3 Methodology

3.0 Methodology Introduction

Regional analyses of income, employment, and business establishment growth typically apply one or a combination of methods to explain why some locations or regions are more likely to grow or decline compared to other places. These analytical methods can be broadly categorized as 'quantitative' or 'qualitative', but there is clearly overlap between these two types of methods. Findings from good, theoretically informed quantitative models may supplement inductive conclusions from qualitative studies. Solid inductive qualitative analyses, in turn, make generalizations deduced from quantitative models even more compelling.

The approach used in this study supplements a quantitative analysis of income, employment, and business establishment growth in the ARC region based on *secondary* data sources with *primary* information collected through surveys and focus groups in 10 counties in the region. The approach combines (1) spatial regression analysis, (2) case-cohort identification of paired counties, and (3) a survey of individual and structured focus groups in the paired counties. Specific details of each methodology are described in the three subsections that follow. The first subsection (3.1, "Analysis of Regional Growth in the ARC Region, 2000–2007") discusses the empirical model used to analyze growth of the aforementioned economic indicators in the ARC region. The second subsection (3.2, "Selection of County Cohorts") discusses the methodology used to select county cohort pairs. The final subsection (3.3, "Survey Design and Structured Focus Groups") discusses the methodology used to conduct the surveys and the focus groups.

3.1 Analysis of Regional Growth in the ARC Region, 2000–2007

Quantitative or 'econometric' approaches typically used to describe or explain the factors driving growth include regressions analyses, input-output modeling, or computable general equilibrium modeling. The quantitative approach used in this research applies regression analyses to understand the factors associated with growth in the ARC region from 2000 – 2007. Particular attention is given to the role of industry clusters on growth during this time period.

Regional economists and policy makers also increasingly emphasize the identification of industry clusters as an important component of regional development strategies (Barkley and Henry, 1997; Porter, 1998; Stimson, Stough, and Roberts, 2006; St. John and Pouder, 2006; Feser, Renski, and Goldstein, 2008). A regional adjustment model is used to test this industry cluster hypothesis by estimating the conditional influence technology clusters identified by Feser and Isserman (2009) (Appendix 7.1, "Basic Technology Industry Clusters") had on economic growth in the Appalachian region from 2000 – 2007. Regional adjustment models have been used in a wide variety of empirical applications studying jobs and population migration dynamics and conditional economic growth (e.g., Carlino and Mills, 1987; Carruthers and Vias, 2005; Carruthers and Mulligan, 2007; Lambert et al., 2007; Pede, Florax, and de Groot, 2006), explain changes in county per capita income (Monchuk, et al., 2007), or firm entry-exit (Brown, Lambert, and

Florax, 2010). Economic growth is hypothesized to move toward an unknown future state of spatial equilibrium where income, jobs, and employers are distributed such that individual utility or firm profits are maximized with respect to location.[16]

Local determinants typically hypothesized to influence job, business establishment, and income growth include demographic characteristics, settlement patterns, growth momentum, industry structure, infrastructure, human and social capital, and physical and natural amenities. A log-linear model was used to specify the regional adjustment model used in the regression analysis (Appendix 7.2.1, "Growth Regression Model"). The variable names and summary statistics are described in Appendix 7.4.2.

Growth determinants include the change in real per capita income ($\Delta pci_{2000-2007}$), change in employment ($\Delta emp_{2000-2007}$), and the change in business establishments ($\Delta estabs_{2000-2007}$); all in natural log ratios with the initial (terminal) years of 2000 (2007). The initial year for employment was normalized by county area; more densely settled areas tend to be correlated with employment density (Carruthers and Vias, 2005). These areas are attractive to business for their higher level of services, but employment density entails more than urban amenities. Rural jobs have tended to concentrate in densely populated areas, and employment density itself may suggest a relative shortage of land and higher housing prices (McGranahan, Wojan, and Lambert, 2010).[17] Previous research also found that job growth in rural locations is faster in more densely settled rural areas and in sprawling urban areas perhaps because of the availability of larger labor pools (McGranahan and Wojan, 2007).

Access advantage to economic centers is measured by three variables. The percent of workers commuting outside a county (*percomm*) is expected to be positively associated with growth, given continuing advantages of urban proximity. Counties with relatively low unemployment (or higher employment rates) are expected to grow faster than counties with fewer available jobs (*emprt*). Employment rates are obtained from the Regional Economic Information System (REIS) files from the Bureau of Economic Analysis (BEA, 2007) and the percent of workers commuting to other counties is taken from the 2000 Census.

---

[16] But growth trajectories need not be similar across an entire region, and multiple but unstable equilibriums are possible. Baldwin et al. (2003) typify development in terms of regional adjustment as fits and starts occurring much in the same way biological species evolve through 'punctuated equilibrium' (Gould and Eldridge, 1977).

[17] Conceivably, one could include the (2000) business establishment density, employment density, and per capita income as initial conditions pertaining to, for example, job growth. However, the correlation between the initial conditions of each outcome variable was substantial (all above 0.60), which is not too surprising. Employment typically follows business establishments growth, business establishment density tracks consumer demand which is related to per capita income, and earnings per job closely follow per capita income. In sum, little additional information would be gained by including all of these variables as initial starting points in each outcome equation. To avoid potential problems that could arise from multicollinearity between these factors, only the initial variables corresponding with the base of the change variables were included in each equation, and each equation was estimated separately. However, higher incomes might provide resources for new business startups, but higher labor costs may be associated with slower rates of job creation (McGranahan, Wojan, and Lambert, 2010). Thus, for the employment growth equation, the natural logarithm of median household income (in 2000) (lnmedhhi) was included to control for potential income effects on job growth.

Industry structure and composition are measured by the percentage of manufacturing establishments with less than 10 employees (*perestab20*) and the percentage of manufacturing establishments with more than 100 employees (*perestab100*). Both variables intend to capture effects due to agglomeration economies and economies of scale internal to the firm (Lambert, Brown, and Florax, 2010). We also included two broad indicators of industry structure; the percent employed in agriculture, forestry, and mining (*peragmi*, NAICS 11 and 21) and the percent employed in manufacturing (*permanf*, NAICS 31 – 33). Counties dependent on resource-based industries and manufacturing may exhibit slower growth (McGranahan, Wojan, and Lambert, 2010) than counties with more post-industrial economic activities.

Change measures from the previous decade (1990 – 2000) were included for business establishments ($\Delta estab_{9000}$), employment (Δemp_{9000}) and population (Δpop_{9000}). Including these variables suggests hypotheses about growth sustainability (McGranahan, Wojan, and Lambert, 2010). Population growth in the preceding decade may also be indicative of more favorable demand conditions. The relationship between growth in business establishments and jobs has also been found to have some lag (Fritsch and Mueller, 2004). Change in establishments over the previous decade should be related to changes in jobs where growth was sustained. These variables were calculated as the logged ratio of the end-of-the-decade to start-of-the-decade measures.

Demographic variables include the percent of the population age 20 – 64 (*perpop2064*), a proxy for labor availability (2000), and the proportion of the population over 65 (*perpop65up*), both measured at the beginning of the decade. Some counties in the Appalachian region have become magnets for retirees (Lambert et al., 2007; Clark et al., 2009). Retirees may be inclined to start small businesses, but with no intention of becoming major employers (Rogoff, 2008). The expected relationship is ambiguous. The proportions of the Black (*perblk*), Hispanic (*perhsp*), and Native Americans (*peramind*) populations were also included because these groups and Whites may have different opportunities to participate in different job markets or develop new businesses (McGranahan, Wojan, and Lambert, 2010).

Human capital, often found associated with economic growth, is represented by the percent of the population with bachelor's degrees (*perhsdip*), and the percent of persons working in creative occupations (*percc*). In earlier decades, many rural areas with low education attracted businesses offering low-skill, low-wage jobs, but many of these firms tended to relocate operations off–shore in the 1990s or adopted new technologies requiring higher skilled labor (Johnson, 2001). Demand markets may also harbor a relatively larger stock of creative individuals capable of solving difficult supply problems or combining old ideas in new ways, which may influence growth in businesses and per capita income. We include the percent of persons in creative occupations (Wojan and McGranahan, 2007) to control for stock of local talent and intellectual capacity.

Natural amenities and public land availability may play a role with respect to income growth and jobs by attracting new businesses and people to locations with wilderness or scenic environments (Deller et al., 2001; McGranahan, 2008). In low-amenity areas, growth may occur mostly through changes in demand for producer services from the local economic base or the expansion of local colleges or universities (McGranahan, Wojan, and Lambert, 2010). However, high-amenity areas, like many areas of the Appalachia region, may also be remote and difficult to access. A natural amenity index (*amenity*) was included to measure the relationship between economic growth and locations rich in natural amenities (McGranahan, 1999). The variable is an aggregate index of sunlight, humidity, and temperature; topography; and water resources. The percent of the county in public land was also included to control for the effects of public access to un-built areas on growth (*landpub*).

Dummy variables indicating the presence of an interstate (*interstate*) or an Appalachian Development Highway (*adhs*) were included to control for the influence of transportation infrastructure on job and income growth. The expected sign is generally ambiguous. Good roads may be attractive to prospective firms, which may increase the likelihood of attracting new investment and jobs. However, roads may also encourage out commuting followed by growth elsewhere (Kahn, Orazem, and Otto, 2001).

3.1.1 Industry Cluster Concentration Measures

Industry clusters are built around export oriented firms that bring new wealth into a region and help drive regional economic growth (Barkley and Henry, 1997; Stimson, Stough, and Roberts, 2006). Gibbs and Bernat (1997) characterized industry clusters as businesses in similar industries seeking comparative advantage by locating near raw materials, demand centers, or labor markets. Industry clusters also influence competition by fostering innovation, research, and development, which in turn support future productivity growth by stimulating business formation. These exchanges encourage additional rounds of interaction, which advance core industry sectors and reinforce the cluster (Porter, 1996, 1998). The resulting agglomeration of competing but collaborating industries in a well-defined region is arranged into horizontal and vertical relationships with similar resource and/or labor needs (Fujita and Thisse, 2002). The basic technology clusters identified by Feser and Isserman (2009) are based on 1997 benchmark input-output account tables of the US economy. Feser and Isserman identified value chains which are groups of industries with highly similar sales and purchases patterns. While a variety of measures could be used to proxy industry concentration (Goetz, Shields, and Wang, 2009), the relationship between technology clusters and job and income growth was measured using a industry concentration index (CI) calculated as;

$$CI_{i2000}^{k} = \left(s_{ik} / \Sigma_j s_{ij}\right) / \left(\Sigma_i s_{ik} / \Sigma_i \Sigma_j s_{ij}\right),$$

where k is a technology cluster based on Isserman and Feser's (2009) results, and s is the number of establishments in technology cluster k (Appendix 7.1, "Basic Technology Industry Clusters")[18]. The index is similar to a location quotient. Location quotients may be the most commonly used measure for identifying clusters (Shields, Barkley, and Emery, 2009), and are typically used in economic base analyses to compare local economic composition to other economies. Location indices can also be useful for characterizing counties with comparative advantage in a given industry (Shaffer, Deller, and Marcouiller, 2004).[19] The

---

[18] Location quotients could be calculated using sector employment data, but disclosure issues precluded this convention. Guimarães, Figueiredo, and Woodward (2009) also suggest there may be some statistical advantages in using establishment counts rather than employment in terms of confounding effects that may arise arising from scaling issues.

[19] Advantages, disadvantages, and the assumptions behind the location quotient are summarized by Shields, Barkley, and Emery (2009).

traditional location quotient would use employment data to construct the concentration index. But publicly available data tracking sector-specific employment is subject to disclosure. However, the number of firms in a given sector is fully available in the County Business Pattern files. When the measure exceeds 1.0, a region is considered to be competitive (self-sufficient, or "exporting") with respect to that sector. We maintain no priors on the expected relationships the technology clusters might have on income and job growth.[20]

Of interest is the extent to which the initial level (or "stock") of a particular technology cluster is associated with job or income growth. Because the CI measure is an index, explaining the log-linear relationship between growth and the sector concentration index as an elasticity (as opposed to simple marginal effects) has some advantage. The issue would be trivial if the concentration index was in natural logs. However, some establishments belonging to a cluster classification were not observed in a county. This has implications with respect to calculation of the elasticities corresponding with each sector.[21] The Smooth Transition Regression (STAR) model of Pede (2010) and Pede, Florax, and Holt (2009) allows for parameter variation across space. The elasticities associated with the concentration indices were mapped, and the resulting patterns subsequently analyzed using a Local Index of Spatial Association (LISA), the Local Moran's I statistic (Anselin, 1995). The resulting LISAs identify the "core" counties of a technology cluster are areas where the elasticity associated with a cluster are, on average near neighboring counties where the growth indicator-concentration index is also relatively high. In this analysis, counties shaded "red" indicate these regions, which may be loosely interpreted as regions that exhibit comparative advantage with respect to a given sector.

Three important modeling concerns arise considering that (1) county eligibility to participate in ARC programs is based mainly on historical and political concerns; (2) the main focus of the application is on the performance of ARC counties; and (3) the economies of ARC counties are tied to wider regional economies. These facts preclude isolating the differential growth of ARC (*n* = 420) and non-ARC counties (*n* = 650) to local determinants, while simultaneously allowing for arbitrary correlation between ARC and non-ARC members. Dummy variables indicating ARC county inclusion (*arc*) and non-ARC counties (*nonarc*) were

---

[20] Feser, Renski, and Goldstein (2008) found that clustering did not guarantee employment growth, but was associated with new businesses formation from 1998–2002 in the Appalachian region.

[21] The percent change in the economic growth indicator given a 1% change in the concentration index is approximated as

$$\eta_i^k = CI_{i2000}^k \cdot \theta_k ,$$

which is the contribution of an industry sector in 2000 to the predicted value of growth until 2007. Elasticities of the index with respect to growth can always be written as (Chiang, 1984),

$$\eta_i^k \approx \frac{\partial \ln\left(y_{i2007}/y_{i2000}\right)}{\partial LQ_{i2000}^k} \frac{\partial CI_{i2000}^k}{\partial \ln CI_{i2000}^k} ,$$

because $\partial \ln z / \partial z = 1/z$, the percent change in the economic growth indicator given a 1% change in the concentration index is approximated as $\eta_i^k = CI_{i2000}^k \cdot \theta_k$, which is the contribution of an industry sector in 2000 to the predicted value of growth until 2007.

interacted with the local determinants, allowing for slopes and intercepts to vary between ARC and non-ARC counties (Appendix 7.2.1, "Growth Regression Model"). Thus our regressions analysis allows us to focus on ARC counties specifically, but recognizing that these counties are connected to a wider regional economy by allowing for geographic dependence between ARC and non-ARC counties through the spatial process models.

3.1.2 Spatial Regression Model and Growth Regimes

We hypothesize that growth in jobs, business establishments, and real per capita income may be simultaneously determined by job or income growth in neighboring counties; a county with a given change in employment or income growth (y_i) may be surrounded by j other counties with similar growth rates, e.g., $\sum_{j=1,i\neq j}^{n} w_{ij}y_j$, suggesting information spillovers, thick labor markets, or forward-backward economic linkages across space (Anselin, 2002; Moreno et al., 2004). Most studies incorporating spatial dependence typically use a spatial process model attributed to Whittle (1954) in which an endogenous variable specifies interactions between spatial units plus a disturbance term. Anselin and Florax (1995) call this a spatial lag autoregressive (SAR) model (Appendix 7.2.2, "Spatial Process Model"). A more general spatial process model permitting spatial correlation between SAR disturbances is the Autoregressive-Autocorrelation model (ARAR, Anslein, and Florax, 1995), which is considered here as the "null model" explaining growth (Appendix 7.2.2, "Spatial Process Model")

The reduced form of the SAR-type models suggests that the calculation of the marginal effects is more complicated than ordinary estimates due to the spatial lag multiplier. LeSage and Pace (2009) identify two methods whereby the marginal effects implied by SAR-type models can be calculated. The first method interprets the lag spatial multiplier as in infinite geometric series (see also Anselin and Lozano-Gracia, 2008). For example, the "direct effect" of a covariate (β) is the impact it has on a given spatial unit. In the limit the "total" marginal effect is $\beta^{Total} = \beta(1-\rho)^{-1}$. The "indirect effect" is the difference between the total and direct effect, or the impact neighboring locations (on average) have on a given spatial unit. The second approach partitions the marginal effects of the SAR-type models into neighborhood order effects, such that the effect decays over space moving away from a target county to neighboring counties (LeSage and Pace, 2009, page 40). [22]

The economic geography literature generally predicts that friction caused by uneven trade costs due to geographic features may induce the concentration of firms and jobs across space into one or a few regions (Brakman, Garretsen, and van Marrewijk, 2001 This research applies a relatively new class of spatial econometric models to explain per capita income, jobs, and business establishment growth using the Smooth Transition Regression (STAR) model of Pede (2010) and Pede, Florax, and Holt (2009). The STAR model provides a direct test for identifying these so-called endogenous structural breaks, and the extent to which access to agglomeration economies affect the relationship between local resource constraints and

---

[22] *For example, $A^{-1}(I_n;\beta_k) = [I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \rho^4 W^4 + \rho^5 W^5 + \ldots \rho^q W^q]\beta_k$, where the order q refers to the location itself (q = 0), the impact of the neighbors (q = 1), the impact of the neighbors of the neighbors (q = 2), etc.*

growth (Appendix 7.2.3, "Endogenous Growth Regime Specification"). The STAR spatial regression model is modified to accommodate endogenous regime-switching potential (Appendix 7.2.1, "Growth Regimes and Spatial Process Models").

3.2 Selection of County Cohorts

The goal of the county cohort selection exercise was to identify counties that were closest in terms of local characteristics to the targeted counties in 1960. Thus, the by selecting counties that were similar in characteristics to the target counties in 1960, questions can be asked as to why counties that were very similar in 1960 may have diverged over time (to 2010). In this study, a goal was to survey counties that were similar in certain characteristics, but for other reasons that may not be immediately evident in secondary data sources were qualitatively different.

Target counties were largely selected on the basis of their economic performance between 1960 and 2010. The Appalachian Regional Commission's (ARC) economic status index was used to measure the economic position of counties in 1960 and 2011. The index is composed of the average of three economic indicators: per capita income, the unemployment rate, and the poverty rate (all relative to the US average). This period was selected because (1) the ARC was not established until 1965, and (2) the 1960 and 2011 indices are comparable across years because they are constructed using the same variables but in different periods.

Two arbitrary cut-offs were considered to isolate counties that had improved their economic status substantially since 1960. The first criterion was that candidate counties were considered those with status indices greater than or equal to 200 or twice the value of the US average of the index during this time period. The second criterion was based on the status index value in 2011. In all, 11 counties were identified using this criterion (see table below). In 2011, 7 of these counties were labeled as "Transitional", 3 were "At-Risk", and one was categorized as "Distressed."

The second criteria considered candidate counties to be those with ARC distress indices greater than 200 in 1960, but had improved their index position by 2011 to less than 150. The second criteria resulted in 23 candidates. In 2011, two were considered "At-Risk", one was "Competitive", and 20 were "Transitional".

The matching procedure applied Isserman and Rephmann's (1995) method to counties that were cohorts (in quantitative aspects) to the candidate counties determined by the selection criteria outlined above. Measures of the similarity between counties were based on a set of key characteristics, as measured by the Mahalanobis d^2 statistic. The smaller the distance, the more similar the match, based on the characteristics examined. These matches were based on the components making up the ARC status index in 1960 (county per capita income relative to the average per capita income in the US, the county level unemployment rate relative to the average unemployment rate of the US in 1960, and the county level poverty rate relative to the average US poverty rate in 1960), and the US Census population density of 1960. Of the possible matches for each target county, the top five ranking counties were selected. The matching counties are therefore the counties that shared similar local economic conditions (in terms of employment, per capita income, poverty, and population density) in 1960. Counties were matched on a subregional basis following the 2009 ARC convention. The matching algorithm was conducted once for each candidate group; n = 11 under the first criteria, and *n* = 23 under the second criteria. Additional details of the cohort county selection process are provided in Appendix C, "County Cohort Selection."

The statistical matches that resulted from the matching procedure were then culled further using qualitative criteria. First, non-eligible counties were eliminated from the pool of potential cohorts. These included target counties that were currently at risk and matching counties that were no longer classified as distressed. ARC and project staff then used a consensus approach to the selection of cohort pairs. Several factors were evaluated in making these decisions, including the following:

- Geographic distribution of sites
- Proximity of sites
- Area and population
- Local economic and natural resources
- Transportation infrastructure
- ARC investments and interventions
- Historical levels of economic performance
- Project logistics

During this cohort selection process, concerns arose regarding two problematic subregions. In the Northern subregion, distressed county choices were very limited, and the favored nonperforming match, Monroe County, Ohio, had recently transitioned off the distressed counties list. Because of this, and after discussions with ARC staff, it was decided to select the next closest distressed county, Morgan County, Ohio, a county located on the northern border of the North Central subregion.

In the Central subregion, which includes the Coal Belt of Eastern Kentucky, the opposite situation occurred. In that area – the worst performing portion of Appalachia – there is a scarcity of target counties successfully transitioning from distress. The best performing county, in fact, appeared to be Pike County, Kentucky – a county that had made remarkable progress but was currently classified as at-risk. Given this rate of performance and Pike County's relative success in the midst of the region's most distressed area, staff elected to use Pike as the performing county in the Central subregion. As a result of this consensus approach, ARC and project staff concluded the following cohort pairings:

- Morgan County, Ohio, and Greene County, Pennsylvania (Northern)
- Calhoun County, West Virginia, and Pendleton County, West Virginia (North Central)
- Bell County, Kentucky, and Pike County, Kentucky (Central)
- Johnson County, Tennessee, and Avery County, North Carolina (South Central)
- Noxubee County, Mississippi, and Lawrence County, Alabama (Southern)

3.3 Survey Design and Structured Focus Groups

Once identified, the ten project case study counties were then surveyed to help determine possible sources of growth and to help identify potential barriers to economic growth. A comprehensive survey was developed to measure a number of important factors and perception. Among the topics covered by the survey were local infrastructure, local governments, community organizations, public facilities, schools, environmental impacts, and social capital. In addition to traditional scaled responses, the survey also included several open ended questions. Some of these questions, such as "list three words to describe your county," were designed to help determine local attitudes and perceptions. Others such as "list your county's three top attractions for tourists and visitors" helped identify county-specific strengths and provided information for subsequent site visits.

From the outset, the survey delivery method was a topic of considerable discussion. Project staff acknowledged the benefits of a web-based survey approach. Internet surveys offer considerable benefits in terms of delivery time, data collection, and data analysis. Staff, however, also recognized the constraints present in many Appalachian communities. Rural Appalachian counties often have limited access to the Internet and in many areas few residents can obtain services at all. Staff also understood the digital divide that exists both geographically and generationally in the region. These limitations are often compounded in economically distressed communities.

For these reasons, the project staff feared that an Internet survey might, in effect, deny participation amongst the regions poorest and most isolated areas, communities whose voice was already largely unheard. While an Internet survey offered appealing cost and time savings, it was an approach that might further disenfranchise the region's most vulnerable residents and rob the study of vital data. Staff thus elected to conduct a traditional mail survey, which, while more time consuming and cumbersome, offered the advantage of accessibility to all residents of the region.

The project staff distributed 1,000 surveys among the ten counties, mailing 100 surveys to each case study county. Survey recipients included local officials, business owners, educators, health care providers, ministers, nonprofit organizations, and other key stakeholders and community leaders. Surveys were tracked using codes affixed to each questionnaire and responses were anonymous. Surveys were distributed according to the following process:

1. Recipients were first mailed a postcard, explaining the survey and informing them that they would receive the survey questionnaire in the next few days.
2. About 3 days later, recipients received a survey packet containing a questionnaire, cover letter, and a pre-paid return envelope.
3. After two weeks, non-respondents received a reminder postcard, asking them to return the completed survey.
4. After about a month, additional surveys were distributed to counties with low response rates.

To encourage participation, respondents were given the opportunity to participate in a drawing for an Apple iPad2 (value approximately $500). Survey packets included a reply card, to be returned with the completed survey, for a random drawing. To maintain anonymity, completed surveys and iPad reply cards, which had names and contact information were separated once survey packets were opened. All the iPad reply cards were placed in a box until all surveys had been returned. Almost 75% or participants elected to participate in the drawing, which was won by a participant from Bell County, Kentucky.

The overall return rate for all counties was exactly 25% (250 completed surveys were returned and tabulated). Return rates for individual counties varied considerably from a high of 38% (Morgan County, WV) to a low of 15% (Lawrence County, AL). Approximately 5% of surveys were returned by the postal service as undeliverable. One survey was returned with no responses. Another recipient returned two identical surveys, one of which was discarded.

Survey responses were compiled using Survey Crafter Pro, a survey development and analysis program. Responses were assessed and cross tabulated by county, ARC Subregion, and ARC economic status (distressed versus non-distressed). A survey form and survey results can be found in Appendices E and F.

3.3.1 Historical, Descriptive, and Demographic Data

While developing the survey, the project staff also created a descriptive profile of each community using a wide range of data sources. These included census records, school test scores, county literacy reports, county health profiles, environmental risk assessments, county highway maps and traffic counts, and USDA food environment statistics. Important data sets, including educational attainment rates, income statistics, No Child Left Behind test scores, and historical population trends were then charted and analyzed using Excel spreadsheet software.

Project staff also conducted a Lexis-Nexis news search for each county and key communities. This search compiled news stories from the counties and helped identify potential local issues. This search was supplemented by Internet searches using Google and Google Earth.

3.3.2 Site Visits, Interviews, and Structured Focus Groups

From the outset, an important assumption of this study was that some information that was essential to understanding local conditions could not be readily detected using statistical models, secondary data, or primary data collected through surveys. Significant descriptive and qualitative information can only be collected by visiting the community, observing local conditions, and by actually talking directly with local residents. For these reasons, the project team conducted site visits to each of the case study counties where they collected observational data, conducted interviews, and facilitated local focus group discussions.

Project staff conducted site visits from November 2010 until June 2011. Each day-long visit included a morning interview with local elected and appointed officials and an afternoon focus group with an assortment of community stakeholders. The morning meeting typically included discussions with local elected and appointed officials, Chamber of Commerce leadership, local economic development leadership, and other participants invited by local officials. Afternoon focus groups included a wider range of community stakeholders, including educators, housing officials, business owners, public health representatives, university extension staff, local professionals, and representatives from faith and nonprofit organizations.

Discussions covered both county-wide issues, such as access to healthcare, wastewater issues, and broadband; and region-wide issues, such as coal production (Central subregion), geologic hydrraulic fracturing (North subregion), and race relations (South subregion). Larger macro-economic issues were broached as well, with often healthy discussion around globalization, aging populations, and the national economy as it relates to the county.

Discussions usually began with broad topics, like, "How do you feel your county is doing," and then focused on specific concerns such as education, substance abuse, or entrepreneurship. Responses and reactions were recorded in detail on paper for later analysis. While direct quotes were gleaned from these talks for analysis and for use in this report, they were not, in the interest of confidentiality, attributed to any particular participant.

During the site visits, the project staff also collected observational data. The site visit team, for example, noted road conditions, driving times, the quality of local tourism assets, cell phone access, and local fuel prices. When possible, the team also visited local attractions and local businesses. Team members also photographed site communities for later analysis. Select photographs that were taken during site visits are included in the report and are the work of study authors, unless otherwise noted.