Industries & Sprawl

Measuring the Effect of Labor Demand Structure on Urban Form Using Landscape Metrics

Travis Maverick Green

MCRP Candidate 2012

Department of City and Regional Planning

University of North Carolina—Chapel Hill

T. William Lester, Advisor and First Reader

Nikhil Kaza, Second Reader

Abstract

Using industry data from the Bureau of Labor Statistics and landscape metrics to describe county growth patterns, this study finds that as counties move towards employment in industries with lower mean incomes, total urbanized area in a county grows more rapidly. The development also becomes more contiguous. As number of workplaces in a county increase, so does total urbanized area. As counties shift towards industries with flatter income distributions and low rates of college education, total urbanized area also grows but it also remains more fractured. The ratio of large-to-small employers did not appear to have pronounced effects on rate of growth or on contiguity of urban form. Based on these industry attributes, counties with particularly large ratios of their workforces in accommodation and food services, construction, retail trade, and administrative, support, waste management are at increased risk of sprawl. On the other hand, counties with increased ratios of health care, information, finance and insurance, educational services, and professional, scientific, technical services are at decreased risk of sprawl.

1.   Introduction

Apple and the City of Cupertino began holding public meetings to discuss the environmental impact study for a new headquarters for the corporation. The tech company has outgrown its existing office park, and the company hopes to build a new facility to house 13,000 employees while also staying close to the institutions that helped catalyze its creation. The proposed facility will require demolition of 2.6 million square feet of outdated office space. Apple prefers to stay inside the city and has elected not to build at a greenfield site at the city's periphery...

In an attempt to attract and retain high tech jobs, central North Carolina's governments established Research Triangle Park in 1959 on a vacant 7,000-acre strip of land between Raleigh-Durham and Chapel Hill. The Park has become an epicenter of high tech growth as the area's tobacco and furniture industries decline. The Park now hosts 38,000 full time employees in 22.5 million square feet of office space. Wake County's core has shifted to the West as new employment opportunities grow up around this park...

Between 2001 and 2006 Las Vegas housing construction boomed as the market for vacation properties and investment properties shot up during the housing bubble. Ten percent of all jobs in the city were dedicated to construction--70% higher employment than the national average. Suburban subdivisions sprawled across the desert. Total urbanized area in Clark County, the principal county for Las Vegas, grew by 55 square miles in just five years...

* * *

These three examples—Cupertino, Raleigh-Durham, and Las Vegas—suggest a link between urban growth and the structure of labor demand. But how are they linked exactly? Are the kinds and quality of jobs related to how a city grows? Can employment in one industry create one form of growth while employment in another industry produces an entirely different form?

The three examples above reveal potential form characteristics associated with various industries, observations that, if validated, may be indispensable to urban planners. In Cupertino, the rapid growth spurred by high tech industry led to significant growth challenges. Apple evolved from a garage business to a multi-national corporation all within a 2-mile-radius. The realities of the industry mean that such a trajectory for a firm is possible. Increasing density agglomerations may be a growth characteristic found in cities with large employment location quotients of high tech research and development industries. In Raleigh-Durham, the shift from

furniture and tobacco manufacturing employment to university-based research employment changed the location of workplaces, commuting patterns, average household incomes, average employees per workplace, and the educational qualifications of a substantial portion of the region's population. The new growth characteristic emerging in Raleigh-Durham may be that of a hub and spoke system emanating from key centers needed to drive employment in these sectors. In Las Vegas, the construction industry was fundamental to the built environment of city. Here, the growth is itself the product of a dominant industry in the community. Cities with larger residential construction industries may see sprawling urban growth.

The purpose of this paper is to empirically test the role of industries, or more precisely, the role of selected "industrial attributes" in shaping two characteristics of urban expansion: rate of growth and fragmentation. Do different industries encourage more rapid urban expansion? And can certain industries, as characterized by various "industrial attributes", lead to fragmented growth as opposed to contiguous urban growth? To foreshadow my conclusion, industries can impact urban form through income, number of establishments, income inequality and relative use of college-educated workers. As counties move towards employment in industries with lower mean incomes, total urbanized area in a county grows more rapidly. The development also becomes more contiguous. As number of workplaces in a county increase, so does total urbanized area. As counties shift towards industries with flatter income distributions and low rates of college education, total urbanized area also grows but it also remains more fractured.

In this introductory section, I will expand on my definitions of urban growth and local economies by discussing some of the literature. I will also wrestle with the important question of causality and offer an argument for why one can see form of growth as a product of economy and not the other way around. Finally, I will conclude this section with an outline of the paper's subsequent chapters.

Urban Growth and City Form

In this empirical study I will use quantitative landscape metrics, a way of mathematically characterizing landscapes, to describe city form and urban growth. Quantitative descriptions of city form are severely limited in what they communicate about human experiences of the built environment. Being clear about what we can and can't measure with quantitative analysis is an important prerequisite to using

the techniques. The key attributes of urban form and growth that can be understood using landscape metrics include size, complexity of shape, fragmentation and overall density. These attributes say very little about what is happening at the core of cities; the technique favors analysis of what is happening at the periphery of urban areas. Therefore, for the purposes of this paper, "city form" will primarily be reduced to understanding if various industries produce growth and if the growth is contiguous or fractured.

Sprawl, a form of growth, is a complicated term that will appear throughout this analysis. Scholars have not settled on a common definition of sprawl. Squires, in a volume he edited in 2002, argues that urban sprawl, or uneven growth, is characterized by "low-density, automobile-dependent, exclusionary new development on the fringe of settled areas often surrounding a deteriorating city." Squires's work focuses on sprawl as a U.S. phenomenon produced by decades of federal housing and transportation policies favoring suburbs over older, denser cities. In adopting such a definition, Squires connects growth at the fringe with decline in the center; it connects new development with racially charged disinvestment at the core. By comparison, Galster, in a 2001 article, promotes separating sprawl and its causes from its consequences and advocates for a more

Figure 1: Leapfrogging sprawl in Wake County, North Carolina
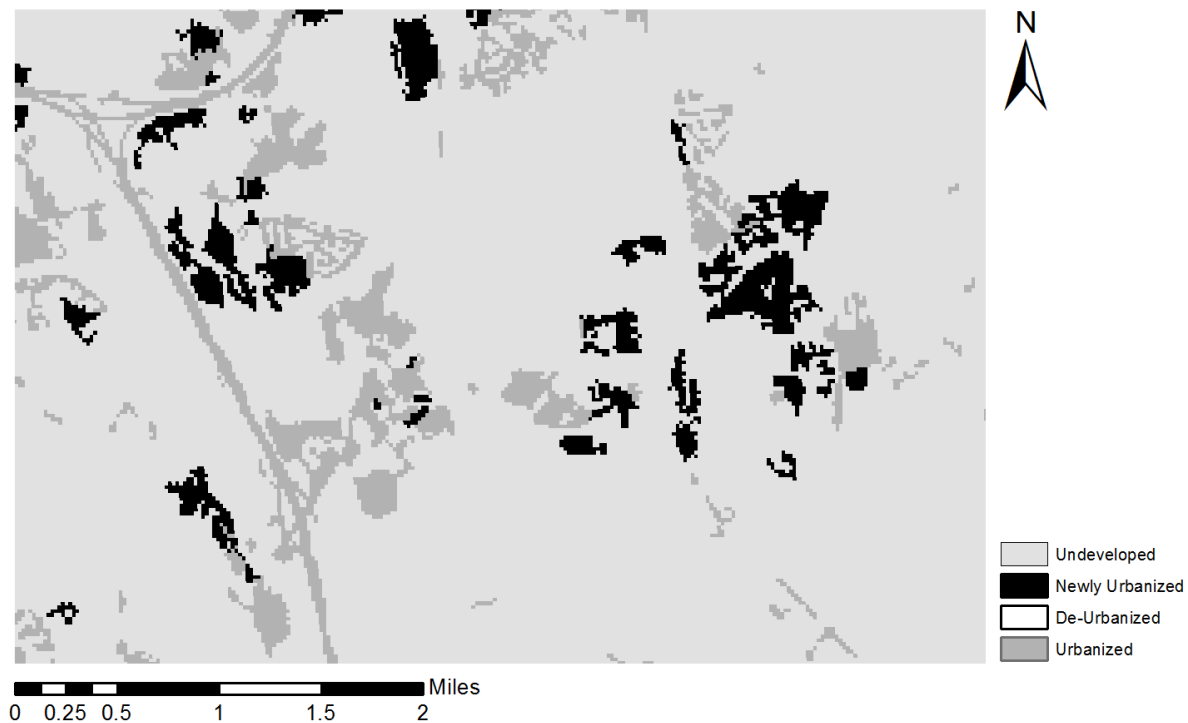
technical definition. He favors using measurements in eight distinct dimensions of land use patterns—density, continuity, concentration, clustering, centrality, nuclearity mixed uses, and proximity—to quantitatively asses whether a community is sprawling. For Galster, sprawl is something that is simply measured.

There are also many conflicting understandings of the causes of sprawl. While Squires favors the theory that sprawling development has been spurred on by federal policies, Beauregard (2001), counters the claim with empirical evidence and says that scholars should look elsewhere to explain the phenomenon. In a land cover study, Burcheld (2006) found that "ground water availability, temperate climate, rugged terrain, decentralized employment, early public transport infrastructure, uncertainty about metropolitan growth, and unincorporated land in the urban fringe" increase sprawl. Brueckner (2000) argues that sprawl has been primarily caused by population growth, rising incomes and falling transportation costs. Wu (2001) explains sprawl as a product of household consumption preferences favoring various environmental and social amenities more available in suburbs. Finally, Mieszkowski (1993) explains sprawl as a product of the natural evolution of urban form downplaying the role of any one historical event.

For the purposes of this paper, the term "sprawl" will be used without the socio-economic associations proposed by Squires. Instead, like Galster, it will be used to describe a collection of attributes of development at a city's edge. Although some may disagree about whether simple residential expansion at a city's periphery constitutes sprawl, most scholars have agreed that "leapfrog" development constitutes sprawl (see Figure 1). Leapfrogging occurs when developers skip over vacant land adjacent to existing urban areas and develop a disconnected urban area. Sprawl, in this analysis, will be used to describe counties experiencing leapfrog development. Spatially, this will be described as (1) rapid growth and (2) low rates of defragmentation. While urban development consumes more land in almost every county from year to year, sprawling counties do so more quickly. Similarly, urban development tends to become more contiguous as urban areas merge into each other. Sprawling development, by this paper's definition, does so less rapidly. It remains or becomes more fragmented.

Local Economies and Spatial Economics

To this point, spatial economics is a field dominated by economic theory of firm locational decisions, a related but causally inverted concept from what I will be

testing. Economists have predominately written about firm decisions. Alfred Weber described firm locational decisions as a product of two inputs: transaction costs and labor costs. Weber's Least Cost Theory predicts that firms locate to minimize both labor and transaction costs. Walter Christaller and, later, August Losch described Central Place Theory recognizing that firms are spatially distributed to capture markets. Because of travel costs for consumers, firms appreciate a locational monopoly when they are a certain distance from their nearest competitor. More recently, Fujita (1999) summarized decades of economic thought on spatial economies, a work for which Krugman won the Nobel Prize in economics in 2008. Collectively, these works of urban economics explain how firms locate to reduce transaction costs, capture spillover effects, and share labor pools.

These theories, in turn, have shaped economic development practice. Michael Porter writes about clusters and encourages cities to recruit businesses by capitalizing on economic advantages of agglomeration economies. Richard Florida discusses the locational preferences of the "creative class," suggesting that city form is itself a commodity sought by this new techno-elite, the most important resource in the new economy.

Other scholars have taken an alternative approach arguing that space is not just a variable to which firms respond, but a product of firms themselves. Storper and Walker (1983) pioneered this rethinking of the orthodox approach to spatial economics using labor. They argue that it is an oversimplification to say that firms simply minimize costs through locational decisions. Businesses generate unique labor markets so as to capitalize on a "spatial division of labor." (33) This paper will build on the assumptions of Storper and Walker arguing that firms co-produce their environments.

Finally, I must address a second spatial question. To what area should one ascribe the impacts of an industry? It may be obvious how jobs in a given sector in Houston impact Houston and not Portland. But how do you draw distinctions between San Diego and Los Angeles? Minneapolis and St. Paul? Do the Federal agencies in Washington have spatial consequences for Baltimore?

This dilemma does not have a satisfying answer. At what level should we call local economies distinct? One possibility would be using metropolitan areas which have a common housing and labor pool. The people who work in a metropolitan area live there for the most part. This would allow for better measurement of secondary effects of industries like how income in a given industry impacts the housing decisions of its employees. The disadvantage of using metropolitan areas, however, is

the significant exclusion of major sectors like agriculture and mining, two sectors that may have large impacts on urban form. Second, this study is principally concerned with the impacts of *industries* not housing and therefore prioritizes the location of workplaces over residences. The data sources for this study report employment based on the county of the workplace and not the employee's home. Recognizing that this geography is problematic, I have selected to use county data.

Causality

Do industries *produce* city form, does city form *attract* firms, or is there no relation between industries and city form? As mentioned above, most spatial economics literature presumes that firms view their locational decision as a maximization problem suggesting that industries respond to form instead of producing it. In this paper I will argue the opposite. In the vignettes discussed above—Cupertino, Raleigh-Durham, and Las Vegas—we saw how industries may be seen as producing space. In Cupertino, the success of a major corporation in concert with regional spillover economies have led to plans to restructure the city's center. In Raleigh-Durham, the creation of a suburban industrial park grew the high tech sector that changed the fundamental shape and commute patterns of the region. Las Vegas's leisure economy led to a vacation home boom that sent the city sprawling in all directions.

Of course, the argument that cities can be a product of their economies is not new. Adam Smith argued that the formation of towns and cities was primarily an economic activity driven by the need for specialized labor. "The scattered families that live at eight or ten miles distance from the nearest of them must learn to perform themselves a great number of little pieces of work, for which, in more populous countries, they would call in the assistance of those workmen." As the nature of work changed, so did cities.

Looking at a list of modern industries (see Table 1), one can hypothesize how each may produce distinct forms just as through history different economic needs have created different kinds of cities. Port towns built shipping centers along coastlines. Capitals typically maintain large amounts of public space around government office buildings. Historically, manufacturing has built high-density housing around their facilities that require thousands of workers. For modern industries that use land as inputs, agriculture, mining, construction, real estate, it is easier to make these hypotheses. It is possible that other industries impact form as well through incomes, establishment size, educational needs, and other industrial attributes.

Table 1: 2007 NAICS (North American Industry Classification System) Sector Codes

NAICS CODE	SECTOR NAME
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)

The Work Ahead

In this section I have presented a research question: does industrial change influence city form? I have clarified that city form will be characterized using quantitative landscape metrics measuring rate of growth and degree of fragmentation. Rapid growth and lack of substantial defragmentation will be described as "sprawl". I will use counties as the unit of analysis. The remaining sections of this paper will be divided as follows: In the methodology section I will define and discuss the six landscape metrics used for this analysis: total patch area, number of patches, patch size, standard deviation of patch size, mean shape index, and fractal dimensional index. I will also explain the independent variables including categorization of employment, which will describe "industry" in the model. In the data section I will detail the sample and how it was selected. In the results section I will share the results of various regressions. In the discussion and conclusion sections I will expand on the results of the regressions and look closely at the three counties introduced at the beginning of this paper. Also in the conclusion, I will expand on my hypothesis that industries produce distinct forms and discuss which industries may be producing sprawl based on those attributes.

2.   Methodology

To measure industry's impact on urban form I will measure form and industry changes from 2001 to 2006. I will then conduct a basic employment analysis on counties that are sprawling to determine how they may differ from the entire sample. Then, using linear regression I will measure the statistical relationships between industrial characteristics and urban form. Finally, I will use logistic regression to measure the likelihood of various industrial characteristics producing sprawl.

In my analysis, I will measure forms of growth using landscape metrics. I will use levels of employment in various sectors to describe industrial character. The models will then be tested using various statistical techniques as well as by comparing the difference between industry inferred economic characteristics and observed characteristics.

Landscape Metrics

Landscape metrics are created by first classifying predetermined units of area into various categories of interest. This land cover map usually starts as a satellite image that is then broken into a grid of pixels (see **Figure 2**). Adjacent pixels with the same value are called a "patch". In an urbanized area map, land that has been urbanized is given one value, "1" or "black", and land that is un-urbanized is given another, "0" or "white". A contiguous urbanized area in then called an "urbanized patch". These maps then can be run through a geospatial statistical software program that generates descriptive landscape metric values for a given area.

Evelin Uuemaa (2009) conducted a review of landscape metrics and their various uses in geo-spatial research. The study found that from 1994 to 2008, scholars published 331 articles using landscape metric techniques. A significant number of these articles were dedicated to studies of biodiversity and habitat. Landscape metrics have also been used for estimating water quality and measuring landscape pattern responses to natural disasters. More recently, landscape metrics have been used to study urbanization patterns mostly as it relates to impacts on ecosystems. For instance, papers have tried to measure how infrastructure has resulted in habitat fragmentation.

The literature indicates several concerns with the use of landscape metrics. First, the growing popularity of the technique paired with the rise of numerous software

packages to aid researchers in producing metrics has created a fragmented assortment of metrics many of which communicate the same thing. To test the value of individual metrics, Hargis (1998) created artificial landscapes with various controls representing different fragmentation processes. Hargis found that different landscape measures were highly correlated and linearly associated. She also documented some blind spots in the technique, i.e. none of the measures were able to distinguish between dispersed and aggregated patches. Such statistical relationships, the risk for autocorrelation and simultaneously the risk for blind spots in data analysis, must be taken into account when selecting appropriate metrics to describe space.

The second key issue identified by Uuemaa is the modifiable areal unit problem which covers "how the grain size, zoning and areal extent of investigation influence results and how to determine their optimal values for each particular case." (7) For the most part, this second problem will go unaddressed by this research project. Two things can be said however in defense of the selected scale of the study. First, I have

Figure 2: Characterizing urbanization using landscape metrics

not selected boundaries or a resolution to achieve preferred results; the boundaries and resolution were selected because of data availability. I will be using the 2001 and 2006 data files from the National Land Cover Database that divides the United States into 30-meter classified pixels. I have broken this map by current county boundaries, a unit for which employment data is readily available. Second, it is hard to imagine circumstances under which the chosen metrics for this study would vary greatly at smaller resolutions of study. The modifiable unit problem has great consequences for habitat where contiguity of territory has very real impacts on the survivability of creatures. Understanding how the species of study interacts with the environment should guide the decision regarding units. In this case, the relevant level of study is more subjective and should reflect the ways humans interact with the built environment. The provided data set with 30-meter resolution passes the gut test of appropriate resolution for study: smaller units would be prone to more errors and would provide unneeded detail; larger units would conceal processes of interest to regional decision makers.

Finally, another criticism of the technique comes from Li (2004) who points out that there were high expectations for spatial analysis, but that it has done little to "improve our understanding and prediction of ecological processes". Li concludes that statistical description of space is just that. It must be paired with meaningful analysis of how such characterizations of form relate to underlying ecological or urban forms. In other words, studies with metrics should be paired with additional analyses of form. The metrics used in this paper are two-dimensional and binary: urban or not-urban. A 30-meter urbanized patch in Manhattan is likely very different from a 30-meter urbanized patch Harrisburg. For this paper I will have limited opportunities to go beyond the statistics, but I will briefly discuss how urbanization trends are impacting the three counties detailed in the introduction.

Landscape Metrics Describing Urban Growth

To aid in the selection of appropriate metrics, I relied on the work of Seto (2005) and Kaza (2012). Seto measured the growth pattern of Chinese cities using three buffer zones over a ten-year period. Kaza's paper, using Seto's paper as a guide, looked at the linkages between socioeconomic variables and urban fragmentation in the United States. Both Seto and Kaza used the same set of six variables to characterize binary urban development: total urbanized area, number of urbanized patches, mean urbanized patch area, standard deviation of urbanized patch area, urbanized patch

edge density, and mean fractal dimensional index of urbanized patches. I use similar metrics.

These six metrics can be divided into three categories: measures of absolute size, measures of relative size and complexity of urban form. Measures of absolute size, number of patches and total area, increase with increased urbanization. Seto writes, "As urban growth occurs, total urban area continually increases due to the highly non-reversible nature of urbanization. The number of urban patches metric is a measure of discrete urban areas in the landscape and is expected to increase during period of rapid urban nuclei development, but may decrease if urban areas expand and merge into continuous urban fabric." (877)

The measures of relative size are mean urban patch area and standard deviation of patch area. "Decreasing values of mean urban patch size implies that new urban centers are growing faster than existing urban areas. That is, urban growth occurs more as a process of new and multiple urban nuclei formation than of envelopment or annexation." (877) Standard deviation of patch area measures the breadth or diversity of sizes of urbanized patches. Decreasing values indicate that patches are becoming more uniform in size.

The third set of metrics measure complexity of urban form. Mean urbanized patch fractal dimension "describes the degree to which the shape of an urban area is irregular or complex. The more irregular the shape of the urban area, the higher the fractal dimension". (878) The second measure of complexity is mean shape index, a slight revision from Seto's preferred edge density. Mean shape index measures a patch's likeness to a shape with a very simple and low edge to area ratio. "In this index low values reflect the compactness of individual developments." (Kaza, 7)

How can these six metrics describe specific forms of growth? As mentioned in the introduction, this study is limited to describing two principal characteristics of growth: rate of urbanization and rate of defragmentation. I have labeled a community that is both growing rapidly and experiencing little defragmentation as "sprawling". I will measure rate of growth as the percent total urban area increase and fragmentation as the percent change in number of urbanized patches.

Measuring Industrial Character

I will measure industrial character using raw private sector employment data available from the Quarterly Census of Employment and Wages and distributed by

the Bureau of Labor Statistics. This data is collected at the establishment level. Each establishment is categorized into one of nineteen sectors, or NAICS codes, based on its primary product. The Bureau of Labor Statistics collects this data from the states and then distributes the data publically at the county level.

One option would be to characterize industrial character in regression models by employment data alone. I suspected and later confirmed that this would be problematic because of high levels of autocorrelation between levels of employment in some industries. Including 19 separate employment values in a regression can be difficult to interpret. Instead of using raw employment data, I chose to re categorize employment in various industries into five "industry attributes" that I hypothesized, based on previous sprawl literature and based on intuition, may be impacting urban form. The five variables are:

Mean Income Percent Change – Using industry employment totals for each county, I generated a weighted average of "mean industrial incomes" in 2001 and 2006. I inflated the 2001 income and measured the percent change. Profiling the dominant theoretical frameworks for suburbanization, Mieszkowski and Mills (1993) argue that historically incomes have been a cause of sprawl. "As new housing is built at the periphery, high income groups who can afford larger and more modern housing settle there." (136) As employment shifts to industries with higher mean incomes, we might expect total urbanized area to increase.

Income Inequality Percent Change – Using a matrix of occupations available from the Occupational Employment Statistics Survey from the Bureau of Labor Statistics, I divided employment in each industry into a subset of occupations. From this list of occupations I derived a rough measure of income inequality based on income distributions within occupations themselves. To generate this metric, I calculated a weighted average industry 90th percentile income and a weighted average industry 10th percentile income. I divided the two figures for 2006 and 2001, inflating the 2001 figure to predict industry spread of wages. I then calculated the percent change in income distribution for each county from 2001 to 2006. Mieszkowski and Mills predict that increased income heterogeneity will also lead to sprawl. "Those who move to the suburbs often seek to form homogenous communities, for several reasons. There is the preference for residing among individuals of like income, education, race, and ethnicity. By residing in income-stratified communities, the affluent avoid local redistributive taxes. Homogenous community formation is also motivated by varying demands for local public goods, caused by income and taste differences." (137)

Portion of Population with College Degrees Percent Change – Like income inequality, industries have different arrangements of occupations requiring different skill levels. I summed the occupations in each industry to predict a percentage of jobs requiring college education. I then calculated the percent change from 2001 to 2006. Predicting the impacts of labor skill on urban growth is a little more difficult. Like with income inequality, educational attainment may produce more pressure at the urban periphery. "Homogenous groupings enhance the quality of education, as there is evidence that peer-group effects are important in the production of educational achievement." (Mieszkowski and Mills, 137) I find this argument not all that compelling. Instead, I prefer looking at education as an indicator of high value labor inputs. Higher value inputs, using a Weber analysis, creates a tighter network of firms around those inputs. I predict that increasing percent of college-educated employees will discourage sprawl and urban growth. Others, like Florida (2002), have argued that "the creative class", a more educated group of workers, now prefers denser center city living.

Number of Establishments Percent Change – Industries may be predisposed to certain size establishments. Using industry averages, I calculated the number of establishments in a county using employment data. I then measured the percent change from 2006 to 2001. Like changes in numbers of housing units, I predict that an increase in the percent of establishments will result in sprawling urban form.

Ratio of Super Establishments Percent Change – Some industries have higher ratios of super establishments, or establishments with 100+ employees. Using employment data, I calculated the ratio of super establishments to non-super establishments in each county. I then measured the percent change from 2006 to 2001. Super establishments will likely favor more compact urban development as they tether employees to more compact employment centers.

For a summary of hypotheses on how these variables impact sprawl, see Table 2.

Table 2: Hypothesized Industry Impacts on Sprawl

INDUSTRY VARIABLE	ANTICIPATED SIGN
Mean Income Percent Change	+
Income Inequality Percent Change	+
Portion of Population with College Degrees Percent Change	-
Number of Establishments Percent Change	+
Ratio of Super Establishments Percent Change	-

In addition to the five industry metrics, three control variables will complete my models for sprawl. Index of Relative Rurality (detailed for 2000) measures a county's degree of rurality on a scale of 0 to 1 with 1 representing a perfectly rural county and 0 representing a perfectly urban county (Isserman 2005). I included average July temperature to control for the Sun Belt effect, or the population and employment shift of the country from the Northeast towards the Southwest. Finally, I included percent change in number of housing units as a stand-in for change in population. I selected change in housing units because it more accurately represents the spatial consequences of an additional household, but it also allows for a more direct analysis of the impacts of housing versus the impacts of employment.

Comparing Observed Variables and Industry Implied Variables

The industry implied variables described above capture how employment shifts between industries impact expected labor force and workplace attributes. A 1% increase in mean income, for example, describes a shift in employment from industries that pay lower national mean incomes to industries that pay higher national mean incomes. *It does not mean that local incomes actually increase.*

Why might industry implied income shifts be different from real income shifts? One explanation is that industry shifts may have a gradual effect on local incomes. As industries expand in a region they may, at the beginning, pay the prevailing local wage and gradually shift towards industry averages. Another possible explanation is that industries vary from region to region, and pay higher wages in one place than in another place. A third explanation is that many local factors—beyond industrial character—may be impacting these variables. For instance, all industry incomes may be higher in one area because of higher cost of living. In this third case, the challenge becomes isolating the portion of change that can be attributed to shifts in the structure of labor demand from other local factors. Designing the statistical models with industry implied variables rather than locally observed variables is one way to isolate for regional heterogeneity in factors that might be impacting incomes, income inequality, college attainment, number of establishments and ratio of super establishments beyond the industries themselves. What happens, then, if the coefficients in the statistical models change sign when using industry implied versus locally observed variables? Further study may be required to validate the implied model's validity.

Statistical Analysis

Before creating statistical models, I will first conduct a basic employment analysis of sprawling counties. How do they differ from the rest of the sample? Following this analysis I will conduct two tests. The first test requires six regressions and measures the impacts of industry attributes on the six individual landscape metrics. In a second model I will conduct a logistic regression to determine if individual industrial attributes collectively impact landscape metrics to generate fragmented urban growth or "sprawl".

3. Data

Two-hundred fifty-six counties in the lower-48 possessed all the data required to be a part of the study. These counties constituted 44.6% of the total population of the lower-48, and included a variety of urban area types from older mature population cores to actively developing urban areas. These 256 counties possessed full employment data for the 19 sectors of interest in both 2001 and 2006, the single most limiting factor on the sample size. The 256 counties also had complete land area calculations, housing unit data from 2000 and 2005, and complete temperature data. Several parishes in Louisiana were excluded because of the catastrophic impacts of Katrina.

A majority of these counties, 223, are part of metropolitan areas. By Census classification, the most (133) are classified as "mixed rural". Sixty six are classified as urban, 56 are mixed urban and only 1 is completely rural. At least one county is included from 45 of the 48 contiguous states. No counties are included from Delaware, Vermont, the District of Columbia or Virginia. In terms of 2006 population, 74 counties had populations over 500,000. Eighty-two have populations between 200,000 and 500,000, and 100 have populations below 200,000.

Another interesting measure of the data set is the relative "developmental age" of the counties. One can draw some conclusions about the form of a county based on when it went through rapid development. Kaza (2011) introduced such a measure in his paper regarding landscape metrics. Kaza's technique divided counties into four categories based on when they achieved 75% of their 2000 population. If it was achieved prior to 1950, counties were labeled "very mature". Counties reaching the 75% threshold between 1950 and 1970 were labeled "mature", between 1970 and 1990 were labeled "developed" and counties that have not reached 75% of current population by 1990 were labeled "developing". Kaza found this variable to be well correlated with the total area, number of patches, average size and standard deviation of area of urbanized patches. In my dataset, 29 counties are "developing", 112 are "developed", 42 are "mature" and 48 are "very mature". The developmental age of 25 counties cannot be determined because of limitations of Census data and changing county boundaries.

The dependent variables show some interesting trends (for a summary see Table 3). For instance, change in total urbanized area is exclusively greater zero. In other words, in no county in the study group did the total urbanized area decrease. The dataset's mean change in total urbanized area was 4.4%. The largest increase in total

urbanized area was an increase of 23.3% (Jackson, MS). San Juan County, NM, (Farmington) by comparison had the smallest increase of 0.2%

The change in number of patches has a mean of -5.6% indicating that counties, for the most part, have a decreasing number of urbanized patches. The majority of urbanized areas in the sample are becoming more contiguous. This however varied from 53% decrease in Salt Lake County, UT to a 57.6% increase in Taylor County, TX. Mean urbanized patch area typically increased (the mean was 12%), although this too varied from large decreases, 35.2%, in Taylor County, to a 134% increase in Salt Lake County, UT where the large change in number of patches merged into one large patch. The mean county had a slight increase, 8.3%, in standard deviation of urbanized patch area.

Counties typically had a slight increase, 0.5%, in urbanized mean shape index, meaning form was becoming more irregular. The largest increase was Salt Lake County, UT, 10.4%, where the new emerging urban shapes were much more irregular. Taylor County, TX, had the largest drop of 13.9% where new urbanized areas were relatively square-like.

Table 3: Variable Mean, Standard Deviation, Minimum, and Maximum

	Source	Mean	Std. Dev.	Min	Max
Metrics - Percent Change					
Total Area	KAZA	4.37%	0.038	0.20%	23.31%
Number of Patches	KAZA	-5.58%	0.097	-53.46%	57.56%
Mean Patch Area	KAZA	12.00%	0.157	-35.20%	134.31%
Standard Deviation of Patch Area	KAZA	8.30%	0.089	-18.96%	60.17%
Mean Shape Index	KAZA	0.46%	0.021	-13.92%	10.36%
Mean Fractal Dimensional Index	KAZA	-0.54%	0.009	-4.77%	0.73%
Industry Implied Independent Variables - Percent Change					
Mean Income	QCEW	0.08%	0.011	-3.13%	5.18%
Income Inequality	QCEW, OES	-0.15%	0.005	-1.53%	1.44%
Portion of Population with College Degrees	QCEW, OES	0.59%	0.010	-2.78%	4.32%
Number of Establishments	QCEW, SBA	8.96%	0.112	-12.04%	49.32%
Ratio of Super Establishments	QCEW, SBA	-0.74%	0.034	-13.51%	9.10%
Observed Independent Variables - Percent Change					
Mean Income	QCEW	0.05%	0.052	-14.65%	19.37%
Income Inequality	CENSUS 2000, ACS 2005	9.80%	0.118	-21.57%	66.56%
Portion of Population with College Degrees	CENSUS 2000, ACS 2005	5.44%	0.047	-7.76%	19.39%
Number of Establishments	QCEW	8.72%	0.094	-14.47%	36.01%
Ratio of Super Establishments	CBP	0.22%	0.116	-27.59%	84.38%
Control Variables					
July Temperature	ICPSR	75.7	5.966	55.9	91.2
Index of Relative Rurality (2000)	KAZA	0.239	0.104	0.021	0.531
Percent Change in Number of Housing Units (2000 to 2005)	ICPSR	8.50%	0.064	0.30%	29.90%

Finally, fractal dimensional index of urbanized patches tended to decrease meaning counties were becoming more regular and grid-like in urbanized form. The largest decrease, 4.8%, occurred in Webb County, TX. The largest increase, just 0.7%, occurred in Fort Bend County, TX in Houston where the form became more complex.

As you can see from some of these extreme values, some counties appear to be outliers in the dataset. I did not drop these outliers, however, because in most of these counties, the urban form is changing because of economic activities. It might be expected that urban form would be most impacted by industries that use land as inputs for production (like agriculture and resource extraction). So while it may be a worthwhile exercise to repeat this study and exclude these outliers, I have intentionally left them in the study because it reflects a more realistic diversity in economic activity.

4. Results

The results of my data analysis are divided into four sections. First, I conducted some initial data analysis to determine how the economies of sprawling counties were changing relative to the entire sample. Generally, employment grew more rapidly in sprawling counties. Declining industries declined less rapidly and growing sectors grew more rapidly. Health care, retail trade and accommodation and food services were the biggest job gainers, and manufacturing jobs were better preserved in sprawling counties. Second, I conducted regressions to evaluate how individual landscape metrics were impacted by the independent industry variables. From 2001-2006, *decreasing* industry incomes increased urban growth and encouraged urban defragmentation. *Decreasing* industry income inequality both increased urbanized area and increased urban fragmentation. Third, I conducted a logistic regression to determine how independent industry variables collectively impacted landscape metrics to specifically produce sprawling form. *Decreasing* industry income inequality and *decreasing* industry college education attainment rates had a higher likelihood of producing sprawl. Finally, I reviewed how the landscape metrics and employment changes in the three counties mentioned in the introduction.

Which counties are sprawling, and how is their employment changing?

I identified five counties with populations over 500,000 (2005 population) that were rapidly urbanizing but not defragmenting, our selected indicator for sprawl (see Table 4 for a complete list). Arapahoe (part of Denver) and El Paso (Colorado Springs) counties in Colorado, Hidalgo (McAllen) and Travis (Austin) counties in Texas, and Wake County (Raleigh) in North Carolina went through substantial fragmented urbanization from 2001-2006. In addition to these larger counties, I identified 20 counties with populations between 200,000 and 500,000 and 29 counties with populations between 50,000 and 200,000 that showed similar growth patterns.

The fragmented-growth counties contained 9.3% of the sample's employment in 2001 and 9.7% (5,161,000 jobs) of the employment in 2006 and were therefore growing faster (8.2% increase instead of 3.8% in the whole sample). In 2001, these counties were underrepresented in agriculture (a location quotient of 0.67), utilities (0.68), and management of companies (0.75) and overrepresented in retail trade (1.16), health care and social assistance (1.17) and mining (2.24, something that will be discussed later). All other sectors had location quotients of between 0.85 and 1.15, or were roughly on-par with employment expectations.

Table 4: Sprawling counties

COUNTY	CORE BASED STATISTICAL AREA	2005 POPLTN	COUNTY	CORE BASED STATISTICAL AREA	2005 POPLTN
Aiken County	Augusta-Richmond County, GA-SC	150,181	Madison County	Jackson, MS	84,286
Arapahoe County	Denver-Aurora, CO	529,090	Marathon County	Wausau, WI	128,941
Boone County	Cincinnati-Middletown, OH-KY-IN	106,272	Marion County	Ocala, FL	303,442
Caddo Parish	Shreveport-Bossier City, LA	251,309	Merrimack County	Concord, NH	146,881
Canadian County	Oklahoma City, OK	98,701	Midland County	Midland, TX	121,371
Cape Girardeau County	Cape Girardeau-Jackson, MO-IL	71,161	Missoula County	Missoula, MT	100,086
Cass County	Kansas City, MO-KS	94,232	Monroe County	Bloomington, IN	121,407
Charleston County	Charleston-North Charleston, SC	330,368	Montgomery County	Montgomery, AL	221,619
Cleveland County	Oklahoma City, OK	224,898	Montgomery County	Houston-Sugar Land-Baytown, TX	378,033
Comal County	San Antonio, TX	96,018	Northampton County	Allentown-Bethlehem-Easton, PA-NJ	287,767
Craven County	New Bern, NC	90,795	Orange County	Poughkeepsie-Newburgh-Middletown, NY	372,893
Dauphin County	Harrisburg-Carlisle, PA	253,995	Placer County	Sacramento--Arden-Arcade--Roseville, CA	317,028
Deschutes County	Bend, OR	141,382	Potter County	Amarillo, TX	119,852
Dutchess County	Poughkeepsie-Newburgh-Middletown, NY	294,849	Pulaski County	Little Rock-North Little Rock, AR	366,463
El Dorado County	Sacramento--Arden-Arcade--Roseville, CA	176,841	Rankin County	Jackson, MS	131,841
El Paso County	Colorado Springs, CO	565,582	Rockingham County	Boston-Cambridge-Quincy, MA-NH	295,076
Escambia County	Pensacola-Ferry Pass-Brent, FL	296,772	Sangamon County	Springfield, IL	192,789
Flathead County	Kalispell, MT	83,172	Sebastian County	Fort Smith, AR-OK	118,750
Forrest County	Hattiesburg, MS	75,095	Shelby County	Birmingham-Hoover, AL	171,465
Guilford County	Greensboro-High Point, NC	443,519	Spartanburg County	Spartanburg, SC	266,809
Hamilton County	Chattanooga, TN-GA	310,935	Travis County	Austin-Round Rock, TX	888,185
Hidalgo County	McAllen-Edinburg-Mission, TX	678,275	Tuscaloosa County	Tuscaloosa, AL	168,908
Jasper County	Joplin, MO	110,624	Wake County	Raleigh-Cary, NC	748,815
Kootenai County	Coeur d'Alene, ID	127,668	Warren County	Cincinnati-Middletown, OH-KY-IN	196,622
Lackawanna County	Scranton--Wilkes-Barre, PA	209,525	Webb County	Laredo, TX	224,695
Lake County	Orlando-Kissimmee, FL	277,035	Williamson County	Nashville-Davidson--Murfreesboro, TN	153,595
Linn County	Cedar Rapids, IA	198,903	Yavapai County	Prescott, AZ	198,701

Figure 3: Method for choosing sprawling counties – above average rates of growth and below average rates of defragmentation.
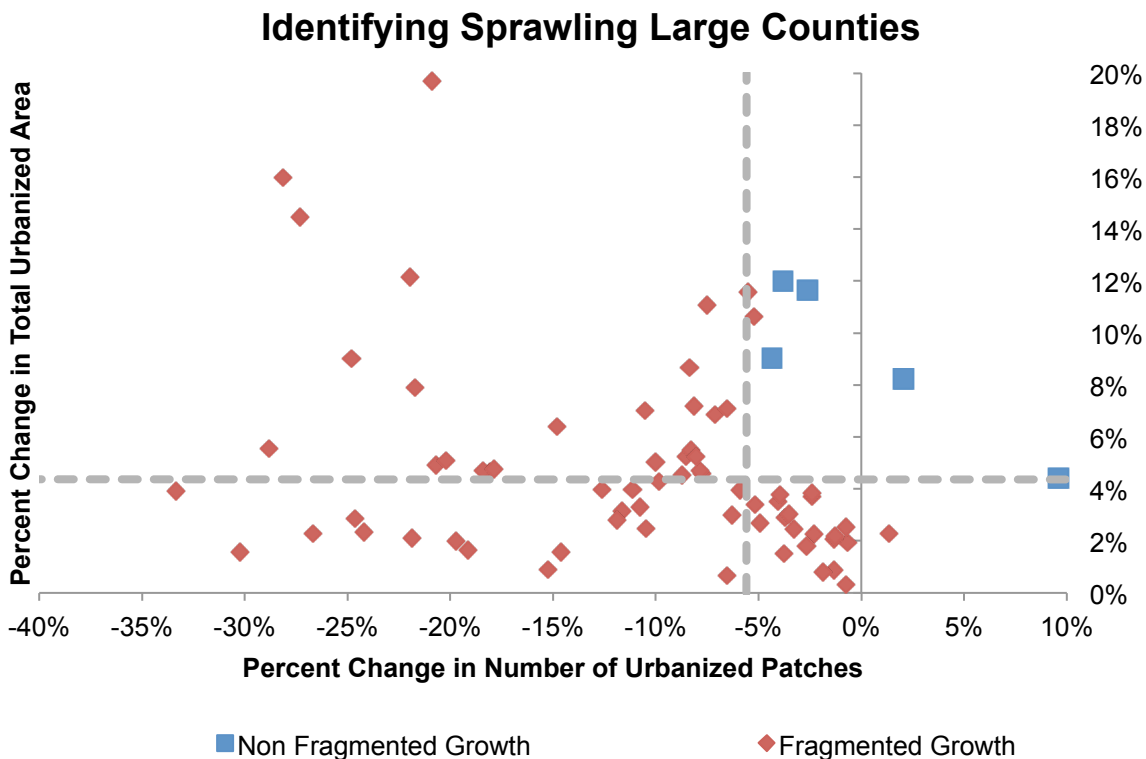

Identifying Sprawling Large Counties

Percent Change in Total Urbanized Area (y-axis)

Percent Change in Number of Urbanized Patches (x-axis)

■ Non Fragmented Growth   ◆ Fragmented Growth

Table 5: Employment changes from 2001-2006

NAICS	Sprawling County Employment				All County Employment			
	2001	2006	Change	%	2001	2006	Change	%
11	26,674	24,355	-2,319	-8.69%	428,305	419,846	-8,459	-1.97%
21	27,374	33,532	6,158	22.50%	210,737	242,506	31,769	15.08%
22	22,374	22,277	-97	-0.43%	239,054	213,245	-25,809	-10.80%
23	342,772	385,607	42,835	12.50%	3,258,224	3,722,509	464,285	14.25%
31-33	647,230	577,940	-69,290	-10.71%	7,117,527	6,021,467	-1,096,060	-15.40%
42	244,541	253,597	9,056	3.70%	2,804,372	2,840,993	36,621	1.31%
44-45	723,983	769,724	45,741	6.32%	6,979,100	7,153,604	174,504	2.50%
48-49	195,686	207,973	12,287	6.28%	1,992,678	2,007,986	15,308	0.77%
51	175,808	149,162	-26,646	-15.16%	1,820,419	1,577,677	-242,742	-13.33%
52	247,718	291,225	43,507	17.56%	2,757,777	2,967,689	209,912	7.61%
53	89,307	96,293	6,986	7.82%	1,037,251	1,115,089	77,838	7.50%
54	279,987	318,739	38,752	13.84%	3,506,924	3,731,998	225,074	6.42%
55	54,935	69,025	14,090	25.65%	855,041	905,874	50,833	5.95%
56	338,753	388,775	50,022	14.77%	3,937,423	4,259,107	321,684	8.17%
61	69,619	81,654	12,035	17.29%	880,408	1,051,598	171,190	19.44%
62	558,977	670,647	111,670	19.98%	5,952,662	6,776,844	824,182	13.85%
71	80,131	87,666	7,535	9.40%	884,444	952,372	67,928	7.68%
72	463,949	545,112	81,163	17.49%	4,780,323	5,351,362	571,039	11.95%
81	182,689	188,101	5,412	2.96%	2,037,620	2,143,744	106,124	5.21%

The fragmented-growth counties saw a rather unique employment change pattern from 2001 to 2006 (see Table 5). Although agriculture was already underrepresented in the sprawling counties, it declined much faster (-8.7% versus -2.0%). Management of companies grew much faster at 25.7% versus 6.0% in the entire sample. Most employment losses in utilities were prevented in sprawling counties. They only declined at -0.4% rather than at -10.8% in all counties. Accommodation and food services (17.5%), health care (20.0%), administrative and support (14.8%), mining (22.5%), and finance and insurance (17.6%), grew between 5 and 10% faster than all other counties.

The most consequential employment changes can be revealed by predicting how a county's employment should change if it followed the trends in the entire sample. I found that sprawling counties do not follow the employment changes in the rest of the sample. Sprawling counties were better at preserving manufacturing jobs, saving 30,380 more jobs than expected. Health care added 34,276 more jobs than expected. Retail trade added 27,639 more jobs and accommodation and food services added 25,741 more jobs than expected. These employment shifts have interesting socio-economic consequences for these counties that might be impacting their urbanization. The additional growth in retail trade and accommodation services, for instance, may be dropping countywide mean incomes (see Figure 4).
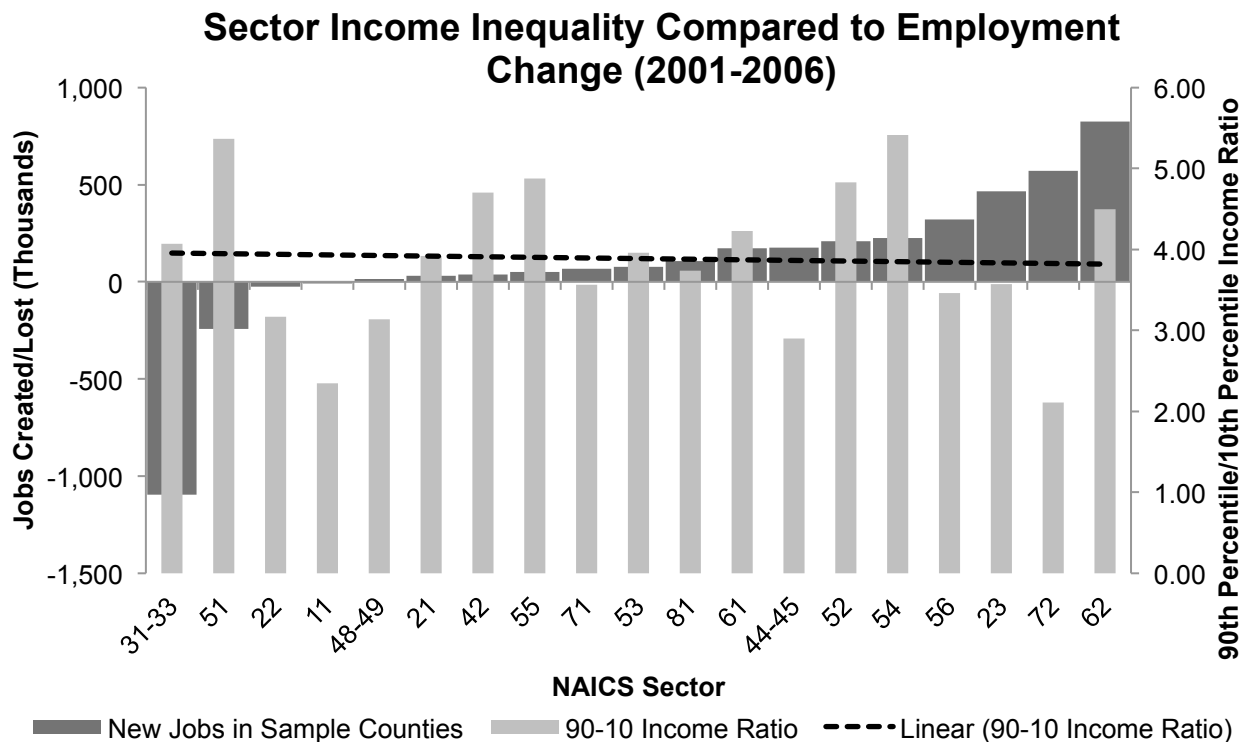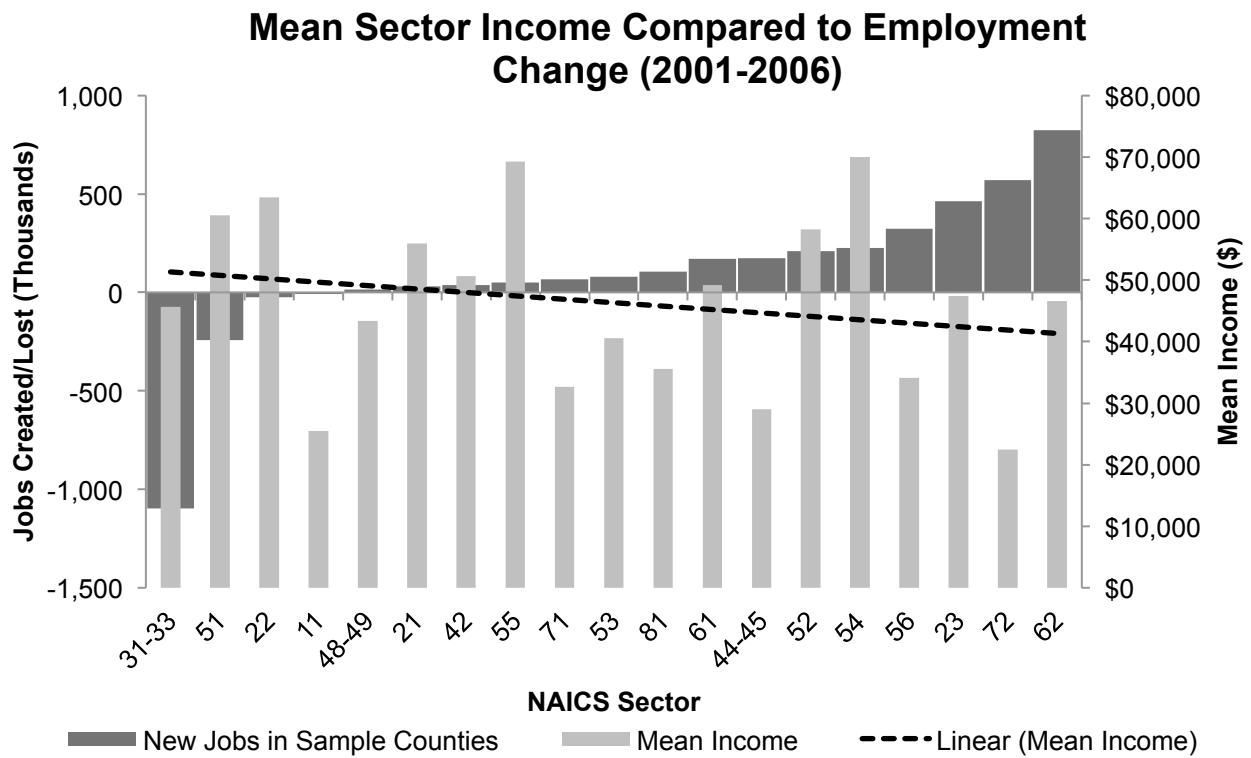
Table 6: Industry variables in whole sample vs. sprawling counties

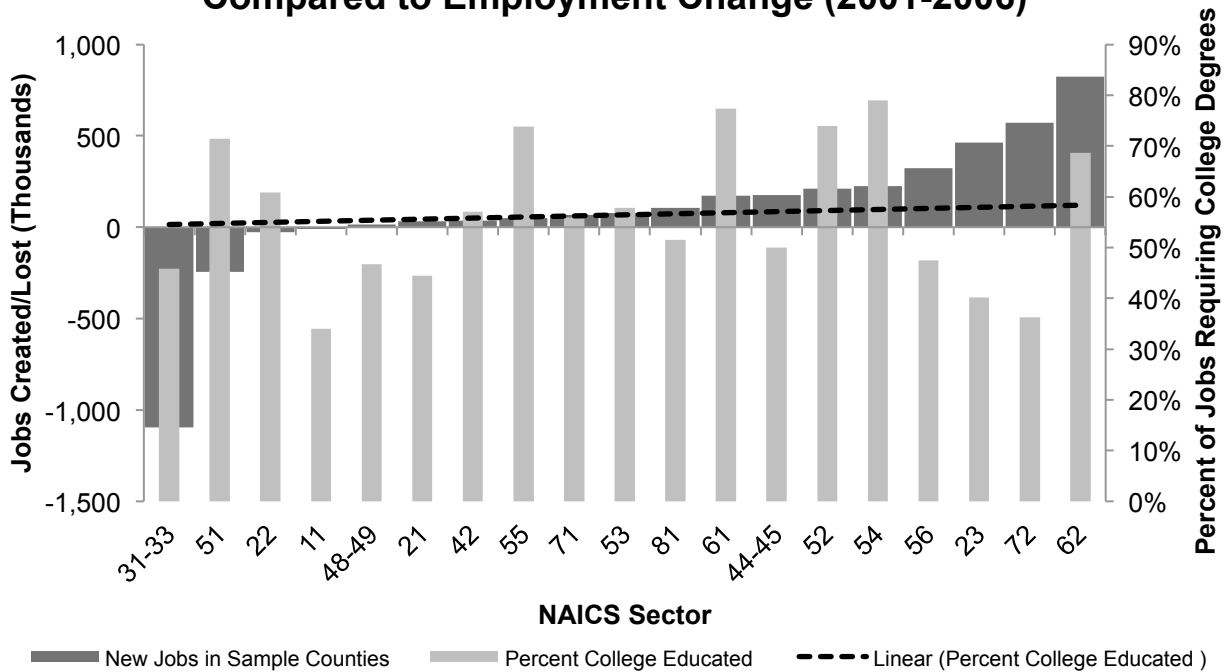|	Sprawling Counties				All Counties			
	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Industry Variables								
Mean Income	0.14%	0.013	-3.13%	2.78%	0.06%	0.011	-2.99%	5.18%
Income Inequality	-0.29%	0.005	-1.53%	1.20%	-0.11%	0.005	-1.30%	1.44%
Portion of Population with College Degrees	0.57%	0.012	-2.78%	4.32%	0.60%	0.009	-2.22%	3.88%
Number of Establishments	13.52%	0.125	-6.21%	40.01%	7.74%	0.105	-12.04%	49.32%
Ratio of Super Establishments	-0.80%	0.032	-10.34%	6.36%	-0.72%	0.034	-13.51%	9.10%
Observed Variables								
Mean Income	-0.40%	0.054	-13.58%	16.65%	0.17%	0.052	-14.65%	19.37%
Income Inequality	10.86%	0.160	-12.82%	66.56%	9.52%	0.105	-21.57%	37.81%
Portion of Population with College Degrees	5.26%	0.050	-4.18%	18.77%	5.48%	0.046	-7.76%	19.39%
Number of Establishments	13.41%	0.100	-0.64%	36.01%	7.46%	0.089	-14.47%	34.21%
Ratio of Super Establishments	0.84%	0.123	-18.15%	44.56%	0.05%	0.114	-27.59%	84.38%

These two sectors are the lowest and third lowest, respectively, in terms of sector mean income. Health care, by comparison, is roughly average. We might also be observing a shift to sectors with more "super establishments", or workplaces with more than 100+ employees. Retail trade ranks 5th, accommodation and food services ranks 6th and finance and insurance ranks 4th for number of super establishments, and all four sectors added disproportionately more jobs in sprawling counties.

More in-depth data analysis, however, complicates these preliminary observations of how industrial composition may be impacting these counties and thereby impacting their spatial growth patterns. Just looking at averages suggests that based on the dominant sectors of employment, fragmented growth counties will have large increases in incomes, have greater decreases in income inequality, have smaller increases in number of college educated employees, have significantly greater increases in number of establishments and have slightly greater decreases in the ratio of large to small establishments (see Table 6). These differences in averages should be seen skeptically because in all instances, the standard deviations of the observations significantly overlap. The industrial differences of these counties may be intriguing, but more statistical rigor must be employed to really demonstrate a connection between changes in employment and sprawl.

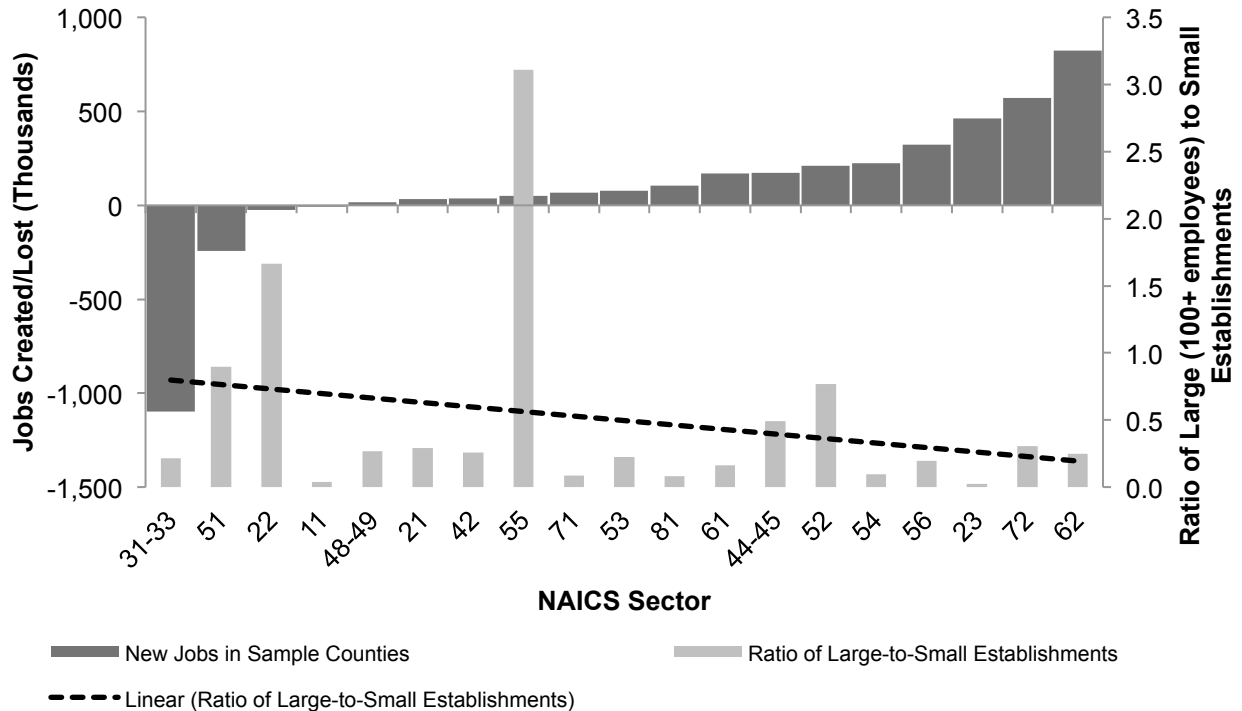**Figure 4: Comparing sector employment change and industry attributes**

Mean Sector Income Compared to Employment Change (2001-2006)

Legend: New Jobs in Sample Counties — Mean Income — Linear (Mean Income)

Sector Income Inequality Compared to Employment Change (2001-2006)

Legend: New Jobs in Sample Counties — 90-10 Income Ratio — Linear (90-10 Income Ratio)

Sector Educational Attainment
Compared to Employment Change (2001-2006)

Legend: New Jobs in Sample Counties — Percent College Educated — Linear (Percent College Educated)

X-axis (NAICS Sector): 31-33, 51, 22, 11, 48-49, 21, 42, 55, 71, 53, 81, 61, 44-45, 52, 54, 56, 23, 72, 62

Left Y-axis: Jobs Created/Lost (Thousands) — 1,000 to -1,500
Right Y-axis: Percent of Jobs Requiring College Degrees — 0% to 90%

Ratio of Large-to-Small Establishments
Compared to Employment Change (2001-2006)

Legend: New Jobs in Sample Counties — Ratio of Large-to-Small Establishments — Linear (Ratio of Large-to-Small Establishments)

X-axis (NAICS Sector): 31-33, 51, 22, 11, 48-49, 21, 42, 55, 71, 53, 81, 61, 44-45, 52, 54, 56, 23, 72, 62

Left Y-axis: Jobs Created/Lost (Thousands) — 1,000 to -1,500
Right Y-axis: Ratio of Large (100+ employees) to Small Establishments — 0.0 to 3.5

Average Employees Per Establishment
Compared to Employment Change (2001-2006)

Are landscape metrics tied to changes in employment?

Sprawling counties appear to have some economic differences from other counties, but can we measure a statistically significant relationship? To test this question, I conducted six regressions measuring the relationship between the five industry constructed variables and the six landscape metrics, both discussed in the methodology section (see Table 7). These regressions were meant to test how employment shifts, as measured by characteristics of various sectors, change urbanization.

I then conducted a second set of regressions that replaced the "industry implied" attributes such as "industry implied number of establishments" and with observed economic characteristics. As mentioned in my methodology section, one concern with using simply observed values for these attributes is that many factors—beyond industrial character—may be impacting income, income inequality, etc. etc. However, the industry implied model depicted above can and should be compared to the actually observed variables. If the coefficients in the models change sign or are no longer statistically significant, further study may be require to validate the models.

Of the industry related independent variables, change in mean income was the most significant in the most models. The coefficient was statistically significant in models of change of total urbanized area, change of number of urbanized patches, change in mean urbanized patch area and change in standard deviation of urbanized patch area. This change in mean income coefficient implies that as local employment composition shifts towards industries paying higher mean incomes, urbanized area grows less rapidly, number of urbanized patches increase, mean urbanized patch area falls and standard deviation of urbanized patch area falls. The models measuring observed change in mean incomes confirmed these results. The coefficients had similar signs and in all cases, except for change in total urbanized area, were also statistically significant.

Table 7: Regressions of industry implied and observed variables

	% Chg Total Urbanized Area		% Chg Number of Urbanized Patches		% Chg Mean Urbanized Patch Area	
	Industry	Observed	Industry	Observed	Industry	Observed
July Average Temperature	1.27	2.66***	1.26	1.24	-0.55	-0.21
2000 Index of Relative Rurality	-4.83***	-3.83***	6.98***	7.11***	-6.27***	-6.45***
Change in Housing Units	6.89***	4.51***	-1.79*	-0.27	3.13***	1.04
Mean Income	-2.45**	-1.48	1.99**	1.87*	-2.05**	-1.98**
Income Inequality	-1.87*	0.37	-1.9*	-0.55	0.89	0.5
Portion of Population with College Degrees	0.34	-0.54	-1.22	-1.32	1.12	0.41
Number of Establishments	2.14**	1.35	0.79	-1.52	0.07	1.47
Ratio of Super Establishments	0.78	-0.55	-0.94	1.83*	1.27	-1.43
Constant	-0.25	-1.32	-2.64***	-2.78***	2.14**	2.05**
Model R-squared	0.5129	0.4671	0.2064	0.2213	0.2486	0.2553

	% Chg Std Dev of Urbanized Patch Area		% Chg Mean Urbanized Patch Shape Index		% Chg Mean Urbanized Patch Fractal Dimensional Index	
	Industry	Observed	Industry	Observed	Industry	Observed
July Average Temperature	-0.46	0.06	-0.34	-0.44	-0.36	-0.81
2000 Index of Relative Rurality	-6.53***	-6.45***	-0.74	-0.53	4.28***	4.31***
Change in Housing Units	4.35***	2.24**	1.51	0.01	0.7	-0.07
Mean Income	-2.19**	-1.62	-0.44	-0.45	0.51	1.64
Income Inequality	-0.2	0.81	2.96***	0.92	1.32	0.5
Portion of Population with College Degrees	0.71	0.39	0.16	0.97	-1.14	-0.32
Number of Establishments	0.42	1.42	-0.48	1.76*	-0.62	0.8
Ratio of Super Establishments	1.02	-1.05	2.52**	-1.96*	-0.12	-0.76
Constant	1.9*	1.61	0.55	0.46	-0.98	-0.87
Model R-squared	0.3142	0.3125	0.061	0.0616	0.1052	0.0911

*** 99% significance
** 95% significance
* 90% significance

Three landscape metrics had statistically significant coefficients for industry implied changes in income distribution. As county structure of labor demand shifted towards industries with larger income gaps, change in urbanized area decreased relatively, number of urbanized patches decreased, and mean urbanized patch shape index increased. This last model, however, had very low explanatory value. In all three cases, observed shifts in income gaps did not confirm the industry-implied results. This may be an artifact of the discrepancy in how the income inequality variables were constructed, an issue I will discuss later.

Number of college-educated employees, industry implied and observed, appears to have little or no impact on the six landscape metrics I measured. This may be because the major components of educational attainment are already captured by income.

An increase in industry implied number of establishments positively impacted change in total area. The coefficient was not statistically significant in any other model, and the observed change in number of establishments was not statistically significant in any model, except for change of mean urbanized patch shape index, a low explanatory value model.

Finally, the industry implied presence of super employers appears to only have an impact on mean shape index, again a low explanatory value model. The sign on the coefficient flipped in the observed model. The observed presence of super employers also appears to have had a positive impact on the number of urbanized patches.

The control variables, 2000 Index of Relative Rurality and percent change in housing units, both appear to be highly significant in most models. July average temperature, an attempt to control for the "sunbelt factor", appears to have only been significant in the observed change of total urbanized area.

Can we predict sprawl using employment data?

The initial analysis of sprawling counties and the subsequent linear regressions demonstrated that some attributes of industrial shifts can have consequences for individual landscape metrics. But can a collection of industry attributes impact urbanized growth and rate of fragmentation simultaneously thus producing sprawl? To test this final question, I labeled the fragmented-growth counties (counties that had high rates of growth and low rates of defragmentation) with a dummy variable. Using the sprawl dummy variable as the new outcome, I conducted a logistic regression (see Table 8).

The model had limited explanatory power (a pseudo R2 of 0.171, the sample had 256 observations). More rural counties (having a higher IRR in 2000) had an increased likelihood of sprawling. A one percent increase in housing units produced a 2.66% increase in likelihood of sprawling, falling within the 99% confidence interval. A 1% increase in income equality increased the likelihood of sprawling by 3.90%. This coefficient was significant at the 99% level. Also, a 1% increase in the industry implied percentage of college educated employees decreased the likelihood of sprawling by 2.04%. This final variable was significant at the 95% level. None of the other industry variables had statistically significant coefficients. It should also be noted that the logistic model, when replicated with observed variables, did not produce statistically significant coefficients for any of the industry variables.

Table 8: Logistic regression predicting likelihood of sprawl, robust SE

	Industry	Observed
July Average Temperature	0.99	2.60***
2000 Index of Relative Rurality	3.94***	4.14***
Change in Housing Units	2.66***	0.3
Mean Income	0.86	-1.34
Income Inequality	-3.90***	0.75
Portion of Population with College Degrees	-2.04**	-1.02
Number of Establishments	0.03	1.32
Ratio of Super Establishments	-1.39	-0.1
Constant	-2.68***	-3.97***
Model Pseudo R-squared	0.171	0.124

*** 99% significance
** 95% significance
* 90% significance

Cupertino, Raleigh, and Las Vegas

In the introduction I profiled three counties that I believed helped make the case the economies were instrumental in shaping urban form. Before we look at the wider data set, it may be worthwhile to look at what happened in these three counties.

Clark County's (part of Las Vegas, Nevada) urbanized area grew the most, 16.0%. Wake County (Raleigh, North Carolina) grew by 11.7%, and Santa Clara's (San Jose, California) urbanized area only grew by 0.3%. With respect to mean urbanized shape index, all counties, taking into account area variations, saw increases in mean edge to interior ratios. In other words, shapes became less square like. Santa Clara became more fractal-like, while both Clark and Wake counties became less fractal-like. Mean urbanized patch area and mean standard deviation of urbanized patch area increased

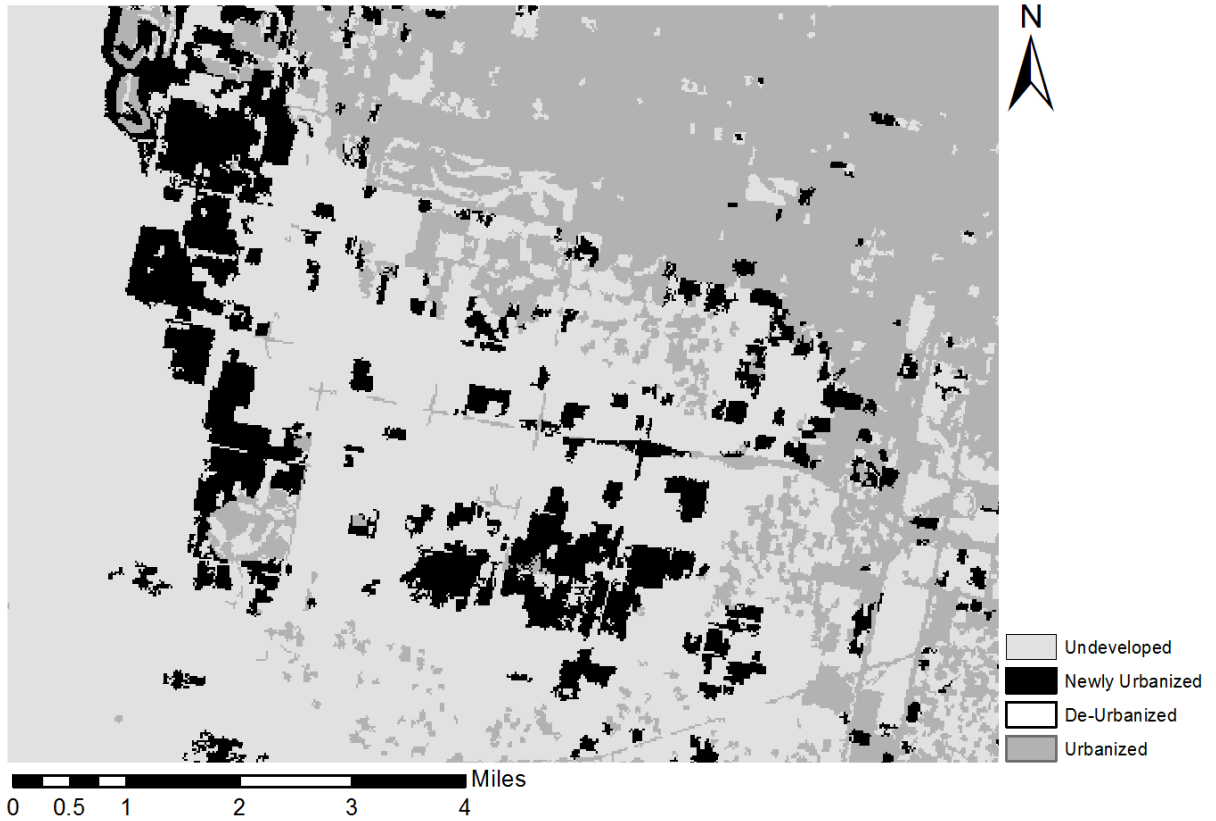**Figure 5: Clark County, Nevada urbanization pattern**

Legend:
- Undeveloped
- Newly Urbanized
- De-Urbanized
- Urbanized

Miles
0 0.5 1 2 3 4

Figure 6: Santa Clara County, California urbanization pattern

Legend:
- Undeveloped
- Newly Urbanized
- De-Urbanized
- Urbanized

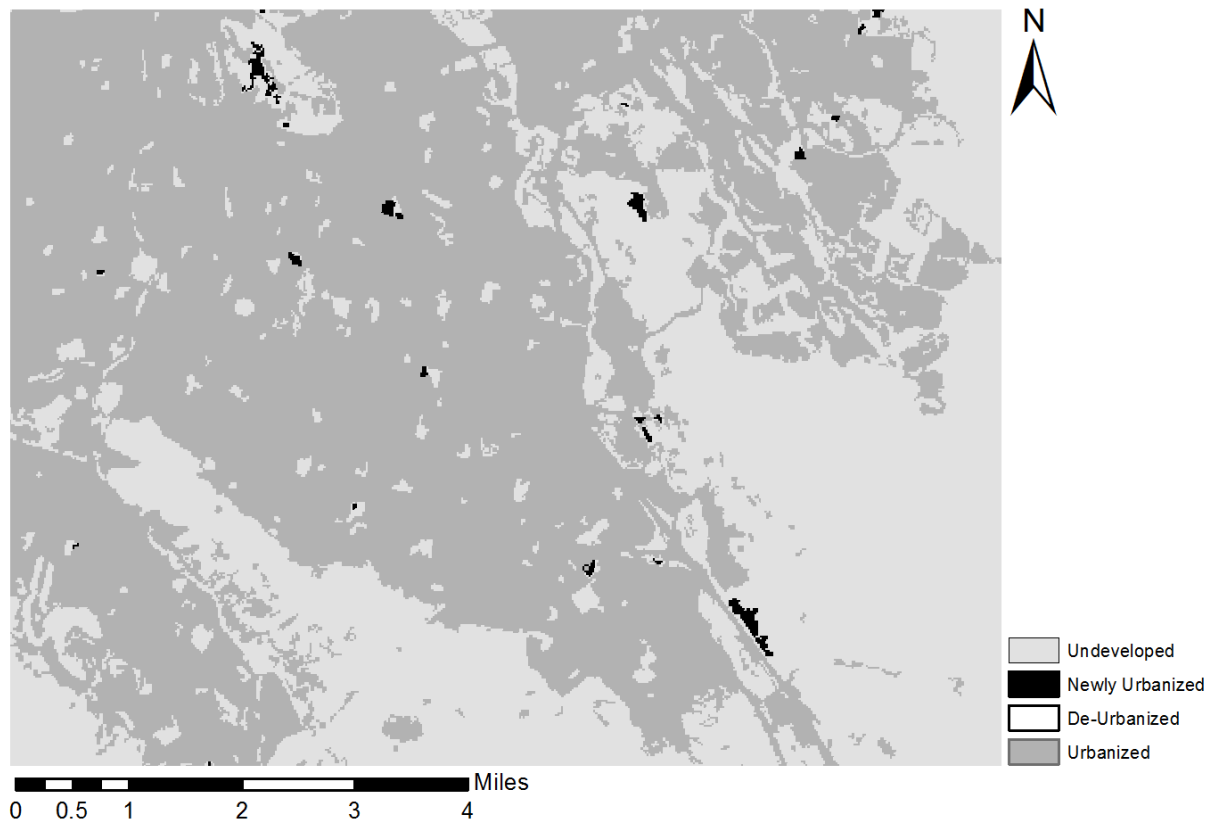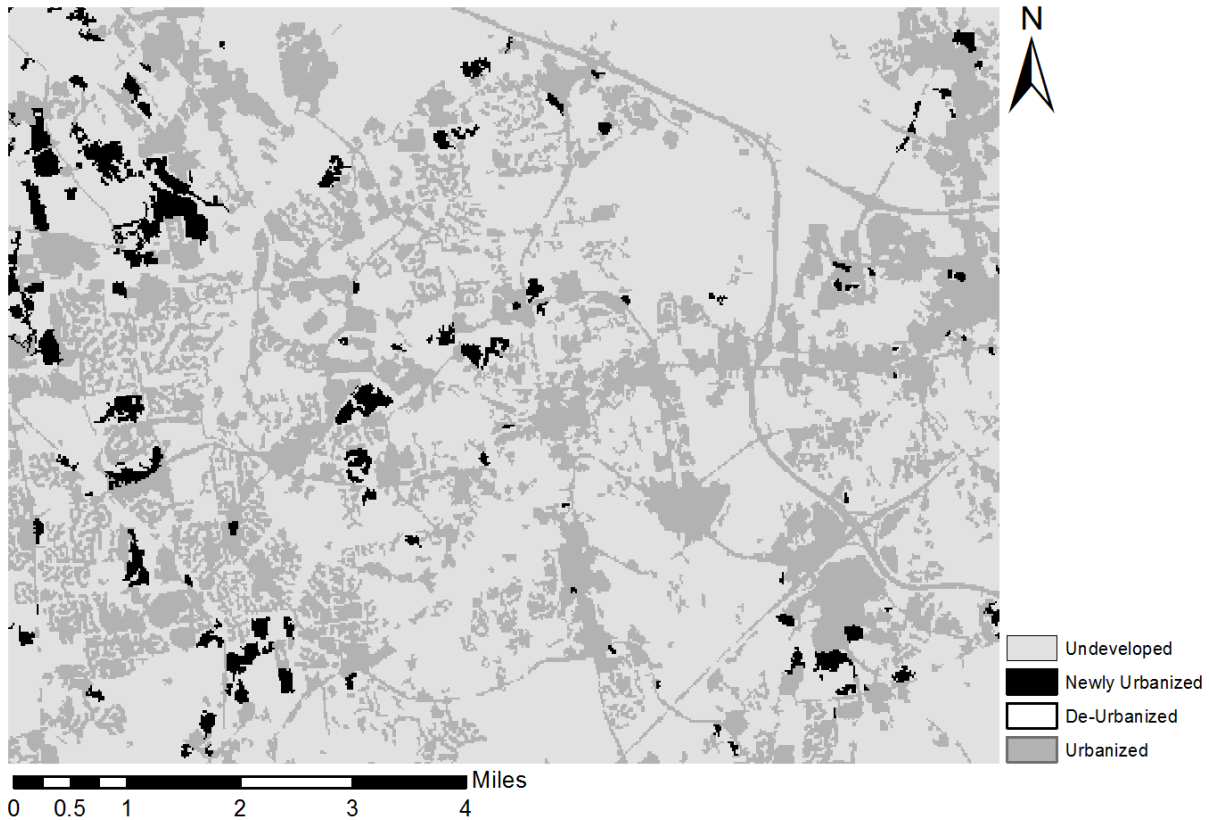Miles
0 0.5 1 2 3 4

Figure 7: Wake County, Nevada urbanization pattern

in all three counties. This suggests that larger, older patches are the focus of most growth rather than new nuclei. The main difference between the three counties, however, is what is happening to the number of urbanized patches in each. Santa Clara County is seeing a small reduction in number of patches, roughly 0.7%, which is consistent with its relatively low level of new development. Clark County has seen a precipitous drop in number of urbanized patches: number of patches in the county decreased by 28.1% in five years. This is consistent with growth connecting these patches. Wake County, by comparison, has seen only a 2.6% reduction in urbanized patches despite its remarkable growth. This difference is what led me label the county as a sprawling county.

How might economic forces be shaping the forms of these three counties? The rapid growth of Wake and Clark counties means that new establishments are being built and peripheral development is being influenced by the needs of these new workplaces. Santa Clara's relatively modest development means that the county's form has been already largely established. Even if the industrial character of the county were to change, the form may not change unless it is subjected to additional growth beyond the existing urbanized footprint. The three largest job growth industries in Wake County, the one sprawling county of the three, were health care,

accommodation and food services, and professional, scientific, technical services. The three largest job growth industries in Clark County were construction, accommodation and food services, and retail-trade. Santa Clara grew the most in health care, other services, and finance and insurance.

For Further Investigation

There are several caveats that should be placed on the above results. First, insomuch as industrial attributes are predicted to change urban form through housing markets, it should be again stressed that some employees live in one county and work in another. Further analysis should be conducted to determine how cross-border commutes may be changing these growth patterns. Second, these two dimensional landscape metrics have severe limitations. Future studies should attempt to combine these two dimensional descriptions of space with richer and more complex descriptions of form. Third, income inequality has many ways of being computed, and each technique has limitations. Using the method selected for this paper (the 90-10 income ratio), accommodation and food services was designated the most income "equitable" industry. There are some who make take issue with this calculation. Also because of data limitations, the income inequality measures for the industry implied variable and the observed variable were calculated somewhat differently. Again, further analysis should be done to see if the results can be replicated with different measures of inequality like a GINI coefficient. Finally, although industry averages were used to construct the industry attribute data variables, more time should be dedicated to analyzing how "inherent" these attributes are to given industries.

5.   Discussion and Conclusion

The models used in my analysis of urban form point to several important results. First, increasing industry income is associated with less rapid, fragmented, urbanized growth, not sprawl. This result goes against the findings in literature up to this point. An alternative way of stating this finding is that decreasing mean incomes creates rapidly expanding, contiguous urban growth. There are several theories that may explain why this may be. First, traditional models of cities presume the most affordable housing can be found at a city's periphery where land is most affordable and where residents are willing to substitute longer commute times for housing costs. This theory is called the bid-rent model. Decreasing mean income in a county is likely to generate outward growth pressures. In this scenario, industrial character influences urban form through housing choices.

A second hypothesis is that outward development is cyclical, and that as cities experience massive development pressures, they first go through increased perimeter expansion. The observation that incomes are causing this change may be caused by the fact that the fastest growing industries, at this time, have lower incomes than other sectors. In other words, the observation is a coincidence and is not causal. Had the 2000s been defined by marked job growth in industries with higher mean incomes, we may have found that higher mean incomes are associated with expanding development.

Another key finding of my analysis is that as number of workplaces increases, total urban area increases. Increases in both the number of housing units and workplaces create urban growth in counties.

The most curious finding, however, is the relationship between income inequality and sprawl. In the employment derived model, decreasing income inequality produces increased urban growth and increasing urbanized patches. This suggests that decreased income inequality may produce sprawl. This result is confirmed with the logistic regression.

Why might a county's economic shift towards occupations with decreased income inequality put a community at risk of fractured-growth? Sprawl might be related to the size of a community's middle class. Communities tending towards industries with flatter income distributions may be producing leapfrogging development because of the sizable housing demands of this population. This answer, however, is not entirely satisfactory. It does not explain why greater income equality produces fragmented

development. Greater understanding of the causal mechanisms would be required to more confidently state this conclusion. It is also important to point out that while the industry employment shifts were significant (i.e. shifts towards industries with greater income equality), counties with greater observed income equality were themselves not statistically more likely to be undergoing leapfrogging development. Again, this may be caused by how these two variables were constructed, but again further analysis is required to explain that difference.

Finally, growth in industries with lower rates of educational attainment is associated with sprawl. I had originally hypothesized that those industries with specialized labor inputs would form tighter networks based on the Weber theory of firm location. In other words, the increasing value of the labor inputs would draw urban forms closer. These results are consistent with that observation.

Based on the income inequality and college attainment findings, it is possible to highlight a set of industries that might be producing sprawl. The most likely sprawl causing industries will have 1) large employment changes, 2) low levels of income inequality, and 3) low levels of college attainment. Based on these three standards, four industries are likely to be associated with sprawl (see Table 9):

- accommodation and food services
- construction
- retail trade
- administrative, support, waste management

Table 9: Industries ranked by industry attributes

NAICS	Change in No. of Jobs	Rank	Mean Income	Rank	90-10 Income Ratio	Rank	Percent College Educated	Rank	Super Establish Ratio	Rank	Average Jobs Per Estab	Rank
11	-8,459	1 - Lowest	$25,530	2	2.35	2	34%	1 - Lowest	0.04	2	7	2
21	31,769	4	$55,910	14	3.92	10	45%	4	0.29	13	27	15
22	-25,809	3	$63,400	17	3.17	5	61%	13	1.67	18	37	17
23	464,285	16	$47,470	11	3.57	8	40%	3	0.02	1 - Lowest	9	4
31-33	-1,096,060	19	$45,680	9	4.07	12	46%	5	0.21	8	40	18
42	36,621	5	$50,660	13	4.70	15	57%	11	0.26	11	14	7
44-45	174,504	11	$28,980	3	2.90	3	50%	8	0.49	15	14	8
48-49	15,308	2	$43,400	8	3.13	4	47%	6	0.27	12	20	11
51	-242,742	14	$60,500	16	5.37	18	71%	15	0.90	17	24	13
52	209,912	12	$58,250	15	4.83	16	74%	17	0.77	16	13	6
53	77,838	8	$40,540	7	3.96	11	58%	12	0.23	9	6	1 - Lowest
54	225,074	13	$69,960	19	5.42	19	79%	19	0.09	5	9	5
55	50,833	6	$69,280	18	4.88	17	74%	16	3.11	19	62	19
56	321,684	15	$34,140	5	3.46	6	47%	7	0.20	7	26	14
61	171,190	10	$49,200	12	4.22	13	77%	18	0.16	6	35	16
62	824,182	18	$46,540	10	4.49	14	69%	14	0.25	10	21	12
71	67,928	7	$32,670	4	3.56	7	56%	10	0.09	4	16	9
72	571,039	17	$22,440	1 - Lowest	2.11	1 - Lowest	36%	2	0.31	14	18	10
81	106,124	9	$35,610	6	3.74	9	52%	9	0.08	3	7	3

Using the same results, one can hypothesize which industries may prevent sprawl:

- health care
- information
- finance and insurance
- educational services, and professional, scientific, technical service

These industries have 1) large employment changes, 2) high levels of income inequality, and 3) high levels of college attainment (see Table 9)

Final Thoughts

In early sprawl literature, sprawling growth was primarily attributed to housing decisions, federal policies and desires for socio-economic homogeneity, not industry labor demand. This paper suggests there are some industries that are impacting urban form primarily through their labor force demand. As counties shift in employment towards industries with lower mean incomes, total urbanized area in a county grows more rapidly. The urbanized area also becomes more contiguous. Just as with housing, as number of workplaces in a county increase, so does total urbanized area.

In counties with industries with flatter income distributions, total urbanized area grows but also remains more fractured. The same leapfrogging-sprawl is associated with industries with lower levels of college attainment.

Based on these industry attributes, counties with particularly large ratios of their workforces in accommodation and food services, construction, retail trade, and administrative, support, waste management are at increased risk of sprawl.

These results demonstrate that a comprehensive understanding of sprawl should include further analysis regarding the spatial forms of labor demand.

6. Bibliography

Beauregard, R. A. 2001. Federal policy and postwar urban decline: A case of government complicity? *Housing Policy Debate* 12 (1): 129-51.

Brueckner, J. K. 2000. Urban sprawl: Diagnosis and remedies. *International Regional Science Review* 23 (2): 160-71.

Burcheld, M., H. G. Overman, D. Puga, and M. A. Turner (2006). Causes of sprawl: A portrait from space. *Quarterly Journal of Economics* 121 (2), 587{633.

Buyantuyev, A., J. Wu, and C. Gries (2010). Multiscale analysis of the urbanization pattern of the phoenix metropolitan landscape of USA: time, space and thematic resolution. *Landscape and Urban Planning* 94 (3-4), 206{217.

Fujita, Masahisa, Paul R. Krugman, and Anthony Venables. 1999. *The spatial economy: Cities, regions, and international trade*. Cambridge, Mass.: MIT Press.

Florida, R. L. 2002. *The rise of the creative class: And how it's transforming work, leisure, community and everyday life*. New York: Basic Civitas Books.

Galster, G., R. Hanson, M.R. Ratcliffe, H. Wolman, S. Coleman, and J. Freihage. 2001. Wrestling sprawl to the ground: defining and measuring an elusive concept. *Housing Policy Debate*, 12(4): 681-718

Gordon, P., and H. W. Richardson. 2000. Critiquing Sprawl's critics. *Cato Policy Analysis* No. 365.

Hargis, C. D., J. A. Bissonette, and J. L. David (1998). The behavior of landscape metrics commonly used in the study of habitat fragmentation. *Landscape Ecology* 13 (3), 167{186.

Hietel, E., R. Waldhardt, and A. Otte. 2005. Linking socio-economic factors, environment and land cover in the German highlands, 1945-1999. *Journal of Environmental Management* 75 (2): 133-43.

Irwin, E. and N. Bockstael (2007). The evolution of urban sprawl: evidence of spatial heterogeneity and increasing land fragmentation. *Proceedings of the National Academy of Sciences* 104 (52), 20672.

Isserman, A. (2005). In the national interest: Dening rural and urban correctly in research and public policy. *International Regional Science Review* 28 (4), 465{499.

Kaza, N. 2011. The Changing Urban Landscape of the United States. Unpublished.

Li, H. and J. Wu (2004). Use and misuse of landscape indices. *Landscape Ecology* 19, 389{399.

Mieszkowski, P., and E. S. Mills. 1993. The causes of metropolitan suburbanization. *The Journal of Economic Perspectives* 7 (3): 135-47.

Seto, K. C. and M. Fragkias (2005). Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. *Landscape Ecology* 20 (7), 871{888.

Squires, G. D. 2002. *Urban sprawl: Causes, consequences, and policy responses.* Washington, D.C.: Urban Inst Pr.

Storper, M., and R. Walker. 1983. The theory of labour and the theory of location. *International Journal of Urban and Regional Research* 7 (1): 1-43.

Uuemaa, E., M. Antrop, J. Roosaare, R. Marja, and U. Mander (2009). Landscape metrics and indices: An overview of their use in landscape research. *Living Reviews in Landscape Research* 3 (1).

Wu, J. J. 2001. Environmental amenities and the spatial pattern of urban sprawl. *American Journal of Agricultural Economics* 83 (3): 691-7.