RURAL-URBAN WAGE DICHOTOMY: THE PREVAILING DIFFERENTIAL

A Thesis

Presented in Partial Fulfillment of the Requirements for the

Degree of Master of Science

with a

Major in Bioregional Planning & Community Design

in the

College of Graduate Studies

University of Idaho

by

Jalal Jahir

March 2014

Major Professor: Philip Watson, Ph.D.

Authorization to Submit Thesis

This Thesis of Jalal Jahir, submitted for the degree of Master of Science with a Major in Bioregional Planning and Community Design titled “Rural-Urban Wage Dichotomy: The Prevailing Differential”, has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: ______________________________ Date: ______________
(Philip Watson, Ph.D.)

Committee Members: ______________________________ Date: ______________
(Tim Frazier, Ph.D.)

______________________________ Date: ______________
(Christopher McIntosh, Ph.D.)

Department Administrator: ______________________________ Date: ______________
(Tamara Laninga, Ph.D.)

Discipline’s College Dean: ______________________________ Date: ______________
(Larry Makus, Ph.D.)

Final Approval and Acceptance

Dean of the College Of Graduate Studies: ______________________________ Date: ______________
(Jie Chen, Ph.D.)

Abstract

The growing income inequality in the U.S. is widening the gap between rural and urban returns to labor. Most income inequality studies to date have focused either on international or inter-state wage differences. There is a continued need for improved understanding of the increasing rural urban wage differentials at a localized (i.e. county level) geography. This study identifies a method of delineating the factors contributing to the rural urban wage gap. Both Fixed Effect and Ordinary Least Square (OLS) models are utilized in defining the county wage functions for the contiguous U.S. counties. The Blinder Oaxaca decomposition method is employed to decompose the contribution of the explanatory variables of the wage models in explaining the rural urban wage differentials. The decomposition results of the county wage models suggest that physical, social, natural, economic and location attributes explain a major portion, but not all, of the rural-urban wage differentials. Controlling for industry mix variables only slightly improves the model. However, controlling for variation in human capital endowments (e.g. expected education) yields a model that explains almost all of the wage differences between urban and rural regions. The results inform that identifying and focusing on the key reasons behind the growing differences in rural-urban average returns are likely to help formulate proper policies for rural poverty alleviation.

Acknowledgements

I would like to thank all the members of my outstanding graduate committee Dr. Philip Watson, Dr. Tim Frazier and Dr. Christopher McIntosh. Particularly, I would like to express my deepest gratitude and appreciation to my mentor and major advisor, Dr. Philip Watson, who provided me with research fund as well as the intellectual challenge and motivation to extend and improve my academic and professional experiences. I would also like to thank former director of the Bioregional Planning program, Dr. Steven Hollenhorst, who also provided me with advises and help during my entire time at the university. Special thanks to all the faculty, staff and students of the Bioregional Planning program as well as the Department of Agricultural Economics and Rural Sociology, the Department of Geography and the Department of Conservation Social Science, from whom I have enjoyed support in diverse ways. I would like to express my gratitude to the Government of the People's Republic of Bangladesh for providing me with the support until my bachelor's degree from Jahangirnagar University and helping me prepare for advance education. Finally, I appreciate the material and spiritual support from my family, parents, siblings, other family members, and friends who have always believed in me.

Dedication

This work is dedicated to my parents for all the sacrifices they have made in order to make me who I am today. Especially I would like to dedicate this work to my maternal grandmother as an appreciation for her endless love, support, prayers, and for being on my side at every step of my life; whom I couldn’t see for the last time as she passed away during my stay at the university, thousands of miles away. This is for you, Nanu…

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Persistent wage differences exist between rural and urban settings and there are growing indications that the United States is experiencing an increase in income inequality. The increase began in the 1980s and by the first decade of the 21st century nearly 50 percent of the nation's income was received by the top earners (Peters 2013). The concentration of income and wealth over the past decades has largely been in urban areas, and rural areas are falling further behind their urban contemporaries (Anderson et al. 2013). This study attempts to identify and explain the wage differentials between employees in urban and rural communities. Changes in the composition of human capital and change in industry clusters are identified as leading contributors to the rising income inequality (Partridge et al. 1996). The primary hypothesis is that, controlling for existing industry mix, social variables, locational attributes and policy differences, the variation in average returns would be equalized across space. The Blinder-Oaxaca decomposition method is then employed to explain wage differentials between employees in rural and urban counties. The model presented here investigates the effects of human capital, existing industry mix, policy variation, social capital, financial capital, and natural attributes on rural-urban differentials in average returns. The results of this study can assist policy makers and development planners in formulating policy strategies leading to the effective development of rural areas and an improved rural-urban relationship.

While this study does not seek to identify the complex cluster of factors that could potentially explain the rural-urban wage differential, it does attempt to identify the existence and quantify the magnitude of the differential. Peters (2013) argued that though it is important to emphasize inequality research on ‘‘who gets what and why’’, it is also important to concentrate on ‘‘where and when”. While this study is time invariant, this study does focus on ‘where and why’ there is rising income inequality in the U.S. The models are intended to estimate the effects of human capital (e.g. expected education), existing industry mix, and rural-urban characteristics on the wages across the U.S. counties. This study contributes to the literature by using multiple data sources to estimate the county level wage model, decompose the elasticities of control variables on the rural-urban wage differentials, and finally identifying the magnitude of the explained and unexplained portion of the rural-urban wage differentials (wage differentials hereafter), as well as the significances of the estimations.

Chapter 2: Literature Review

Workers in urban areas with high spatial density of economic activity have been shown to earn higher wages than their counterparts in rural and less densely populated regions (Andersson et al. 2013). Glaeser and Maré (2001) report that wages of urban workers in the United States are roughly 33 percent higher than those of their non-urban counterparts. The empirical regularities of the higher urban wage are generally referred to as the urban wage premium (UWP). Andersson et al. (2013) argue that one strand of inquiry on UWP focuses on the difference of UWP between workers with different sets of human capital. Bacolod et al. (2009) show evidence of spatial heterogeneity of UWP and its dependability on human capital (e.g. skills). Andersson et al. (2013) conclude that who you are is more important than where you live in explaining spatial wage disparities, thus the main reason why workers in denser regions earn more is simply that they are different from the workers in more rural regions. This urban wage premium is the major source of the income/wage inequality. Using data from US metropolitan areas to compare wages between different skill groups, Fallah et al. (2011) find that greater market potential is positively associated with greater shares of skilled workers (higher level of human capital), thus wage inequality is positively linked to the intensity of high-skilled workers.

In a recent study Peters (2013) reports that by the first decade of the 21st century nearly 50 percent of the nation’s income was received by the top earners, approaching inequality levels not seen since the late 1920s. Though social scientists have begun to document the causes of rising income inequality (Peters 2011), the bulk of these works have focused on national and state-level analyses, and the conclusions from these studies are evident in most states (Partridge et al. 1996; Lynch 2003). More recently, studies on income inequality have begun at more localized levels, especially at the county level (Levernier et al. 1998; Moller et al. 2009). Previous research has clearly demonstrated that income inequality persists in the United States across regions over time. This body of work has established that inequality can be explained by differences in economic structures (i.e. industry mix), individuals (i.e. human capital), natural resources (i.e. natural amenities), geography (i.e. urban/rural), and history (Morrill 2000; Lobao and Saenz 2002; Lobao 2004). Spatial wage disparities can result from spatial differences in the skill composition of the workforce, in non-human endowments (natural amenities), and in local interactions (industry mix). Spatial sorting of labor forces refers to the selection of locations and explains the wage gap as more productive workers being more prone to locate in denser regions. A general finding is that spatial sorting (i.e. rural-urban) of workers is the main source of the UWP (Combes et al. 2008; Andersson et al. 2013).

The centrality of space and place has always been taken for granted in geography and regional science (Goodchild et al. 2000). The functional difference between urban and rural areas is the distinction between spatially extensive industries and occupations, mainly agriculture, and spatially intensive industries and occupations (Stewart 1958). Attention of economists and geographers has been recently refocused on the contribution of agglomeration economies on the process of local economic growth. Economic landscapes are increasingly being shaped by a complex mixture of forces operating simultaneously at the regional level sharing a common paradigm: the structural shift from manufacturing to services. The main effect of such phenomenon in space is that, urban areas lose manufacturing capacity to become more service oriented (spatially intensive). On the other hand, peripheral areas (rural counterparts) become potential locations both for manufacturing (spatially extensive) and service (Paci and Usai 2008) where agriculture is a predominant economic sector. Though the importance of industry mix in rural-urban characterization on local economic growth is evident, Fallah et al. (2011) argue that wage inequality is more tilted in favor of human capital rather than industry composition.

When investigating human capital externalities (HCE) and urban wage premium (UWP), Heuermann et al. (2010) report that high urban wages may simply compensate for high urban housing prices and costs of living (COL). The notable increases in wages resulting from the presence of human capital (i.e. education), HCE (workers are more productive in human-capital-intensive environments) have a role to play as a driving-force behind the UWP. Bollinger et al. (2011) argued that significantly upgraded human capital (education) may prevent the wages from falling further behind, that is widening inequality overall.

The Blinder-Oaxaca decomposition (Blinder 1973; Oaxaca 1973) divides the product differential between two categories into a part that is explained by group differences and a residual part that cannot be accounted for by such differences in the model estimators. Jann (2008) argues that besides the labor market and discrimination studies, this technique can be used to study group differences in any continuous outcome variable.

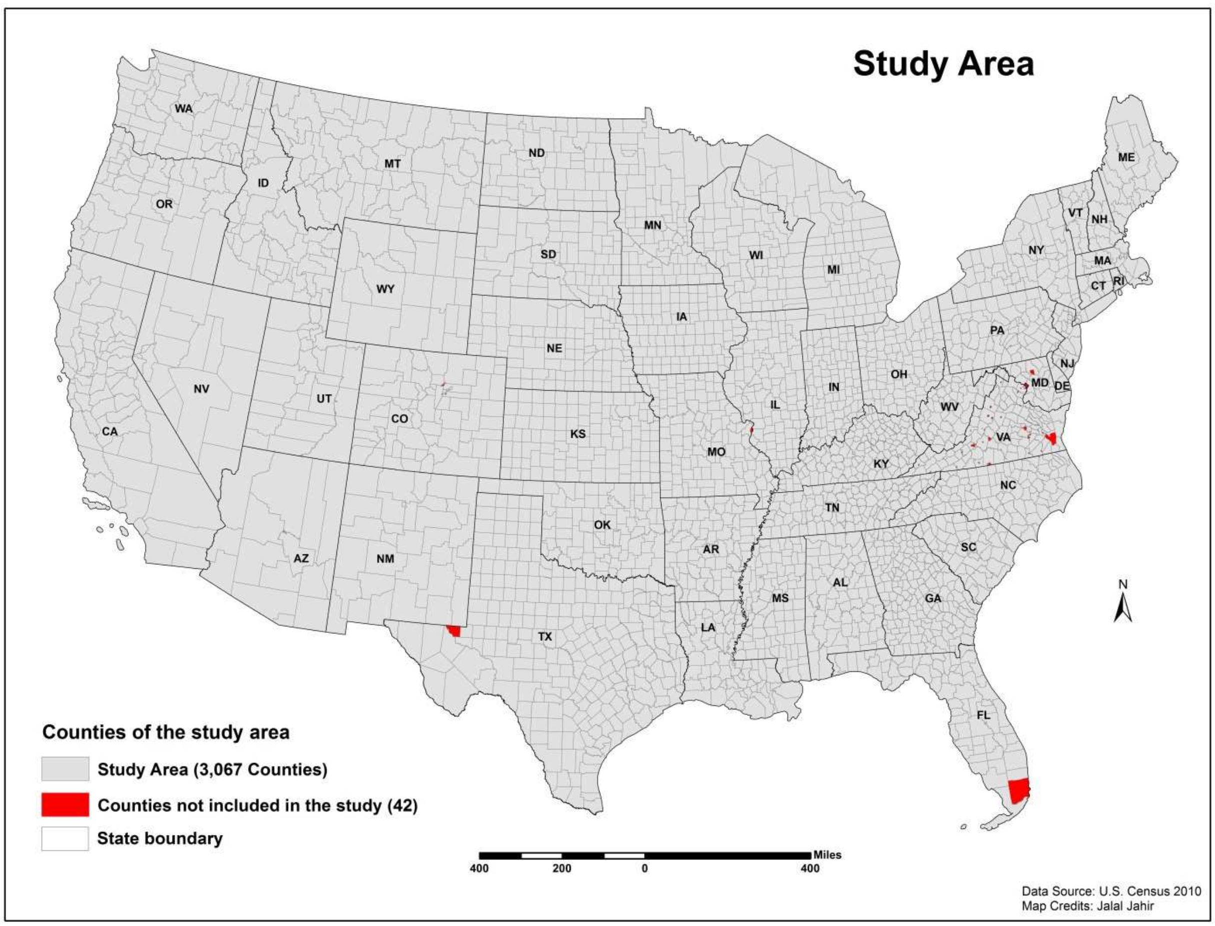
Although there are several studies to date dealing with wage inequality and its relation to location attributes, industry mix, human capital, spatial agglomeration economies, natural amenities, etc., the majority of them focus on geographical extents such as international or either urban metropolitan areas or rural areas. While attempting to decompose the wage differentials, focus has been on gender differences. The interest of this study is in identifying a decomposition method for wage differentials and application of that technique to estimate how much and which of the determining factors are contributing to explaining the wage differential. The contribution of this study to the contemporary literature is introducing the Blinder-Oaxaca decomposition method as a means to decompose the effects of explanatory variables on wage differentials. This study attempts to employ the empirical wage decomposition method to identify how much of the wage variation can be explained by influencing factors such as rural-urban characteristics, state based economic and development policies, existing industry mix and human capital (expected education). The results would be useful to focus attention on particular variables, and where to focus attention. Is it the varying industry mix that is playing the dominant role in wage differentials, or is it the composition of human capital that is determining who gets the wage premium? The answers to these questions are discussed in the following chapters.

Chapter 3: Data and Methods

Data

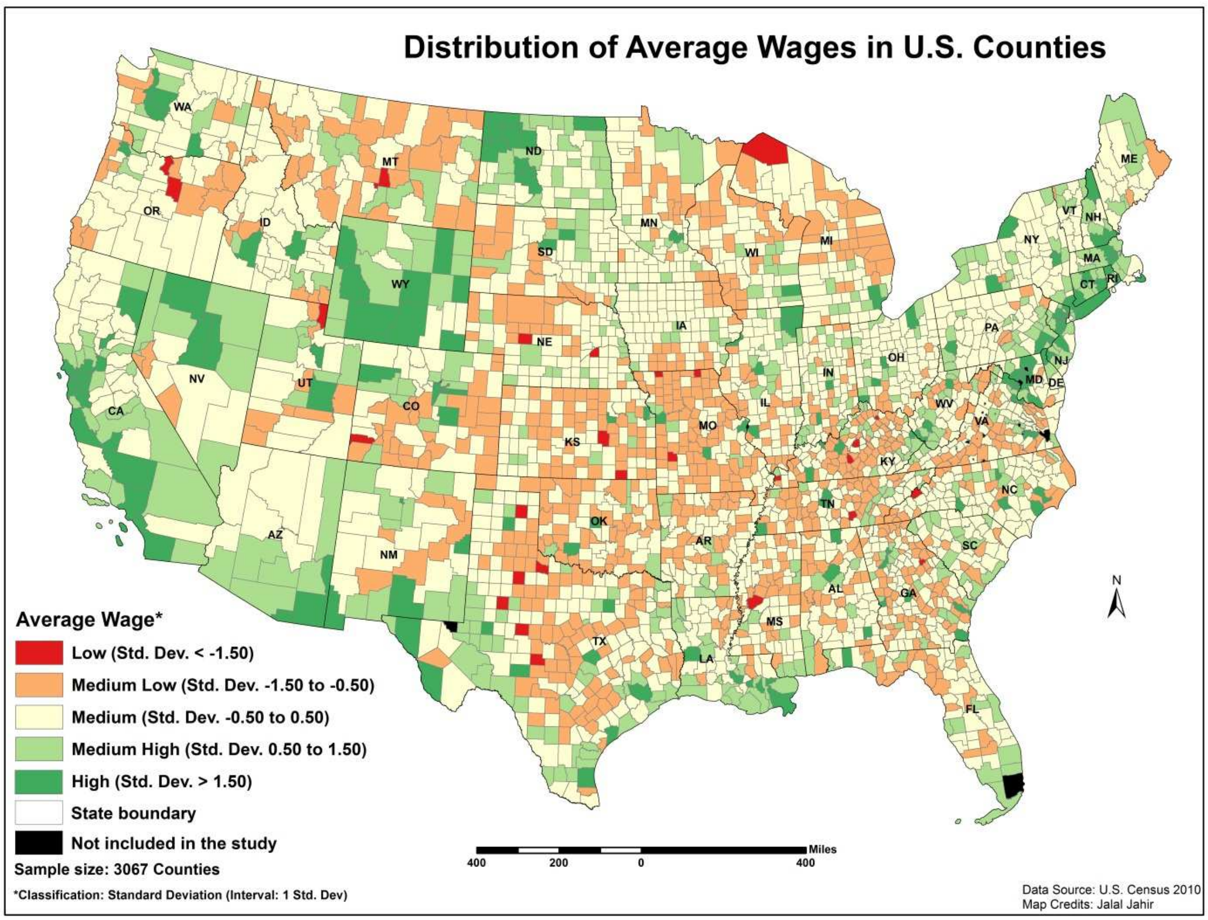
This study utilized county level data for 3,067 contiguous U.S. counties. After starting with all (3,109) counties in the 48 contiguous states, due to inconsistency in the data from multiple sources, the sample size was reduced to 3,067 counties for which values of all the vectors were present. Peters (2013) argues that the counties are ideal units of analysis to study wage differentials because their boundaries are relatively stable over time, and there is a wide range of data available at the county level (Curtis et al. 2012; Slack and Myers 2012).

Figure 3-1: The study area delineating counties included in the analysis



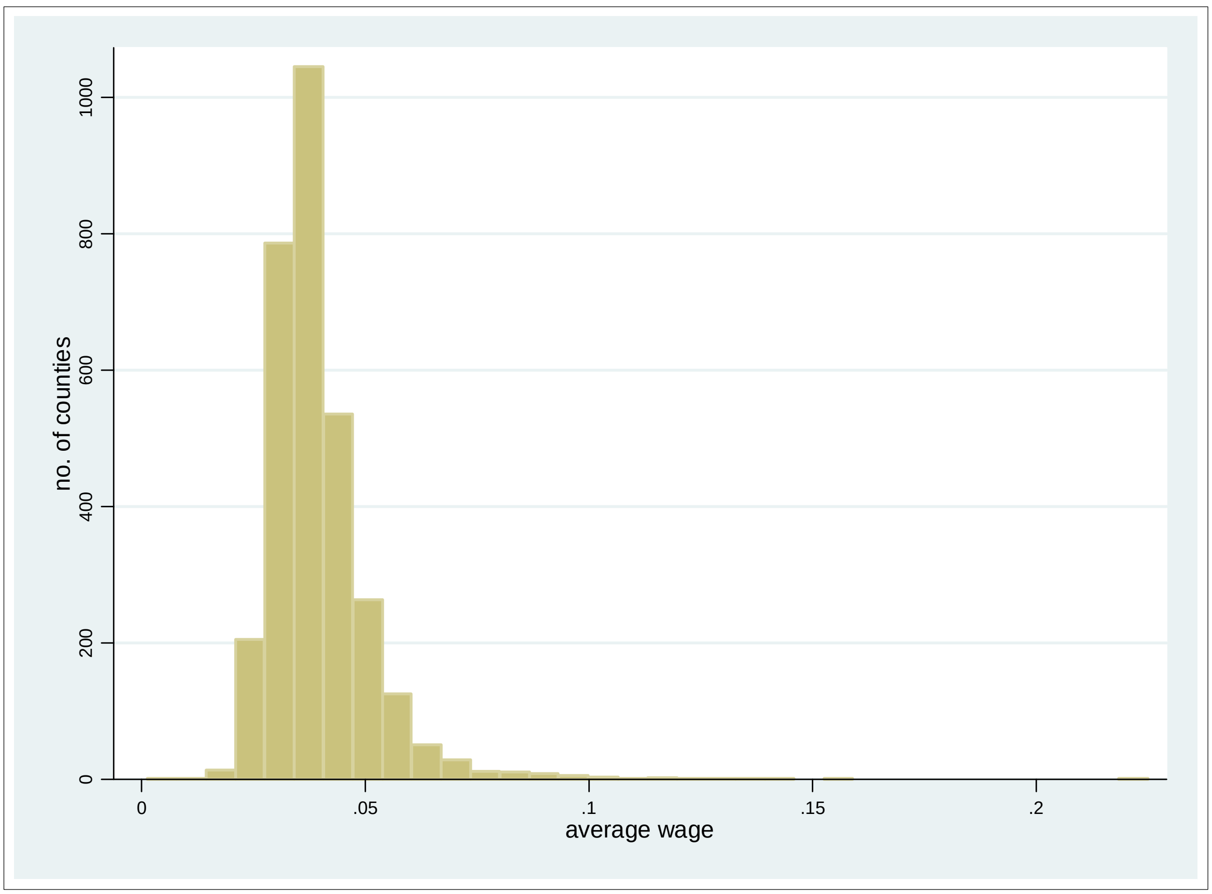
County level existing industry mix, expected education share and average wage data came from IMPLAN input–output regional modeling system (MIG 2000). The 86 industry mix vectors are the portion of county employees involved in each of the 86 industry sectors. The 12 expected education vectors represent the expected county share of employees with 12 different education levels given the existing industry mix. A detailed description of data particulars can be found in the IMPLAN Data Guide (Olson and Lindall 1999). IMPLAN constructs county level accounts based on a variety of data sources including the U.S. Census Bureau, U.S. Bureau of Economic Analysis (BEA), and ES-202 employment data.

Figure 3-2: The distribution of average wages across contiguous U.S. counties



The primary sources of data controlling for urban rural characteristics (Rural-Urban Continuum Code, Creative Class, and Natural Amenities Rank) were from USDA Economic Research Service (ERS). The Social Capital Index was from Rupasingha and Goetz 2008. Detailed methodology of construction of social capital index can be reviewed at Rupasingha et al. 2006. The values of agricultural land per acre are from the 2007 Census of Agriculture published by the USDA. The median rent data came from the American Community Survey 2011 (five year data) published by U.S. Census Bureau. Variables of this study are listed in Appendix A and the descriptive statistics are in Appendix F.

Figure 3-3: The average wage (in million dollars) vector



Methods

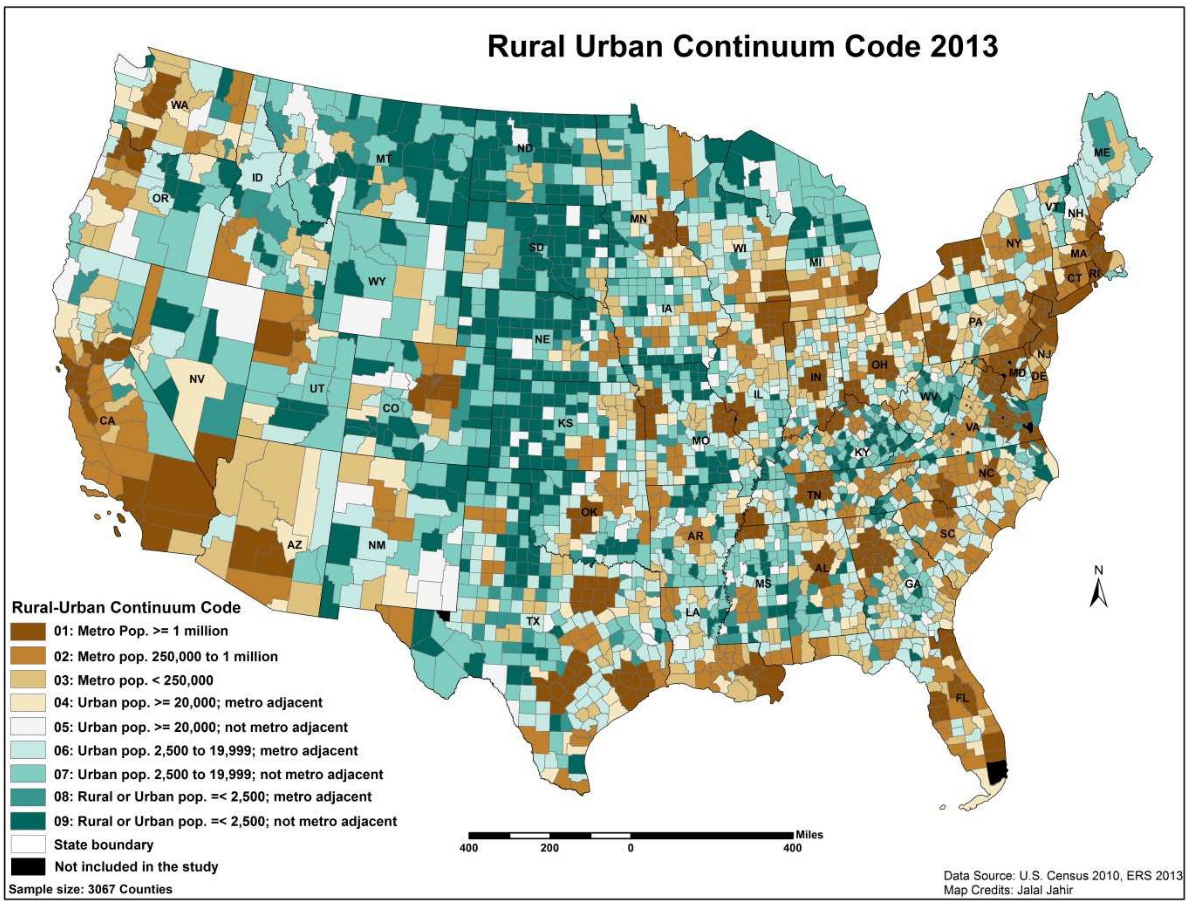
The Ordinary Least Square (OLS) method is utilized for the initial model:

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}) \ \ldots\ldots \ (1)$$

Where, the avg_wage is the county average wage as the dependent variable.

RUCC13 is the Rural-Urban Continuum Code controlling for the ruralness/urbanness. The metropolitan (metro) counties are distinguished by the population size of their metro area, where nonmetropolitan (nonmetro) counties are identified based on the degree of urbanization and proximity to metro areas.

Figure 3-4: The rural urban continuum code 2013 for the study area



ERS further subdivided official metro and nonmetro categories of the Office of Management and Budget (OMB) into three metro and six nonmetro categories; thus each county is assigned one of the nine codes (1 to 9), allowing the researchers to break the county data into finer grains beyond metro and nonmetro to analyze the trends in nonmetro areas resulting from the population density and metro influence (ERS 2013). The lower the code the more urban characteristics are reflected in the corresponding county. Recent studies as well as historical evidence indicate that average wages are considerably higher in urban areas than in rural areas (Heuermann et al. 2010). Thus, the RUCC13 vector is expected to achieve a negative coefficient loading with the increase in the county average wage.

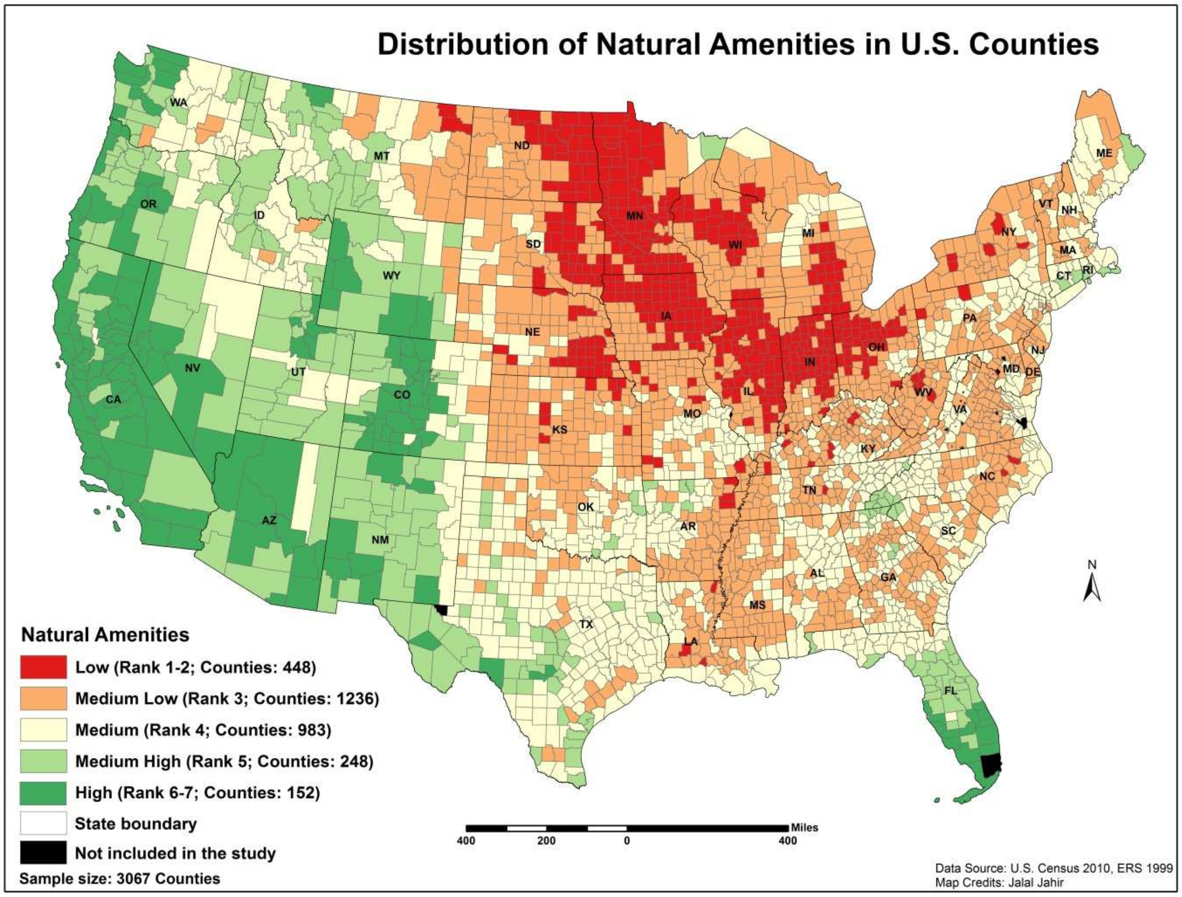
The vector Creative2000 is a proxy for the existing agglomeration of creative class population, involved in economic activity requiring creative thinking and the like with high average returns. Rural communities suffer from brain drain when the young adults migrate to the urban areas in search of higher quality of life and access to more resources. To compete in today's economy, communities need to attract planners, architects, engineers, economists, scientists and people in other creative occupations. This is more eminent for the rural communities which tend to lose much of their talent to the urban counterparts (ERS 2000). The vector with share of county creative class employees controls for the intensity of heterogeneously distributed county occupations requiring creative thinking. Higher share of creative class employees is indicative of a more competitive economy with higher average wage. Existence of creative class in a community has spillover effects. Communities with higher share of creative class population with higher average returns certainly enjoy higher standard of living, access to more resources etc., providing children in these communities with access to higher educational facilities and economic freedom, allowing more engagement in creative class economic activities. This circles in for generations alongside the fact that these communities attract talents from communities with less return to human capital. Whereas, the other counterpart with less share of creative class experiences the opposite as talents migrate out of these communities to the preceding, more urban areas.. This vicious cycle exacerbates the process of the country economic landscape becoming more heterogeneous. In the models, this vector controls for this particular heterogeneity. Neither all urban counties benefit from this heterogeneity, nor all rural counties are deprived of the effects of agglomeration of creative class, but the majority of each group encounters the respective effects. The creative class vector plays a major role in teasing out the effects of these existing heterogeneities from the wage models.

SK09 is the county-level measure of social capital (Rupasingha and Goetz 2008) initially developed in Rupasingha et al. (2006). This variable is the first principal component of five variables, including the number of social capital-generating associations per 10,000 residents (religious organizations, civic organizations, business organizations, political organizations, professional organizations, labor organizations, bowling centers, public golf courses, fitness centers and sports organizations); voter turnout in the presidential election; number of tax exempt non-profit organizations per 10,000; and participation in the decennial Census. The variables capture the sense of belonging to the nation and represent both local and national allegiance and local civic engagement, i.e., the social attributes of counties. Rupasingha et al. 2006 argued that low average wages may lead individuals to work for more hours to secure additional income, leaving them with less time for civic engagement, resulting in less stock of social capital for the community. It is more likely that higher average wage would correspond with higher level of social capital.

By building comprehensive measures of natural amenity and quality of life attributes within the framework of a rigorous theoretical and empirical regional economic growth model, Deller et al. (2001) emphasized that there are foreseeable interactions between amenities, quality of life, and local economic performance. The natural amenities scale is a measure of the physical characteristics of a county area that reflect environmental qualities most people prefer, enhancing the location as a choice to live (ERS 1999). This index serves as a scaled index value of multiple environmental characteristics such as topography, water, temperature, and humidity (Wilson et al. 2006). Natam_Rank is the county rank based on natural amenities index controlling for agglomeration economies due to demand for natural amenities. Marcouiller et al. (2004) describes that local amenities affects local wages (Roback 1988) and the differences in natural amenities generate wage differentials across regions (Roback 1982). These variances imply that some people might enjoy local amenities at the expense of higher rents and lower wages. Wang and Wu (2011) documents that empirical studies have examined the effect of local amenities on regional inequalities in wages, housing prices and human capital accumulation (Glaeser et al. 2001; Deller et al. 2001; Rappaport and Sachs 2003; Wu 2006; Florida et al. 2008); although Deller et al. (2001) points out that the relationships among amenities, agglomerations of economic activities and regional economic development are not well understood.

The vector Aglb_Val is the per acre value of agricultural land and building, a proxy for land value differences in urban and rural areas. Returns to land bids into land rent and agricultural returns are lower than urban land uses. In urban areas, the rent is higher than that of rural areas, thus agricultural uses tend to be demolished by urban land use. This implies that value of agricultural land is higher in urban areas and proximities than in rural areas. Med_Rent is the median rent as percentage of household income, controlling for the difference in living expenses between urban and rural areas. Both Aglb_Val and Med_Rent variables are to control for local economic attributes at the county level.

Figure 3-5: The counties ranked according to natural amenities index



Exploratory Spatial Data Analysis (ESDA) techniques are employed to identify spatial pattern, if any, for the dependent variable as well as for the residuals of regression models (equation 01, 03 & 04). Anselin et al. 2007 describes ESDA as a division of exploratory data analysis (EDA) concentrating on the unique characteristics of spatial data- particularly on spatial autocorrelation and spatial heterogeneity (Anselin 1999a; Anselin 1999b; Goodchild et al. 2000). Spatial autocorrelation, a statistical test of non-randomness, measures the extent to which the occurrence of an event in a spatial unit constrains, or makes more probable, the occurrence of an event in a neighboring spatial unit. Moran's I (Moran 1950) is one of the oldest indices of spatial autocorrelation which is still a de facto standard for determining spatial autocorrelation. Moran's I is calculated from the following equation:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}) \sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Where, n is the number of counties, y_i is the variable value at a particular county, y_j is the variable value at another county, $\bar{y}$ is the mean of the variable and w_{ij} is the spatial weight applied to the comparison between county i and j (O'Sullivan and Unwin 2010). Inspired by Tobler's "first law of geography" (Tobler 1970) that "Everything is related to everything else, but near things are more related than distant things", O'Sullivan and Unwin (2010) describes that autocorrelation is likely to be most pronounced at short distances. Thus both the global (test for clustering) and the local Moran's I statistics (test for clusters) for average wage as well as residuals from equation 01 are derived using first order Queen's contiguity weighting matrix and maximum allowed permutation of 99999 in Geoda routine (Anselin 2005).

Fixed Effect method is employed to control for the across state policy differences. The main purpose in using the fixed-effects model is to derive estimates free from selection bias (England et al. 1988). By including fixed effects (state dummies), the model is controlling for the average differences across states in any observable or unobservable predictors, such as differences in public policy, development strategy etc. In studying the Kuznets curve (Kuznets and Jenks 1953; Kuznets 1955) and income inequality in U.S. counties, Nielsen and Alderson (1997) utilized state fixed effects in modeling the impact of unmeasured state-level variables on income variation in counties. They argued that, counties are subsets of the larger units (i.e. the states) and therefore it is the states rather than the counties that have a substantial degree of political autonomy which can influence the distribution of income in numerous ways (Jacobs 1985).

The fixed effect coefficients soak up all the across-state action and greatly reduced the threat of potential omitted variable bias. Between-state variation is very likely to be contaminated by unmeasured state characteristics that are correlated with average wage. By restricting to the within-state variation, that contamination is eliminated and it is much more likely to get unbiased estimates (Allison 2005; Borenstein et al. 2010; Clarke et al. 2010). The fixed-effect model is as follows:

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}, \boldsymbol{i.state\_1\ to\ 48}) \quad \ldots\ldots (2)$$

Where the i.state_1to 48 are the state dummies.

Then in equation 02 existing industry mix variables and expected education share variables are introduced to construct wage-industry mix model and wage-education share model consecutively:

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}, \boldsymbol{i.state\_1\ to\ 48}, \boldsymbol{ind01\ to\ 86}) \quad \ldots\ldots (3)$$

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}, \boldsymbol{i.state\_1\ to\ 48}, \boldsymbol{E01\ to\ 12}) \quad \ldots\ldots (4)$$

Where, ind01 to 86 are the industry mix variables for 86 industry sectors and E01 to 12 are the education variables for expected share of different levels of education given the existing industry mix. The industry mix vectors are representative of existing county industry clusters which would in turn identify if and how the wage differences are taking place due to differences in county industry mix. The expected education share vectors contain the expected level of employee share in different education levels given the existing county industry mix. The knowledge, skills and ability of the labor force likely depend on the level of education and experience. The expected education level of a given industry mix delineate the combination of knowledge, skills, abilities, education and experience required for each industry in the county industry mix. Thus expected education levels for a given industry mix are more a representation of expected combination of human capital per se, rather than just the education attainment component of the broader human capital attributes. These vectors would inform if, and if yes, how much of the wage differential could be explained by the expected education levels.

To analyze mean product differences between groups, the counterfactual decomposition technique, which is known in the literature as the Blinder-Oaxaca decomposition (Blinder 1973; Oaxaca 1973), is widely used. This method divides the product differential between two groups into a part that is ‘explained’ by group differences (characteristics effects) and a residual part (coefficient effects) that cannot be accounted for by such differences in the model determinants. This ‘unexplained’ part is often used as a measure for discrimination (Blinder 1973), but it also subsumes the effects of group differences in unobserved predictors (Jann 2008). Considering the average returns example of this study in modeling the decomposition method (modified from (Blinder 1973)). The wage model contrasting rural-urban dichotomy (equation 1) rewrites:

$$Y_i = \beta_0 + \sum_{j=1}^{n} \beta_i X_{ji} + \mu_i$$

Where, Y_i is the level of county average wage, $X_{1i},\ldots,X_{ni}$ are n observable characteristics used to explain Y_i. As the interest is in comparing between urban and rural counties; the above equation is reconstructed as follows:

For urban counties,

$$Y_i^u = \beta_0^u + \sum_{j=1}^{n} \beta_j^u X_{ji}^u + \mu_i^u$$

For rural counties,

$$Y_i^r = \beta_0^r + \sum_{j=1}^{n} \beta_j^r X_{ji}^r + \mu_i^r$$

The wage difference due to the attribute difference:

$$\sum_j \beta_j^u \bar{X}_j^u - \sum_j \beta_j^r \bar{X}_j^r$$

This portion is attributable to the average wage differences between urban and rural areas due to differences in the model predictors.

The amount of difference captured by the shift coefficient: $\beta_0^u - \beta_0^r$

This portion is typically attributed to discrimination. The decomposition process could carry further as the explained portion comes from both the differences in the coefficient and the differences in the average characteristic. The three fold decomposition (Jann 2008) estimates group differences in the predictors (endowment effects), contribution of differences in the coefficients and estimate for the interaction term between differences in endowments and coefficients. Primary interest here is not to investigate the interaction with in the predictors; rather more on the difference explained by the predictors as a whole. Thus the two fold decomposition method is applied to identify the outcome difference that is explained by the rural-urban differences in the predictors and the unexplained part. Jann (2008) also documents that although the application of this technique can be found mostly in the labor market and discrimination literature (Cotton 1988; Stanley and Jarrell 1998; Oaxaca and Ransom 1999; Horrace and Oaxaca 2001; Weichselbaumer and Winter-Ebmer 2005) in general, this technique can be employed to study group differences in any outcome variable. A structural approach is used to employ the Blinder-Oaxaca decomposition method to equation 1, 2 and 3 & 4 simultaneously, to estimate how much of the wage difference is explained by the rural-urban control variables (equation 1), by across state differences in development policy (equation 2), and by industry mix (equation 3) and education share (equation 4). The ‘Oaxaca’ package in Stata statistical routine is used to reveal the decomposition estimates.

Chapter 4: Results and Discussion

Controls for the Rural-Urban Dichotomy

The initial OLS model is utilized to see how the distribution of average wage among counties reacts to the factors distinguishing between urban and rural counties in the country. Elasticities are reported as the coefficients of the independent variables. Interests are in interpreting and measuring the effects of independent variables on average wage (Appendix: B). Elasticities of an independent variable measures the impact of change in an independent variable on the expected change in the dependent variable in a regression model. For instance, the elasticity of independent variable *X* on dependent variable *Y* can be computed by taking the partial derivative of E(*ln(Y)*|*ln(X))* with respect to ln(*X)* (Wang 2007).

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}) \ldots\ldots \text{(1)}$$

The model predicts that the value of rural urban continuum code (RUCC13) reduces as average wage increases ($p<0.05$). This is expected and as evident in previous literature, the lower the RUCC13, the more urban the county and thus, the higher the value of average wage. As expected, the creative class vector has a positive interaction ($p<0.05$) with average wage. Communities with a higher share of creative class economic activities compared to places with less agglomeration of creative class groups enjoy higher returns and introduce more heterogeneity to the economic landscape. The social capital variable (SK09) also demonstrates a significant ($p<0.05$) positive relation to average wage. In counties with high average wage, people are perhaps more prone to civic engagement in comparison to counties with lower average returns, thus building a high stock of social capital. Though natural amenities index carries a negative coefficient, it is not significant ($p>0.05$), implying that counties with low natural endowments have high average wage. The value of agricultural land and buildings per acre (Aglb_Val) is significant ($p<0.05$) and positive; which implies that value of agricultural capital is high in counties (mostly urban) with high average wage. The Med_Rent variable, the median rent as percentage of household income, is significant ($p<0.05$) and negative. This documents that median rent doesn't increase at a rate the same as average wage across regions, thus with increasing returns the share of wage designated for rent decreases. The model describes about 22 percent of the variability ($p<0.001$) among the explanatory variables.

The state fixed effects

State fixed effects are introduced to improve the robustness of the initial model (equation 1). State fixed effects are a proxy for economic policy, development patterns that are unique to each state, and have effect on county average wage.

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}, \boldsymbol{i.state\_1\ to\ 48}) \quad \ldots\ldots \text{(2)}$$

Inclusion of state fixed effects improves the robustness of the model with an increase in the adjusted R^2 to 0.31 ($p<0.05$), meaning now approximately 31percent of the variability in average wage across counties is explained. The nature and direction of the controls for rural-urban characteristics aren't diverted except that the rural-urban continuum code becomes less significant ($p<0.05$) and the effect of natural amenities become significant ($p<0.05$) (Appendix: C).

Now, after controlling for urbanness, creative class share, social capital stock, natural amenities, land value and rent differentials; existing industry mix vectors and expected education vectors are introduced to the state fixed effect model (equation 2) to identify separately how the model responses to these groups of vectors.

The industry mix model

Although, controlling for existing industry mix variation across counties does not change the direction of the control variables, both social capital stock ($p = 0.98$) and natural amenities ($p = 0.61$) became insignificant (Appendix: D). Of the 86 industry sectors, the elasticity of health care stores, religious and civic organizations, general merchandise stores, personal and laundry services, food and beverage stores are not significant ($p>0.05$) on county average returns. Information services, broadcasting, petroleum industry and rental & leasing services have the highest significant ($p<0.05$) elasticity respectively on the county average returns.

$$\boldsymbol{avg\_wage} = \int(\boldsymbol{RUCC13}, \boldsymbol{Creative2000}, \boldsymbol{SK09}, \boldsymbol{Natam\_Rank}, \boldsymbol{Aglb\_Val}, \boldsymbol{Med\_Rent}, \boldsymbol{i.state\_1\ to\ 48}, \boldsymbol{ind01\ to\ 86}) \quad \ldots\ldots (3)$$

This indicates that industry mix teases out the effects of social capital and natural amenities on average returns. With significant ($p<0.05$) robustness, the industry mix model explains approximately 55percent of the variability of average returns across counties.

The education share model

When expected education share vectors are introduced to the state fixed effects model both county social capital stock ($p>0.05$) and effects of natural amenities ($p>0.05$) on average wage became insignificant; but unlike the industry mix model, rural-urban distinction vector (RUCC13) was also insignificant ($p>0.05$) (Appendix: E).

$$avg\_wage = \int(RUCC13, Creative2000, SK09, Natam\_Rank, Aglb\_Val, Med\_Rent, i.state\_1\ to\ 48, E01\ to\ 12)$$

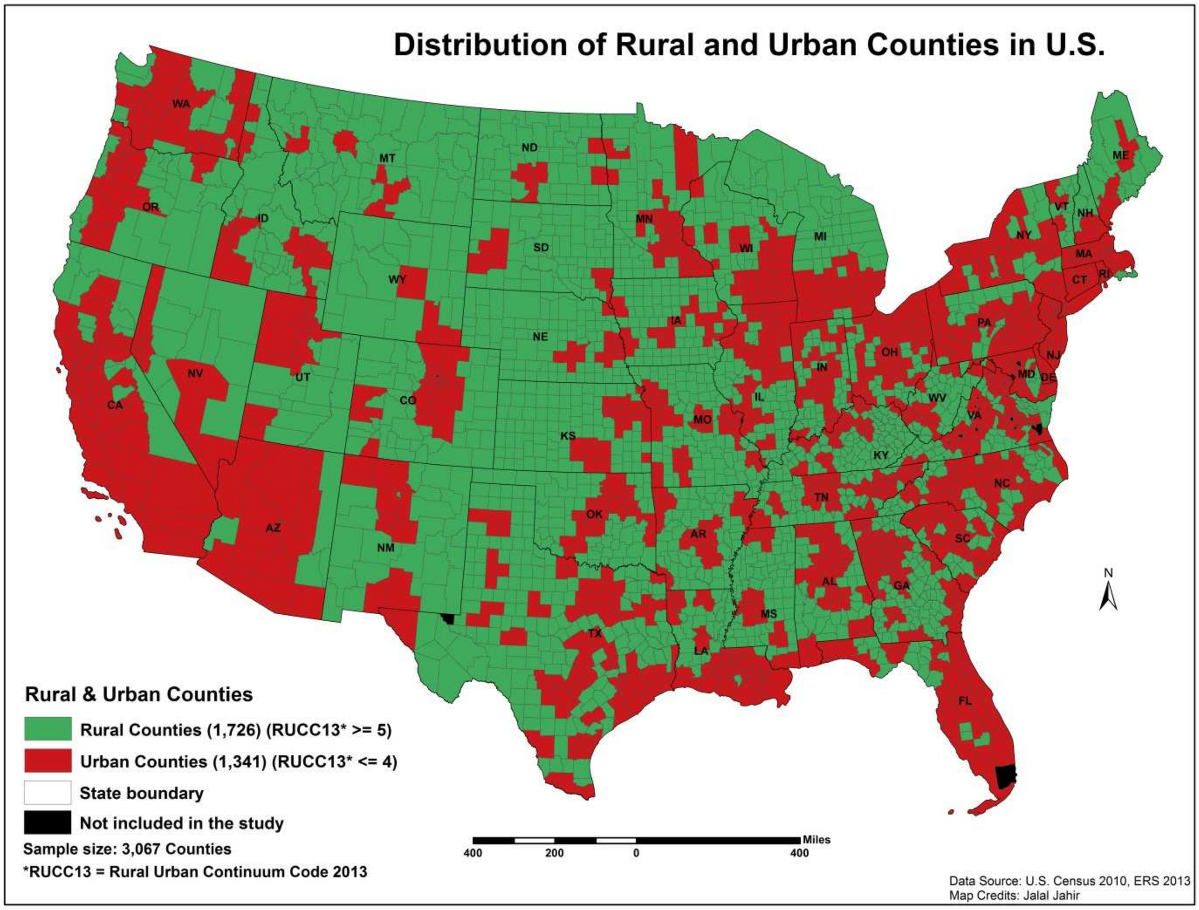
…… (4)

People with a high school diploma or less have the two highest (56.58 and 23.13 respectively) significant (p<0.05) elasticities on county level average wage. People with bachelor's degree have the third highest influence in determining county average wage. The effect of post baccalaureate certificate is the lowest on county average wage among 12 education segments and also not significant (p>0.05). This indicates that high school diploma is the leading player as a level of education in the county economy. This also points out that it is likely that the combination of skills, ability and knowledge required for most of the industries in the existing mix could be achieved with a high school level education. The general indication is that even though with industry mix variable in the model, rural-urban distinction has significant effect on the variation of average returns across counties: education share vector captures that particular variation. This provokes the inquiry if the effect of human capital on the average return is more prominent than county industry mix. In response to this query, the effects of the vectors introduced so far are sequentially decomposed.

The Blinder-Oaxaca decomposition

To decompose the effects of the vectors in equation 1, 2, 3 & 4, the counties are grouped into either urban or rural. The rural urban continuum code 2013 (ERS 2013) is used in creating a new vector 'urban', where counties with urban population of 20,000 or more (RUCC13 code 1-4) are coded as 1=urban and otherwise 0=rural (RUCC13 code 5-9). This groups 1,341 counties as urban and 1,726 counties as rural out of the sample of 3,067 counties in 48 states of the contiguous U.S. The two-fold decomposition method of the Blinder-Oaxaca technique is applied using Stata statistical routine where output reports the mean predictions by groups and their difference (Jann 2008).

Figure 4-1: The urban and the rural counties in the U.S.



Empirical results show that variables contrasting between urban and rural counties (equation 1) explain most of the differentiation in average returns between urban and rural counties (Table: 4-1). The rural urban continuum code, creative class vector, county stock of social capital, natural amenities, land value and proxy for living expense altogether account for about 82.37 percent of the variation in average wage between urban and rural counties. On one hand, this supports the selection of urban rural contrasting vectors for the model; on the other hand, it implies that there is about 18percent of the variation still unexplained by the model. Introduction of state fixed effects (equation 2) to control for across state variation in policy strategies and to account for unobserved variables (state specific) which have effects in determining average wage structure, yielded improvement in the robustness of the model as well as improvement in explaining rural urban wage variation. The state fixed effects model explains 87.52 percent of the variation in average wage due to urban rural distinction. This indicates that state level policy variation across states is responsible for 5.15 percent of the total rural-urban differences in returns. Yet about 13percent of the variation remains unexplained. The effects of the explanatory variables in the industry mix model (equation 3) and education share model (equation 4) are further decomposed to investigate separately which variable, the existing industry mix or the expected education (i.e. human capital), explains more of the remaining 13 percent of the variation in average returns between urban and rural areas.

Table 4-1: The Blinder-Oaxaca decomposition results

Eq.	**Description**	**Explained**	**p-val**	**Unexplained**	**p-val**	**Improvement in the Explained part**	
01	Controls for Rural-Urban Dichotomy	82.37%	p<0.05	17.63%	p<0.05	-	
02	State Fixed Effects	87.52%	p<0.05	12.48%	P>0.05	5.15%	from eq. 01
03	Existing Industry Mix	88.89%	p<0.05	11.11%	p>0.05	1.37%	from eq. 02
04	Expected Education Share	99.95%	p<0.05	0.05%	p>0.05	12.43%	from eq. 02

Introduction of industry mix vectors in equation 2 further improves the robustness of the model (adjusted R^2= 0.55, $p<0.001$) as well as a small improvement in explaining the wage variation. It was expected that agglomeration economies, that is the clustering of industries, would account for, if not all, a majority of the percent unexplained variation. But it turned out that, industry mix variation across counties is not a major player in the rural-urban wage inequality as the existing industry mix variation is only responsible for about 1.37 percent of the total differences in returns; the remaining 11.11 percent unexplained variation is not statistically significant ($P>0.05$). Expected education attributes are then introduced in equation 2 as a proxy for county human capital variation. The expected education share model (equation 4) with significant ($p<0.05$) adjusted robustness of 0.43, is able to explain almost all the variation (99.95 percent) of average wage between urban and rural counties. The unexplained portion of 0.05 percent is also statistically insignificant for the expected education share model.

The implications of the decomposition results are three fold: first physical, social and natural distinctions between rural and urban counties account for the majority (82.37 percent) of the rural-urban variation ($p<0.05$) of average returns; secondly, across state policy differences explains a significant ($p<0.05$) but small portion (5.15 percent) of this variation; third, returns to human capital (i.e. expected education) explains more (12.43 percent, $p<0.05$) of the wage differentials than average returns to county industry mix (1.37 percent, $p<0.05$).

The physical, social and natural attributes are more the unchangeable conditions, with few exceptions. These attributes are likely to take longer period of time to change in response to any policy inference. The access to natural amenities is entirely a fixed attribute; other than policies to retain the existing resources, it is not possible to create access to natural amenities for communities that don’t have any. The accumulation of social capital is dependent on local norms and cultures; and increase in the stock of county social capital is likely a gradual process. As like the expected change in the rural-urban physical distinctions, social capitol change may have implications in the long run. The land value and median rent vector also move along with the gradual process of change in the rural-urban physical distinction. Thus the attributes contrasting between urban and rural counties, although these explain most of the variation in the returns, is less likely to have implication in reducing the rural-urban wage gap from the policy perspective compared to the state policies, the existing industry mix, and the education share scenario.

The state fixed effects explain a portion of the wage variation. As argued in the previous literature, being a subset of a state, the counties are influenced to a certain level by state policy-strategies. With each state having distinct economic development policies and a view to local community development, most likely all share the similar goal of sustainable economic growth and efficient utilization of accessible resources. This simplifies the task of calling for harmony in inter-state development policies. On the other hand, regular dialogue, exchange of thoughts and greater familiarity of neighboring communities' economic interest and resources will likely create a more favorable economic environment; which in turn will play a more effective role in reducing the wage gap.

According to the structural approach in decomposing the effects of the vectors, the existing industry mix vectors contribute less, compared to other sets of attributes, in explaining the prevailing wage gap. The effect of economies of scale, agglomeration of economies and clustering of industries in local and regional economic growth is well documented in contemporary literature. Thus the preliminary assumption was that the likelihood of the industry mix vectors explaining the greater portion of the wage gap was higher than the human capital. But the results suggest that the existing industry mix difference is not likely the major reason resulting in the wage gap. Unlike in the state fixed effects models, when the industry mix vectors are explaining the varying average returns, the county stock of social capital no longer remains as a significant ($p>0.05$) predictor. This implies that without the industry clustering, community social capital is positively related to average returns. Greater stock of social capital means more mutual trust among community members which is important in promoting new business which leads to creation of new jobs, more sharing of information which lead to more access to job leads. But when the industry mix is specified, the effect of social capital on returns reduces and the variation of county average returns becomes more attributable to the industry clusters.

The education share variables delineate the expected education for the existing county industry mix. In other words, the education vectors are more of the depictions of qualifications for employment in the industries. Thus the expected education vectors are likely the proxy for the human capital required to have access to jobs in the existing county industry mix. It is likely that the education vectors, as these are more of the representation of the expected human capital rather only educational attainments, would be able to explain the rural-urban wage gap significantly ($p<0.05$). As expected the education vectors explained the majority of the variation in rural-urban differences in average return. After controlling for rural-urban dichotomy, across state policy differentials, social, economic and natural attributes; the models predict that similar level of education achieves more returns in urban areas than the rural counterpart. The land value and median rent variables controls for the differences in expenses in rural and urban areas, strengthening the argument of higher average returns to similar human capital in urban than in rural counties. The increasing difference in returns to human capital between urban and rural regions motivates the growth of income inequality, resulting in an increased number of deprived people in the community with lack of access to resources. Thus another indication of the result is the call for labor force improvement in rural areas as well as policy development to attract higher skilled labor into the rural regions. The decomposition estimates would be helpful for policy makers to determine where (state strategy, industrial policy or human capital development) and how much (which factor is most important) attention is required to address the issue of the rural-urban wage gap; in other words, in formulating policies for rural poverty alleviation.

The test for spatial dependence: Exploratory Spatial Data Analysis

Spatial autocorrelation analysis involves tests and visualization of both the global Moran's I statistic, which is the test for clustering as well as the local Moran's I statistic, which is the test for clusters. The global spatial autocorrelation only identifies the overall clustering, not the location of clusters or outliers, nor the significance of such clustering. The Local Moran statistic, Local indicator of spatial autocorrelation (LISA) (Anselin 1995), addresses this issue by providing a means to assess significance of the ''local'' spatial patterns, where classes like 'high–high' and 'low–low' indicates significant local clusters and 'high–low' and 'low–high' indicates the local spatial outliers (Anselin et al. 2007). The global test is visualized by means of a Moran scatter plot (Anselin 2005) where the slope of the regression line corresponds to Moran's I. The local analysis is based on the Local Moran statistic (Anselin 1995), which is visualized in the form of the significance and the cluster maps (Anselin et al. 2006). The global and the local Moran's I statistics for both the average wage and the residuals from equation 01 are estimated using the first order Queen's contiguity weighting matrix in Geoda routine (Anselin 2005).

Figure 4-2: The global Moran's I scatter plots for average wage and residual of equation 01

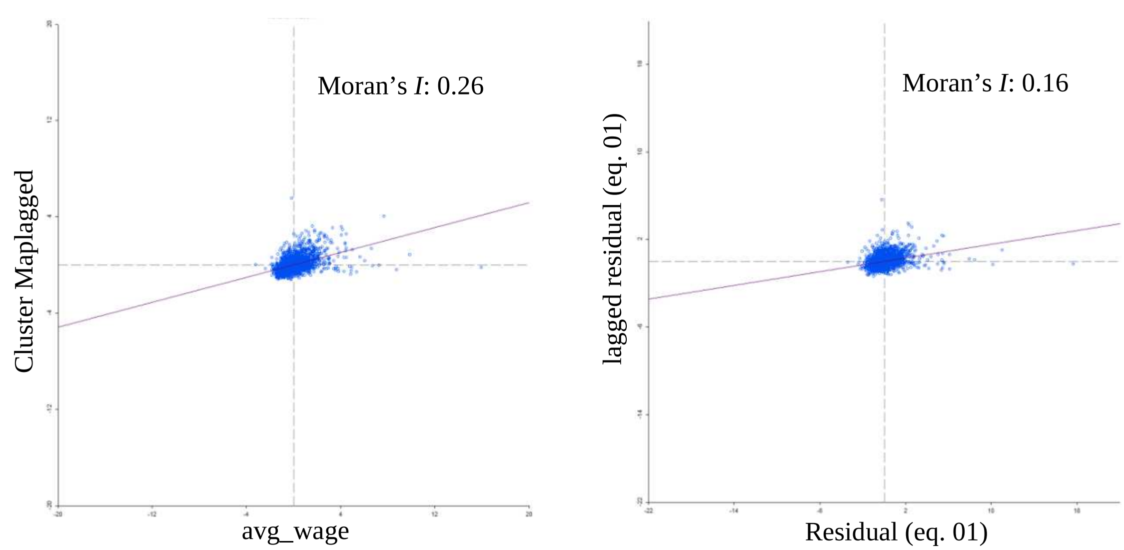


Figure 4-3: The LISA cluster map of the average wage vector

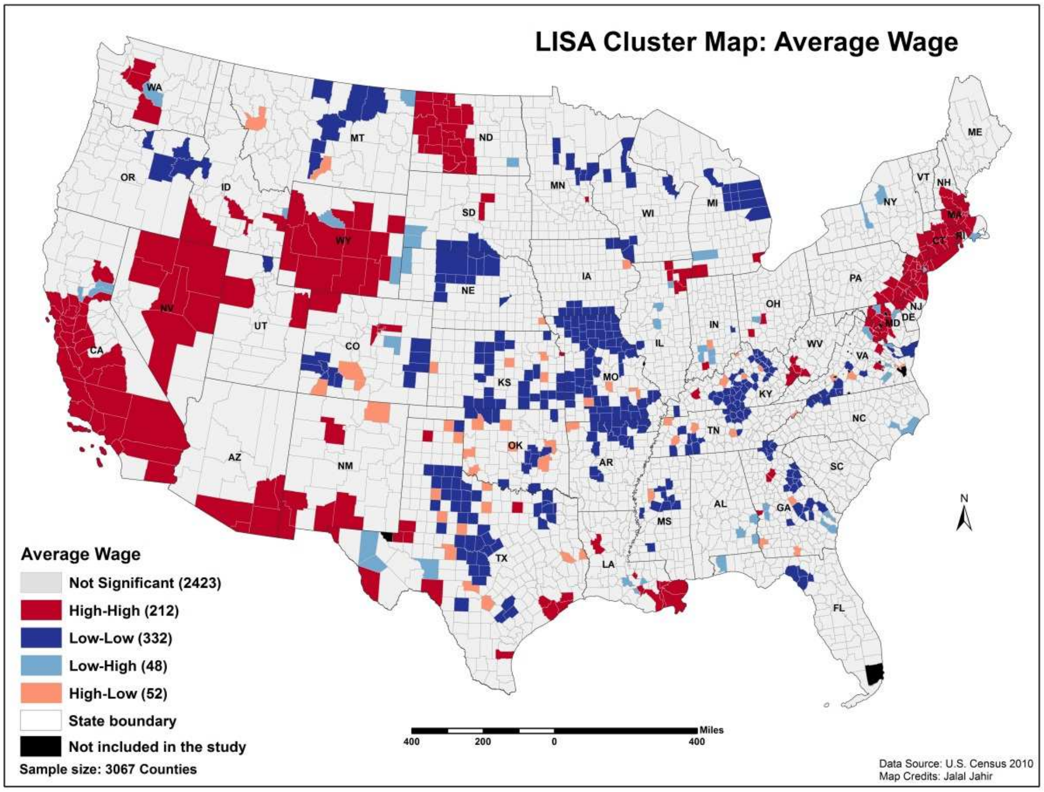
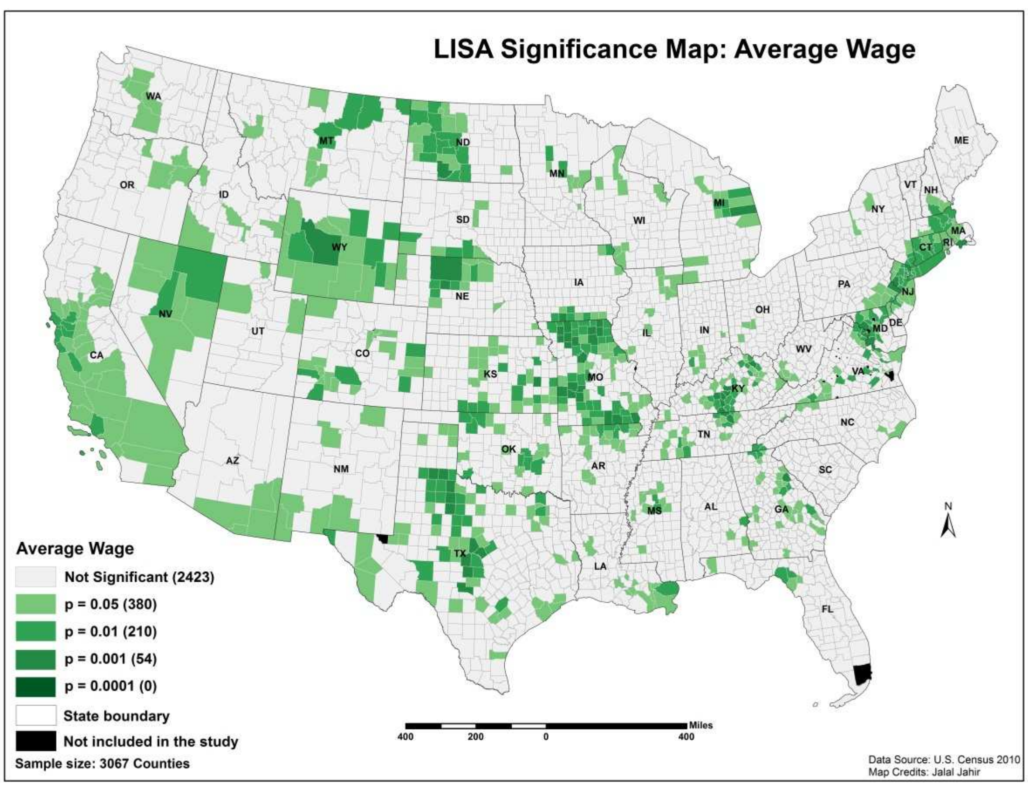
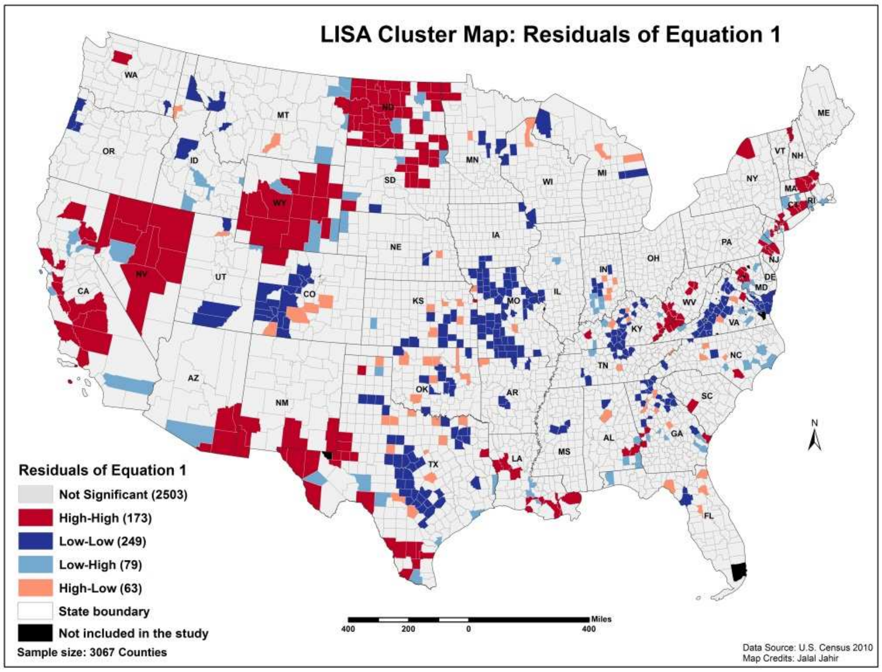


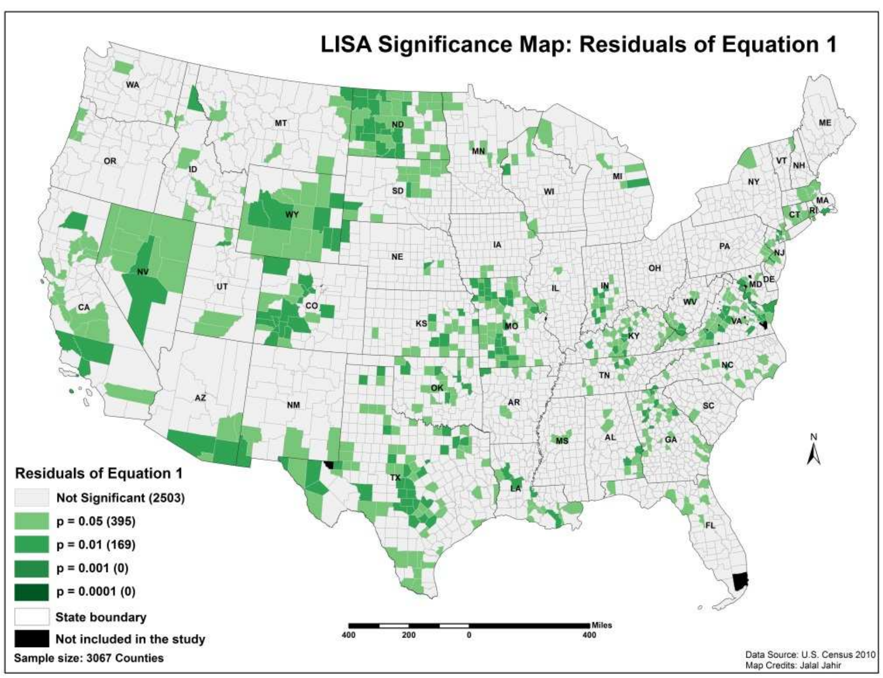
Figure 4-4: The LISA significance map of the average wage vector



The global Moran's I for the average wage vector is 0.26 ($p<0.05$) which indicates there is positive clustering among the observations. The LISA significance map (fig 4-2) for average wage shows that among 3067 counties, clustering of about 79% are not significant and only 644 counties show significant clustering ($p<0.05$) with 54 of them highly significant ($p<0.001$). The LISA cluster map (fig 4-3) indicates among the counties with significant clustering, 212 counties with high average wage, of which 157 counties are urban, tend to cluster around each other. The largest clusters consist of all urban counties and are along the west coast in the state of California (29 counties) and along the east coast in the State of New York, New Jersey, Maryland etc. (67 counties). The other large clusters are in North Dakota (18 counties), in Wyoming (16 counties) and in Nevada (08 counties). 332 counties with comparatively low average wage clusters around each other of which 289 counties are rural. There are 48 counties with low average wage tend to cluster near counties with high average wage, which is due to spatial spillover effect of economic activities and benefit of spatial agglomeration. On the other hand, 52 counties with high average wages tend to cluster near counties with comparatively high average wages. These 48 and 52 counties are local spatial outliers (Anselin et al. 2007) with scattered distribution across space. The OLS regression (equation 01) captures about 39% of the spatial clustering effects resulting in a global Moran's I of the residual of 0.16 ($p<0.05$). The LISA significance map (fig 4-4) shows that counties with insignificant clustering increases from 79% to 82% and only 564 counties shows significant ($p<0.05$) clustering with none of the counties as highly significant ($p<0.001$). The LISA cluster map (fig 4-4) of the residual shows that, most of the counties from the two largest clusters (previous) became insignificant, reducing the number of high-high clustering counties from 212 to 173. Also the number of counties with low-low clustering is reduced from 332 to 249. There is an increase in the number of spatial outlier, from 100 to 142 counties, indicating that with the correction for spatial dependence, 42 counties from different clustered becomes spatial outliers.

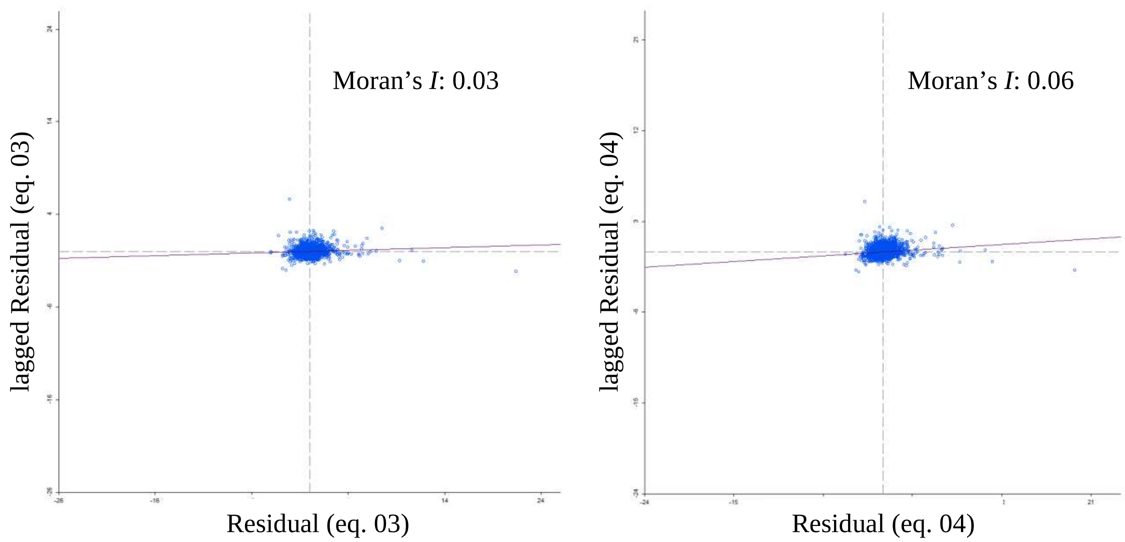
Figure 4-5: The LISA cluster map and the significance map of residual of equation 01





The residuals of the industry mix model with the state fixed effects (equation 03) and the education model with the state fixed effects (equation 4) are further investigated to reveal if there is any improvements (reduction of spatial dependence/clustering) in the spatial autocorrelation indexes. In case of the industry mix model with state fixed effects, the global Moran's I value is 0.03 ($p<0.05$), an 81% reduction of spatial dependence from equation 1. Also number of counties with significant clustering decreased from 564 to 335.

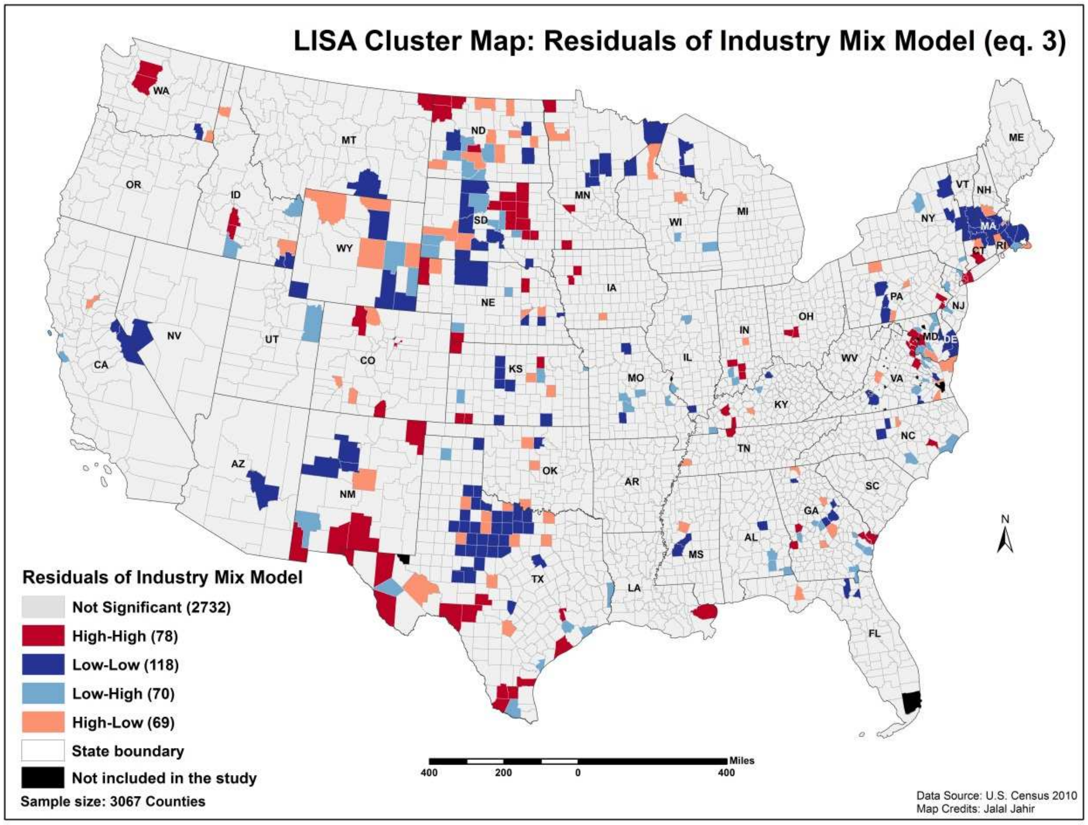
Figure 4-6: The global Moran's I scatter plots of residuals from equation 03 and equation 04



The residuals of the education model shows a global Moran's I index of 0.06 ($p<0.005$), also about 63% reduction of spatial dependence from equation 1. Significant ($p<0.05$) clustering is not evident in about 87% of the counties.

Thus with the global Moran's I indexes of residuals close to randomness (<0.1) and notable decrease in the number of counties with significant clustering, it is evident that both the industry mix model and the education model have captured most of the spatial dependence and thus it is highly unlikely that the remaining low level of spatial dependence have vastly contributed in biasing the estimations.

Figure 4-7: The LISA cluster and the significance maps of the residuals of the industry mix model (eq. 03)



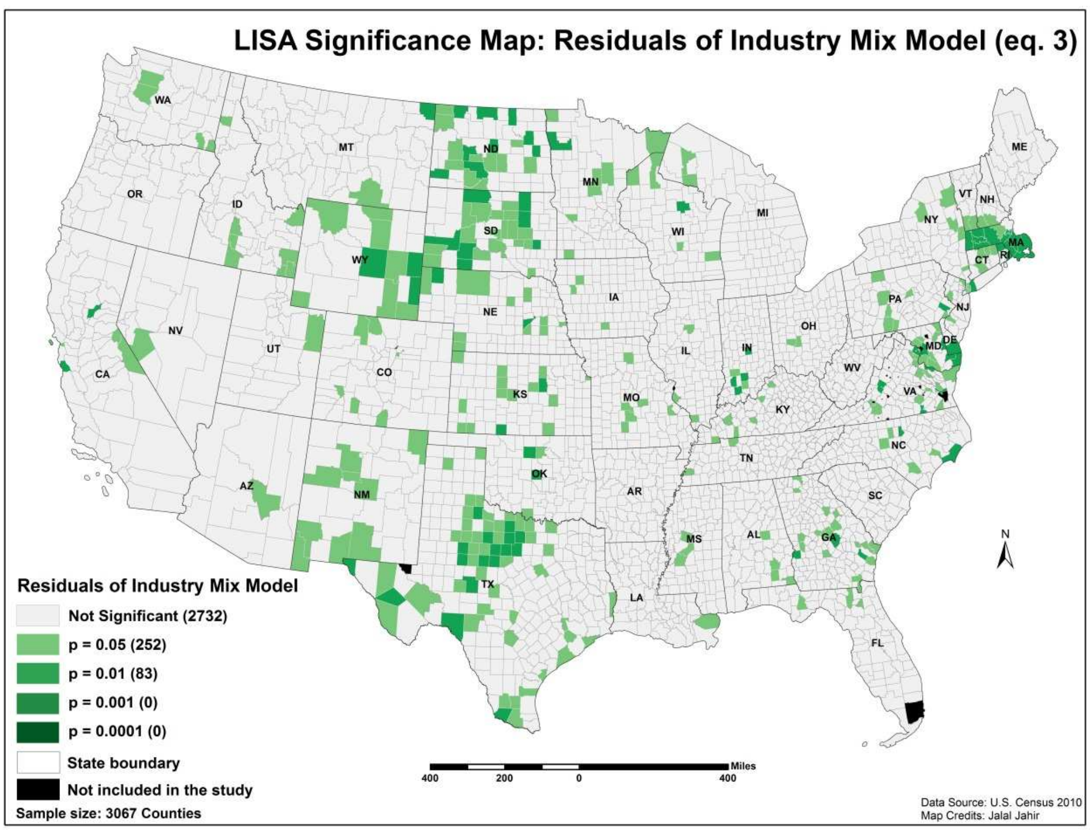
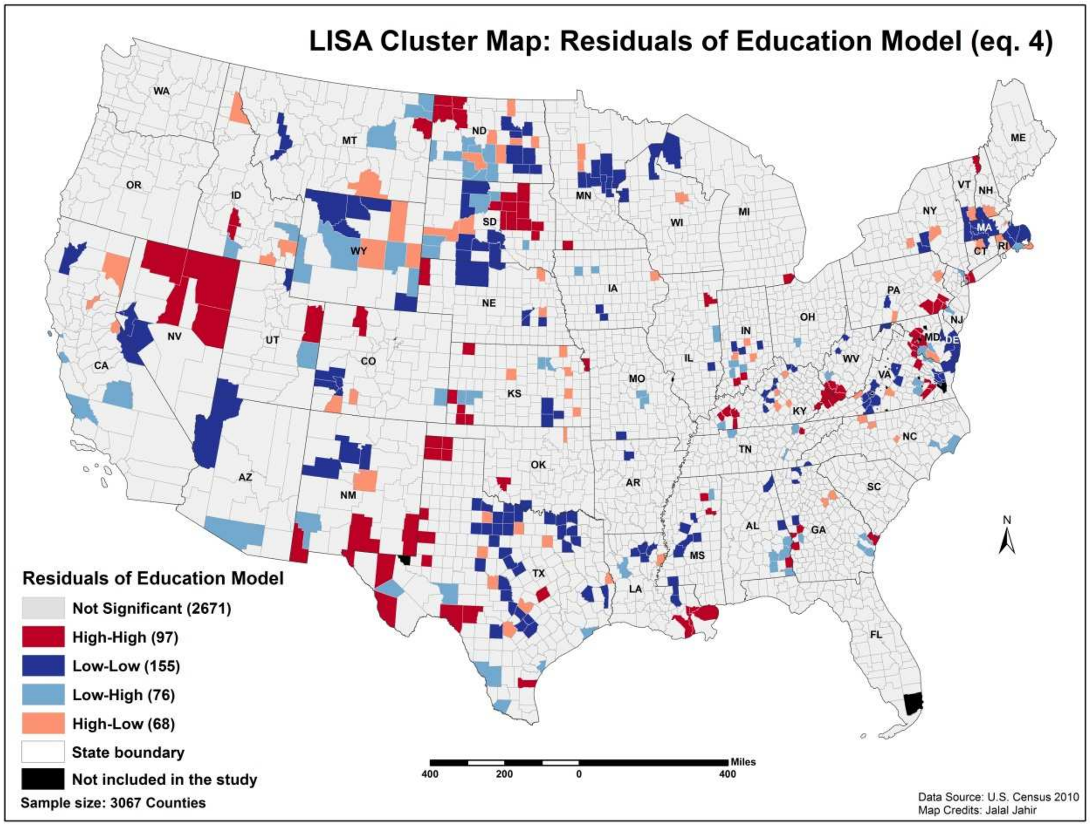
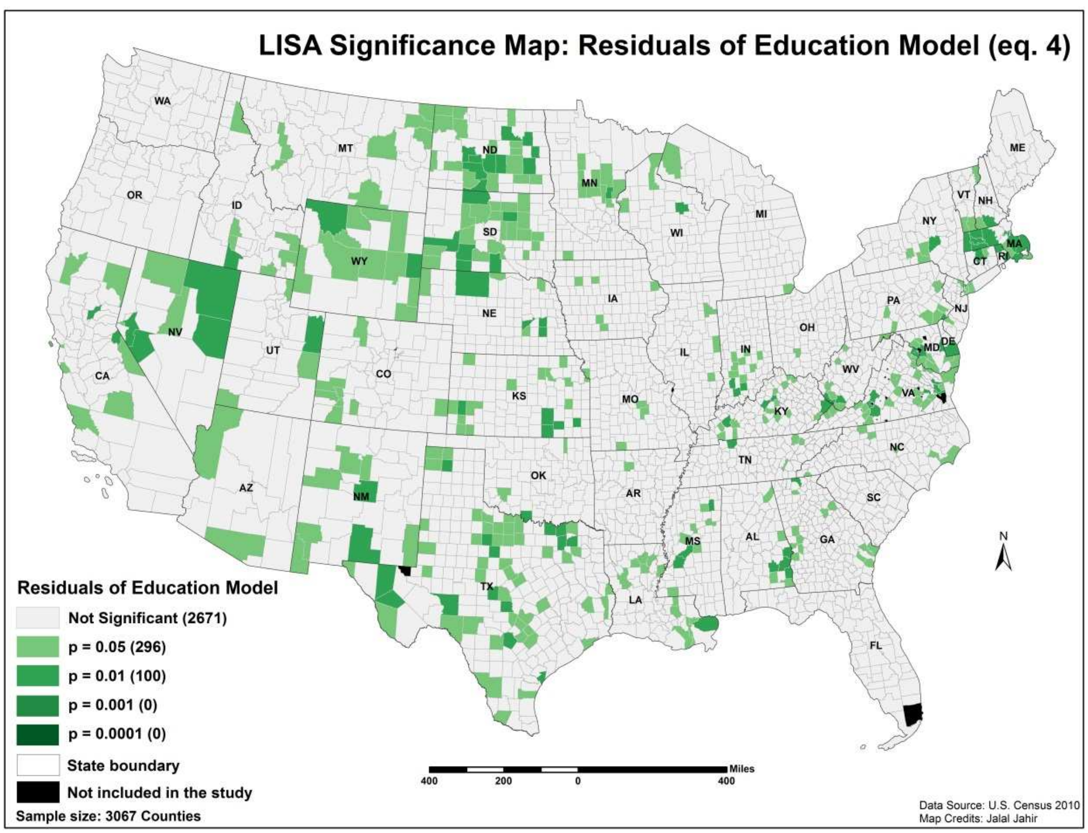


Figure 4-8: The LISA cluster map and the significace map of the residuals of the education model (eq. 04)





Chapter 5: Conclusion

Expectation was that controlling for physical, natural, social and economic distinction would tease out the rural-urban differences in average returns. But the rural-urban physical distinction, rural-urban differences in social capital stock, natural endowments, land value differences and variation in living expenses altogether explain the majority of the rural-urban wage differences, but not all. Inclusion of across state policy variability resulted in an improvement in the model capacity to explain the wage inequality, but there still remains an unexplained share of the variation. Further it was identified that the county industry mix explains only a small portion of the remaining variability but county education share are able to explain approximately 12 percent by itself and about 99.05 percent of the rural-urban wage disparity as a whole. As the education share vectors are representative of expected education given the industry mix, these vectors are more reflective of the composition of human capital required to meet the demand of existing industry mix. This implies that differences in the industry mix, agglomeration economies and spatial economic spillover effects between rural and urban areas are less influential than the inconsistency in the returns to human capital (expected education). Even after controlling for rural-urban segregation, state based policy variances, social, economic and natural attributes; the models predict that similar level of education achieves more returns in urban areas than the rural counterpart. The literatures unanimously support the findings as Mouw and Kalleberg (2010) documents that changes in education level and industry mix is related to the structure of wage inequality and also Moller et al. (2009) finds that higher levels of education reduce inequality, especially high school degree. The models in this study indicates that industry mix, a place based phenomenon is likely less influential in explaining wage differentials compare to the human capital. Fallah et al. (2011) also finds that wage inequality is more responsive to the human capital rather than the industry composition; which also resembles to the conclusion of Andersson et al. (2013) who argued that the qualities of the labor forces are more important than where the labor forces are located in explaining spatial wage differentiations. The findings of this study trigger the need for improved rural-urban relationships by addressing unbiased development policy formulation. The growing difference in returns to human capital between urban and rural regions is contributing to the growing income inequality, hence fostering increase in the number of people with lack of access to resources. The increasing rural-urban gap in average returns to similar qualities is resulting in increase in land value, high commodity prices, health care etc. that is increase in cost of living. These effects don't stay confined to urban boundaries, though extreme in urban core and proximities; rather spread over to rural regions where people find it more difficult to cope with the increase in expenses with less returns compared to urban counterparts. The results of this study inform where to pay attention when seeking rural poverty alleviation. On one hand, with further improvement in industry mix, rural areas need policy attention to improve the wage structure to reduce distinction between urban counterparts. These require policy initiatives like import substitution and export promotion. On the other hand there is a need for the policies which will train or attract higher skilled workers, rather than the policies which simply "create jobs". These policies include technological advancement etc. which also foster the policy initiatives mentioned above. As documented by Partridge and Rickman (2005), both completion of high school degree and obtaining associate degree reduce poverty in high-poverty rural counties. Creating more traditional jobs without improving the quality of the workforce is likely to exacerbate the rural-urban wage gap.

This study invites further research on identifying more vectors for urban rural distinction, examine possible missing (observable/unobservable) variable biases (e.g. county fixed effect models), investigation of the effects of other human capital (e.g. skills, abilities, knowledge) separately on average returns, incorporation of the spatial autoregressive modeling techniques for further model improvements and inclusion of both the county human capital stock and the industry mix attributes in the same model to analyze the model performances.

References

Allison PD (2005) Fixed effects regression methods for longitudinal data: Using SAS. SAS Institute

Andersson M, Klaesson J, Larsson JP (2013) The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies? Pap Reg Sci n/a–n/a. doi: 10.1111/pirs.12025

Anselin L (1999a) The future of spatial analysis in the social sciences. Geogr Inf Sci 5:67–76.

Anselin L (1999b) Interactive techniques and exploratory spatial data analysis. Geogr. Inf. Syst. Princ. Tech. Manag. Appl. Eds P Longley M Goodchild Maguire Rhind Camb. Geoinformation Int

Anselin L (2005) Exploring Spatial Data With GeoDa: A Work Book. Spatial Analysis Laboratory, University of Illinois. Cent. Spatially Integr. Soc. Sci.

Anselin L (1995) Local Indicators of Spatial Association—LISA. Geogr Anal 27:93–115. doi: 10.1111/j.1538-4632.1995.tb00338.x

Anselin L, Sridharan S, Gholston S (2007) Using Exploratory Spatial Data Analysis to Leverage Social Indicator Databases: The Discovery of Interesting Patterns. Soc Indic Res 82:287–309. doi: 10.1007/s11205-006-9034-x

Anselin L, Syabri I, Kho Y (2006) GeoDa: An Introduction to Spatial Data Analysis. Geogr Anal 38:5–22. doi: 10.1111/j.0016-7363.2005.00671.x

Bacolod M, Blum BS, Strange WC (2009) Skills in the city. J Urban Econ 65:136–153. doi: 10.1016/j.jue.2008.09.003

Blinder AS (1973) Wage Discrimination: Reduced Form and Structural Estimates. J Hum Resour 8:436–455. doi: 10.2307/144855

Bollinger C, Ziliak James P., Troske KR (2011) Down from the Mountain: Skill Upgrading and Wages in Appalachia. J Labor Econ 29:819–857. doi: 10.1086/660773

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2010) A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 1:97–111. doi: 10.1002/jrsm.12

Clarke P, Crawford C, Steele F, Vignoles AF (2010) The choice between fixed and random effects models: some considerations for educational research.

Combes P-P, Duranton G, Gobillon L (2008) Spatial wage disparities: Sorting matters! J Urban Econ 63:723–742. doi: 10.1016/j.jue.2007.04.004

Cotton J (1988) On the Decomposition of Wage Differentials. Rev Econ Stat 70:236–243. doi: 10.2307/1928307

Curtis KJ, Voss PR, Long DD (2012) Spatial variation in poverty-generating processes: Child poverty in the United States. Soc Sci Res 41:146–159. doi: 10.1016/j.ssresearch.2011.07.007

Deller SC, Tsai T-H (Sue), Marcouiller DW, English DBK (2001) The Role of Amenities and Quality of Life In Rural Economic Growth. Am J Agric Econ 83:352–365. doi: 10.1111/0002-9092.00161

England P, Farkas G, Kilbourne BS, Dou T (1988) Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects. Am Sociol Rev 53:544–558. doi: 10.2307/2095848

ERS (2013) USDA Economic Research Service - Rural-Urban Continuum Codes. http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx. Accessed 18 Feb 2014

ERS (2000) USDA Economic Research Service - Creative Class County Codes. http://www.ers.usda.gov/data-products/creative-class-county-codes.aspx. Accessed 18 Feb 2014

ERS (1999) USDA Economic Research Service - Natural Amenities Scale. http://www.ers.usda.gov/data-products/natural-amenities-scale.aspx. Accessed 18 Feb 2014

Fallah BN, Partridge MD, Olfert MR (2011) New economic geography and US metropolitan wage inequality. J Econ Geogr 11:865–895. doi: 10.1093/jeg/lbq016

Florida R, Mellander C, Stolarick K (2008) Inside the black box of regional development—human capital, the creative class and tolerance. J Econ Geogr 8:615–649.

Glaeser EL, Kolko J, Saiz A (2001) Consumer city. J Econ Geogr 1:27–50.

Goodchild MF, Anselin L, Appelbaum RP, Harthorn BH (2000) Toward Spatially Integrated Social Science. Int Reg Sci Rev 23:139–159. doi: 10.1177/016001760002300201

Heuermann D, Halfdanarson B, Suedekum J (2010) Human Capital Externalities and the Urban Wage Premium: Two Literatures and their Interrelations. Urban Stud 47:749–767. doi: 10.1177/0042098009352363

Horrace WC, Oaxaca RL (2001) Inter-Industry Wage Differentials and the Gender Wage Gap: An Identification Problem. Ind Labor Relat Rev 54:611–618. doi: 10.2307/2695992

Jacobs D (1985) Unequal Organizations or Unequal Attainments? An Empirical Comparison of Sectoral and Individualistic Explanations for Aggregate Inequality. Am Sociol Rev 50:166–180. doi: 10.2307/2095407

Jann B (2008) The Blinder-Oaxaca decomposition for linear regression models. Stata J 8:453–479.

Kuznets S (1955) Economic Growth and Income Inequality. Am Econ Rev 45:1–28.

Kuznets S, Jenks E (1953) Shares of Upper Income Groups in Savings. Shares Up. Income Groups Income Sav. NBER, pp 171–218

Levernier W, Partridge MD, Rickman DS (1998) Differences in Metropolitan and Nonmetropolitan U.S. Family Income Inequality: A Cross-County Comparison. J Urban Econ 44:272–290. doi: 10.1006/juec.1997.2070

Lobao L (2004) Continuity and Change in Place Stratification: Spatial Inequality and Middle-Range Territorial Units*. Rural Sociol 69:1–30. doi: 10.1526/003601104322919883

Lobao L, Saenz R (2002) Spatial Inequality and Diversity as an Emerging Research Area*. Rural Sociol 67:497–511. doi: 10.1111/j.1549-0831.2002.tb00116.x

Lynch RG (2003) Estimates of Income and Income Inequality in the United States and in Each of the Fifty States: 1988–1999. J Reg Sci 43:571–588. doi: 10.1111/1467-9787.00311

Marcouiller DW, Kim K-K, Deller SC (2004) Natural amenities, tourism and income distribution. Ann Tour Res 31:1031–1050. doi: 10.1016/j.annals.2004.04.003

MIG (2000) IMPLAN Professional Version 2.0 Social Accounting and Impact Analysis Software. Minn. IMPLAN Group Inc Stillwater Minn.

Moller S, Alderson AS, Nielsen F (2009) Changing Patterns of Income Inequality in U.S. Counties, 1970–2000. Am J Sociol 114:1037–1101. doi: 10.1086/592200

Morrill R (2000) Geographic variation in change in income inequality among US states, 1970–1990. Ann Reg Sci 34:109–130. doi: 10.1007/s001680050129

Mouw T, Kalleberg AL (2010) Occupations and the Structure of Wage Inequality in the United States, 1980s to 2000s. Am Sociol Rev 75:402–431. doi: 10.1177/0003122410363564

Nielsen F, Alderson AS (1997) The Kuznets Curve and the Great U-Turn: Income Inequality in U.S. Counties, 1970 to 1990. Am Sociol Rev 62:12–33. doi: 10.2307/2657450

O'Sullivan D, Unwin DJ (2010) Geographic Information Analysis, 2nd Edition. John Wiley & Sons, Inc., New York

Oaxaca R (1973) Male-Female Wage Differentials in Urban Labor Markets. Int Econ Rev 14:693–709. doi: 10.2307/2525981

Oaxaca RL, Ransom MR (1999) Identification in Detailed Wage Decompositions. Rev Econ Stat 81:154–157. doi: 10.1162/003465399767923908

Olson D, Lindall S (1999) IMPLAN Professional Version 2.0 Social Accounting and Impact Analysis Software—User's Guide, Analysis Guide, and Data Guide. Minnesota IMPLAN Group, Inc

Paci R, Usai S (2008) Agglomeration economies, spatial dependence and local industry growth. Rev Déconomie Ind 87–109. doi: 10.4000/rei.3917

Partridge MD, Rickman DS (2005) High-Poverty Nonmetropolitan Counties in America: Can Economic Development Help? Int Reg Sci Rev 28:415–440. doi: 10.1177/0160017605278998

Partridge MD, Rickman DS, Levernier W (1996) Trends in U.S. income inequality: Evidence from a panel of states. Q Rev Econ Finance 36:17–37. doi: 10.1016/S1062-9769(96)90027-7

Peters DJ (2013) American income inequality across economic and geographic space, 1970–2010. Soc Sci Res 42:1490–1504. doi: 10.1016/j.ssresearch.2013.06.009

Peters DJ (2011) Place-Based Income Inequality Clusters in the Rural North Central Region, 1979–2009. Appl Econ Perspect Policy 33:222–240. doi: 10.1093/aepp/ppr007

Rappaport J, Sachs JD (2003) The United States as a coastal nation. J Econ Growth 8:5–46.

Roback J (1988) Wages, rents, and amenities: differences among workers and regions. Econ Inq 26:23–41.

Roback J (1982) Wages, Rents, and the Quality of Life. J Polit Econ 90:1257–1278.

Rupasingha A, Goetz SJ (2008) US county-level social capital data, 1990-2005. Northeast Reg. Cent. Rural Dev. Penn State Univ. Univ. Park PA

Rupasingha A, Goetz SJ, Freshwater D (2006) The production of social capital in US counties. J Socio-Econ 35:83–101.

Slack T, Myers CA (2012) Understanding the geography of food stamp program participation: Do space and place matter? Soc Sci Res 41:263–275. doi: 10.1016/j.ssresearch.2011.10.001

Stanley TD, Jarrell SB (1998) Gender Wage Discrimination Bias? A Meta-Regression Analysis. J Hum Resour 33:947–973. doi: 10.2307/146404

Stewart CT Jr (1958) The Urban-Rural Dichotomy: Concepts and Uses. Am J Sociol 64:152–158.

Tobler WR (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. Econ Geogr 46:234–240. doi: 10.2307/143141

Wang C, Wu J (2011) Natural amenities, increasing returns and urban development. J Econ Geogr 11:687–707. doi: 10.1093/jeg/lbq020

Wang C-L (2007) Marginal Effects. Encycl. Soc. Sci. Res. Methods. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States, pp 609–610

Weichselbaumer D, Winter-Ebmer R (2005) A Meta-Analysis of the International Gender Wage Gap. J Econ Surv 19:479–511.

Wilson JB, Thilmany D, Watson P (2006) The Role of Agritourism in Western States: Place-Specific and Policy Factors Influencing Recreational Income for Producers. Rev Reg Stud 36:381–399.

Wu J (2006) Environmental amenities, urban sprawl, and community characteristics. J Environ Econ Manag 52:527–547. doi: 10.1016/j.jeem.2006.03.003

Appendix A

List of Variables

Appendix A: List of Variables

Variable	Description	Source
RUCC13	2013 Rural-urban Continuum Codes	ERS 2013
	Metropolitan Counties =	
	1 = Counties in metro areas of 1 million population or more	
	2 = Counties in metro areas of 250,000 to 1 million population	
	3 = Counties in metro areas of fewer than 250,000 population	
	Nonmetropolitan Counties =	
	4 = Urban population of 20,000 or more, adjacent to a metro area	
	5 = Urban population of 20,000 or more, not adjacent to a metro area	
	6 = Urban population of 2,500 to 19,999, adjacent to a metro area	
	7 = Urban population of 2,500 to 19,999, not adjacent to a metro area	
	8 = Completely rural or less than 2,500 urban population, adjacent to a metro area	
	9 = Completely rural or less than 2,500 urban population, not adjacent to a metro area	
Creative2000	2000 Creative Class County Codes	ERS 2000
SK09	2009 Social Capital Index	Rupasingha and Goetz 2008
NatAm_Rank	1999 Natural Amenity Rank	ERS 1999
	Natural amenity rank (1=low; 7=high)	
Aglb_Val	Estimated market value of land and buildings \ Average per acre (dollars)	Census of Agriculture 2007
Med_Rent	Median Gross Rent As A Percentage of Household Income (Dollars)	5 Year ACS 2011, Table: B25071_001
Urban	If RUCC13<5, 1=Urban and if RUCC13>4, 0=Rural	
	List of Education Vectors	MIG 2000
E01	Associate's Degree (or other 2-year degree)	
E02	Bachelor's Degree	
E03	Doctoral Degree	
E04	First Professional Degree	
E05	High School Diploma (or Equivalence)	
E06	Less than a High School Diploma	
E07	Master's Degree	
E08	Post-Baccalaureate Certificate	
E09	Post-Doctoral Training	
E10	Post-Master's Certificate	
E11	Post-Secondary Certificate	
E12	Some College Courses	
	List of Industry Mix Vectors	MIG 2000
Ind111	Crop Farming	
Ind112	Livestock	
Ind113	Forestry and Logging	
Ind114	Fishing- HuntingTrapping	
Ind115	Support Activities for Agriculture and Forestry	
Ind211	Oil and Gas Extraction	
Ind212	Mining (except Oil and Gas)	
Ind213	Support Activities for Mining	

Variable	Description	Source
Ind221	Utilities	
Ind230	Construction of Buildings	
Ind311	Food Manufacturing	
Ind312	Beverage and Tobacco Product Manufacturing	
Ind313	Textile Mills	
Ind314	Textile Product Mills	
Ind316	Leather and Allied Product Manufacturing	
Ind321	Wood Product Manufacturing	
Ind322	Paper Manufacturing	
Ind323	Printing and Related Support Activities	
Ind324	Petroleum and Coal Products Manufacturing	
Ind325	Chemical Manufacturing	
Ind326	Plastics and Rubber Products Manufacturing	
Ind327	Nonmetallic Mineral Product Manufacturing	
Ind331	Primary Metal Manufacturing	
Ind332	Fabricated Metal Product Manufacturing	
Ind333	Machinery Manufacturing	
Ind334	Computer and Electronic Product Manufacturing	
Ind335	Electrical Equipment, Appliance, and Component Manufacturing	
Ind336	Transportation Equipment Manufacturing	
Ind337	Furniture and Related Product Manufacturing	
Ind339	Miscellaneous Manufacturing	
Ind42	Wholesale Trade	
Ind441	Motor Vehicle and Parts Dealers	
Ind442	Furniture and Home Furnishings Stores	
Ind443	Electronics and Appliance Stores	
Ind444	Building Material and Garden Equipment and Supplies Dealers	
Ind445	Food and Beverage Stores	
Ind446	Health and Personal Care Stores	
Ind447	Gasoline Stations	
Ind448	Clothing and Clothing Accessories Stores	
Ind451	Sporting Goods, Hobby, Book, and Music Stores	
Ind452	General Merchandise Stores	
Ind453	Miscellaneous Store Retailers	
Ind454	Nonstore Retailers	
Ind481	Air Transportation	
Ind482	Rail Transportation	
Ind483	Water Transportation	
Ind484	Truck Transportation	
Ind485	Transit and Ground Passenger Transportation	
Ind486	Pipeline Transportation	
Ind487	Scenic and Sightseeing Transportation	
Ind492	Couriers and Messengers	
Ind493	Warehousing and Storage	
Ind511	Publishing Industries (except Internet)	
Ind512	Motion Picture and Sound Recording Industries	
Ind515	Broadcasting (except Internet)	
Ind516	Internet publishing and broadcasting	
Ind517	Telecommunications	
Ind518	Data Processing, Hosting and Related Services	
Ind519	Other Information Services	

Variable	Description	Source
Ind521	Monetary Authorities-Central Bank	
Ind522	Credit Intermediation and Related Activities	
Ind523	Securities, Commodity Contracts, and Other Financial Investments and Relate	
Ind524	Insurance Carriers and Related Activities	
Ind525	Funds, Trusts, and Other Financial Vehicles	
Ind531	Real Estate	
Ind532	Rental and Leasing Services	
Ind533	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	
Ind541	Professional, Scientific, and Technical Services	
Ind551	Management of Companies and Enterprises	
Ind561	Administrative and Support Services	
Ind562	Waste Management and Remediation Services	
Ind611	Educational Services	
Ind621	Ambulatory Health Care Services	
Ind622	Hospitals	
Ind623	Nursing and Residential Care Facilities	
Ind624	Social Assistance	
Ind711	Performing Arts, Spectator Sports, and Related Industries	
Ind712	Museums, Historical Sites, and Similar Institutions	
Ind713	Amusement, Gambling, and Recreation Industries	
Ind721	Accommodation	
Ind722	Food Services and Drinking Places	
Ind811	Repair and Maintenance	
Ind812	Personal and Laundry Services	
Ind813	Religious, Grantmaking, Civic, Professional, and Similar Organizations	
Ind814	Private households	
Ind999	Federal, State, and Local Government (OES Designation)	

Appendix B

Results from Rural-Urban Dichotomy Model (eq. 1)

Appendix B: Results from Rural-Urban Dichotomy Model (eq. 1)

Number of obs.	**3067**					
Prob > F	**0.0000**					
R-squared	**0.2212**					
Adj R-squared	**0.2197**					
Variables	**Coefficient**	**p-val**		**Variables**	**Elasticities**	**p-val**
RUCC13	-0.0003	0.0070		RUCC13	-0.0329	0.0070
Creative2000	0.0798	0.0000		Creative2000	0.3473	0.0000
SK09	-0.0009	0.0000		SK09	0.0002	0.0000
Natam_Rank	-0.0002	0.2170		Natam_Rank	-0.0206	0.2160
Aglb_Val	0.0000	0.0000		Aglb_Val	0.0100	0.0000
Med_Rent	-0.0003	0.0000		Med_Rent	-0.2423	0.0000
constant	0.0370	0.0000				

Appendix C

Results from State Fixed Effects Model (eq. 2)

Appendix C: Results from State Fixed Effects Model (eq. 2)

Number of obs	**3067**					
Prob > F	**0.0000**					
R-squared	**0.3258**					
Adj R-squared	**0.3139**					
Variables	**Coefficient**	**p-val**		**variables**	**Elasticities**	**p-val**
RUCC13	-0.0002	0.0110		RUCC13	-0.0304	0.0110
Creative2000	0.0722	0.0000		Creative2000	0.3143	0.0000
SK09	-0.0014	0.0000		SK09	0.0004	0.0000
Natam_Rank	-0.0006	0.0490		Natam_Rank	-0.0499	0.0490
Aglb_Val	0.0000	0.0000		Aglb_Val	0.0072	0.0000
Med_Rent	-0.0003	0.0000		Med_Rent	-0.1981	0.0000
constant	0.0373	0.0000				
_Istate_1-48	not reported			_Istate_1-48	not reported	

Appendix D

Results from Industry Mix Model (eq. 3)

Appendix D: Results from Industry Mix Model (eq. 3)

Number of obs	**3067**					
Prob > F	**0.0000**					
R-squared	**0.5712**					
Adj R-squared	**0.5509**					
Variables	**Coefficient**	**p-val**		**variables**	**Elasticities**	**p-val**
RUCC13	-0.0002	0.0110		RUCC13	-0.0269	0.0110
Creative2000	0.0319	0.0000		Creative2000	0.1390	0.0000
SK09	0.0000	0.9870		SK09	0.0000	0.9870
Natam_Rank	-0.0001	0.6140		Natam_Rank	-0.0113	0.6140
Aglb_Val	0.0000	0.0000		Aglb_Val	0.0065	0.0000
Med_Rent	-0.0001	0.0130		Med_Rent	-0.0709	0.0130
constant	-0.3932	0.0740				
Ind519	1.6229	0.0010		Ind519	0.0053	0.0010
Ind515	0.7796	0.0020		Ind515	0.0177	0.0020
Ind324	0.6750	0.0030		Ind324	0.0105	0.0030
Ind532	0.6281	0.0050		Ind532	0.0627	0.0050
Ind525	0.6350	0.0050		Ind525	0.0115	0.0050
Ind518	0.6373	0.0070		Ind518	0.0087	0.0070
Ind443	0.6125	0.0120		Ind443	0.0276	0.0120
Ind482	0.5374	0.0150		Ind482	0.0194	0.0150
Ind441	0.5321	0.0160		Ind441	0.1489	0.0160
Ind511	0.5307	0.0180		Ind511	0.0374	0.0180
Ind221	0.5165	0.0190		Ind221	0.0712	0.0190
Ind551	0.5161	0.0190		Ind551	0.0581	0.0190
Ind325	0.5109	0.0210		Ind325	0.0563	0.0200
Ind712	0.5820	0.0210		Ind712	0.0058	0.0210
Ind562	0.5085	0.0220		Ind562	0.0310	0.0220
Ind442	0.5583	0.0220		Ind442	0.0285	0.0220
Ind517	0.4986	0.0240		Ind517	0.0435	0.0240
Ind448	0.4968	0.0250		Ind448	0.0550	0.0250
Ind322	0.4938	0.0250		Ind322	0.0392	0.0250
Ind212	0.4886	0.0260		Ind212	0.0673	0.0260
Ind522	0.4922	0.0270		Ind522	0.0265	0.0270
Ind213	0.4770	0.0300		Ind213	0.0742	0.0300
Ind541	0.4719	0.0320		Ind541	0.4494	0.0320
Ind999	0.4683	0.0330		Ind999	2.1173	0.0330
Ind561	0.4692	0.0330		Ind561	0.4185	0.0330
Ind331	0.4699	0.0330		Ind331	0.0381	0.0330

Ind512	0.5379	0.0340		Ind512	0.0110	0.0340
Ind42	0.4621	0.0360		Ind42	0.3231	0.0360
Ind516	0.5790	0.0360		Ind516	0.0020	0.0360
Ind622	0.4585	0.0370		Ind622	0.2023	0.0370
Ind115	0.4620	0.0370		Ind115	0.1042	0.0370
Ind486	0.4778	0.0370		Ind486	0.0060	0.0370
Ind484	0.4551	0.0390		Ind484	0.2029	0.0380
Ind327	0.4587	0.0380		Ind327	0.0360	0.0380
Ind492	0.4601	0.0390		Ind492	0.0291	0.0390
Ind334	0.4533	0.0400		Ind334	0.0296	0.0400
Ind333	0.4509	0.0410		Ind333	0.0885	0.0410
Ind339	0.4491	0.0420		Ind339	0.0360	0.0410
Ind336	0.4486	0.0420		Ind336	0.0965	0.0420
Ind451	0.4597	0.0420		Ind451	0.0373	0.0420
Ind335	0.4466	0.0430		Ind335	0.0261	0.0430
Ind211	0.4440	0.0440		Ind211	0.0522	0.0440
Ind523	0.4424	0.0450		Ind523	0.1166	0.0450
Ind311	0.4394	0.0460		Ind311	0.1634	0.0460
Ind493	0.4399	0.0460		Ind493	0.0426	0.0460
Ind314	0.4450	0.0460		Ind314	0.0097	0.0460
Ind713	0.4376	0.0470		Ind713	0.1030	0.0470
Ind312	0.4405	0.0470		Ind312	0.0105	0.0470
Ind332	0.4350	0.0480		Ind332	0.0979	0.0480
Ind337	0.4360	0.0480		Ind337	0.0354	0.0480
Ind481	0.4363	0.0480		Ind481	0.0125	0.0480
Ind323	0.4367	0.0490		Ind323	0.0257	0.0490
Ind721	0.4328	0.0500		Ind721	0.0895	0.0500
Ind321	0.4318	0.0500		Ind321	0.0726	0.0500
Ind313	0.4293	0.0510		Ind313	0.0279	0.0510
Ind521	0.4303	0.0520		Ind521	0.1357	0.0520
Ind711	0.4429	0.0520		Ind711	0.0618	0.0520
Ind326	0.4256	0.0530		Ind326	0.0517	0.0530
Ind524	0.4274	0.0540		Ind524	0.1065	0.0540
Ind114	0.4226	0.0560		Ind114	0.0126	0.0560
Ind621	0.4138	0.0600		Ind621	0.3212	0.0600
Ind111	0.4117	0.0610		Ind111	0.5220	0.0610
Ind722	0.4064	0.0650		Ind722	0.5384	0.0650
Ind485	0.4142	0.0660		Ind485	0.0279	0.0660
Ind316	0.4243	0.0670		Ind316	0.0031	0.0670
Ind112	0.3963	0.0720		Ind112	0.3279	0.0720
Ind611	0.3973	0.0720		Ind611	0.1250	0.0720

Ind230	0.3946	0.0730		Ind230	0.6350	0.0730
Ind531	0.3928	0.0740		Ind531	0.2614	0.0740
Ind623	0.3900	0.0760		Ind623	0.2348	0.0760
Ind444	0.3976	0.0770		Ind444	0.0788	0.0770
Ind487	0.3873	0.0780		Ind487	0.0445	0.0780
Ind624	0.3871	0.0790		Ind624	0.1994	0.0790
Ind814	0.3839	0.0810		Ind814	0.0804	0.0810
Ind483	0.3920	0.0870		Ind483	0.0040	0.0860
Ind447	0.3761	0.0900		Ind447	0.1070	0.0900
Ind454	0.3754	0.0900		Ind454	0.1001	0.0900
Ind113	0.3767	0.0900		Ind113	0.0294	0.0900
Ind811	0.3654	0.0970		Ind811	0.1546	0.0970
Ind446	0.3727	0.1000		Ind446	0.0550	0.0990
Ind813	0.3616	0.1000		Ind813	0.2131	0.1000
Ind452	0.3463	0.1160		Ind452	0.1499	0.1160
Ind812	0.3506	0.1180		Ind812	0.0963	0.1180
Ind445	0.3404	0.1230		Ind445	0.1527	0.1230
Ind453	0.3035	0.1720		Ind453	0.0751	0.1720
Ind533	(dropped)			Ind533	0.0000	.
_Istate_1-48	not reported			_Istate_1-48	not reported	

Appendix E

Results from Education Share Model (eq. 4)

Appendix E: Results from Education Share Model (eq. 4)

Number of obs	**3067**					
Prob > F	**0.0000**					
R-squared	**0.4383**					
Adj R-squared	**0.4263**					
Variables	**Coefficient**	**p-val**		**variables**	**Elasticities**	**p-val**
RUCC13	0.0000	0.9970		RUCC13	0.0000	0.9970
Creative2000	0.0356	0.0000		Creative2000	0.1550	0.0000
SK09	-0.0002	0.4060		SK09	0.0000	0.4060
Natam_Rank	-0.0005	0.0570		Natam_Rank	-0.0457	0.0570
Aglb_Val	0.0000	0.0000		Aglb_Val	0.0069	0.0000
Med_Rent	-0.0003	0.0000		Med_Rent	-0.2203	0.0000
constant	-5.8478	0.0000				
E05	5.9683	0.0000		E05	56.5797	0.0000
E06	5.8545	0.0000		E06	23.1302	0.0000
E02	6.1638	0.0000		E02	21.8733	0.0000
E11	5.8117	0.0000		E11	14.8958	0.0000
E12	5.0480	0.0000		E12	11.9555	0.0000
E01	6.7507	0.0000		E01	11.6631	0.0000
E07	6.8400	0.0000		E07	5.2640	0.0000
E03	6.2100	0.0000		E03	2.0569	0.0000
E09	6.5403	0.0000		E09	1.0500	0.0000
E04	3.9023	0.0000		E04	0.7910	0.0000
E08	0.9930	0.3860		E08	0.2039	0.3850
E10	(dropped)			E10	0.0000	.
_Istate_1-48	not reported			_Istate_1-48	not reported	

Appendix F

Descriptive Statistics

Appendix F: Descriptive Statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
avg_wage	2013 Rural-urban Continuum Codes	3067	0.039	0.012	0.001	0.225
RUCC13		3067	5.018	2.696	1.000	9.000
Natam_Rank	1999 Natural Amenity Rank	3067	3.492	1.047	1.000	7.000
Med_Rent	Median Gross Rent As A Percentage of Household Income (Dollars)	3067	28.387	4.714	10	50.000
Aglb_Val	Estimated market value of land and buildings \ Average per acre (dollars)	3067	3497.358	9822.823	0.000	457143
Creative2000	2000 Creative Class County Codes	3067	0.172	0.059	0.039	0.541
SK09	2009 Social Capital Index	3067	-0.011	1.331	-3.941	17.553
urban	If RUCC13<5, 1=Urban and if RUCC13>4, 0=Rural	3067	0.437	0.496	0.000	1.000
	List of Education Vectors					
E01		3067	0.068	0.007	0.040	0.112
E02		3067	0.140	0.017	0.075	0.334
E03		3067	0.013	0.003	0.004	0.054
E04		3067	0.008	0.001	0.003	0.021
E05		3067	0.374	0.016	0.196	0.449
E06		3067	0.156	0.025	0.035	0.287
E07		3067	0.030	0.005	0.015	0.063
E08		3067	0.008	0.001	0.004	0.019
E09		3067	0.006	0.001	0.003	0.016
E10		3067	0.002	0.001	0.001	0.007
E11		3067	0.101	0.007	0.068	0.172
E12		3067	0.093	0.003	0.071	0.105
	List of Industry Mix Vectors					
Ind111		3067	0.050	0.066	0.000	0.490
Ind112		3067	0.033	0.044	0.000	0.507
Ind113		3067	0.003	0.008	0.000	0.094
Ind114		3067	0.001	0.006	0.000	0.174
Ind115		3067	0.009	0.016	0.000	0.225
Ind211		3067	0.005	0.013	0.000	0.199
Ind212		3067	0.005	0.027	0.000	0.781
Ind213		3067	0.006	0.022	0.000	0.295
Ind221		3067	0.005	0.013	0.000	0.380
Ind230		3067	0.063	0.033	0.000	0.263
Ind311		3067	0.015	0.034	0.000	0.385

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Ind312		3067	0.001	0.005	0.000	0.205
Ind313		3067	0.003	0.013	0.000	0.331
Ind314		3067	0.001	0.004	0.000	0.070
Ind316		3067	0.000	0.002	0.000	0.042
Ind321		3067	0.007	0.015	0.000	0.190
Ind322		3067	0.003	0.011	0.000	0.185
Ind323		3067	0.002	0.006	0.000	0.151
Ind324		3067	0.001	0.003	0.000	0.077
Ind325		3067	0.004	0.012	0.000	0.165
Ind326		3067	0.005	0.011	0.000	0.146
Ind327		3067	0.003	0.007	0.000	0.114
Ind331		3067	0.003	0.012	0.000	0.353
Ind332		3067	0.009	0.016	0.000	0.393
Ind333		3067	0.008	0.016	0.000	0.380
Ind334		3067	0.003	0.009	0.000	0.211
Ind335		3067	0.002	0.009	0.000	0.153
Ind336		3067	0.008	0.022	0.000	0.311
Ind337		3067	0.003	0.013	0.000	0.344
Ind339		3067	0.003	0.008	0.000	0.226
Ind42		3067	0.028	0.018	0.000	0.273
Ind441		3067	0.011	0.006	0.000	0.065
Ind442		3067	0.002	0.002	0.000	0.046
Ind443		3067	0.002	0.002	0.000	0.032
Ind444		3067	0.008	0.004	0.000	0.054
Ind445		3067	0.018	0.007	0.000	0.068
Ind446		3067	0.006	0.003	0.000	0.057
Ind447		3067	0.011	0.008	0.000	0.146
Ind448		3067	0.004	0.005	0.000	0.111
Ind451		3067	0.003	0.003	0.000	0.100
Ind452		3067	0.017	0.012	0.000	0.087
Ind453		3067	0.010	0.006	0.000	0.086
Ind454		3067	0.011	0.008	0.000	0.235
Ind481		3067	0.001	0.008	0.000	0.180
Ind482		3067	0.001	0.005	0.000	0.153
Ind483		3067	0.000	0.002	0.000	0.082
Ind484		3067	0.018	0.016	0.000	0.303
Ind485		3067	0.003	0.004	0.000	0.104
Ind486		3067	0.000	0.002	0.000	0.082
Ind487		3067	0.005	0.018	0.000	0.212
Ind492		3067	0.002	0.005	0.000	0.087

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Ind493		3067	0.004	0.010	0.000	0.293
Ind511		3067	0.003	0.003	0.000	0.045
Ind512		3067	0.001	0.001	0.000	0.026
Ind515		3067	0.001	0.001	0.000	0.012
Ind516		3067	0.000	0.001	0.000	0.041
Ind517		3067	0.003	0.005	0.000	0.093
Ind518		3067	0.001	0.002	0.000	0.037
Ind519		3067	0.000	0.000	0.000	0.010
Ind521		3067	0.012	0.006	0.000	0.172
Ind522		3067	0.002	0.005	0.000	0.156
Ind523		3067	0.010	0.013	0.000	0.173
Ind524		3067	0.010	0.009	0.000	0.106
Ind525		3067	0.001	0.003	0.000	0.091
Ind531		3067	0.026	0.022	0.000	0.304
Ind532		3067	0.004	0.004	0.000	0.060
Ind533		3067	0.000	0.001	0.000	0.027
Ind541		3067	0.038	0.033	0.000	0.893
Ind551		3067	0.004	0.008	0.000	0.136
Ind561		3067	0.035	0.025	0.000	0.506
Ind562		3067	0.002	0.005	0.000	0.125
Ind611		3067	0.012	0.015	0.000	0.240
Ind621		3067	0.031	0.019	0.000	0.280
Ind622		3067	0.017	0.019	0.000	0.199
Ind623		3067	0.024	0.017	0.000	0.263
Ind624		3067	0.020	0.014	0.000	0.138
Ind711		3067	0.006	0.008	0.000	0.158
Ind712		3067	0.000	0.001	0.000	0.024
Ind713		3067	0.009	0.014	0.000	0.429
Ind721		3067	0.008	0.016	0.000	0.410
Ind722		3067	0.052	0.022	0.000	0.206
Ind811		3067	0.017	0.008	0.000	0.144
Ind812		3067	0.011	0.006	0.000	0.071
Ind813		3067	0.023	0.012	0.000	0.192
Ind814		3067	0.008	0.014	0.000	0.321
Ind999		3067	0.178	0.077	0.013	0.898