INCOME VOLATILITY AND COUNTY SIZE: A STUDY OF NORTH CAROLINA COUNTY POPULATION SIZE AND PER CAPITA PERSONAL INCOME VOLATILITY

by

Allison Carman

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Economics

Charlotte

2016

Approved by:

____

Dr. Hwan C. Lin

____

Mr. Azhar Iqbal

____

Dr. Lisa Schulkind

ABSTRACT

ALLISON CARMAN. Income volatility and county size: a study of North Carolina county population size and per capita personal income volatility. (Under the direction of DR. HWAN C. LIN)

This investigates the relationship between North Carolina counties and patterns in per capita personal income volatility over various samples ranging from 1969 to 2014. The overarching hypothesis is that people in smaller counties experience increased personal income volatility compared to people in larger counties. Using real, annual, county-level data, several regressions are performed to identify patterns in the size-volatility relationship. The explanatory variables in these models include two proxies for county size as well as employment types and distance. Additionally, Granger causality tests are employed to investigate whether large counties impress their supposed smoother volatility on surrounding counties.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

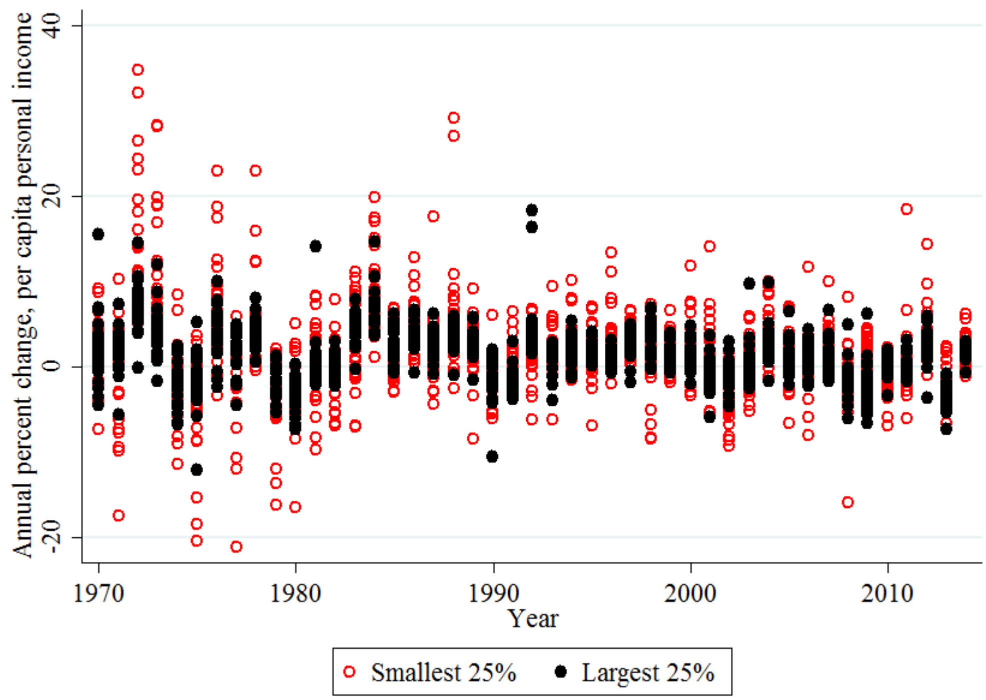

Figure 1.1: Small versus large counties: true population, 1970-2014.

Economy size has been a focus in economic research because of its policy implications. Specifically, there may be inequality between the two extremes: small and large economies. With county size as a motivating factor, this analysis focuses on the patterns and potential causes of per-person personal income volatility in North Carolina counties.

The per-person component of the data is determined using two per capita measures: the first is a true per capita measure, relying on annual population data; the other utilizes the labor force as the sample population. Volatility is measured as the standard deviation of annual percentage change observations of per-person personal income. Five samples are used, all within the range of 1969 to 2014, and volatility is computed over the length of each sample.

For a sense of the general problem, see Figure 1.1. This plot depicts the annual percentage change in real per-person personal income values in the largest and smallest counties in North Carolina from 1970 to 2014[1]. Here, large and small counties are respectively defined as those that fall above the 75th percentile and at/below the 25th percentile of average county-to-state annual population ratios.

There appears to be higher volatility in small counties (marked in red) than in large counties (in black). Not only are several outliers exhibited by small counties, but the overall distribution is much wider than the tight range that is found in observations from large counties. A preliminary regression quantifies these patterns by measuring personal income volatility as a function of dummy variables that indicate whether a county is small (in the lowest 25%) or large (the highest 25%). The excluded dummy variable is that of the average county size (the middle 50%). The results show that small counties clearly experience increased volatility, but there is less of a predictable impact on expected volatility in above-average sized counties. The question is whether this relationship can be consistently quantified and whether the observed volatility is also explained by factors other than a county's population.

The statistical analysis consists of three main steps. First, a simple, pooled ordinary least squares regression model is developed to measure the relationship between county per capita personal income volatility and county size. There are two proxies for county size, so there are two sets of results for this regression. One of these proxies is annual county population, in hundreds of thousands, where population is measured as the labor force or the true population, depending on the sample used. The other size proxy is the annual percentage of the total North Carolina population or labor force. These two proxies are used to ensure that results are robust and that disproportionate growth does not skew the outcome. The results consistently show that county size is a significant determinant of in-

[1]For the plots that correspond to the remaining four samples, see Appendix A.

come volatility. Furthermore, as anticipated, the two variables have a negative relationship, showing that there is a size-volatility tradeoff.

The next step in the quantitative process employs a multiple regression model that encompasses information other than county size. Annual data on employment composition are included to provide an indication of the impact that industrial composition has on volatility. Additionally, information is included on the distance between a county and the nearest economically important county. This follows the notion that rural counties are at a disadvantage and experience heightened volatility because of their isolation. Again, there are two sets of results because of the two size proxies. As with the previous regression analysis, the results are similar for both size proxies. County size is no longer a strong determinant of volatility, but the employment variables show that farm employment and nonfarm proprietors employment generally increase volatility, while farm proprietors employment, private nonfarm employment, and government employment all consistently decrease volatility. Additionally, an increased distance from economic centers is effectively shown to have a positive impact on volatility.

The final aspect of the analysis encompasses Granger causality tests to measure causality between changes in per-person personal income in economically central counties and their surrounding counties. The results show no clear pattern of personal income changes radiating outward from central counties, nor is there evidence of personal income changes moving into central counties from surrounding counties. Combined with the regression results, the relationship between distance and volatility is not consistent. However, it is possible that being very far from an economic hub is consequential, but being near an economically central county is neither beneficial nor detrimental. This aligns with the preliminary dummy-variable regression that shows small counties to have consistently above-average volatility and larger counties to have wholly unpredictable volatility in comparison to the average-sized counties.

Aggregating these results shows that qualities generally associated with small counties, such as more farm employment and a rural location, lead to heightened volatility, but size is not a thoroughly significant determinant of volatility. One is able to conclude that it is not population itself, but other elements of small counties, that results in exacerbated per-person personal income volatility.

CHAPTER 2: PREVIOUS LITERATURE

There is much literature on income growth and its causes, but the focus on income volatility is limited. Nonetheless, a few consistencies can be observed in this area of research. First, there is significant interest in the relationship between a geographical division's industrial structure and its income patterns. Industrial diversity is shown to lead to more stable growth in wages and employment (Felix (2012)), as well as stability in personal income growth (Cortes *et al.* 2013). Grennes *et al.* (2010) find similar results regarding per capita personal income volatility. With a more specific focus, Shaffer (2009) shows that the presence of more, and larger, firms reduces personal income volatility. The general pattern is that stability is a direct result of industrial diversity, as specialization proves more risky. This idea even translates into portfolio theory, as Spelman (2006) finds that concentration in one industry (portfolio) leads to heightened income volatility (risk).

An additional hypothesis is that location affects growth and, therefore, volatility. A distance component is often included in regressions (Grennes *et al.* (2010), Cortes *et al.* (2013), Cortes *et al.* (2015)), as it is expected that rural areas lack accessibility to, and are independent of, economically central areas. To test whether these economic centers have an effect on other local areas, Voith (1998) measures the relationship between urban and suburban cities, finding that income growth in urban areas has a positive effect on suburban income growth. Presumably, this effect decreases when an area is located further away from a central city, hence the literature's interest in distance.

These theories about causes of volatility (industrial specialization, increased distance from urban areas) are not always examined separately. Cortes *et al* (2013) explore the relationship between personal income growth and several potential predictors, including industrial composition, distance, and regional control variables. They study a similar relationship in Cortes *et al* (2015), where the dependent variable is personal income volatility. From these two papers, they find that distance causes growth to decrease and volatility to increase.

Another important component of Cortes *et al.* (2015) is that of population. They find that micropolitan statistical areas with smaller populations experience elevated personal income volatility. This is similar to Easterly and Kraay (2000), who find that small nation states exhibit increased volatility of annual growth rates.

This study aims to merge these hypotheses—that both industrial composition and location with respect to central economic areas are factors that contribute to personal income volatility patterns—and apply them on the county level in North Carolina. This research contributes to the existing literature by providing a unique combination of geographical setting, dependent variable, and predictive methodology.

CHAPTER 3: DATA

3.1 Sources

The data used in this analysis come from three sources. The first, and most used, source is the Bureau of Economic Analysis (BEA). Through the Economic Profile (series CA30), the BEA provides annual, county-level data on employment, population, and income from 1969 to 2014; these are nominal data in the level form. From the Economic Profile, data on population and personal income are used. According to the BEA's *Local Area Personal Income Methodology*, the population component of this data set comes from the Census Bureau's annual midyear population estimates, and the personal income estimates are mostly based on administrative-records data. The data are in thousands of dollars.

An additional data set, Total Full-Time and Part-Time Employment by Industry (series CA25 and CA25N), is extracted from the BEA and contains county-level information on total employment by industry. It is worth noting that the industry classification system changes from the Standard Industrial Classification (SIC) to the North American Industry Classification System (NAICS) during the span of the data set. However, because the utilized statistics aren't industry-specific (such as "Retail" or "Forestry"), they aren't impacted by the shift in classification methodology, which is first implemented in 2001. The statistics used are annual, county-level, general employment indicators: farm proprietors employment; nonfarm proprietors employment; farm employment; private nonfarm employment; federal, civilian employment; and state and local employment. Total annual county employment is also used. Several additional statistics accompany these data, but they are excluded to avoid collinearity. The excluded statistics are total wage/salary employment, which is the complement to proprietors employment, and military employment, which is the complement to federal and state/local employment. Data on total annual employment are also extracted from this source.

The next source used is the Bureau of Labor Statistics (BLS). The BLS provides annual, county-level data on the labor force for years 1990-2016; the data are in level form and are not seasonally adjusted. The "Regional and State Employment and Unemployment" news release from the BLS indicates that the labor force is measured on a place-of-residence basis and consists of the civilian noninstitutional population age 16 and above.

The BLS also provides annual averages for the Consumer Price Index for All Urban Consumers (CPI-U). The CPI-U data begin in 1913 and continue through present-day; these data are used to adjust for inflation, which is discussed in the next section.

The final data source is the National Bureau of Economic Research (NBER), which provides a County Distance Database. According to the NBER, the distances are great-circle distances calculated using the Haversine formula. These data are for both the 2000 Census and the 2010 Census, for which the distances are different.

3.2 Transformations

3.2.1 Inflation

To capture a more accurate value of income volatility, the personal income data are adjusted for inflation. Using the CPI-U (obtained from the BLS), the level-form data are transformed using constant 2014 dollars. This follows recommendations in the BLS report "Math calculations to better utilize CPI data." The dollars are adjusted to the end-of-sample values for consistency between samples. Adjusting for inflation is important because income may appear to be volatile when, in fact, the volatility stems from small per-person incomes in earlier years versus larger per-person incomes in more recent years, a common effect of inflation. Without this adjustment, observed volatility may be the result of a change in spending power throughout the sample.

3.2.2 Per-Person

Inflation-adjusted income values in thousands of dollars do not offer much comparability in the level form. For instance, counties with more people (and, most likely, more total income) may experience a million-dollar rise in total personal income and be generally unaffected, while that same increase in a 100-person county would be outlandish. Therefore, the data are transformed into per capita values. This is done by dividing the total personal income in a county by its population for the same year. This step yields more opportunity for comparison. However, it is important to consider the two relevant ways to calculate per capita measures in this context.

The first is a true per capita value: each county's income measure is divided by its total population. This sort of per capita measure is useful as a quality-of-life indicator: does everybody have a more dependable income in one county compared to another? Using true per capita values facilitates a benchmark of the financial resources one would expect to be available to each resident of the county, despite their age or ability to work. Such an index looks at the well-being of every member of the population.

Another way to measure per capita personal income is by using the labor force rather than the population. This provides estimates of personal income for each able worker and will be hereafter referenced as "per-worker" values. This version of per capita measurement allows for the comparison of worker well-being: can a person expect certain income stability (or lack thereof) by working in a particular county?

3.2.3 Annual Percentage Change

Because high- or low-paying employment can prove more concentrated in certain areas, inter-county comparisons are difficult in the level form. Thus, the per-person income values are transformed into a percentage change from the previous year; this sacrifices one observation. Note that this transformation does not take away from the economic intuition of the analysis: the question is not whether some counties have higher incomes; rather, the interest is in whether different counties experience different income fluctuations, which would be most conveniently measured in the percent-change format.

3.3 Samples

As previously indicated, the BEA Economic Profile (overall population; personal income) spans from 1969 to 2014, while the BLS labor force values are only available for years 1990-2016. This makes direct comparison impossible without sacrificing over twenty observations. To facilitate comparison and robustness, five samples are used. The first sample, "(1)," is the BEA data from 1969 to 2014 (1970-2014 in percent-change). This ensures the use of all available information. The second sample, "(2)," is the BEA data for years 1990-2014. The data are trimmed to this sample so as to allow for a direct comparison with the third sample, "(3)," the full BLS data set (1990-2014). The fourth sample, "(4)," is included to provide relevance to current times and consists of the last ten observations of the BEA sample (when the data are in level form, this sample covers years 2004-2015; in percent-change form, years 2005-2014). The final sample, "(5)," is the same as the fourth, just with BLS data. In sum, there are five sample sets: the general population percent-change samples are provided for years 1970-2014, 1991-2014, and 2005-2014, while the labor force population percent-change samples cover years 1991-2014 and 2005-2014.

Using these population samples specifically allows for two comparisons of economic implications. First, one can compare the results that use the same data over different time periods: do the true per capita data exhibit different patterns when the sample size is shorter/longer? How about the labor force data? Additionally, one can compare the results observed from the different population measures over the same sample years: in overlapping samples, what are the differences resulting from the true population versus the labor force?

Another benefit of using multiple samples is that there are implicit robustness checks. If a result exists in one sample but not the rest, it probably isn't representative of true patterns. However, if a result appears in all five samples, it is worth investigating. Thus, most of the econometric methods employed in this paper encompass all five samples.

CHAPTER 4: METHODOLOGY

4.1 Proxies

In this case, a county's annual population or labor force is used as the determinant of its size. These data are used in both the level form (in hundreds of thousands of people) and as a percentage of the total North Carolina population or labor force, respectively, for the given year. The percentage ratios are multiplied by 100 to provide a true percentage.

4.2 Volatility Measure

Volatility is measured as the standard deviation of the annual percent-change in per-person personal income for each county over the given sample. This is the typical way to measure volatility in this area of research[1]. Volatility is the only dependent variable in the regression analysis component of this study, and it differs from sample to sample (visible in Table 4.1) because it is calculated over each sample's time span.

Table 4.1: Summary statistics: per-person personal income volatility.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Mean	3.8716	2.7753	3.6627	2.6019	3.3263
SD	1.5465	0.9907	1.4033	1.0632	1.6039
Min	2.2839	1.5075	1.6794	1.1935	1.0348
Max	10.5866	7.7846	10.2089	9.4308	11.0731

4.3 Distance

The earlier-described NBER data on county distances are used to measure the potential impact of being near an economic hub. These data are included to facilitate an analysis

[1] See Cortes *et al.* (2015), Easterly and Kraay (2000), Felix (2012), Grennes *et al.* (2010), and Spelman (2006).

of the effects of living near (or far away from) a county with a substantial economy. In the case of North Carolina, Mecklenburg county, which contains Charlotte, is a driving economic force, as well as the Research Triangle, which primarily consists of Durham, Orange, and Wake counties because of Duke University, the University of North Carolina at Chapel Hill, and North Carolina State University, respectively. A close proximity to one of these economically central counties could have an impact on personal income volatility.

To test this hypothesis, a county's distance from the nearest economic point of interest is measured. The NBER data set features the distance between each possible county combination in North Carolina. However, the only combinations of economic interest are those that include one of the four above-mentioned economically significant counties. All other observations are dropped for irrelevance.

After filtering out the irrelevant county pairings, the remaining distances are compared for each county. The distance observation of interest is that which is smallest. Consider Cabarrus county: because it borders Mecklenburg county, that distance is likely more relevant than the distances to Durham, Orange, or Wake counties, so these last three values are removed from the data set.

These calculations are made for both the Census 2000 and the Census 2010 distance data sets. The distance values for each census year differ slightly, most likely from changing county borders. However, the nearest economically significant county assigned to each county does not change. Also, North Carolina does not gain or lose any counties during any of the samples. Thus, to create a more comprehensive view of a county's distance to the nearest economically central county, the average is taken between the 2000 and 2010 distance computations. Note that the four counties considered economic centers are assigned a value of zero so that they will still be included in regression analysis, but the distance variable will drop out by default. All of the utilized distance observations can be found in Table C.1 in Appendix C.

4.4 Employment Statistics

As previously indicated, several employment statistics are included in a portion of this analysis. These values are originally in the form of the total number of employees per category, per county, per year. The data are not adjusted for changes in population, which means that population growth trends are possible. Thus, the employment data are transformed into an annual percentage of total county employment. This eliminates any trend-like population effects, as a change in population does not directly affect the change in the proportion of people employed in each category.

This use of employment statistics is similar to the USDA's Economic Research Service (ERS) county typology, which determines industrial dependence based on a county's employment composition (ERS (2016)). However, the ERS does not determine these classifications annually, which is necessary for this analysis.

Table 4.2 provides insight to the general employment climate: because the time span is shortening from the top of the table to the bottom, one can observe the evolution of the workforce. For instance, as the sample size becomes shorter and earlier observations are dropped, the mean percentages of farm and farm proprietors employment decrease, while nonfarm proprietors and private nonfarm employment both exhibit an increase in their employment shares; the average government shares of employment are relatively stable. Note that these values don't depend on population or the labor force—they are calculated in terms of employment.

4.5 Types of Data

In all, each sample's data set contains both panel and cross-sectional data. The panel data included are annual per-person personal income values, head-counts of each county's population or labor force, the corresponding percentage of the state's total population or labor force, and employment composition percentages. The NBER average distance is the only time-invariant predictor that is used. Each observation contains a county's name, Fed-

Table 4.2: Summary statistics: employment.

	Obs	Mean	SD	Min	Max
Years 1970-2014					
Farm proprietors	4500	4.604527	5.24274	0	41.14151
Nonfarm proprietors	4500	17.11884	7.697058	3.432057	55.59748
Farm	4500	7.281659	8.095584	0	56.64107
Private nonfarm	4500	76.44205	11.53376	20.1894	92.27259
Federal, civilian	4500	1.174346	1.625053	0.1637609	16.71965
State and local	4500	13.15201	5.550839	3.032251	46.82147
Years 1991-2014	Obs	Mean	SD	Min	Max
Farm proprietors	2400	2.666462	2.557801	0	22.40437
Nonfarm proprietors	2400	20.1172	8.369385	6.61383	55.59748
Farm	2400	4.1491	3.849882	0	25.38642
Private nonfarm	2400	79.05556	9.022716	37.05478	91.61767
Federal, civilian	2400	1.037277	1.406901	0.1637609	11.91797
State and local	2400	14.14864	5.634625	5.803383	45.17919
Years 2005-2014	Obs	Mean	SD	Min	Max
Farm proprietors	1000	2.05753	1.793627	0	8.866316
Nonfarm proprietors	1000	23.47681	8.976641	8.509226	55.59748
Farm	1000	3.198775	2.733801	0	13.40652
Private nonfarm	1000	79.73134	8.43873	41.18712	91.421
Federal, civilian	1000	0.9564874	1.345123	0.1810865	9.723082
State and local	1000	14.60068	5.776776	5.94569	45.17919

eral Information Processing Standard (FIPS) code for reference purposes, annual percent-change in per-person personal income, two size proxies, BEA employment statistics and NBER distance information.

4.6 Visual Assessment

As previously explained, the dependent variable in question is the volatility of per-person personal income. For context, if the standard deviation of the annual percent-change in per-person personal income in a county is 15, a person can expect to see a very large fluctuation in the percent change of their personal income from year to year. That is, it is not uncommon to experience a 15% increase or decrease annually in personal income, making it very unpredictable. If that same measure is 2 in another county, there is a lot more smoothness from period to period, so a person shouldn't expect to see their personal income rise or fall very much. The latter is preferable and indicates stability, as a standard deviation measures the deviation from the mean.

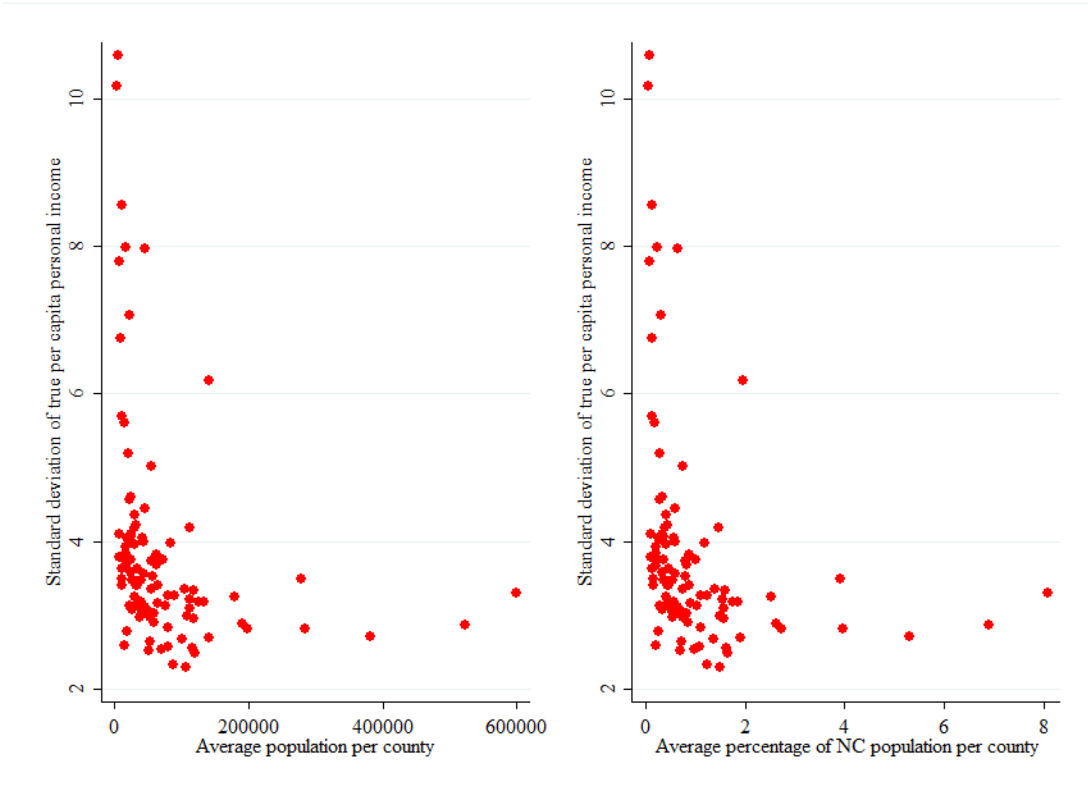

Figure 4.1: Personal income volatility: true population, 1970-2014.

The plot featured in the introduction shows that there is generally an inverse relationship between size and volatility: more heavily populated counties are characterized by smoother income values, while the opposite is true for several smaller counties. However, the introductory plot uses time as the independent variable and the percent-change in income as the dependent variable. This analysis is interested in showing a similar pattern while using a size proxy as the independent variable and volatility itself as the dependent variable.

See Figure 4.1 for observations of personal income volatility in the full, true per capita sample. These observations provide a rudimentary argument for the hypothesis that income volatility and county size are inversely related. There is much vertical movement near the y-axis in both subplots, showing that lower population values are generally associated with higher standard deviations of personal income. However, additional exploration is necessary. It is worth noting that, although each subplot features a different size proxy, the hierarchy of county size generally remains consistent between both size measures (county population in the level form and the percentage of the state's total population); thus, the subplots appear nearly identical.

4.7 Regression Analysis

4.7.1 Size and Personal Income Volatility

To quantify the size-volatility relationship, a simple pooled regression model is constructed:

$$\text{SD}_i = \beta_0 + \beta_1 \text{Size}_j,$$

where *i* represents the sample used to calculate volatility and *j* indicates that there is more than one size proxy to choose from. With *Size* as the only independent variable, a pooled regression[2] is modeled for each sample. This provides a very basic perspective on how a county's size is related to the volatility of per-person personal income. Note that pooled regression is mandated by the time-invariance of the dependent variable.

Similar to categorizations in Spelman (2006), an additional model is developed to provide insight regarding the role of each population extreme: the top and bottom 25%. As with the introductory example, a county is considered "Small" if it falls at or below the 25th percentile of the average of counties' annual percentages of the total North Carolina population; a county is considered "Large" if its population lies above the 75th percentile. This regression features dummy variables, and if neither dummy variable is satisfied, the county is considered to be an average size. The model is as follows:

$$\text{SD}_i = \beta_0 + \beta_1 \text{Small}_i + \beta_2 \text{Large}_i,$$

where *i* indicates that this analysis is performed over each sample and the percentile calculation results may vary. The resulting coefficients, exhibited graphically in Figure 4.2,

[2] All regressions in this analysis use White's robust standard errors.

are expected deviations from the volatility observed in average-sized counties. The coefficients, listed in Table 4.3, are statistically significant in all but one case, which is the coefficient on *Large* in the second sample. It is clear that small counties consistently exhibit higher volatility, while large counties are much less predictable in their volatility patterns. This concurs with related literature, where the focus is generally on small-economy volatility rather than large-economy stability.

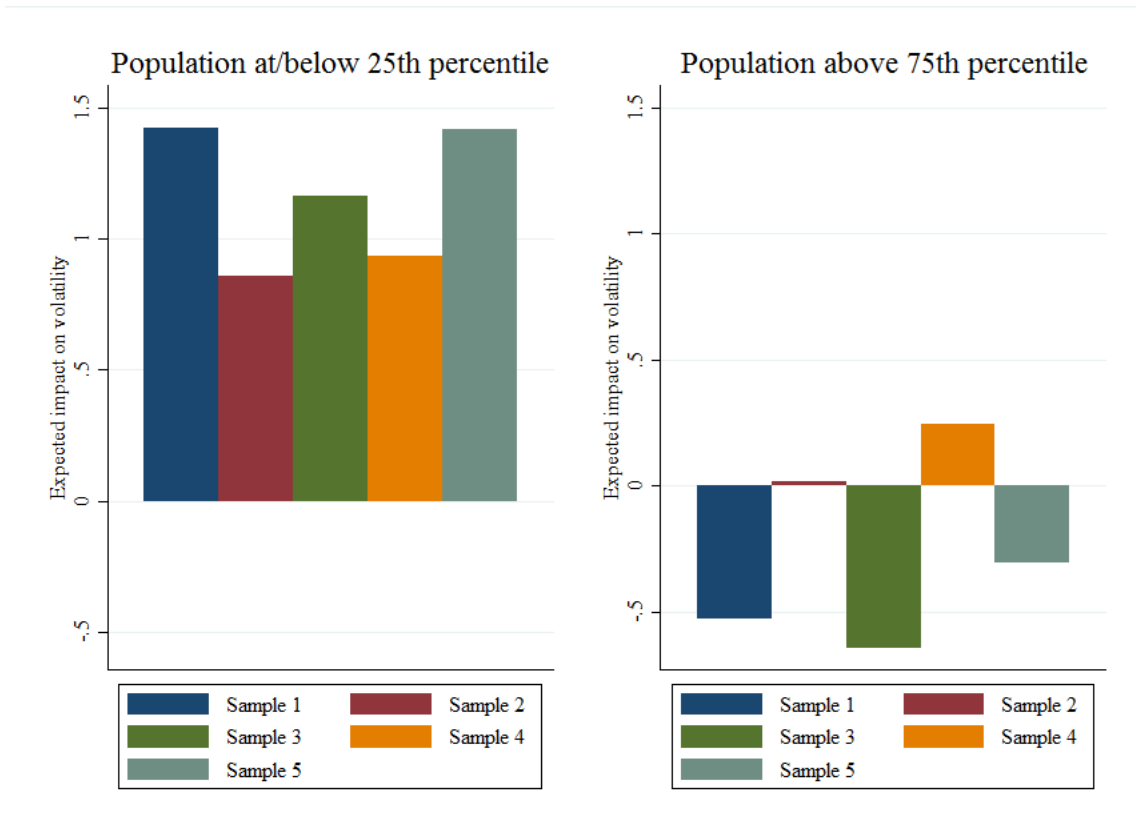

Figure 4.2: Expected volatility impact of large and small populations.

4.7.2 Multiple Regression Model

To simultaneously use all of the data's predictive power, a multiple regression model is developed. This pooled ordinary least squares regression makes use of variables other than the county size proxies, although *Size* is still included. First, the model features the quasi-industrial classification data used in this paper: farm proprietors employment and farm employment indicate whether a county has a high or low focus on farming; nonfarm

Table 4.3: Dummy-variable regression: county size and volatility.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Small	1.424***	0.855***	1.162***	0.936***	1.418***
	(0.0694)	(0.0585)	(0.0727)	(0.109)	(0.156)
Large	-0.527***	0.0152	-0.643***	0.243***	-0.305***
	(0.0301)	(0.0356)	(0.0447)	(0.0493)	(0.0660)
Constant	3.647***	2.558***	3.533***	2.307***	3.048***
	(0.0203)	(0.0211)	(0.0381)	(0.0251)	(0.0500)
N	4500	2400	2400	1000	1000
adj. R^2	0.222	0.139	0.217	0.130	0.175
F	452.6	110.5	417.0	44.59	66.07

Legend: * $p<0.05$; ** $p<0.01$; *** $p<0.001$; standard errors in parentheses.

proprietors employment and private nonfarm employment indicate the concentration on industries other than farming; and federal, civilian employment as well as state and local employment represent whether a county relies on government employment. Additionally, the distance to the nearest economically important county is included as an explanatory variable.

$$\begin{aligned} \text{SD}_i = \beta_0 + \beta_1\text{Size}_j + \beta_2\text{FProp} + \beta_3\text{NFProp} + \beta_4\text{Farm} \\ + \beta_5\text{PrivateNF} + \beta_6\text{Federal} + \beta_7\text{State} + \beta_8\text{Distance} \end{aligned}$$

In the above equation, *FProp* and *NFProp* indicate farm proprietors and nonfarm proprietors employment, respectively, *Farm* represents farm employment, *PrivateNF* represents private nonfarm employment, *Federal* indicates federal, civilian employment, *State* represents state and local employment, and *Distance* indicates the average between the 2000 and 2010 distance measures to the nearest economically important county. Again, the subscript *i* shows that volatility is calculated over each sample, and *Size* is given the subscript *j* because models are separately performed using the two size proxies.

Because there are five samples and two proxies for county size, ten regressions are conducted. Pooled OLS is selected as the best method for two reasons. First, the dependent variable is time-invariant within each sample. Thus, fixed effects would not make sense. Further, this study aims to see whether county size and other factors are invariably associated with different levels of personal income volatility. Using pooled regression analysis allows population and employment data to be seen as truly independent factors, and the idea that these variables may experience growth between the beginning and the end of the sample only enhances this independence. The results of each regression are discussed in the next section.

4.8 Granger Causality

This analysis has hypothesized that the distance from major economic counties may be a major factor in personal income volatility. This theory assumes that population is simply a proxy for another quality found in more volatile counties.

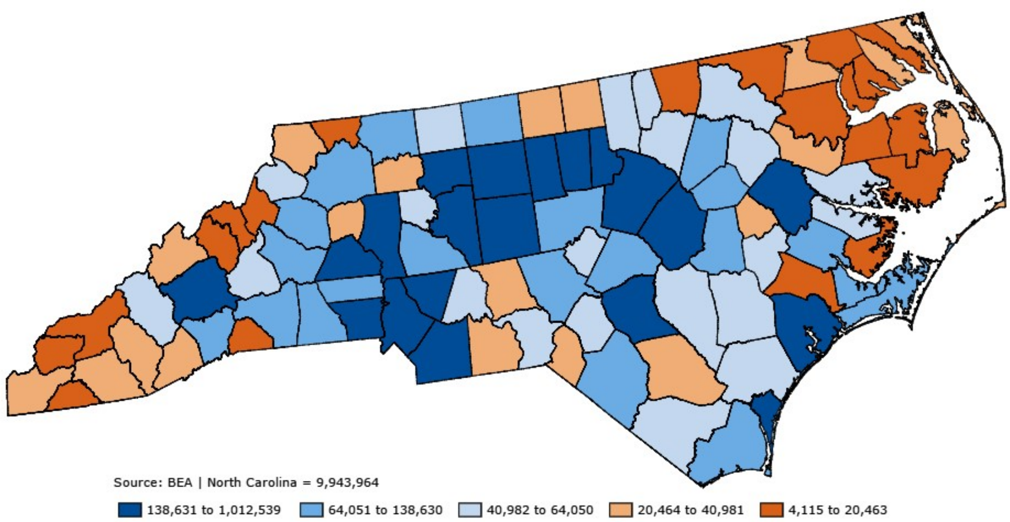

Figure 4.3: Population by county, 2014.

Figure 4.3 shows that, in general, areas with high populations are clustered in the center of the state. However, toward the state's borders, the observed populations get smaller and smaller. Perhaps larger counties are less volatile because of an interdependence, as they are generally in close proximity to one another. Furthermore, if volatility patterns are similar and interdependent in the counties of concern, it follows that this relationship should also be apparent in the underlying annual percent-change data. For instance, if a county experiences a large increase in income one year, does a surrounding county exhibit a similar pattern soon thereafter?

To test this dependence, several Granger causality tests are employed. These tests measure two types of influence: do the more central, economically important counties set the trend in personal income changes for surrounding counties, or does this occur in the opposite direction? While the former is most probable, it is not inefficient to simultaneously test both directional possibilities.

There are 22 counties involved in this testing procedure. Four of these counties have been discussed as economically important: Mecklenburg, Durham, Orange, and Wake. Determining their neighboring counties is more subjective. The Charlotte Chamber of Commerce considers Mecklenburg's surrounding counties to be Cabarrus, Gaston, Iredell, Lincoln, and Union—the counties that touch Mecklenburg more than tangentially. On the other hand, the Research Triangle area is defined as more of a region, so all counties considered part of the region are included in the analysis: Chatham, Edgecombe, Franklin, Granville, Harnett, Johnston, Lee, Moore, Nash, Person, Vance, Warren, and Wilson.

It is important to recall that no time-series analysis has been implemented prior to this section; all variables in the preceding regression analysis are pooled panel data or cross-sectional observations. However, Granger causality innately requires formal time-series analysis, as it questions whether one variable can incite changes in another across time periods. Therefore, only the larger three samples are used. The two samples that cover only years 2005-2014 are excluded, as they may lack meaning in this context after subtracting the necessary degrees of freedom. In sum, the final data set used in this analysis consists of 22 counties over three samples.

Following the methods employed by Silvia *et al.* (2014), there are three main steps involved in testing for Granger causality. First, it is necessary to check that the data are stationary. This is especially important given the time-series nature of the analysis rather than the previous pooled regression analyses. Each series is tested using the Augmented

Dickey Fuller (ADF) unit root test, where the null hypothesis is that there is a unit root (the data are nonstationary). In full, 66 ADF tests are performed: one for each county in each of the three samples. The null hypothesis of a unit root is rejected with 95% confidence in all 66 series. These results meet expectations: the data are in year-over-year percent-change form, which is one common way to avoid a unit root problem. Additionally, the data are inflation-adjusted, so trend possibilities are not as high as they would be using real dollars.

Once the data are confirmed to be stationary, several vector autoregressive (VAR) models are tested. Each model includes a central county and one of its surrounding counties. This means, for example, that there are five models dedicated to testing Mecklenburg county's relationship with each of its five neighbors. The surrounding counties that are part of the full Research Triangle region are tested against the nearest central county, which was determined when forming the distance data set for use in the previous section's multiple regression. Thus, a county that is part of the Triangle but is not Durham, Orange, or Wake county is paired with whichever of the three it is nearest.

For each sample, 18 VAR models (for the 18 surrounding counties) are tested at 4 different lag values. The appropriate lag order must be established prior to determining Granger causality results, as results may be incorrect if the wrong lag is used in causal analysis. To determine lag order, the Schwarz Bayesian criterion (SBC) is used for model selection. For each VAR model, the lag order that results in the lowest SBC is chosen and the others are eliminated. It is important to understand the context of this model selection. Because the data consist of annual observations, implementing a VAR model with four lags implies that changes in per-person personal income still have an impact four years later. This is not as logical as a model with one lag, which is more probable and implies a shorter memory of changes in personal income.

After the correct lag is chosen, causal analysis can be accurately performed and interpreted. Two Granger causality tests are conducted within each model. The first test features the central county (Mecklenburg, Durham, Orange, or Wake) as the first group and the surrounding county as the second group; the second test has these roles reversed. The null hypothesis is that the second group does not Granger-cause the first group. Therefore, the null hypothesis of the first test suggests that changes in per-person personal income in a surrounding county do not cause similar changes in the nearest central county. The second test's null hypothesis is that changes in per-person personal income do not move outward from a central county to nearby counties. Again, it is expected that the central counties are influential, not influenced, so the second test is theoretically more likely to be rejected in favor of the central counties setting the standard for changes in personal income. The results are discussed in the next section.

CHAPTER 5: RESULTS

5.1 Regression Analysis: Size and Personal Income Volatility

Tables 5.1 and 5.2 display the results of simple regression analysis[1] between personal income volatility and each county size proxy. Except for in the fourth sample, the resulting coefficients on *Size* are all positive and statistically significant. The coefficients are similar between size proxies, which indicates that counties have generally grown at proportionate rates; this similarity also shows that a 1% increase in the percentage size proxy is similar to a 100,000-person increase in the level-form size proxy.

Table 5.1: Simple regression: population percentage as size proxy.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (%)	-0.334***	-0.0418**	-0.245***	0.0295	-0.160***
	(0.0181)	(0.0129)	(0.0207)	(0.0210)	(0.0288)
Constant	4.205***	2.817***	3.908***	2.572***	3.487***
	(0.0328)	(0.0271)	(0.0374)	(0.0459)	(0.0659)
N	4500	2400	2400	1000	1000
adj. R^2	0.077	0.003	0.069	0.001	0.025
F	339.9	10.45	139.5	1.967	31.02

Legend: * p<0.05; ** p<0.01; *** p<0.001; standard errors in parentheses.

The results show that, on the average, as a county's size increases, there is a negative effect on the expected volatility of per-person personal income. This aligns with expectations. Table 5.3 displays the expected impact on volatility based on a one-standard-deviation[2] increase in each population variable. It is clear that the size-volatility relationship is most prominent in the first sample, followed by the third and fifth samples. This shows that the

[1] All regression computations were performed in StataIC 14 using Windows 10.

[2] See Table 5.8 for sample standard deviations.

Table 5.2: Simple regression: annual population as size proxy.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (#)	-0.401***	-0.0465**	-0.563***	0.0313	-0.349***
	(0.0273)	(0.0149)	(0.0506)	(0.0221)	(0.0630)
Constant	4.161***	2.814***	3.899***	2.572***	3.486***
	(0.0329)	(0.0268)	(0.0375)	(0.0458)	(0.0659)
N	4500	2400	2400	1000	1000
adj. R^2	0.068	0.003	0.067	0.001	0.025
F	215.5	9.658	123.6	1.993	30.63

Legend: * p<0.05; ** p<0.01; *** p<0.001; standard errors in parentheses.
Note: annual population is measured in hundreds of thousands.

trimmed true-population samples do not exhibit as much of a size effect as does the full sample; additionally, the labor force data allow the size effect to be more pronounced than the true population over the same sample periods. These magnitude disparities correspond with higher mean values of volatility in the first, third, and fifth samples[3].

Table 5.3: One-standard-deviation increase (from Tables 5.1 and 5.2).

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (%)	-0.4293***	-0.0579**	-0.3677***	0.0437	-0.2572***
Size (#)	-0.4011***	-0.0556**	-0.3616***	-0.0438	-0.2571***

Legend: * p<0.05; ** p<0.01; *** p<0.001; standard errors in parentheses.

5.2 Regression Analysis: Multiple Regression Model

For each sample, a multiple regression model is conducted using county size, employment composition, and distance to the nearest economically central county as independent variables. As with the simple regression model, the results are extremely similar across size proxies, showing that population growth has generally occurred proportionately across counties. In total, only five coefficients do not test significant: for both proxies, size is

[3]See Table 4.1.

Table 5.4: Multiple regression: population percentage as size proxy.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (%)	-0.0311**	0.119***	0.00544	0.175***	0.152***
	(0.0107)	(0.00903)	(0.0141)	(0.0180)	(0.0202)
Farm prop	-0.303***	-0.358***	-0.290***	-0.431***	-0.459***
	(0.0124)	(0.0234)	(0.0326)	(0.0705)	(0.0855)
Nonfarm prop	0.0528***	0.0304***	0.0221***	0.0147***	0.0332***
	(0.00271)	(0.00156)	(0.00286)	(0.00396)	(0.00710)
Farm	0.221***	0.277***	0.290***	0.333***	0.453***
	(0.00987)	(0.0199)	(0.0243)	(0.0611)	(0.0706)
Private nonfarm	-0.0531***	-0.0635***	-0.0373***	-0.0473***	-0.0157*
	(0.00266)	(0.00224)	(0.00351)	(0.00485)	(0.00656)
Federal	-0.0696***	-0.125***	-0.0964***	-0.196***	0.0191
	(0.00838)	(0.00886)	(0.0212)	(0.0202)	(0.0364)
State	-0.0304***	-0.0641***	-0.0140**	-0.0378***	0.0244**
	(0.00375)	(0.00331)	(0.00464)	(0.00630)	(0.00855)
Distance	0.00402***	0.000779*	0.00483***	0.00246***	0.0100***
	(0.000457)	(0.000369)	(0.000624)	(0.000706)	(0.00139)
Constant	7.025***	7.852***	5.666***	6.229***	2.009**
	(0.275)	(0.242)	(0.365)	(0.518)	(0.688)
N	4500	2400	2400	1000	1000
adj. R^2	0.480	0.488	0.350	0.280	0.348
F	260.2	226.6	111.7	42.27	27.37

Legend: * $p<0.05$; ** $p<0.01$; *** $p<0.001$; standard errors in parentheses.

not a significant determinant of volatility in the third sample, and the fraction of federal employment is not significant in the fifth sample. When the size proxy is the level form of population, distance is not statistically significant in the second sample.

Tables 5.4 and 5.6 display the multiple regression results for both size proxies—the percentage and level forms of the sample population, respectively. Additionally, Tables 5.5 and 5.7 express the results in terms of a one-standard-deviation[4] increase in each variable. This provides a perspective on the magnitude of each coefficient and the overall impact each variable has on the volatility of per capita personal income.

[4]The sample standard deviations can be found in Tables 4.2 and 5.8.

Table 5.5: One-standard-deviation increase (from Table 5.4).

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (%)	-0.0400**	0.1648***	0.0082	0.2589***	0.2444***
Farm prop	-1.5886***	-0.9157***	-0.7418***	-0.7731***	-0.8233***
Nonfarm prop	0.4064***	0.2544***	0.1850***	0.1320***	0.2980***
Farm	1.7891***	1.0664***	1.1165***	0.9104***	1.2384***
Private nonfarm	-0.6124***	-0.5729***	-0.3365***	-0.3992***	-0.1325*
Federal	-0.1131***	-0.1759***	-0.1356***	-0.2636***	0.0257
State	-0.1687***	-0.3612***	-0.0789**	-0.2184***	0.1410**
Distance	0.1625***	0.0315*	0.1952***	0.0994***	0.4041***

Beginning with population, one can see that its impact on volatility varies across samples, and these variations are similar in both size proxies. First, the coefficients are significant in all but the third sample. Of the significant coefficients, three are positive (samples 2, 4, and 5) while one is negative (sample 1). This mix of coefficient signs and significance does not yield a consistent interpretation about the size-volatility relationship.

Farm proprietors, nonfarm proprietors, farm, and private nonfarm employment types all exhibit significance in every sample, as well as consistent signs across each sample and size proxy. Farm proprietors employment is negative, which means that an increase in this employment's share of the total employment yields a decrease in volatility. This does not make economic sense. Following theory, an employment concentration in one industry leads to a higher dependence on that industry's success and, therefore, higher income volatility. One would expect that a growing fraction of farm proprietors employment would be associated with increased volatility, but the coefficients show otherwise. It is also clear in Tables 5.5 and 5.7 that the expected impact on volatility is not small in magnitude, as a one-standard-deviation increase in farm proprietors employment for the first sample results in an expected decrease of about 1.58 in volatility—more than volatility's standard deviation, visible in 4.1—for both size proxies.

A similar phenomenon occurs with farm wage and salary employment, except the coef-

Table 5.6: Multiple regression: annual population as size proxy.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (#)	-0.0321*	0.131***	0.0134	0.183***	0.329***
	(0.0125)	(0.0105)	(0.0322)	(0.0190)	(0.0442)
Farm prop	-0.302***	-0.359***	-0.290***	-0.432***	-0.459***
	(0.0124)	(0.0234)	(0.0326)	(0.0705)	(0.0855)
Nonfarm prop	0.0532***	0.0296***	0.0220***	0.0145***	0.0332***
	(0.00268)	(0.00155)	(0.00285)	(0.00396)	(0.00710)
Farm	0.221***	0.277***	0.290***	0.333***	0.453***
	(0.00986)	(0.0199)	(0.0243)	(0.0611)	(0.0706)
Private nonfarm	-0.0530***	-0.0636***	-0.0373***	-0.0473***	-0.0157*
	(0.00269)	(0.00221)	(0.00351)	(0.00483)	(0.00656)
Federal	-0.0701***	-0.125***	-0.0965***	-0.197***	0.0191
	(0.00844)	(0.00889)	(0.0212)	(0.0202)	(0.0364)
State	-0.0300***	-0.0648***	-0.0140**	-0.0379***	0.0243**
	(0.00373)	(0.00327)	(0.00463)	(0.00629)	(0.00855)
Distance	0.00409***	0.000719	0.00484***	0.00245***	0.0100***
	(0.000453)	(0.000368)	(0.000624)	(0.000705)	(0.00139)
Constant	6.987***	7.896***	5.666***	6.239***	2.013**
	(0.275)	(0.238)	(0.364)	(0.516)	(0.688)
N	4500	2400	2400	1000	1000
adj. R^2	0.480	0.487	0.350	0.280	0.347
F	255.2	229.2	111.7	41.92	27.36

Legend: * $p<0.05$; ** $p<0.01$; *** $p<0.001$; standard errors in parentheses.
Note: annual population is measured in hundreds of thousands.

ficients have opposite signs. Aligning with theory, an increased focus on farm employment is expected to result in an increase in volatility. In the context presented by Tables 5.5 and 5.7, one can anticipate an effect opposite that of farm proprietors employment, but very similar in magnitude.

This opposite relationship between farm proprietors and farm wage and salary employment results is also visible with nonfarm proprietors and private nonfarm employment data. Again, the proprietors employment type goes against economic theory. It is anticipated that diversification—here best represented by the nonfarm employment categories—is associated with increased stability, as the economic risk is more dispersed. However, nonfarm

Table 5.7: One-standard-deviation increase (from Table 5.6).

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (#)	-0.0321*	0.1567***	0.0086	0.2562***	0.2424***
Farm prop	-1.5833***	-0.9183***	-0.7418***	-0.7748***	-0.8233***
Nonfarm prop	0.4095***	0.2477***	0.1841***	0.1302***	0.2980***
Farm	1.7891***	1.0664***	1.1165***	0.9104***	1.2384***
Private nonfarm	-0.6113***	-0.5738***	-0.3365***	-0.3992***	-0.1325*
Federal	-0.1139***	-0.1759***	-0.1358***	-0.2650***	0.0257
State	-0.1665***	-0.3651***	-0.0789**	-0.2189***	0.1404**
Distance	0.1653***	0.0291	0.1956***	0.0990***	0.4041***

proprietors employment demonstrates a positive relationship with volatility. That is, the results show that more nonfarm proprietors employment leads to more personal income volatility. This effect remains below one-third of the magnitude of farm proprietors employment's volatility impact (see Tables 5.5 and 5.7) . As for private nonfarm employment, the direction of the expected impact follows theory and is negative. Its magnitude, in context, is not consistently larger or smaller than nonfarm proprietors employment.

The results for the government employment categories are less conflicting. In the first four samples, all coefficients are significant and negative; in the fifth sample, the coefficient on *Federal* is not significant, and the coefficient on *State* is positive. The last two samples are comprised of few observations and essentially represent the Great Recession, so it is not unexpected to see results that are unique to the last sample. In general, the results for government employment are robust. The inverse relationship between government employment percentages and volatility makes economic sense: government positions are considered by many to offer financial stability. Tables 5.5 and 5.7 offer no clear position as to which type of government employment provides the greatest impact on stability, as the two fluctuate depending on the sample. However, in comparison to the other employment types, government employment has the impact of lowest magnitude.

Finally, the distance to the nearest economically central county tests significant in nine of the ten regressions and exhibits a positive relationship with volatility. This relationship aligns with expectations, as rural counties are theoretically less economically stable. Distance is measured in miles, hence the small coefficients. When its magnitude is measured in terms of one standard deviation, its impact on volatility is very diverse. For both size proxies, *Distance* has the largest expected impact in the fifth sample, followed by the third sample, then the first—again, these samples also have higher mean values of the dependent variable (see Table 4.1). This shows that the isolation effect is most prominent when using labor force data and when looking at the population sample as a whole. Thus, workers appear to fare worse than the general population when residing in an isolated county. However, the distance effect is not as prominent in magnitude (when measuring a one-standard-deviation increase) as the effect of either farm employment variable.

Table 5.8: Summary statistics: population and distance.

	(1) Population 1970-2014	(2) Population 1991-2014	(3) Labor force 1991-2014	(4) Population 2005-2014	(5) Labor force 2005-2014
Size (%)					
Mean	1	1	1.000002	1	1.000002
SD	1.285207	1.385173	1.5009	1.479676	1.607725
Min	0.0413819	0.0413819	0.0330664	0.0413819	0.0330664
Max	10.18245	10.18245	11.65771	10.18245	11.65771
Size (#)					
Mean	0.7235584	0.8417406	0.4203836	0.9425178	0.457098
SD	1.000233	1.196182	0.6422548	1.399858	0.7366661
Min	0.03748	0.03775	0.01551	0.04105	0.01551
Max	10.12539	10.12539	5.46812	10.12539	5.46812
Distance					
Mean	78.7033				
SD	40.4147				
Min	16.99389				
Max	181.0277				

Note: county distance distributional properties consistent across samples.

5.3 Granger Causality

An examination of the Granger causality[5] test results (see Appendix B) shows that there is no clear, consistent pattern whatsoever. Of the 108 tests conducted, only 13 yield statistically significant results: seven in test 1 and six in test 2. This fairly even split of significant tests makes no statement as to in which direction causality frequently moves. Additionally, no single test is significant across all three samples, nor is a test even significant across two samples! The results appear to be truly random. Thus, robustness is an issue. The only identifiable pattern is in the second sample in Mecklenburg county: half of the tests prove significant, and all but one of the values returned by test 1 are highly significant. This suggests that changes in personal income move from surrounding counties toward Mecklenburg county. Regardless, if there were true causality, one would expect to see it in a variety of samples, which is not the case. It is therefore reasonable to conclude that there is no causal pattern, at least in the location and time periods covered, between per-person personal income values in central counties and their neighboring counties.

[5] All time-series analysis computations were performed in SAS 9.4 using Windows 10.

CHAPTER 6: CONCLUSIONS

The purpose of the preceding econometric analyses is to quantify the volatility of county-level per-person personal income as well as its potential causes. The idea that county population is a significant determinant of volatility is clear in the preliminary regressions and plots. It is most apparent that small size has the greatest impact on volatility, while large size does not show as clear a pattern. However, this may be a result of excluded variable bias. Perhaps county population accompanies another factor that varies with county size. To explore this problem, regression analysis is conducted using information about county population, employment structure, and the distance to the nearest economically important county. It is not clear whether population actually has an effect on volatility, or what that effect is. The other regressors prove to be useful determinants of volatility. The clearest interpretation of the results suggests that rural counties are subject to greater volatility, government employment increases stability, and, suiting expectations, farm and nonfarm wage and salary employment lead to increased and decreased volatility, respectively.

Additionally, several Granger causality tests are conducted to determine whether close proximity is a useful indicator of per-person personal income changes, and therefore volatility, for counties that surround economically central counties. The results prove unpredictable and don't showcase any significant patterns. However, in regression analysis, distance generally displays a positive and significant relationship with volatility. Although the effects of close proximity are ruled out in the Granger causality tests, it is possible that the effects of isolation are still present.

The overall argument remains that certain aspects of small counties lead to increased personal income volatility. One can determine that it is not population size itself that causes increased volatility; rather, it appears that other county characteristics determine personal income volatility. Further exploration is necessary regarding other elements that may vary with county size, such as educational attainment and poverty levels. It is also possible that the observed size-volatility tradeoff results from different sampling practices in different-sized counties; the BEA mentions this in its Local Area Personal Income Methodology. An investigation of sampling practices in smaller counties is necessary.

REFERENCES

Cortes, Bienvenido, Michael McKinnis, and Michael Davidsson (2013) "Industrial Composition, Local Fiscal Policy and Micropolitan Area Economic Growth," *Mid-Continent Regional Science Association Conference Proceedings*, 30-40.

Cortes, Bienvenido, Michael McKinnis, and Michael Davidsson (2015) "Growth and Volatility of Micropolitan Statistical Areas in the U.S.," *International Journal of Business and Finance Research*, Vol. 9(4), 89-102.

Easterly, William and Aart Kraay (2000) "Small States, Small Problems? Income, Growth, and Volatility in Small States," *World Development*, Vol. 28(11), 2013-2027.

ERS. "Documentation." USDA ERS, 11 Oct. 2016. Web.

Felix, Alison (2012) "Industrial Diversity, Growth, and Volatility in the Seven States of the Tenth District," *Federal Reserve Bank of Kansas City Economic Review*, 55-77.

Furceri, Davide and Georgios Karras (2007) "Country size and business cycle volatility: Scale really matters," *Journal of the Japanese and International Economies*, Vol. 21, 424-434.

Grennes, Thomas, Pablo Guerron-quintana, and Asli Leblebicioglu (2010) "Economic Development and Volatility among the States," *Economics Bulletin*, Vol. 30(3), 1963-1976.

Local Area Personal Income Methodology. N.p.: Bureau of Economic Analysis, Nov. 2015. PDF.

Perrins, Gerald, and Diane Nilsen. "Math Calculations to Better Utilize CPI Data." Rep. Bureau of Labor Statistics, n.d. Web.

"Regional and State Employment and Unemployment-August 2016." N.p.: Bureau of Labor Statistics, 20 Sept. 2016. PDF.

Roth, Jean. "County Distance Database." National Bureau of Economic Research, 11 June 2014. Web. 19 Oct. 2016.

Sherrill Shaffer (2009) "Industrial structure and economic stability," *Applied Economics Letters*, Vol. 16(6), 549-555.

Silvia, John, Azhar Iqbal, Kaylyn Swankoski, Sarah Watt, and Sam Bullard (2014). *Economic and Business Forecasting: Analyzing and Interpreting Econometric Results.* Hoboken: Wiley.

Spelman, William (2006) "Growth, Stability, and the Urban Portfolio," *Economic Development Quarterly*, Vol. 20(4), 299-316.

Voith, Richard (1998) "Do Suburbs Need Cities?" *Journal of Regional Science*, Vol. 38, 445-464.

APPENDIX A: PER CAPITA PERSONAL INCOME OVER TIME

Following Figure 1.1 in the introduction, these scatter plots depict the annual percentage change in per capita personal income in the last four samples.

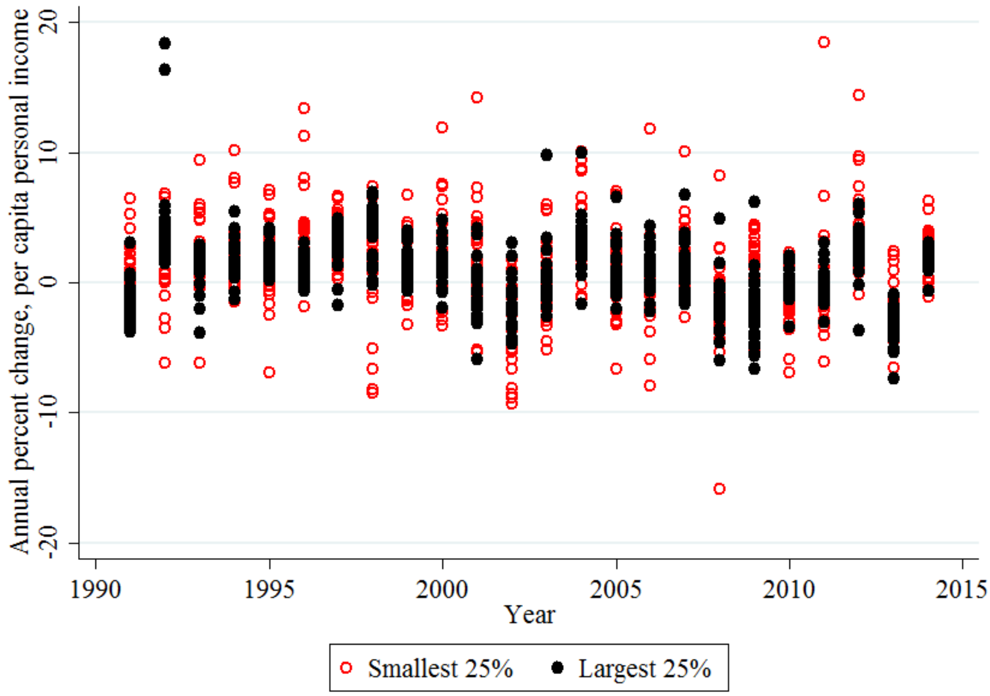

Figure A.1: Small versus large counties: true population, 1991-2014.

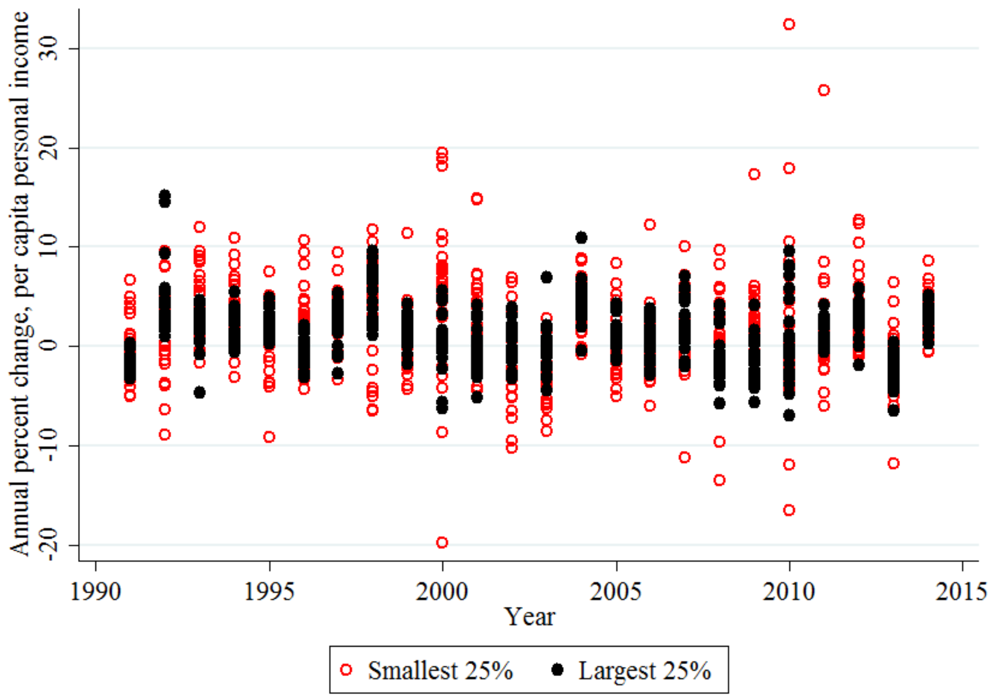

Figure A.2: Small versus large counties: labor force, 1991-2014.

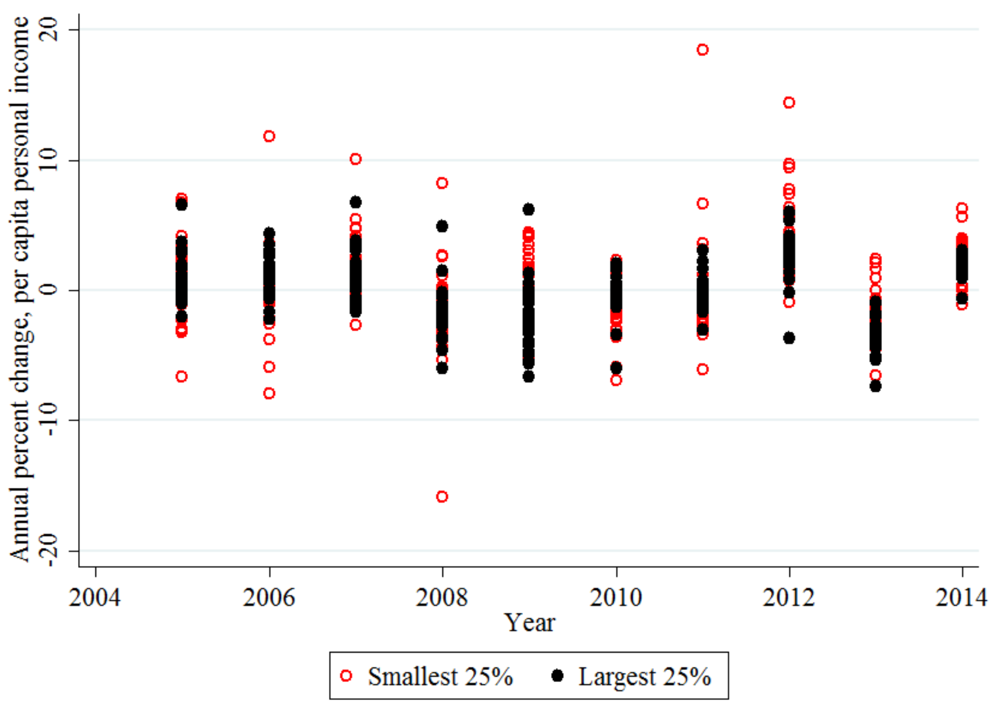

Figure A.3: Small versus large counties: true population, 2005-2014.

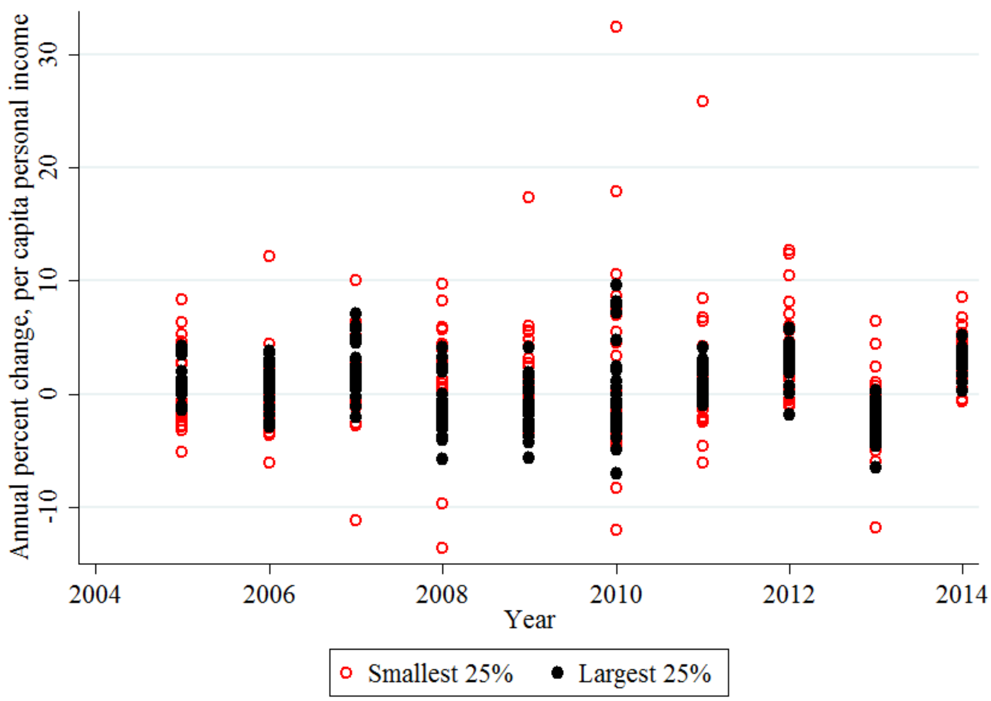

Figure A.4: Small versus large counties: labor force, 2005-2014.

APPENDIX B: GRANGER CAUSALITY

Below are the results of the Granger causality tests, separated by sample.

Table B.1: Granger causality results: true population, 1970-2014.

County			Test 1		Test 2	
Central	Surrounding	Lags	Pr>ChiSq	Decision	Pr>ChiSq	Decision
Mecklenburg	Cabarrus	1	0.1162	Fail	0.9925	Fail
Mecklenburg	Gaston	1	0.1228	Fail	0.4584	Fail
Mecklenburg	Iredell	1	0.0727	Fail	0.1925	Fail
Mecklenburg	Lincoln	1	0.2394	Fail	0.65	Fail
Mecklenburg	Union	1	0.8897	Fail	0.0247	Reject
Durham	Granville	1	0.4171	Fail	0.0007	Reject
Durham	Vance	1	0.6356	Fail	0.0648	Fail
Durham	Warren	1	0.0405	Reject	0.6555	Fail
Orange	Chatham	1	0.3599	Fail	0.7232	Fail
Orange	Moore	1	0.9947	Fail	0.1851	Fail
Orange	Person	1	0.5585	Fail	0.2756	Fail
Wake	Edgecombe	1	0.9761	Fail	0.5824	Fail
Wake	Franklin	1	0.6791	Fail	0.0193	Reject
Wake	Harnett	1	0.6292	Fail	0.2168	Fail
Wake	Johnston	1	0.4683	Fail	0.0309	Reject
Wake	Lee	1	0.5271	Fail	0.5895	Fail
Wake	Nash	1	0.5819	Fail	0.32	Fail
Wake	Wilson	1	0.8378	Fail	0.1132	Fail

Note: using 5% significance as the rejection criterion.

Table B.2: Granger causality results: true population, 1991-2014.

County		Test 1			Test 2	
Central	Surrounding	Lags	Pr>>ChiSq	Decision	Pr>ChiSq	Decision
Mecklenburg	Cabarrus	4	0.0069	Reject	0.013	Reject
Mecklenburg	Gaston	1	0.0032	Reject	0.3046	Fail
Mecklenburg	Iredell	1	0.0091	Reject	0.7261	Fail
Mecklenburg	Lincoln	1	0.6025	Fail	0.803	Fail
Mecklenburg	Union	2	0.0019	Reject	0.0895	Fail
Durham	Granville	1	0.4818	Fail	0.799	Fail
Durham	Vance	1	0.3908	Fail	0.7445	Fail
Durham	Warren	1	0.0957	Fail	0.6773	Fail
Orange	Chatham	1	0.8019	Fail	0.1776	Fail
Orange	Moore	1	0.6768	Fail	0.886	Fail
Orange	Person	1	0.3622	Fail	0.2264	Fail
Wake	Edgecombe	1	0.2983	Fail	0.48	Fail
Wake	Franklin	1	0.1108	Fail	0.6203	Fail
Wake	Harnett	1	0.7983	Fail	0.7086	Fail
Wake	Johnston	1	0.1463	Fail	0.3113	Fail
Wake	Lee	1	0.073	Fail	0.3179	Fail
Wake	Nash	1	0.7373	Fail	0.8086	Fail
Wake	Wilson	1	0.7345	Fail	0.4188	Fail

Note: using 5% significance as the rejection criterion.

Table B.3: Granger causality results: labor force, 1991-2014.

County			Test 1		Test 2	
Central	Surrounding	Lags	Pr>ChiSq	Decision	Pr>ChiSq	Decision
Mecklenburg	Cabarrus	1	0.5186	Fail	0.6557	Fail
Mecklenburg	Gaston	1	0.1446	Fail	0.522	Fail
Mecklenburg	Iredell	1	0.1026	Fail	0.9614	Fail
Mecklenburg	Lincoln	1	0.2182	Fail	0.1754	Fail
Mecklenburg	Union	1	0.4434	Fail	0.5593	Fail
Durham	Granville	1	0.9551	Fail	0.1161	Fail
Durham	Vance	1	0.4371	Fail	0.838	Fail
Durham	Warren	1	0.9564	Fail	0.6945	Fail
Orange	Chatham	1	0.7012	Fail	0.0713	Fail
Orange	Moore	1	0.3235	Fail	0.5405	Fail
Orange	Person	1	0.04	Reject	0.0606	Fail
Wake	Edgecombe	1	0.0758	Fail	0.1259	Fail
Wake	Franklin	1	0.5722	Fail	0.622	Fail
Wake	Harnett	1	0.3247	Fail	0.9892	Fail
Wake	Johnston	1	0.1786	Fail	0.3985	Fail
Wake	Lee	2	0.0436	Reject	0.0135	Reject
Wake	Nash	1	0.911	Fail	0.5116	Fail
Wake	Wilson	1	0.1117	Fail	0.6709	Fail

Note: using 5% significance as the rejection criterion.

APPENDIX C: DISTANCE OBSERVATIONS

The below distances are averages of the 2000 and 2010 NBER values, in miles.

Table C.1: Distances used in analysis, smallest to largest.

County	Distance	County	Distance	County	Distance
Durham	0	Davidson	52.812182	Henderson	92.87513
Wake	0	Montgomery	52.824833	Yancey	94.41296
Orange	0	Moore	56.419015	Bertie	96.237378
Mecklenburg	0	Sampson	57.228532	Pender	98.868193
Alamance	16.99389	Edgecombe	58.679319	Buncombe	98.87238
Cabarrus	18.540163	Burke	58.715363	Onslow	100.34416
Gaston	19.45913	Greene	58.819894	Hertford	100.76001
Granville	22.858814	Caldwell	61.611542	Craven	101.07464
Union	23.178603	Rutherford	62.316341	Beaufort	102.18024
Chatham	24.785655	Forsyth	64.096231	Columbus	105.62137
Johnston	24.952348	Hoke	64.114507	Transylvania	110.37133
Person	25.451176	Yadkin	64.372289	Madison	112.88426
Lincoln	27.48816	Richmond	64.831665	Washington	115.6893
Caswell	27.85388	Halifax	65.931444	Chowan	116.50001
Franklin	27.927989	Stokes	68.27335	Gates	117.06284
Harnett	31.499553	Lenoir	68.924314	Pamlico	118.83989
Rowan	32.103796	Wilkes	69.012805	New Hanover	119.14192
Stanly	33.690625	Duplin	71.475377	Haywood	122.00329
Vance	35.699568	Pitt	72.984294	Brunswick	124.81807
Lee	36.330641	McDowell	74.768619	Perquimans	127.49602
Iredell	37.030368	Polk	76.494205	Jackson	131.22114
Catawba	37.782753	Scotland	80.835219	Carteret	132.71978
Guilford	39.697633	Northampton	81.337908	Tyrrell	137.54485
Cleveland	40.544546	Surry	82.256621	Pasquotank	138.52213
Nash	40.900506	Robeson	82.97188	Hyde	140.85125
Wilson	41.574467	Bladen	83.050992	Camden	144.15098
Randolph	44.737498	Watauga	83.958666	Macon	144.76114
Rockingham	44.925936	Avery	84.516876	Swain	148.94483
Anson	44.931929	Martin	86.332769	Currituck	156.69532
Wayne	47.41508	Alleghany	88.249094	Dare	165.44624
Davie	50.398007	Jones	89.556281	Clay	166.58952
Alexander	50.673001	Ashe	90.291665	Graham	168.45551
Warren	51.805347	Mitchell	90.734678	Cherokee	181.02766
Cumberland	52.502798				