized or Z-score form. As we explain in more detail below in Section 3, the Z-score is a simple numeric transformation that makes it easier to directly compare the distributions of two different variables. For example, it is difficult to directly compare the poverty rate, which is measured as a percentage, with per capita market income, which is measured in dollars per person. With the Z-score transformation, variables that inherently vary a great deal (or statistically those with a large standard deviation), are scaled so as to make them more comparable with variables that do not vary as much from one county to the next. For example, we can directly compare unemployment rates, which may vary nationally from 4% to 18%, to poverty rates that may vary from 5% to 50%.

Figures 2.1a and 2.1b show Z-scores for the rate of population growth between 1996 and 2006. While the center of the nation experienced remarkable population loss, so did a core set of ARC counties that straddle the Kentucky, West Virginia and Virginia borders. A closer look at Figure 2.1b shows that four of these counties are in southwest West Virginia, and one each is located in Kentucky and Virginia. The important role of commuting and attendant urban sprawl is evident from the significant population expansion in eastern Pennsylvania and eastern West Virginia, as well as around the Atlanta metropolitan area. In the national map (Figure 2.1a), significant

population growth stands out in Texas in the Dallas-Ft. Worth mega-region (note the ring around the urban core), San Antonio, and Houston, as well as in selected counties of Colorado, Arizona and Nevada.

The Z-score map for the 2000 adult population share with at least one year of college (Figure 2.2a) shows considerable concentration, with high shares especially in Colorado and in a handful of counties in Wyoming, Montana and Idaho. In the eastern half of the nation, high concentrations of this variable tend to be found in or around major cities. In Figure 2.2b, for the ARC counties, higher shares of adults with at least one year of college are found in large cities or college towns (such as Centre County, PA, home to Pennsylvania State University).

Figures 2.3a and 2.3b, showing z-score maps of the proportion of adults who have completed a four-year college degree, are similar to the immediately preceding maps, and yet there are subtle and important differences. In particular, Figures 2.3a and 2.3b appear to be lighter in color, suggesting that the *relative* concentrations of adults with four years of college in certain counties tends to be greater than is the case for adults with just one year of college. Especially in Figure 2.3b, the relatively small share of adults having completed four years of college in Central Appalachia (and especially southwest West Virginia) is noteworthy.

Relatively high levels of 2006 per capita market income are evident in Figure 2.3a along the Northeast Seaboard, southern Florida, coastal California and in Wyoming. In the latter states, the high income could reflect natural resource-based activities or the high incomes of transplants who earned their wealth elsewhere – such as on Wall Street or in Hollywood. In Figure 2.4b, per capita market income tends to be high around major cities (Pittsburgh, PA and Montgomery, AL). The contiguous counties in Kentucky with low incomes stand out, and relatively strong spillover effects from Lexington (Fayette County) into the adjoining ARC counties are also evident.

A more dispersed pattern of high and low Z-scores appears in the maps showing 2006 self-employment shares of total county employment (Figure 2.5a). It should be noted that this variable captures *non-farm* self-employment or proprietorships. A few counties stand out with high rates of self employment, particularly in the central part of the nation – this is also the region with large population losses shown earlier in Figure 2.1a. For the ACR region, Figure 2.5b reveals relatively low self-employment shares in southwest West Virginia.

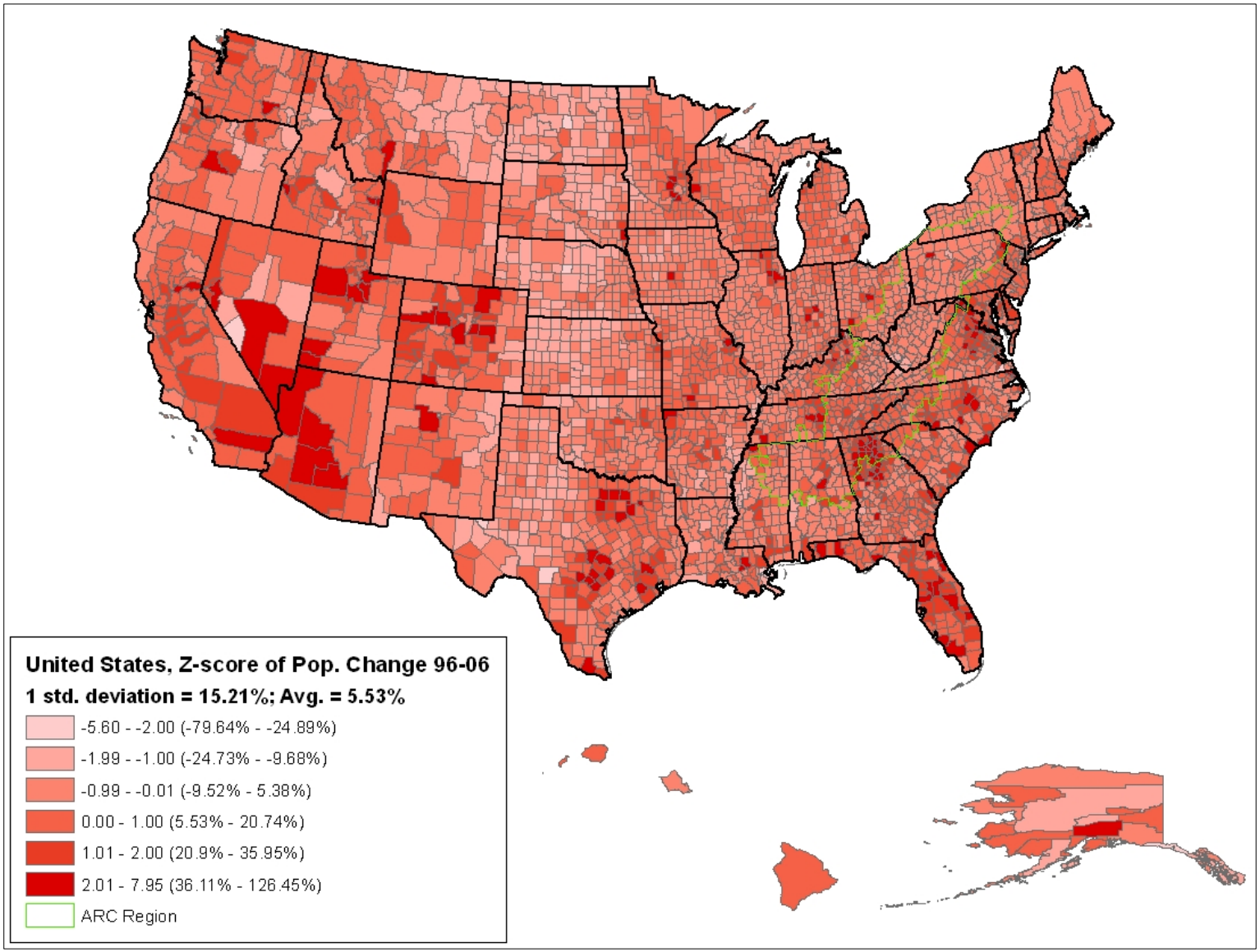
Another key variable, the 2007 employment-to-population ratio (Figure 2.6b), shows high rates in some of the Rocky Mountain states and relatively widespread occurrences of very low rates in the ARC region (as well as other southeast states). This ratio is a powerful predictor of economic well-being and, indeed, some overlap between Figures 2.6a and 2.4a (for per capita market income) is evident. In Figure 2.6b, the low employment/population ratios in some of the eastern Kentucky counties are quite noticeable.

For the unemployment rate (Figure 2.7a), a substantially different pattern emerges as compared to the employment-to-population ratio (shown previously in Figure 2.5a). This different pattern confirms the importance of evaluating the employment/population ratio as a potentially more

reliable alternative to the unemployment rate as a measure of underlying economic distress. Note especially the concentration of low unemployment rates in the northern Great Plains region and the Rocky Mountain states, and the high unemployment rates in California, which entered the current recession at a relatively early date. Michigan, South Carolina and the Mississippi Delta region also have high unemployment rates. In contrast, unemployment rates are very low in Virginia and parts of north central Alabama. In Figure 2.7b, a seemingly sharp divide along state lines appears between Kentucky and Virginia on the one hand, and between Alabama and Mississippi on the other. The counties surrounding metropolitan Atlanta also enjoy comparatively low unemployment rates.

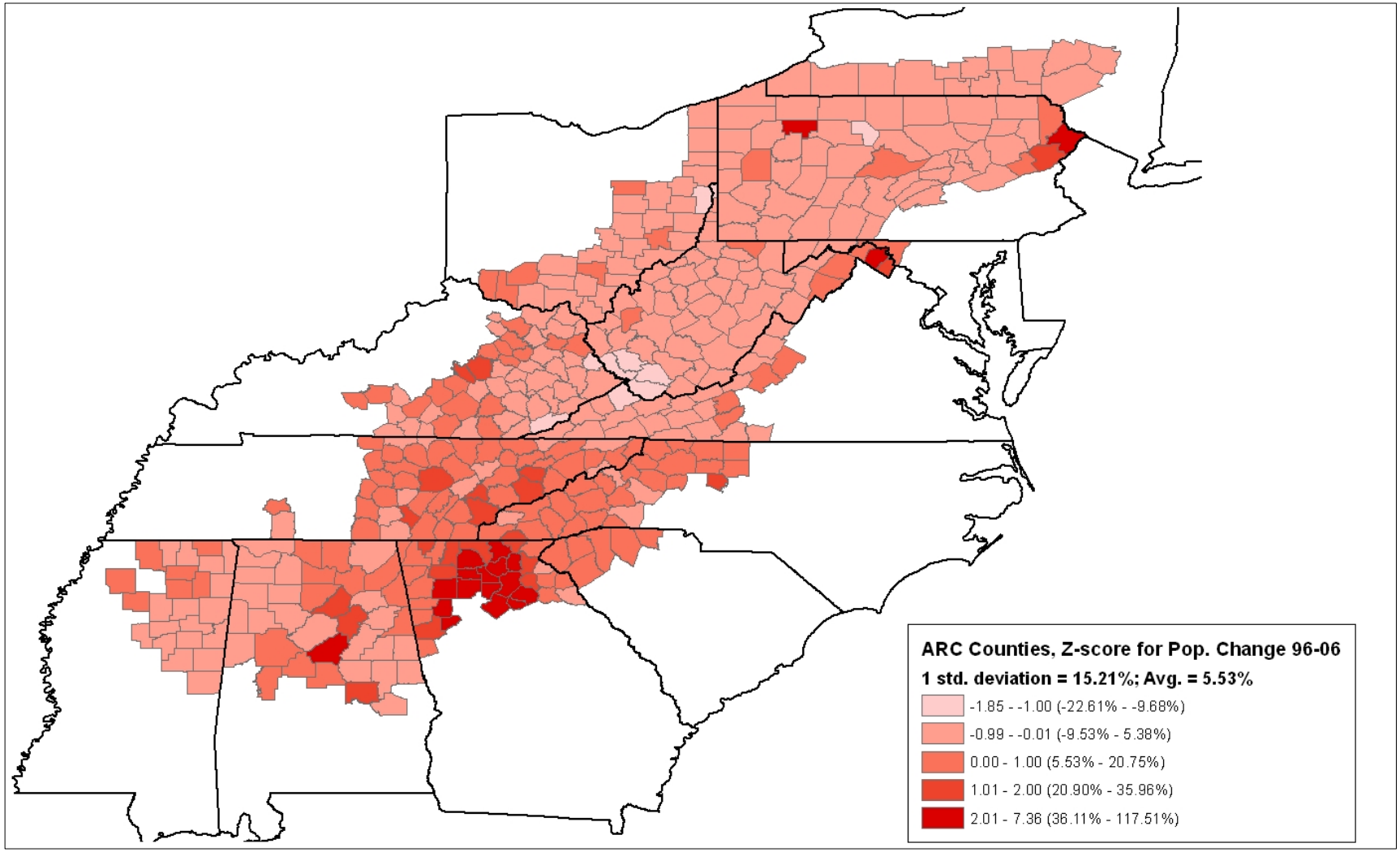
Turning to the last set of maps, Figure 2.8a shows Z-scores for poverty rates across the nation. Again, the central core of the ARC region stands out as does the Mississippi Delta region, southwest Texas and portions of South Dakota. Quite remarkably, there is also a band of very high poverty counties just outside the ARC boundary in Alabama. The map displayed in Figure 2.8b for the ARC region tends to confirm this pattern, but also shows a very high poverty county within the ARC border in Alabama. Also evident are counties with relatively high poverty rates in eastern Kentucky, southwest West Virginia, and Ohio.

***Figure 2.1a*:** County Distribution of Z-scores for Population Growth, 1996-2006



United States, Z-score of Population Growth, 96-06

***Figure 2.1b*:** County Distribution of Z-scores for Population Growth in ARC Counties, 1996-2006



ARC Counties, Z-score of Population Growth, 96-06

***Figure 2.2a*:** County Distribution of Z-scores for 2000 Share of Adults with at least One Year of College Education

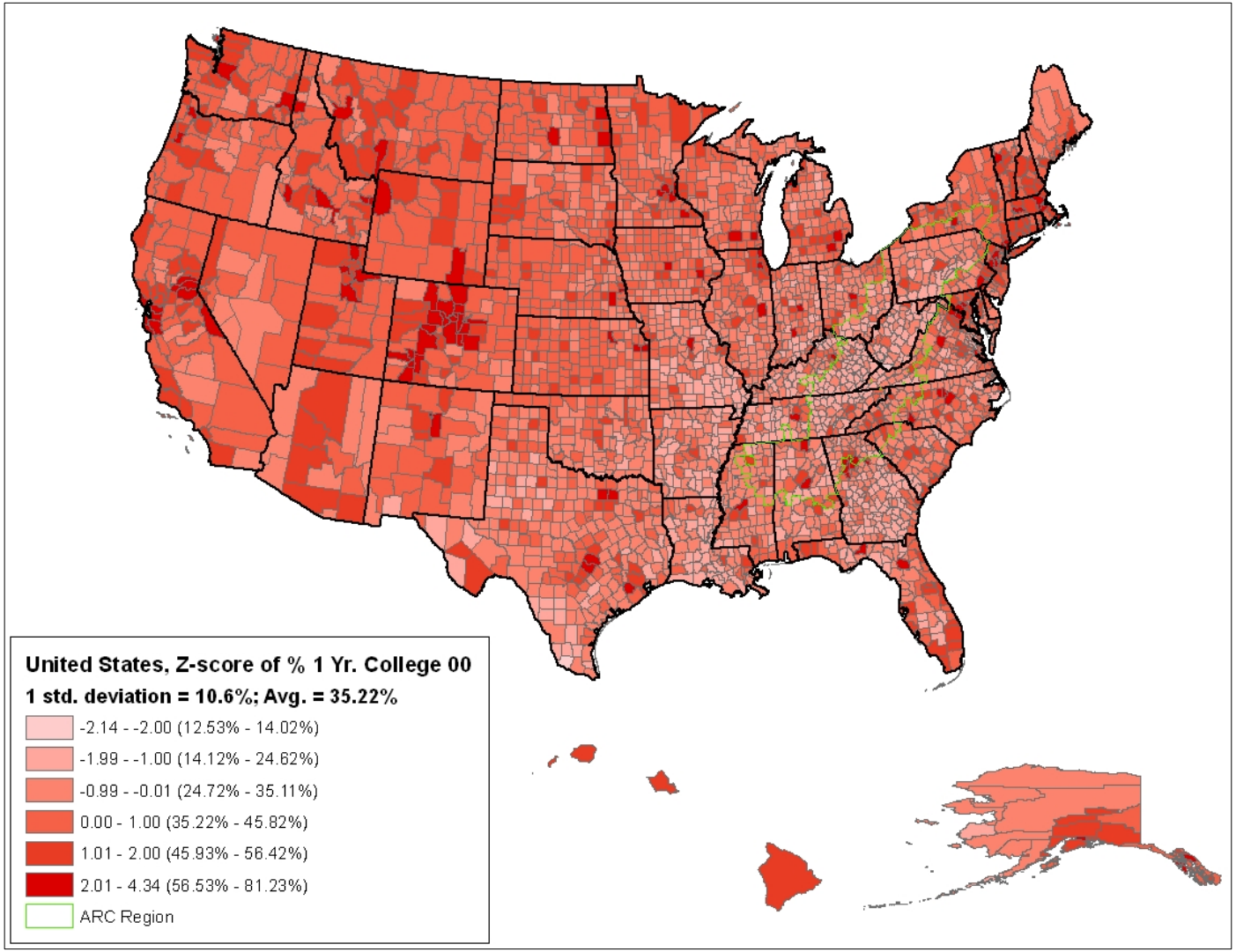


***Figure 2.2b*:** County Distribution of Z-scores for 2000 Share of Adults with at least One Year of College Education, ARC Counties

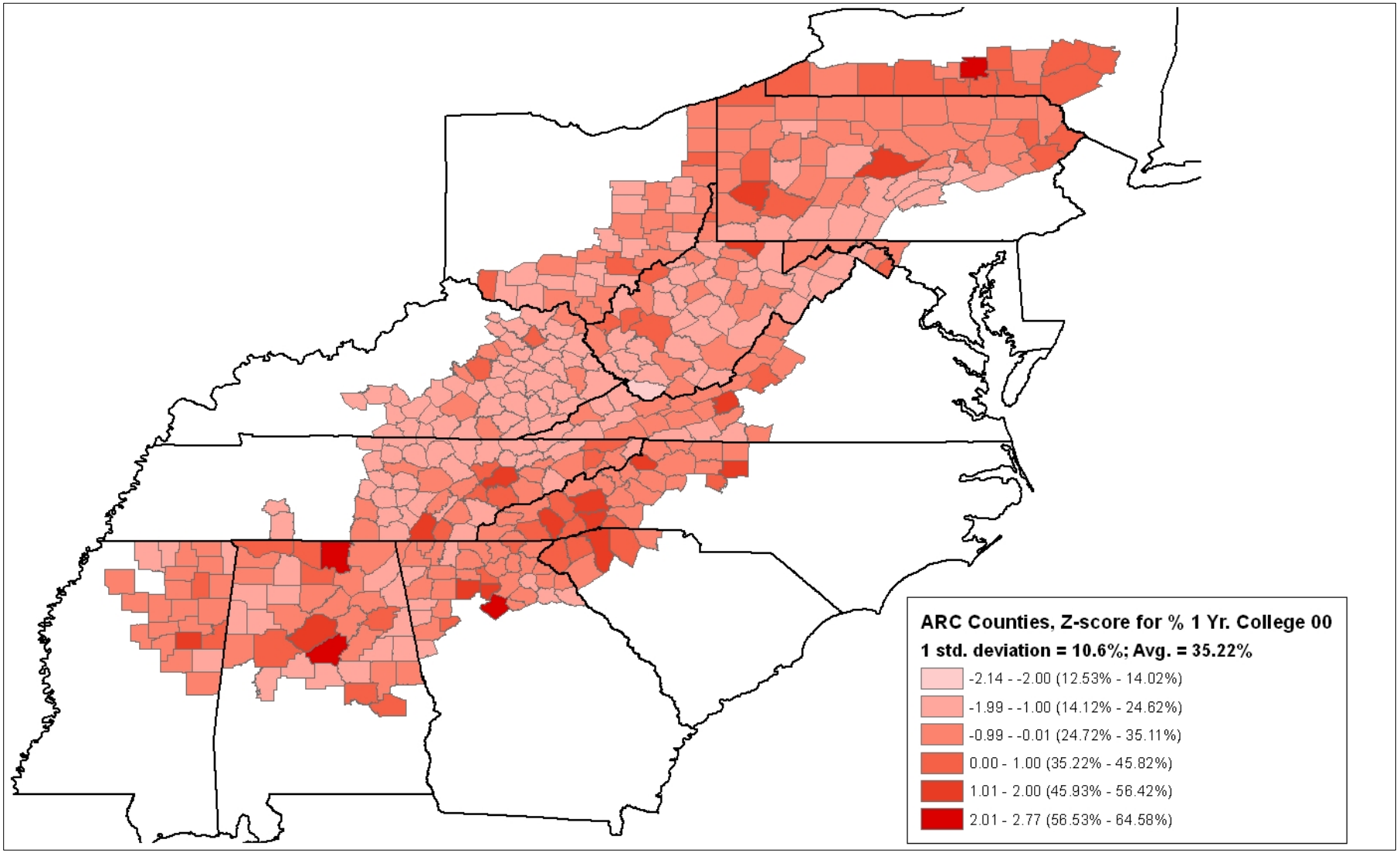


***Figure 2.3a*:** County Distribution of Z-scores for 2000 Share of Adults with a Four-Year College Degree

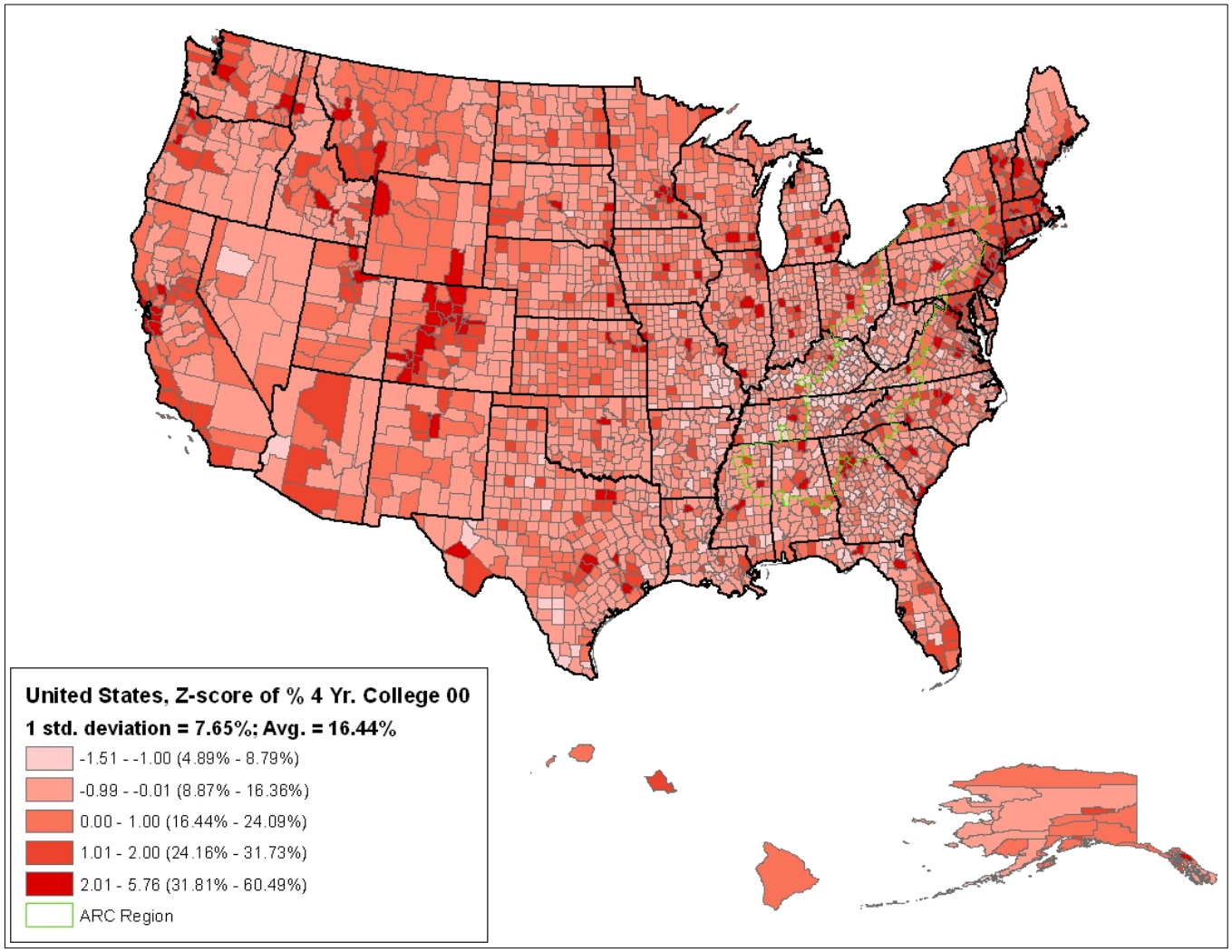


***Figure 2.3b*:** County Distribution of Z-scores for 2000 Share of Adults with a Four-Year College Degree, ARC Counties

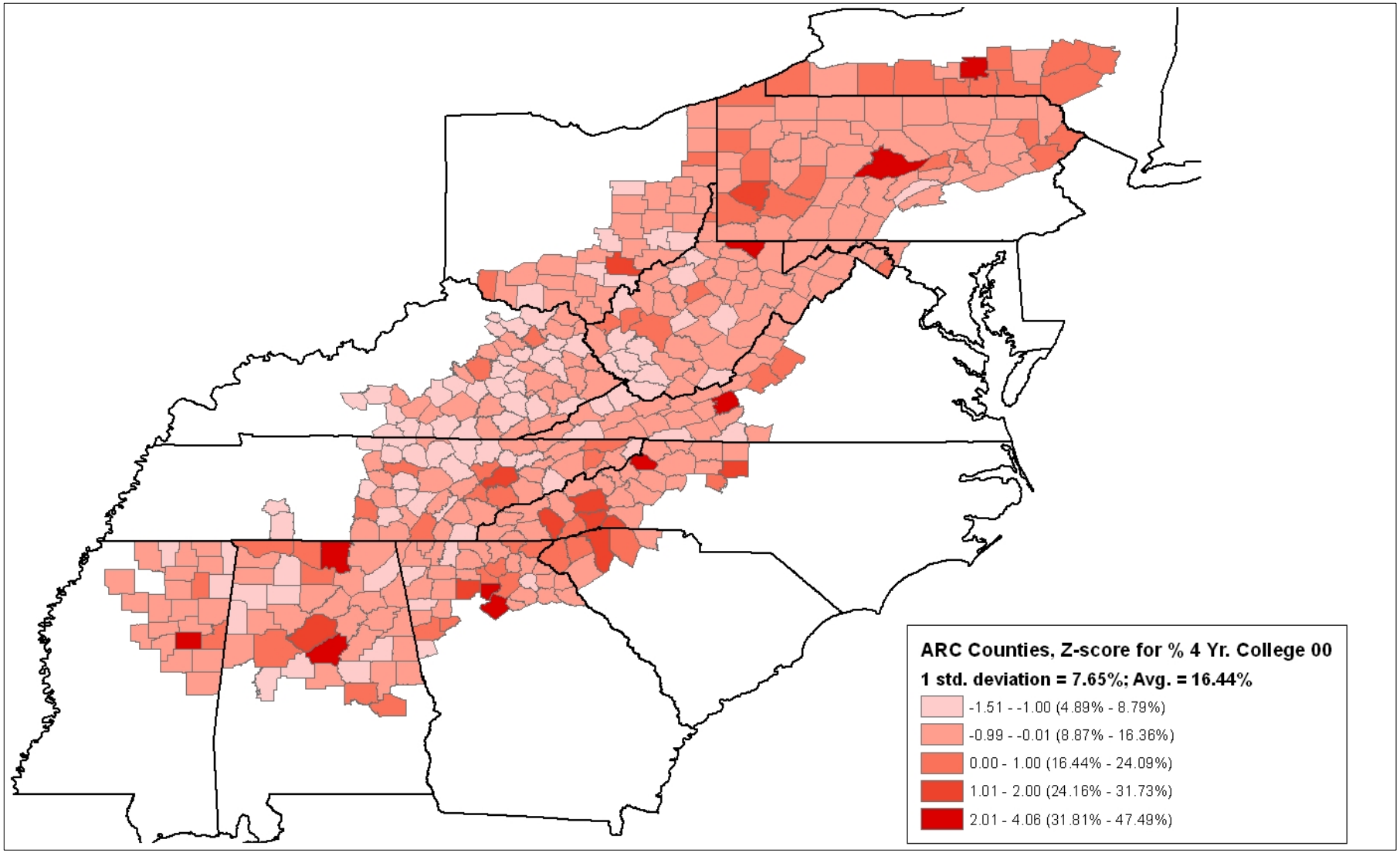


***Figure 2.4a*:** County Distribution of Z-scores for 2006 Per-Capita Market Income

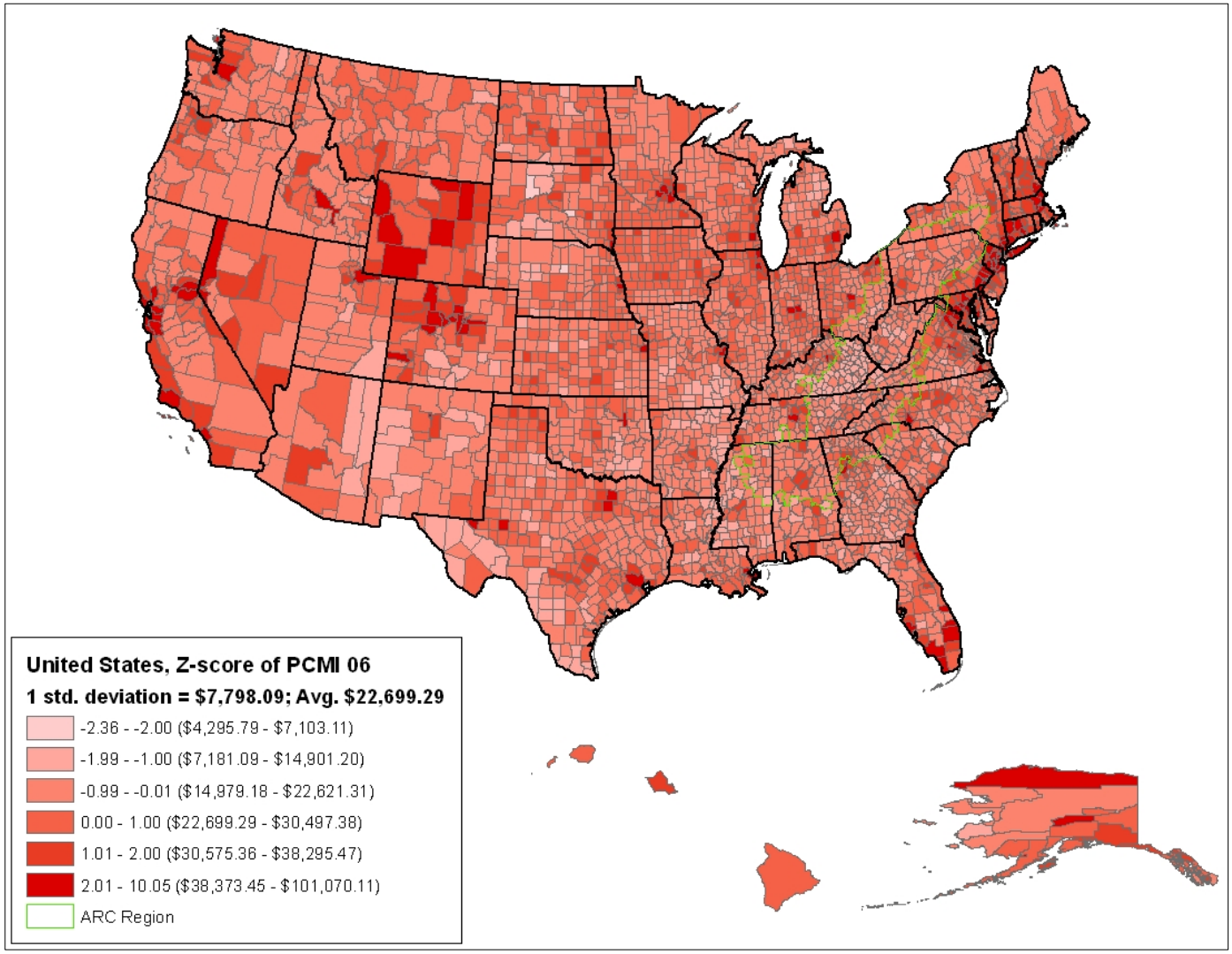


***Figure 2.4b*:** County Distribution of Z-scores for 2006 Per-Capita Market Income, ARC Counties

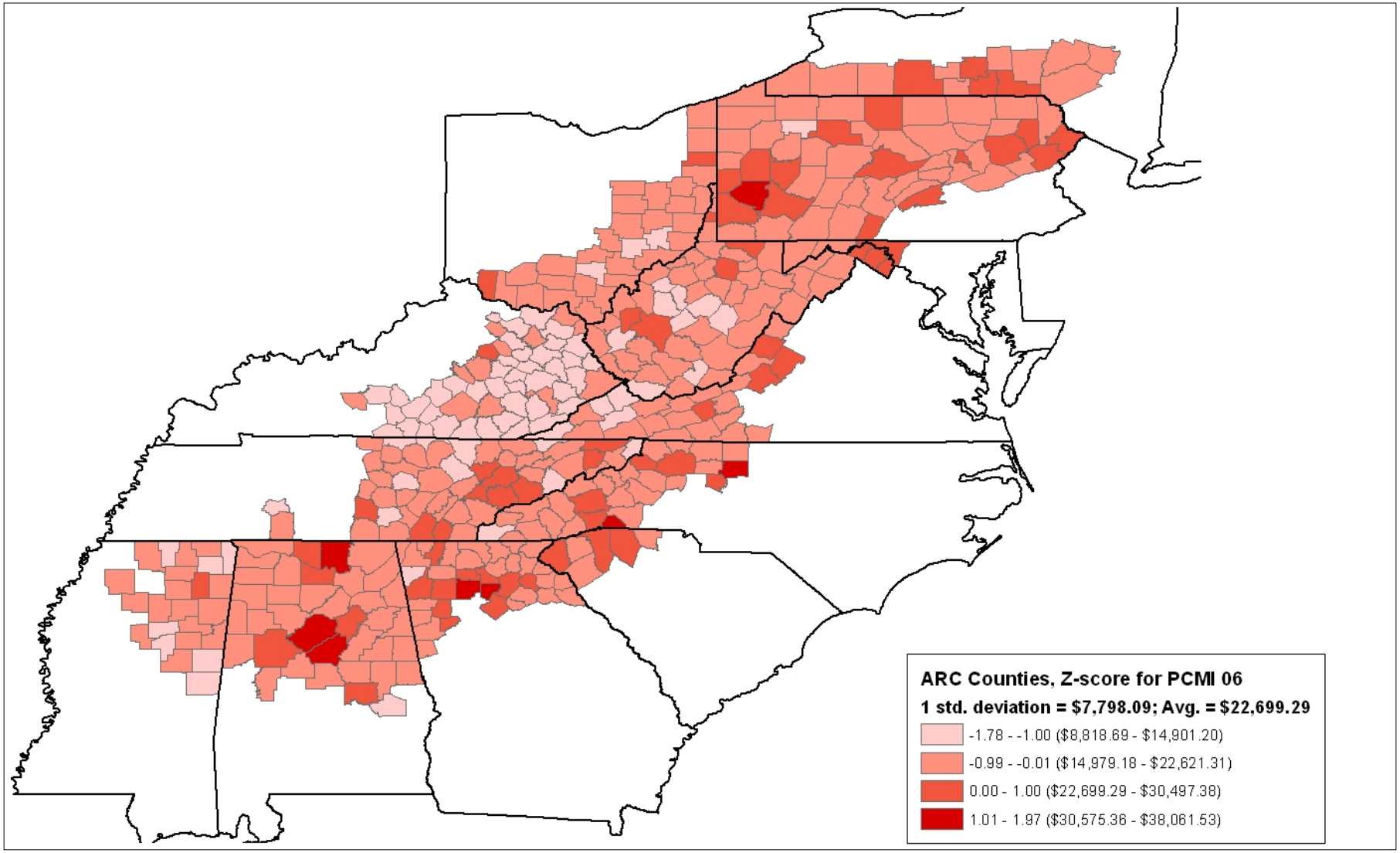


***Figure 2.5a*:** County Distribution of Z-scores for 2006 Nonfarm Self-Employment Share of Total Employment

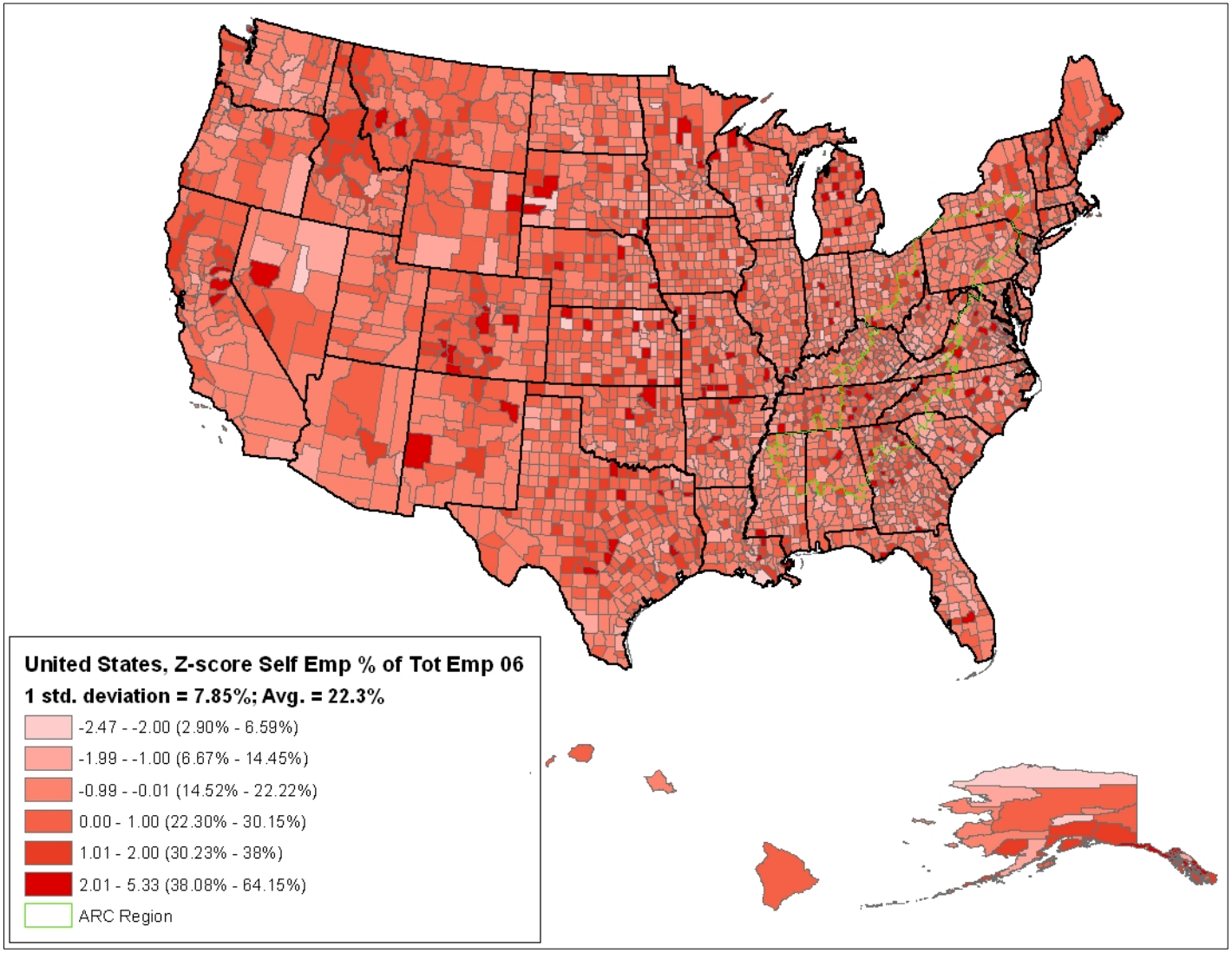


***Figure 2.5b*:** County Distribution of Z-scores for 2006 Nonfarm Self-Employment Share of Total Employment, ARC Counties

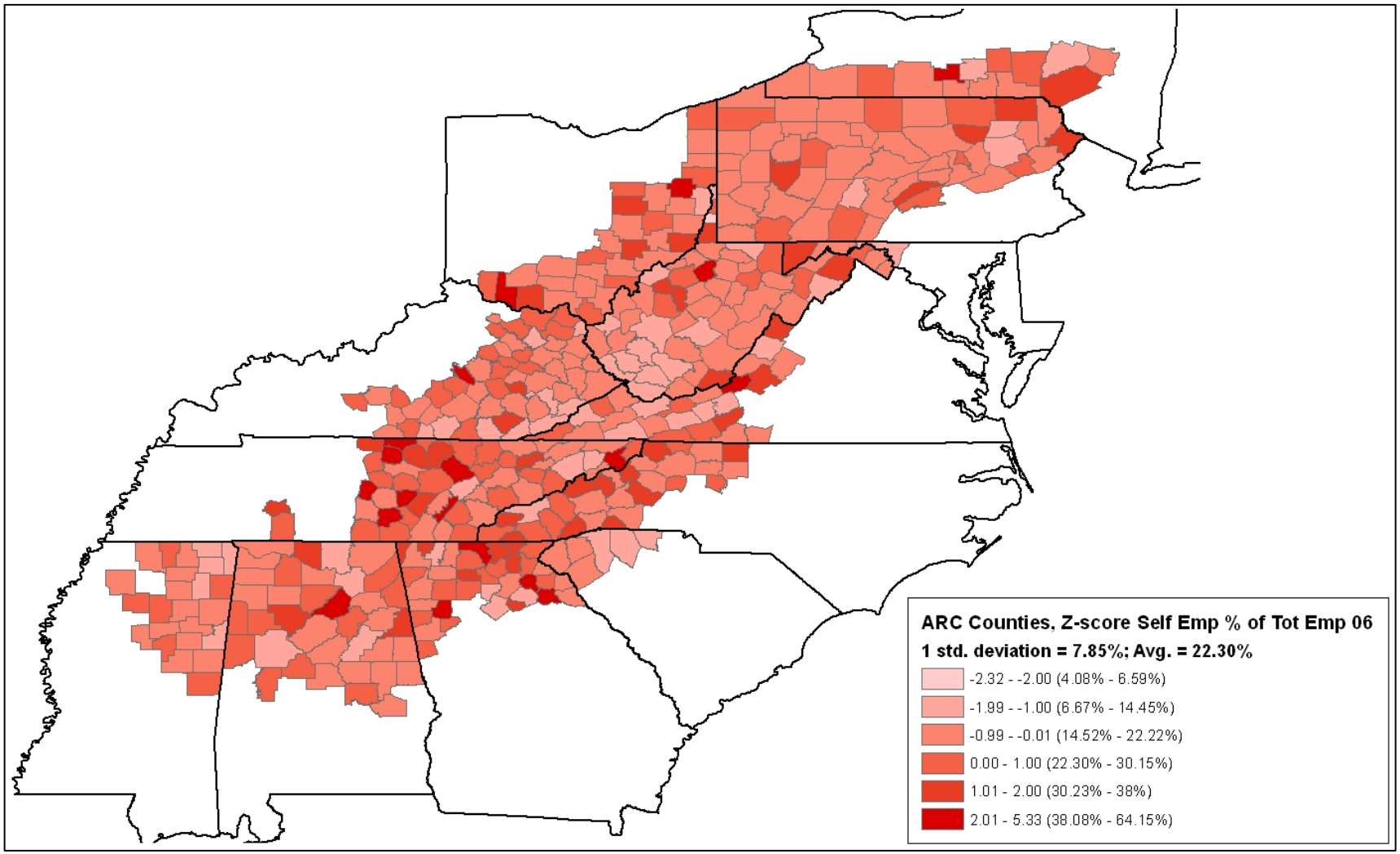


Figure 2.6a: County Distribution of Z-scores for 2007 Employment/Population Ratio

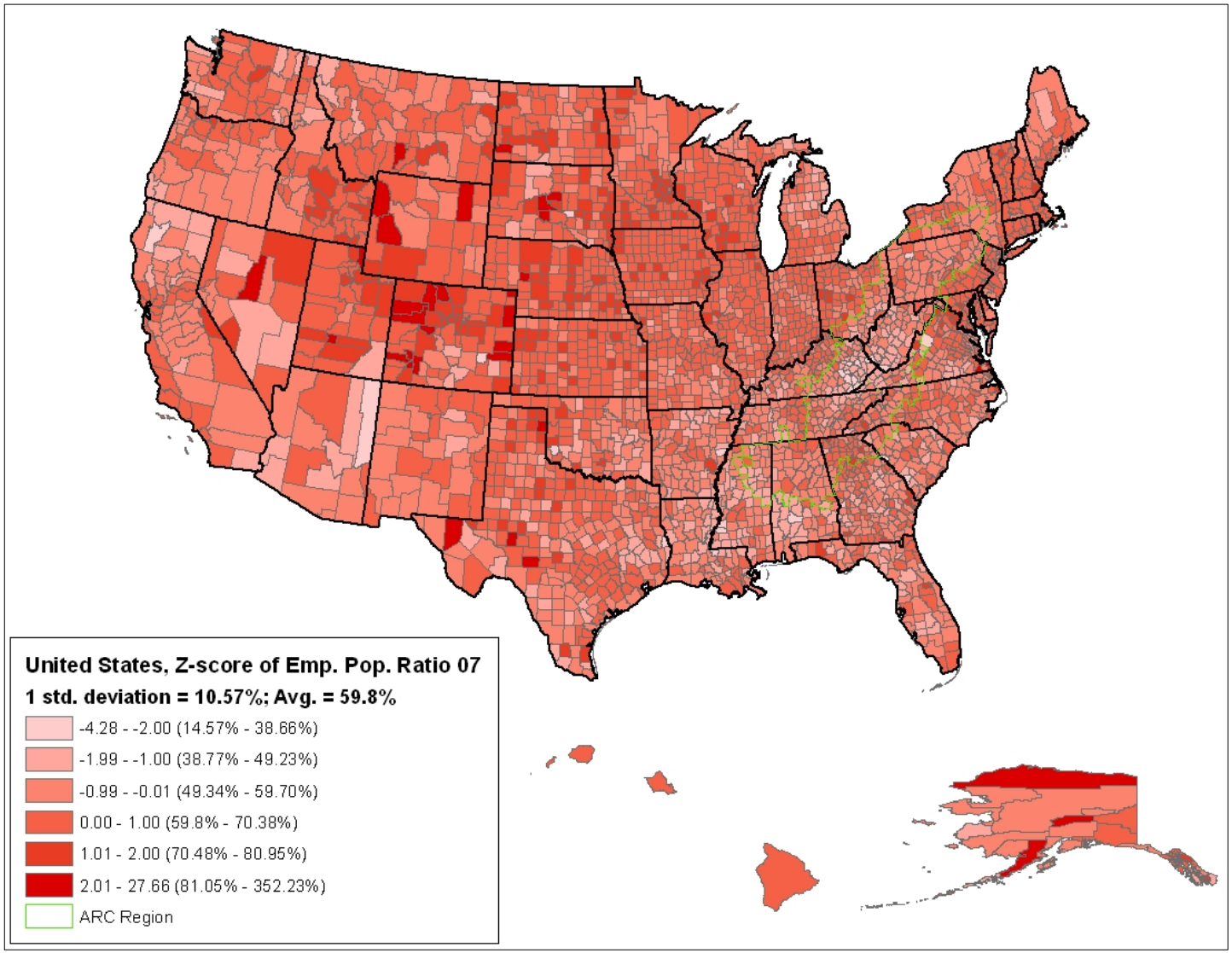


***Figure 2.6b*:** County Distribution of Z-scores for 2007 Employment/Population Ratio, ARC Counties

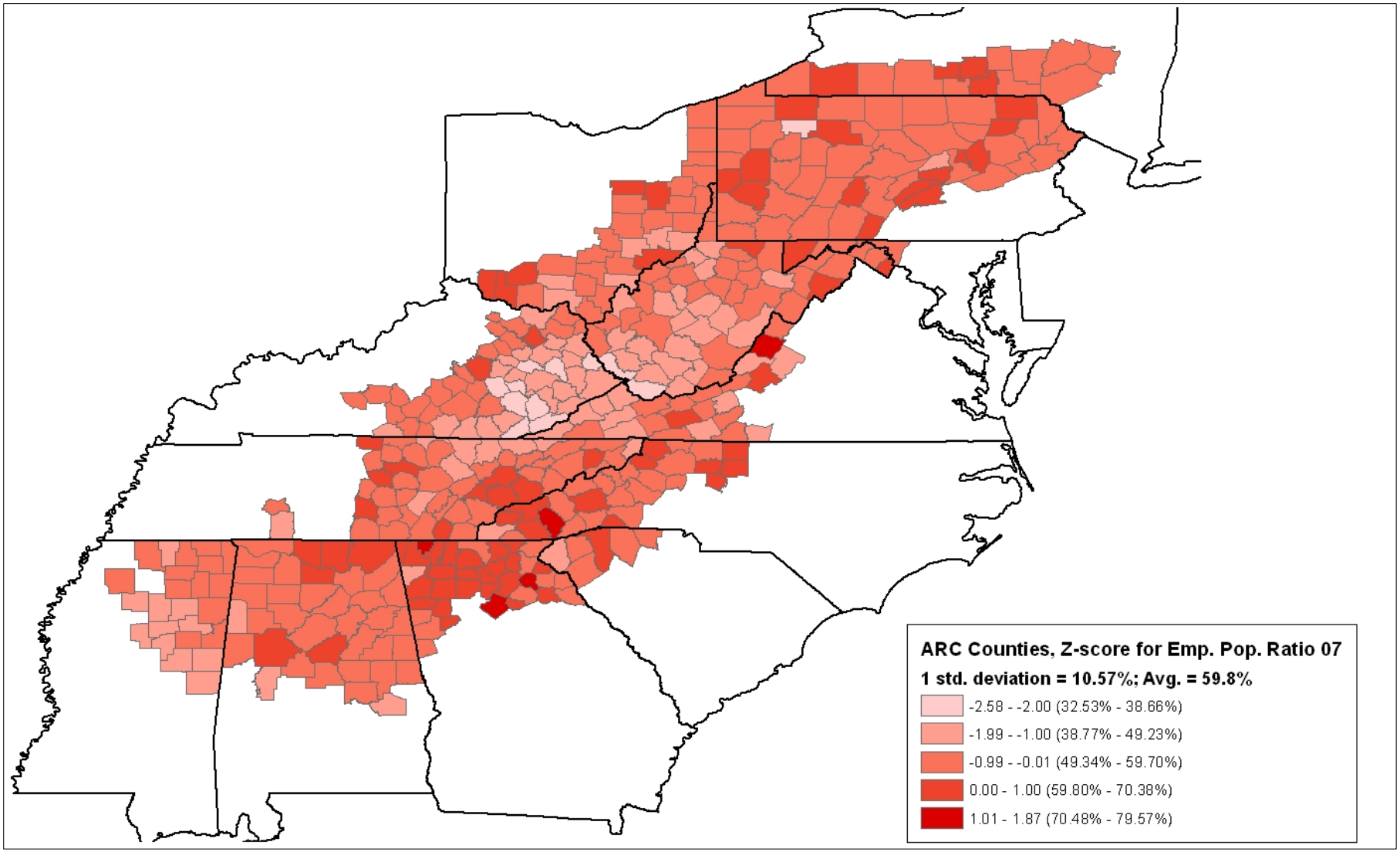


***Figure 2.7a*:** County Distribution of Z-scores for 2007 Unemployment Rate

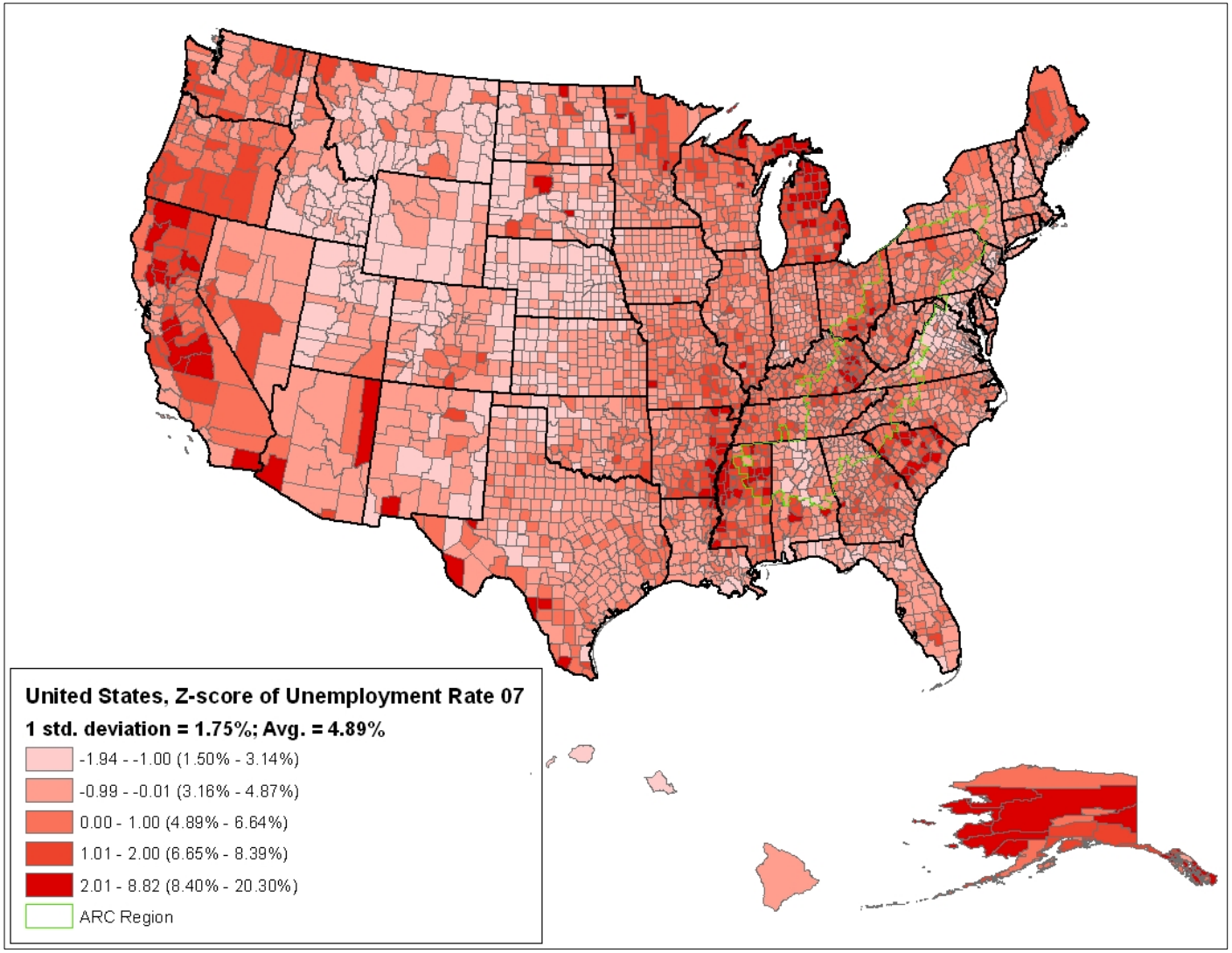


***Figure 2.7b*:** County Distribution of Z-scores for 2007 Unemployment Rate, ARC Counties

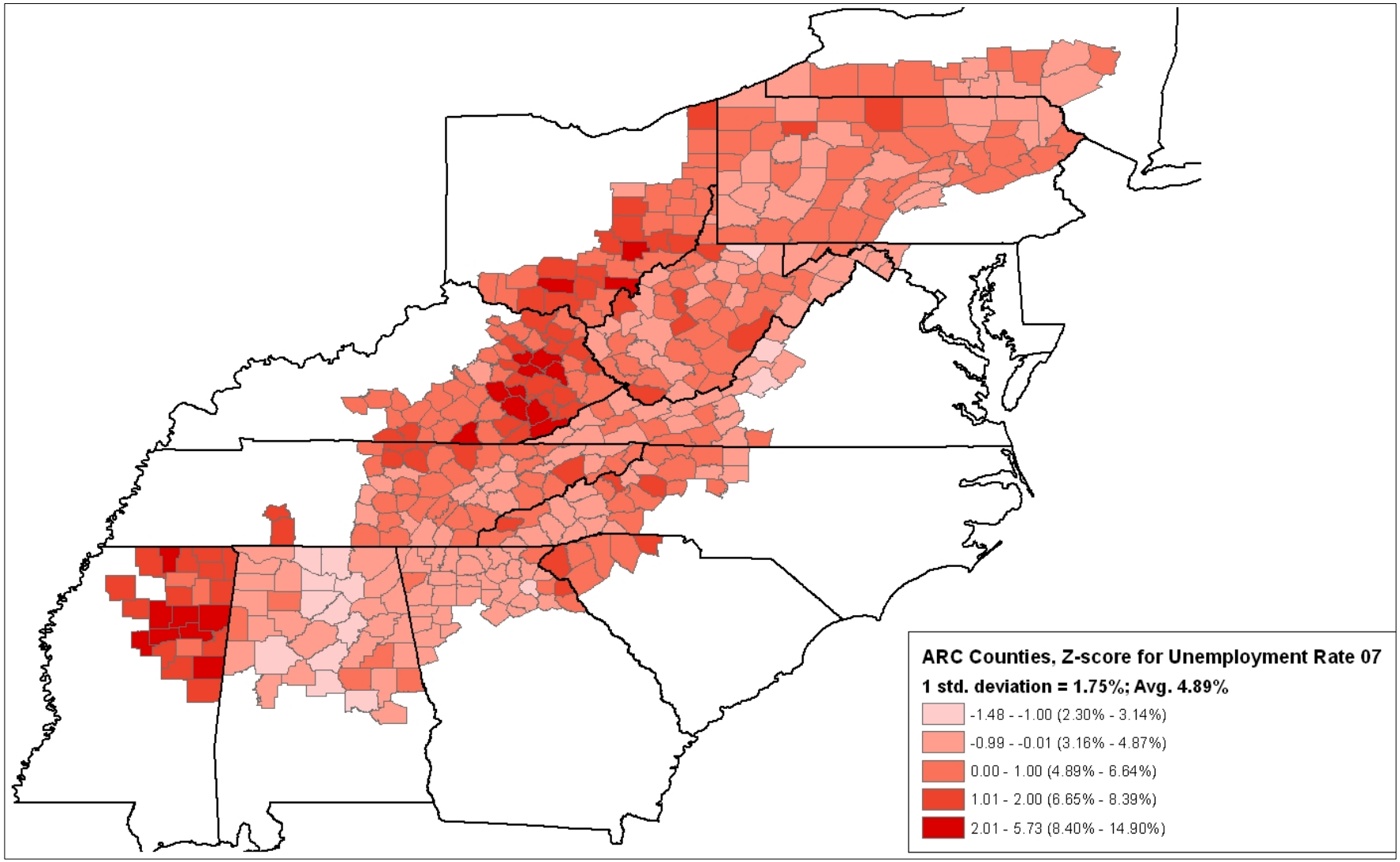


Figure 2.8a: County Distribution of Z-scores for 2007 Poverty Rate

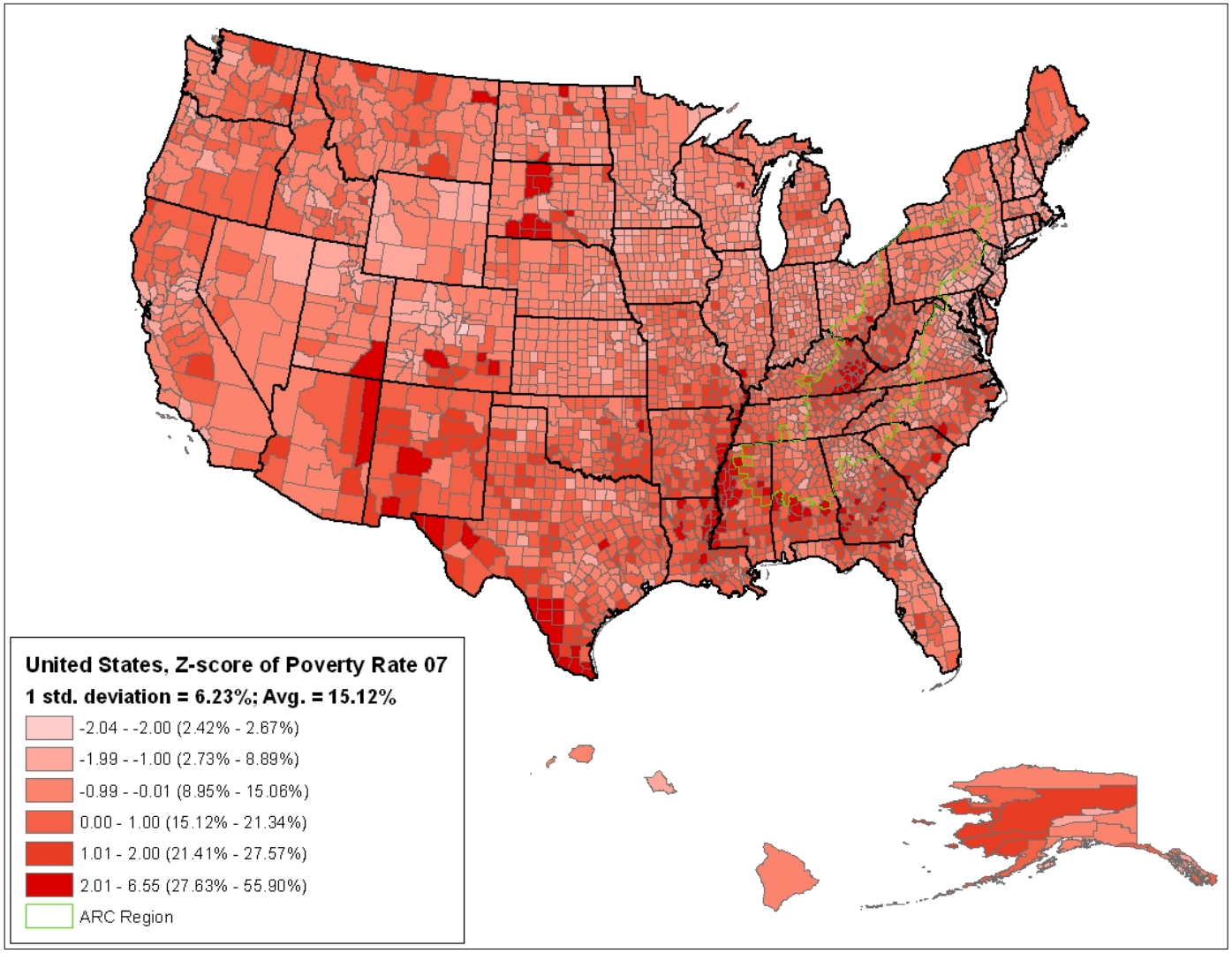
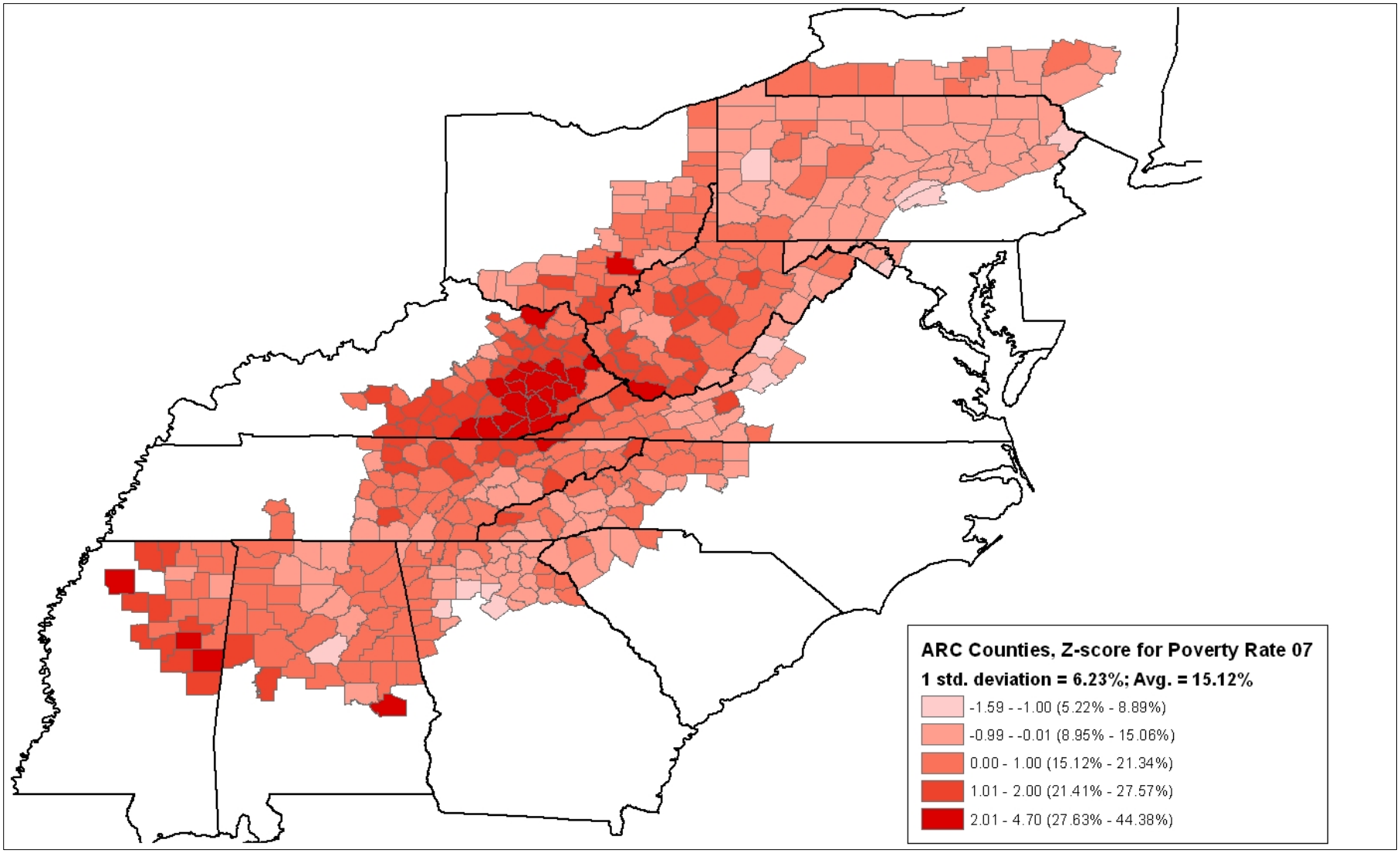



Figure 2.8b: County Distribution of Z-scores for 2007 Poverty Rate, ARC Counties