PLACE AND POVERTY

HOW DOES SPRAWL AFFECT POVERTY RATES IN U.S. CENTRAL PLACES?

A Thesis

Presented to the faculty of the Department of Public Policy and Administration

California State University, Sacramento

Submitted in partial satisfaction of
the requirements for the degree of

MASTER OF PUBLIC POLICY AND ADMINISTRATION

by

Cosmo Q. Garvin

SUMMER
2014

PLACE AND POVERTY

HOW DOES SPRAWL AFFECT POVERTY RATES IN U.S. CENTRAL PLACES?

A Thesis

by

Cosmo Q. Garvin

Approved by:

__________________________________, Committee Chair
Robert W. Wassmer, Ph.D.

__________________________________, Second Reader
Edward L. Lascher, Jr., Ph.D.

____
Date

Student: Cosmo Q. Garvin

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

____________________________, Department Chair ____________________

Robert W. Wassmer, Ph.D. Date

Department of Public Policy and Administration

Abstract

of

PLACE AND POVERTY

HOW DOES SPRAWL AFFECT POVERTY RATES IN U.S. CENTRAL PLACES?

by

Cosmo Q. Garvin

Economists have long been interested in understanding the relationship between economic inequality and urban form. Suburban white-flight, flight from blight, and spatial mismatch between jobs and low-income neighborhoods are believed to contribute to poverty and reduced economic opportunity. Using Census data for all U.S. central places and urbanized areas, along with an array of sprawl measures, this thesis tests for a relationship between sprawl and the change in poverty rates in central places between 2000 and 2010.

The results show urban area density consistently has a negative affect on the change in poverty rates within central places. Results show that for a density increase of 1000 people within an urban area, we expect to see the change in poverty rate in central places decline by 4 to 8 percent, holding other factors constant. The study finds however that other sprawl measures used, such as centrality, or mix of land uses, do not appear to have a statistically significant effect on poverty rates. This study also finds evidence that state and local growth controls may also have a statistically significant relationship to

changing poverty rates. Certain types of growth control may exacerbate poverty rates in local jurisdictions.

_______________________, Committee Chair
Robert W. Wassmer, Ph.D.

____
Date

ACKNOWLEDGEMENTS

Thank you Robert Wassmer for all your help and guidance over the years. Rob was teaching me how to think about urban economics before I ever stepped back onto campus for grad school, and he is a big reason that I joined the program. It has been fascinating and rewarding, and I owe him quite a lot for inviting me along.

Thanks to all of the folks in the CSUS Public Policy program. Ted Lascher was invaluable as second reader and editor, and has made a bigger impression on me than he probably realizes. I'm a better researcher and investigator because of him. Thanks to Suzi Byrd, whose importance to the success of PPA students cannot be overstated. We would be lost.

And finally, thank you my darling wife Rebecca. Who listened to my gripes about preparing data with a minimum of eye-rolling. Whose love and encouragement is the foundation for everything I've ever done that's ever been worth a damn.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Chapter 1.

Introduction

Poverty and the contemporary policy environment

Celebrity statistician Nate Silver recently blogged that mentions of economic inequality have skyrocketed on the major cable news networks between 2013 and 2014 (Silver, 2014). Silver also noted that "inequality" got far fewer mentions on the cable news shows than, say, Malaysian airliner "Flight 370." Nevertheless, the point is well-taken, poverty and inequality are getting more popular attention. Consider the arrival of the "99 percent" and the Occupy Movement on the American political three years ago. Consider the headlines made by the new Pope, Francis, when he began to speak against the growing the gap between rich and poor. President Barack Obama has hammered on themes of inequality, poverty and the decline of middle class to try to advance his policy agenda. And Republican leaders like Mitt Romney have joined in the call for a higher minimum wage. America is in a moment of heightened concern about poverty and economic inequality.

This year, 2014, also marks the 50th Anniversary of the "War on Poverty," and the creation of a whole set of federal anti-poverty policies including food stamps, Medicare and Medicaid, Head Start, and the Job Corps. These programs have had a tremendous impact on American society, but poverty certainly is far from defeated. There are many differences between the policy environment of 1964 and that of 2014. One of them is a more sophisticated understanding of the relationship between urban form, economic segregation and poverty.

That is the subject of this study. I want to find out if urban sprawl is contributing to the increasing economic segregation of American cities, and the effect that sprawl has on local poverty rates. In the next section of this chapter I will introduce some recent academic work making a link between sprawl and poverty, and explain how this recent work builds on long-established theories of urban economics, especially spatial mismatch. Then I will introduce my own hypothesis and outline the remaining chapters in this study.

America's economically segregated cities

In the year 2000, the median poverty rate for U.S. cities was 13.3 percent. By 2010, after a deep recession and weak economic recovery, it had risen to 15.6 percent. Table 1 shows the five U.S. central places with the largest percentage increase in poverty from 2000 to 2010, and the five places with the largest decreases in poverty during that period. (I also included the Census-designated Urbanized Area within which each central place is situated.

Table 1-1. Central places with largest increases and decreases in poverty rate 2000-2010.

Central Place	Urbanized area	Poverty rate in 2000	Poverty rate in 2010	Percentage change in poverty rate
Bradenton City, FL	Sarasota--Bradenton, FL	7.3	16.1	120.54

Wyoming city, MI	Grand Rapids, MI	7.3	16.3	123.28
Brooklyn Park city, MN	Minneapolis--St. Paul, MN	5.1	11.4	123.52
Missouri City city, TX	Houston, TX	3.3	9.1	175.75
Round Lake Beach village, IL	Round Lake Beach--McHenry--Grayslake, IL--WI	5.1	14.7	188.23
Folsom city, CA	Sacramento, CA	7.3	3.7	-49.31
Hidalgo city, TX	McAllen, TX	44.3	24.8	-44.01
Burke CDP, VA	Washington, DC--VA--MD	2.3	1.4	-39.13
Princeton city, TX	McKinney, TX	9.1	5.6	-38.46
Rosemead city, CA	Los Angeles--Long Beach--Santa Ana, CA	22.8	14.2	-37.72

Income inequality has been growing in the United States along with poverty. Economists use the Gini coefficient to describe income inequality within a community. The Gini number ranges from 0 to 1, with 0 representing maximum equality of income distribution within a place (all people have an equal portion of income), and 1 representing maximum inequality (one person has all of the income, and all other people have none). In 1995, the Gini inequality index for the United States stood at .450. By 2010 it rose to .46 and then jumped to .475 in the year 2012. Table 2 shows the U.S. central places with the largest and smallest Gini coefficients in 2010.

Table 1-2. Central places with the highest and lowest Gini index scores

Central place	Urbanized area	Gini score
Naples city, FL	Bonita Springs--Naples, FL	0.599
Morgantown city, WV	Morgantown, WV	0.58
State College borough, PA	State College, PA	0.58
Atlanta city, GA	Atlanta, GA	0.572
College Station city, TX	College Station--Bryan, TX	0.569
Clinton CDP, MI	Detroit, MI	0.339
West Jordan city, UT	Salt Lake City, UT	0.336
Bowie city, MD	Washington, DC--VA--MD	0.33
Burke CDP, VA	Washington, DC--VA--MD	0.329
Levittown CDP, PA	Philadelphia, PA--NJ--DE--MD	0.327

And finally we are becoming more economically segregated by geography. A 2012 study from the Pew Research Center (Fry, 2012) calculated the share of upper income families living in lower income census tracts, and vice versa, and found a growing trend since 1980: The wealthy increasingly live near the wealthy, the poor increasingly live near the poor. A 2013 follow up by The Pew Charitable Trusts (Sharkey, 2013) showed that U.S. metro areas with more concentrated wealth and more concentrated poverty had lower levels of economic mobility, controlling for other factors, and areas that were more economically integrated had higher levels of economic mobility.

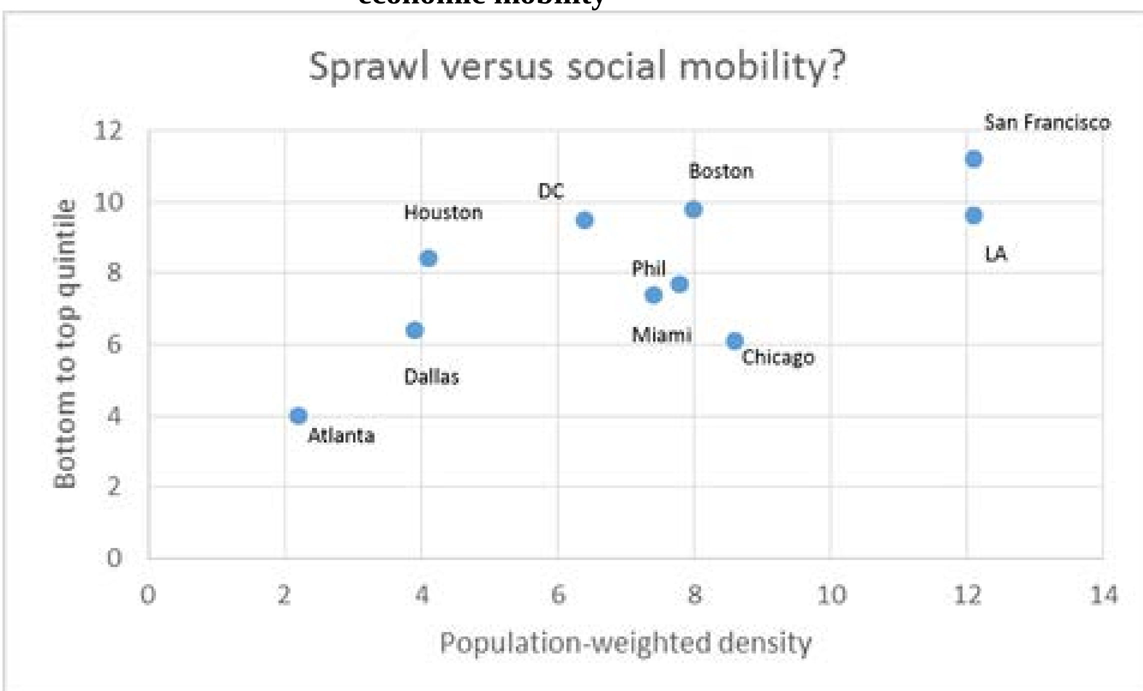
Economic segregation and integration are spatial measures. And the distribution of poor people and rich people in space has much to do with decisions that governments make about land use and urban form. Chetty (2013) found that upward economic mobility was higher in U.S. cities with less sprawl. This research generated great interest among journalists, commentators and researchers interested in the sprawl-poverty connection. But some noted that Chetty's proxy for sprawl—commute times—was perhaps too simple to have much explanatory value. I will return to the problem of measuring sprawl later in this section, and again throughout this study.

Economist Edward Glaeser (2011) has also explored urban form, poverty and urban decline in a book-length treatise. Glaeser makes a strong case for urban density to help lift people out of poverty, writing that, "The best cities have a mix of skills and provide pathways for those who start with less to end up with more" (p. 224). He devotes a chapter to Detroit, which saw its population drop from 1.8 million to 680,000 since 1950. Last year the city became the largest municipal government in the U.S. to declare bankruptcy, and large tracts of the city are abandoned. Glaeser argues that Detroit was hurt not only by the profound changes in the auto industry and American manufacturing and decline of industrial unionism, but it was also crippled by its own land use choices, which favored decentralization, sprawl, and economic segregation.

Sprawl and poverty

Economist Paul Krugman (2013) also draws a connection between Detroit's decline and sprawl. "If you like, sprawl killed Detroit, by depriving it of the kind of environment that could incubate new sources of prosperity." Drawing on Chetty's research on economic mobility, Krugman pointed out a rough correlation between urban density and robust economic opportunity.

Figure 1-1. Krugman's illustration of population-weighted density vs economic mobility



All of this work, Chetty's paper, Glaeser's book and Krugman's back-of-the-envelop scatterplot, all point back to the theory of "spatial mismatch."

Just a few years after the launch of the War on Poverty, economist Kain (1968) introduced the idea of spatial mismatch. He was concerned with racial segregation, and the distance it created between African Americans and available jobs. He argued that along with racist housing policies, jobs were becoming decentralized due to white flight and suburbanization. "The extensive growth of metropolitan areas and the rapid post war dispersal of employment," Kain wrote, have placed black job seekers in a, "more precarious position" (p. 196). Spatial mismatch has been a key concept in the understanding of urban form and economic inequality since Kain.

Some of the most recent and complete research on urban form and economic opportunity comes from Ewing (2013). The researchers employed an updated version of Ewing's well-known sprawl index (2003), and used it to compare U.S. metros using Chetty's economic mobility index. Ewing speculated that sprawl can diminish economic mobility by worsening the spatial mismatch with jobs, reducing access to sources of social capital, like social networks and community organizations, worsening income segregation and the negative externalities associated with pockets of poverty (Ewing 2013). The researchers found that a 10 percent increase in a metro area's sprawl index score (higher numbers indicating less sprawl) led to a 4.1 percent increase in the likelihood of a child moving from the bottom fifth of the local income distribution to the top fifth. Put simply, more sprawl leads to less economic mobility.

But there is disagreement among economists about the costs and benefits of sprawl. Joel Kotkin has written extensively—for a general, non-economist audience—in defense of suburbanization (Kotkin 2013). Kotkin responded to Chetty's research by pointing out that some of the cities ranked best for economic mobility were also relatively low-density places, like Bakersfield, California, and Salt Lake City, Utah. It is important to recognize that Chetty's research, and Ewing's extension of it, found a relationship between sprawl and economic mobility *holding other factors constant*. Kotkin raises an important point. Sprawling places can be places of economic opportunity, and dense places are not necessarily so. Even if sprawl tends to be a drag on economic mobility, many other factors play a part.

Also important, Kotkin points out that suburban development and sprawl has predominated because people like it, or at least they choose it over the alternatives. Kotkin and others are opposed to what they see as policies of “forced densification” and limiting choices in what kinds of communities people can live in. He also notes that the population of children is growing fastest in sprawling southern cities, like Houston, Dallas, Charlotte, Oklahoma City. Even champions of dense vertical cities like Edward Glaeser (2011) acknowledge the lure of suburbia for young families. He lists his own reasons for moving to the suburbs: “Living space, soft grass for spill-prone toddlers, a desire to diversify my life with greater distance from my employer, a fast commute, and good schools.”

All of this squares with what economists have identified as the benefits of sprawl. It provides greater housing choice, offers an escape some of the problems of dense urban living, and allows residents to “vote with their feet” and choose communities that provide the right balance of taxes and government services, like schools, parks and police, and other amenities, like shopping (Wassmer 2008).

The purpose of this paper is not to make a comparison of the costs and benefits of sprawl. Rather it is to learn empirically what effect sprawl has on poverty and income inequality, and to quantify the effect of sprawl on poverty when other known factors are taken into account.

The sprawl-poverty connection needs more research

Although much has been written about the costs and benefits of sprawl, and about the ills of social and economic segregation, there is not much literature that directly tackles the relationship between sprawl and poverty. In fact, although Chetty's research sparked a discussion about sprawl, it came out of a study of the effect of tax expenditures on economic mobility. And I found no empirical study asking what effect sprawl has on poverty rates.

If sprawl does lead to greater economic inequality, then why does it? Sprawl is a many-faceted process—including density, mix of land uses, distribution of jobs, and other factors. Ewing's study leaves some mystery about which ingredients of sprawl are the active ingredients. Is it population density that supports economic mobility? Or the centering of economic activity, or some other aspect of urban form?

Also, sprawl may have different effects on different groups of people within a geographic region. The studies mentioned in this chapter considered the effects of sprawl across metropolitan statistical areas. But the phenomena of interest are likely operating at a different level of analysis, between inner city and suburb, between cities within an MSA, or even between neighborhoods.

And economic mobility is only one way to measure the social effects of sprawl. We could also consider the variation of Gini index scores between sprawling and less sprawling communities, or rates of poverty, or crime rates or school test scores. Each in turn would give us a more complete understanding of the social effects of sprawl, and would suggest what policies, if any, might we adopt.

This study explores one corner of the relationship between sprawl and poverty by measuring the effects of sprawl on established urban places—inner cities, and older suburbs. Specifically, the hypothesis is that cities and other central places in more sprawling urbanized areas experienced greater increases in poverty than cities in less sprawling areas.

The study makes use of relatively simple methods of measuring sprawl, but tests them alongside more complex indices.

This research was prompted in part by the current interest in poverty and economic inequality, and also by the small but growing literature on “first suburbs” and inner ring neighborhoods that are a vast but often overlooked part of our communities. These areas often fall into a “policy blindspot” (Puentes 2006, Lee and Leigh 2007). They don’t receive the same attention and investment as “back to the downtown” neighborhoods. And they don’t receive the same level of investment and amenities as new suburbs. The method of this study is to focus on Census-designated “central places” as the primary unit of analysis—in order to capture the effects of sprawl on both inner cities and older suburbs.

By measuring variation in poverty rates, and by focusing on central places, I believe this study contributes something unique to the understanding of sprawl and income inequality.

Organization of this study

The study is divided into five chapters. This one has introduced the some of the current academic and popular thinking on urban form and economic inequality. The next chapter surveys literature pertinent to any investigation of urban form and economic inequality. I consider methods of measuring sprawl, and methods of measuring the social outcomes of sprawl. I also dig more deeply into the theoretical foundations of why decentralization may exacerbate economic sorting, segregation and spatial mismatch. And I explore some of the literature on older suburbs, to better understand how the processes of economic segregation affect this particular sort of neighborhood.

In the third chapter I describe the data and methods I will use to try and tease out the effects of urban sprawl on poverty and economic segregation. This chapter includes an explanation of the broad causal model of the effect of sprawl on poverty and description and descriptive statistics for the variables that I intend to use in my study. The hypothesis is that poverty increased a greater rate in central places that were situated within more sprawling urbanized areas—accounting for other factors that cause a change in poverty.

The fourth chapter includes a finer specification of the regression model(s) that I used to test my hypothesis. This chapter includes a report of my initial regression results, comparisons of several different models. The models consistently show that there is a statistically significant relationship between the population density in an urbanized area and the change in poverty in central places between 2000 and 2010. There is also evidence to suggest that growth management policies can have an effect on poverty, though sometimes in unintended ways.

The fifth and final chapter contains my conclusions and my assessment of the limitations of this study and recommendations for future research.

Chapter 2.

Literature review

Measuring sprawl and its effects

Suburban sprawl has been blamed for an array of problems. Loss of agricultural land and open space, air pollution, traffic congestion, are among the most commonly cited. But researchers also believe sprawl is a contributor to social ills like economic and racial segregation, lack of civic involvement and decay of inner city neighborhoods.

The purpose of this literature review is to investigate the effect that new suburban sprawl has on older, central places in a community. It proceeds from the question, “In areas with more sprawl, do established neighborhoods and central places fare better or worse?” I am additionally interested in the effects of sprawl on so-called “inner-ring suburbs” or “first suburbs.” Some researchers believe this sort of neighborhood—generally built between 1950 and 1980—is particularly vulnerable to urban decline. Researchers note that first ring suburbs are in a “policy blind spot,” not benefitting from the investment on the urban fringe, or from the “back to downtown” energy that has helped to revitalize some downtowns (Puentes, 2006 and Lee and Leigh 2007).

In order to determine the effect of sprawl on older places, it first must be asked, “what is sprawl?” This study begins with studies that seek to define and measure sprawl. Some methods include studying employment decentralization (Glaeser and Kahn 2001), the rate at which land is consumed relative to population growth (Ewing, Pendall and Chen 2003), and more complicated indexes which measure several facts of sprawl

(Galster 2001). Next, I look more closely at neighborhood decline, and why some neighborhoods experience decay while others thrive. This section reviews some of the fundamental theoretical literature on neighborhood choice (such as Tiebout, 1956), and also asks "what is blight?" and how do researchers quantify and explain the reasons for urban decay.

Then I consider studies that specifically focus on the relationship between suburban sprawl and decline in inner neighborhoods. Among these, Brueckner (2011) explains both sprawl and blight as two results of market failures that reduce the costs of suburban development. The implication is that policies which require suburban development to pay more of its own social costs would discourage both sprawl and blight. Lee (2011) looks at the effect of urban containment policies and finds that more compact cities have less economic disparity between neighborhoods. Finally, I sample those empirical studies that look at the social costs of sprawl, such as income inequality, lack of job opportunity, and crime.

What is sprawl, and how is it measured?

Sprawl is often talked about by policy makers and advocates, but it is difficult to define. The word is variously used as a noun or as a verb, to describe a land use pattern, a kind of ugly style of development, or as the cause of externalities like traffic congestion and air pollution (Galster, 2001).

Some of the most commonly described characteristics of sprawl are low-density growth in residential and commercials uses, rapid growth, decentralization of economic activity over time, lack of mass transit facility, economic segregation, and extensive single-use tracts with little mixing of commercial and residential tracts. The following section considers a small sample of sprawl indexes, which are based on population density and employment centralization, as well as more complex methods that take into account multiple aspects of sprawl. (See Appendix A, Table A-1 for a summary of the methods described in this section.)

Perhaps the most direct method of ascertaining the extent of sprawl is simply to measure the spread of an urbanized area in square miles. Wassmer (2006 and 2008) uses size and population as dependent variables in regressions seeking to determine if growth management policies, or auto-dependency, or other factors such as poverty in central places, are causes of sprawl. Because Wassmer includes population as one of several explanatory variables, it is "controlled for" and the effect of other variables can be explored.

Another simple method of measuring sprawl is to measure the growth of developed land compared to the population growth in an urban area (Fulton 2001). This is the foundation of many studies, but this method is somewhat crude, and leaves out many of the characteristics of sprawl aside from density. "A variable or two cannot adequately represent the inherent complexity of sprawl," (Ewing 2003).

Other approaches try to tease out changes in centrality within an urban or metro area. Wassmer (2000) measured centralization within Census MSAs as central city population divided by regional population. Another Wassmer paper (2001) employed a sprawl index that measured the percent change in urban "fringe," or non-central places within a Census urbanized area, divided by the percent change in population within the urban area.

Other researchers employ more complicated sprawl indices, such as that developed by Galster (2001) and colleagues. This approach considered 13 U.S. Census designated urbanized areas (UA) and eight dimensions of sprawl. Galster used Geographic Information Systems (GIS) to divide each UA into one-mile square grids, and measured eight variables in each. First is density, the number of residential units per square mile of developable land. Second is continuity, whether a census tract is sufficiently dense to be considered a continuation of development. Next, concentration and clustering describe how development is distributed across a UA. Centrality measures the degree to which development is near to the UA's central business district (CBD). Nuclearity describes the degree to which the UA is monocentric, with one center of economic activity, or the degree to which it is polycentric. And finally, mixed uses and proximity describe the degree to which land uses are kept separated. For each of these measurements a Z score (showing the number of standard deviations from the mean) was calculated. The Z scores were then added to give a composite sprawl index for each of the 13 UAs. The lower the score, the greater the extent of sprawl. So Atlanta, Miami and Detroit were the top-sprawling UAs, while New York, Philadelphia and Chicago were less sprawling. One problem with Galster's index is that is limited to a relatively few urbanized areas. To compare a greater number of communities, researchers may need to choose simpler methods.

Glaeser and Kahn (2001) measured decentralization of employment in metropolitan areas. They obtained data from the U.S. Department of Commerce's Zip Code Business Patterns for 1996, then measured the number of and size of firms in each zip code within the 100 largest U.S. metropolitan areas. They then measured the distance of each zip code from the CBD in its metro area—thus getting a distribution of the amount of employment with a three-mile or ten-mile ring of each CBD. Glaeser and Khan found that employment decentralization was correlated with greater political fragmentation. This method allows for comparing metropolitan areas at one time, but says nothing about employment decentralization over time. This method has the advantage of being relatively simple, but there is much it leaves out, such as residential sprawl.

Weitz and Crawford (2012) built on Glaeser and Kahn's employment decentralization method to create their own method of measuring job sprawl. Where Glaeser's model was still largely monocentric, describing employment relatively to single central business district, Weitz and Crawford looked at jobs and housing together. The researchers used Census zip code data and GIS to study the proximity of employment to residential population.

Ewing, Pendall and Chen (2002) used a more complex multidimensional approach that balances population density with the other aspects of sprawl. Using Census data these researchers weighted and combined several variables into four factors: density, land use mix, degree of centering (concentrations of economic activity), and street accessibility (for example street block size, and how connected the street grids are).

These factors were again combined into an overall sprawl index score. One advantage of this approach is that in addition to ranking areas on a scale of more or less sprawl, it also allows the researcher to explore the kind of sprawl that characterizes a community. For example Tucson and Fort Lauderdale are similar in their overall sprawl score. But Tucson has large street blocks and low density housing, while Fort Lauderdale has smaller blocks, higher density residences, with a much less defined city center. (Fort Lauderdale's elongated skyline may also have something to do with the fact that its downtown is just a few blocks from the Atlantic Ocean.) Ewing's sprawl index has been updated and used by other researchers several times to measure different social effects of sprawl, and as mentioned in Chapter 1, most recently used to compare economic mobility and the extent of sprawl in U.S. metro areas.

Cutsinger, et al (2005) refined Galster's earlier sprawl index, extending it to 50 metro areas and introducing the idea of the "extended urban area." The EUA includes the old UA designation in addition to contiguous one-mile square cells that contain at least 60 residential units, where at least 30 percent of workers commute into the UA. The EUA answered earlier concerns that the UA was not truly capturing low-density "leap-frog" development.

Paulsen offered a more recent attempt at addressing the problem of measuring leap-frog development is from Paulsen (2013). He was concerned with the fact that Census definitions of "urbanized area" change every ten years. Paulsen employed an extended urban area, defined by Census blocks with more than 1,000 residents per square mile, and adjacent blocks with at least 500 residents per square mile, allowing for ‘jumps“ over unpopulated areas no more than a half mile wide. Paulsen then used four metrics for measuring sprawl: change in housing unite density within the EUA, marginal land consumed per net new urban housing unit, the housing unit density in newly urbanized areas, and finally the extent of “infill” development, measured as the percentage of net new urban housing units in previously urbanized areas.

This variety of methods used by researchers to measure sprawl suggests how complex the sprawl process is. However, the two most commonly measured aspects of sprawl are density and centrality. Finding the population density within a given geographic area--such as a Census-designated urbanized area—is not complicated, and the data is readily available. For this reason, population density is a good candidate for use as key explanatory variable in my study. At the same time, the widespread publication and use of Ewing’s sprawl measures makes them a strong choice as well. I will use both and see how they compare in predicting change in poverty rates.

The next section considers the process of decline in established and already built-out places. After that, I review the literature on the social costs and benefits of sprawl.

Why do places decline? Blight and residential segregation

Economic explanations for neighborhood decline include economic segregation, housing filtering, and Tiebout sorting.

Housing filtering: Houses over time deteriorate, or become obsolete, or unfashionable. So house that start off as desirable to an affluent occupant will become less desirable over a decade (O’Sullivan 2009). That owner can spend a lot of money maintaining and upgrading. Owners can move on to a higher quality, newer house, perhaps in a different neighborhood. It is the occupant that moves, not the house. The house is re-occupied by someone at a lower income, who can now afford this lower quality house. Through this filtering process, all neighborhoods are in a state of constant decay. Without maintenance and new investment, they physical housing stock will deteriorate, become less desirable, and the social characteristics of the neighborhood will change.

It is not only age that drives the filtering process. As noted in Fusch and Ford's (1983) paper "Architecture and the Geography of the American City," housing materials and architecture (including the overall style of the neighborhood) "play an important role in the scale and rate of socio-economic change" in a neighborhood.

Economic segregation: People are willing to pay a premium to live in a "good neighborhood" and escape a bad one (O'Sullivan 2009). For example, crime rates will influence a family's decision about whether to locate in a neighborhood. O'Sullivan notes that empirical studies show high-income households can and will pay more than low-income households to live in low-crime areas. This causes economic segregation. Schools also drive neighborhood choice. A desireable school will likely put upward pressure on housing prices. High-income families can better afford to locate to the attendance area of a sought-after school. This reinforces income segregation.

Tiebout sorting: A related idea, which is named for economist Charles Tiebout, is that groups of similar-minded people will "vote with their feet" and sort themselves based on their willingness to pay for a package of public goods. This kind of sorting has benefits. You can pick a town or a neighborhood that provides things that are important to you, like schools, or transit, or extra police, or vintage street lamps, or parks. But there is also cost to Tiebout sorting, which is inequity. The willingness to pay also measures the ability to pay. This leads to spatial segregation and economic segregation—some of the most recognized characteristics of sprawl.

Several researchers have done work showing a connection between sprawl and economic segregation. Wassmer (2008) showed that urban containment policies do in fact help reduce the urban area footprint. But he also noted the biggest influences on the size of an urban areas footprint are in poverty and crime, not land use policies. Crime and poverty tend to push out the urban boundary, and cause decentralization. This is called the "flight from blight" explanation of sprawl.

Jargowsky (2002) looked at Census tract data, and variables like median age of homes in census tracts and median income in those tracts, and showed that cities are developed in these ring like patterns. Income tends to fall as you work your way into the inner rings. Since the outer rings are where the job growth is occurring, including all of the growth in low-skilled jobs, there's a spatial mismatch between the poor job seekers in one neighborhood, and the jobs further out in the suburbs. They may not even be able to find out about jobs, or get to the jobs without difficulty. Over the long run, Jargowsky argues those places are not developing human and cultural capital, and that is leading to further poverty. "Sprawl...is clearly part of a larger process that leads to more spatial, racial, social and economic distance between urban and suburban neighborhoods."

The inner ring suburb, a particular kind of urban decline

According to the Brookings institution (Puentes 2006), first suburbs are home to as much as 20 percent of the U.S population. These neighborhoods are close to the urban core, and in many metro areas these are the suburbs that once housed the white-collar workers who would commute downtown. Sometimes first suburbs are known as streetcar suburbs, suggesting their history and also the hope that some policy makers see in their renewal.

These suburbs once fit the stereotype of the 1950s all-American suburb. They were an earlier version of sprawl, and the beachhead for previous waves of flight from blight. Today they are likely to be racially diverse. By some measures, first suburbs have a higher portion of foreign-born residents than their primary cities (Puentes 2006).

The neighborhoods are not just second-hand suburbs where less affluent immigrants can find cheap housing. Because of the role of immigration in inner-ring suburbs, these neighborhoods can be important places in the cultural landscape and economic development of a city (Datel 2008).

Researchers have used a variety of methods to define and describe inner-ring suburbs. (See Appendix A, Table A-2 for a summary of the methods described in this section.) Hanlon (2009) studied Census Designated Places, and identified places where the majority of the housing stock was built prior to 1969. She used factor analysis to show socio-economic differences between 1742 designated places. She then developed a typology of inner-ring suburbs including: ethnic, low-income, vulnerable, old and middle class. The Brookings Institution study (Puentes 2006), compared data in 64 Metropolitan

Statistical Areas. Data from primary cities was extracted. “First suburbs” were defined as areas of counties around primary cities, and this data was separated as well. These data were compared to data from “new suburbs,” defined as remaining area of the MSA surrounding the primary county. Anacker (2012) listed counties with cities of more than 100,000. First suburbs were defined as census tracts outside of places with more than 100,000 people, but still within the primary county. Finally, Lee (2007) defined inner rings of development by census tracts and percentage of housing stock built between 1950-69. This yielded four distinct rings, downtown, inner city, inner-ring, outer-ring.

The methods used by Puentes and Hanlon allow for comparisons across a broad sample of metropolitan areas. Hanlon’s also recognizes the importance of political jurisdictions. However designating entire cities or counties as “rings” of development may not truly capture the nature of development in a community. Lee’s approach of assembling rings from census tracts allows a much finer-grained look at the pattern of development in a community over time. But the difficulty of collecting and normalizing large amounts of census tract data limits this approach to the study of just a few metropolitan areas at a time. For example, Lee (2007) compared just four metro areas: Atlanta, Portland, Philadelphia and Cleveland.

Sprawl and urban decay

Researchers have produced theoretical and empirical work connecting the sprawl and policies toward sprawl with certain social outcomes in inner ring neighborhoods. One common theme in this research is the way in which sprawl—almost by definition—

affects maintenance and re-investment of resources in older neighborhoods. Another theme is the patterns of economic segregation and inequality that are sometimes associations with more sprawling landscapes.

Brueckner (2011) presents a theoretical model that blames a set of market failures for both suburban sprawl and urban blight. Brueckner argues that there are distortions—lack of congestion pricing, lack of adequate pricing for the amenity of suburban open space, and inadequate pricing of infrastructure—which lead to an oversupply of sprawl. By the same token, over-investment on the urban edge means under-investment in the central city.

Lee (2007) discussed a model of urban rings which takes into account the pull of suburbanization, and also acknowledges a "back to the city" trend among certain more affluent professionals, and empty nesters, who are finding that the central city offers a set of amenities and public goods worth paying for. Sometimes the back to the movement follows intentional public policies aimed at revitalizing a downtown. Sometimes the policy follows the movement. This model suggests the creation of a torus, or donut shaped ring of neighborhoods, that are largely missing out on investment and new economic activity happening on the suburban fringe and in the downtowns. This is what the Brookings researchers refer to as a "policy blind spot."

In a later study (2011) Lee hypothesized that cities with urban growth boundaries or more compact growth would have lower levels of socio-economic disparity. He looked at Seattle and Portland, Atlanta and Dallas, Detroit and Cleveland. Using similar census tract data he organized these cities into three rings--inner city, inner ring suburb and outer ring suburb. A Gini inequality index score was given to each ring and for each metro region overall. This was not a regression analysis, and the author acknowledged that two cities like Atlanta and Portland are very different demographically, socially, historically. Some of the differences may be due to “unobserved endogenous factors,” and omitted variables. However, Lee found that Seattle and Portland (both known more for more compact growth) had comparatively lower Gini numbers overall compared to Atlanta, Dallas and Detroit. Atlanta and Dallas also had high inequality between rings of development.

Rico (2013) sampled MSAs for which there was data available on physical blight, including variables like the presence of illegal dumping, broken windows, presence of vacant buildings nearby, percentage of buildings in need of repair. From these Rico created a blight index, which she used as the dependent variable in a subsequent regression study. Explanatory variables included demographic factors like central city median income, population, percentage African American, and percentage Hispanic. The key explanatory variable was whether or not an Urban Containment Policy had been place for a certain number of years. Rico’s regression showed some evidence that UCP’s are associated with less blight. The magnitudes of the effect are difficult to sort out, because the regression coefficients refer to Rico’s blight index. But in one specification, and urban containment policy was shown to be correlated with lower levels of blight to an extent greater than demographic explanatory variables, and greater than crime rates.

The social costs of sprawl

The point of measuring sprawl is in part to understand the social consequences of sprawl and determine what policies, if any, should be applied. This section looks at some ways in which researchers have tried to measure the social effects of sprawl, the most relevant studies concern the effects of sprawl on income inequality, economic opportunity and crime.

There is a large body of literature on the relationship between suburbanization and crime. Shihadeh and Ousey (1996) wrote that suburbanization sorts cities on the basis of race and class, with wealthier white families moving toward the suburban edges. The authors measured the proportion of population within U.S. metro areas that lived in suburbs, according to Census data, and compared that to crime rates from the FBI's Uniform Crime Reporting data. Their regression study found that a one standard deviation increase in suburbanization was associated with a 0.240 standard deviation increase in murder arrests for blacks and a 0.183 increase in arrests of blacks for robbery. Suburbanization had no effect on white arrest rates. The authors concluded that increasing suburbanization exacerbated "decline and disinvestment in black communities."

Sprawl is often explained as product of "flight from blight" and urban problems like crime (Wassmer, 2008). But Jargowsky (2009) suggests the causal arrow can point in the other direction, and sprawl can be a cause of more crime. Comparing county-level data from the FBI Unified Crime Reports program to Census data, Jargowsky found that population density did appear to be associated with crime rates. He reported his results with standardized coefficients, where a change of one standard deviation in the explanatory variable results in a change of one standard deviation in the dependent variable times the value of the coefficient. In this case population density had a coefficient of negative 0.19 for violent crimes per 100,000 residents, and negative 0.46 for property crimes per 100,000.

Massey and Rothwell (2009) looked for evidence that low-density zoning contributed to segregation between black and white residents, by lowering the amount of affordable housing in suburban areas. Such exclusionary zoning is an often-observed feature of sprawl. In this regression study the dependent variables were an index of black-white residential dissimilarity in 49 of the nation's largest metropolitan areas, and an index of black isolation for the same areas. The key explanatory variable was a zoning density measure drawn from previously gathered survey data. The authors found that arenas with higher allowable densities had significantly lower overall black-white segregation and a lower score for black isolation (meaning less isolated).

Researchers have similarly found evidence that sprawl and decentralization is associated with economic inequality. Stoll (2004) noted that African American households remain concentrated within inner-city neighborhoods, but employment opportunity has steadily moved toward the suburban fringe. This "spatial mismatch" between jobs and minority groups has long been considered contributing factor to income inequality. Stoll used an index of "job sprawl" using the U.S. Department of Commerce's1996 Zip Code Business Patterns (similar to Glaeser and Khan, above). Data from the 2000 Census were also used to compare job location to the distribution of blacks, whites and Latinos in U.S. metropolitan areas. A key finding was that a 10 percent increase in the job sprawl was associated with a 3 percent increase in spatial mismatch for blacks.

Yang and Jargowsky (2006) use a neighborhood sorting index to measure economic segregation between neighborhoods in a metropolitan area. The NSI is a ratio of the variance of household income within a Census MSA over the income variance between Census tracts. The NSI is their dependent variable, measuring income inequality, with 1 being completely segregated economically and 0 being unsegregated. The mean for the U.S. in 2000 was .384. The key explanatory variable is a sprawl index that includes measures of: the urban density gradient within an MSA, the MSA population density, the number of local governments per 100,000 households, the homogeneity of housing values in the newest ring of development, and the commute times for central city residents. The authors noted that income inequality decreased overall between 1990 and 2000. But it decreased less in metro areas that experienced a high degree of suburbanization. Density measures had the most significant effects on income inequality. As the density gradient decreased, meaning an MSA became more suburbanized, the NSI increased. However as central city population density increased, so did income inequality, an effect the authors surmised might be due to increasing competition for space.

Wheeler (2006) also studied income inequality between and within census tracts. He found that overall inequality, within all census tracts, rose as overall density of a metro area fell. But in contrast to Yang and Jagorsky, Wheeler found that income inequality between tracts was not significantly associated with sprawl.

And finally the Chetty and Ewing studies mentioned in Chapter 1 drew an explicit connection between economic mobility and sprawl. In particular, Ewing's study (2013) found that a 10 percent increase in a metro area's sprawl index score (higher numbers indicating less sprawl) led to a 4.1 percent increase in the likelihood of a child moving from the bottom fifth of the local income distribution to the top fifth.

The benefits of sprawl

When considering policy, it is important to recognize that suburbanization and sprawl bring benefits as well as social costs. Indeed, some like Kotkin would argue the benefits outweigh the costs. Glaeser (2011) writes that the hunt for quality school may be the single most important factor driving suburban development. Suburbanization may in fact create a feedback loop in which families with the means move to suburban school districts, leaving urban districts behind to cope with greater poverty and budget constraints, thus driving more families to neighboring districts, and so on. The situation may be more desperate for urban schools, but the incentive to move to a suburban district is clear.

Glaeser and many others have noted that another strong incentive for suburban living is that the cost of owning a home is comparatively cheaper in the suburbs. The phrase "drive 'till you qualify" has become short-hand for the choices families must make in balancing housing costs against the costs of long commutes.

Much of the literature reviewed in this chapter defines sprawl in terms of density or centrality. But sprawl can also describe a growth process. Sprawl is by definition "bad" if it means growth that is poorly managed, or wasteful of natural resources and government budgets. But if a suburban jurisdiction is careful in its planning, implements strong habitat conservation efforts, affordable housing plans, and invests in alternative modes of transportation, then the argument for the benefits of suburban development is likely much stronger.

Probably the strongest argument for suburban development is that people want it. Burchell (1997 p. 170) noted, "Americans have repeatedly indication their preference for, and satisfaction with, suburbs and suburban living." That preference is reflected in the processes of sorting and neighborhood choice described above, as those who are able to move to suburbs in search of greater housing opportunity,

Conclusion

Again, the purpose of this literature review is not to try and settle a debate about the overall costs and benefits of sprawl. Instead it is to learn more about how researchers have tried to measure sprawl, and the social effects of sprawl, and in particular the effects that sprawling land use patterns have on older neighborhoods such as inner ring suburbs. It appears that the most commonly measured aspects of sprawl are population and housing density and some measure of the centrality of population and economic activity. I will employ measures of density and centrality in my own study.

Some researchers have used simple indexes for measuring sprawl that allow comparisons of many metropolitan areas. Others like Galster construct more complex indices which capture the many attributes of sprawl but are limited to looking at a relatively few metro areas. And others such as Ewing have developed measures of several attributes of sprawl, which can be applied to dozens of metro areas and hundreds of urbanized areas. As explained in the next chapter, I will use Ewing's sprawl index alongside simpler measures of density in my own regression study.

This literature review also shows that sprawl has been compared to a wide ranging set of social phenomena, from economic mobility to obesity.

The literature also suggests it is possible to measure sprawl and test the effects of sprawl on central places using widely available Census data. Interestingly, I found no empirical studies that directly measured the correlation between sprawl and poverty rates—although this would seem to be a natural question to ask, and detailed poverty data is available from the Census down to the tract level.

In the next section, I will present a model that explains the possible effects of sprawl on poverty in central places, and provide a detailed description of the data I will use to test that model.

Chapter 3.
Methodology

Using regression analysis

The literature review in Chapter 2 lays the theoretical foundation for this study. It suggests there could be a relationship between sprawl and poverty in urban areas. The "flight from blight" explanation of sprawl holds that poverty and attendant social problems drive decentralization of an urban area (Wassmer 2006). The causal arrow may also point in the other direction; development patterns, zoning and other land use policies may exacerbate economic segregation of urban areas. Sprawls may also affect central places through the spatial mismatch of low-skilled jobs in the suburbs and low-skilled workers more centrally located, the decline of local schools, and the loss of investment in local business and amenities.

As mentioned before, I found no studies that directly tested the effect of sprawl or urban decentralization on poverty rates in U.S. cities. The study that came closest to answering my question was Ewing's investigation of sprawl and economic mobility. Economic mobility is a useful concept. But in order to understand poverty, it makes sense to measure poverty. In this chapter I present a theoretical model of the factors which contribute to poverty, and the methods I will use to test the effects of decentralization on poverty rates.

The first section includes a brief description of regression analysis, a statistical tool that can be helpful in this sort of investigation. I will then describe the theoretical model used to explain variation in poverty and income inequality, and how development patterns might affect economic segregation. Next I will describe the dependent variable used to measure the effects of economic segregation, and the explanatory variables used to predict the value of that dependent variable. The last section of this chapter offers the appropriate descriptive statistics for the dataset used, and discusses the correlations between several of the explanatory variables. This chapter lays the methodological groundwork needed before reporting the results my regression analysis in Chapter 4.

Because social phenomena like poverty and inequality are so complex, it is difficult to determine all of the contributing factors or to measure the effect of any one factor. Regression analysis allows us to isolate the influence of one variable while holding all other variables constant (Wheelan 2013). Other techniques, such as correlation analysis, may indicate the strength and direction and significance of a statistical relationship between two variables. But regression analysis allows us to predict the magnitude of effect of one variable on another. It also allows us to control for confounding variables. In this study, confounding variables may be factors such as education level, or the local employment base, or the presence of a large immigrant population, which along with other factors may be contributing to the local poverty rate. It is essential to control for these variables if we are to determine if development patterns truly account for any truly statistically and substantively significant difference in poverty rates between places in more-sprawling and less-sprawling urban areas.

The model and unit of analysis

In this study I am trying to understand the effect that sprawl has on established neighborhoods, in central cities and older suburbs. For this reason, I chose U.S. Census designated "central places" as my unit of analysis. For many years the U.S. Census Bureau designated the largest places in an urban area as central places. I reason that these central places will be most likely to contain the types of older inner city communities that I am interested in studying.

I propose a model in which the change in poverty rate in a given community over time is a function of the local economy, the demographic mix of that place, and the development patterns that lead to a certain level of social and economic sorting. This sorting happens on a scale larger than neighborhoods or census tracts. In the flight-from-blight scenario, families who have the means will move out of central cities to suburbs with greater amenities and fewer poor neighborhoods. Flight from blight concentrates poverty in the central places. However, it also means that those left behind have less access to jobs, quality schools, and other amenities better supported in the more affluent places. If flight-from-blight and sprawl are contributing to poverty and income inequality, then it should be possible to measure greater poverty in places where there is greater sprawl in the surrounding urban area, after controlling for other possible explanations. Also, if sprawl contributes to poverty, that has policy implications. Growth controls and anti-sprawl policies are often proposed as a way to fight traffic congestion, or preserve open space, or curb greenhouse gas emissions. If crafted properly, they may have the added benefit of reducing social inequality.

The first set of explanatory variables in this model concern the local economy, which is affected by factors such as the percentage of the workforce employed in particular industries and education of the workforce. The second set of explanatory variables is demographic, and includes the age mix of the local population, the race and ethnicity of local residents, and the size of the local immigrant population. Finally, this model uses measures of sprawl as key explanatory variables. As suggested by the literature review, local development patterns can be described in terms of the growth rate of the area, the overall population density and centrality—meaning the extent to which residents are more concentrated in central places within a larger urbanized area. I specify the model as follows:

Poverty change = *f* (Local economy$_i$, demographic mix$_i$, development$_i$).

Where i = 1,2,3, ….XXX Central Places in the United States in 2000 and 2010

Local economy$_i$ = *f* (poverty$_i$, unemployment$_i$, educational attainment of residents$_i$, income of residents$_i$, percentage of employment in key economic segments$_i$).

Demographics$_i$ = *f* (race/ethnic mix$_i$, immigration$_i$, percentage of single mothers$_i$, age mix).

Sprawl$_i$ = *f* (density$_i$, centrality$_i$, growth rate$_i$).

Selection of the dependent variable:

The U.S. Census offers a very large collection of data through which to explore the distribution of social phenomena in a given area. I have chosen to measure the change in the poverty rate in U.S. Census-designated central places, between 2000 and 2010, as my dependent variable. The base poverty rate is determined by causal factors operating over decades. By measuring the change in poverty rate between 2000 and 2010, I am hoping to better capture the effects of recent land use patterns. If decentralization and economic segregation are having an effect on economic opportunity in central places, then there should be a statistically significant difference in the change of the poverty rate for places within more sprawling urbanized areas.

As mentioned above, I chose central places because they are the geographical unit that contains the inner cities and older suburbs of concern to this study. I should also note the U.S. Census no longer identifies "central" places within urbanized areas. The places in this study were designated central places in the 2000 Census. The 2010 Census refers to them simply as "places."

Data sources and selection of explanatory variables.

This study seeks to understand the effects of suburbanization on older, established places such as inner cities and inner suburbs. So it is appropriate to measure those effects at the place level. Sprawl can occur inside boundaries of a central place. But to understand how sprawl affects central places, it is necessary to measure sprawl that occurs outside of central places, in the adjacent suburban area. This model supposes that sprawl is "pulling" affluent residents and economic activity away from central places into the surrounding suburbs.

So, although the unit of analysis is the central place, the key explanatory variable is the degree of sprawl occurring in the urbanized area that surrounds each central place.

The Census defines urbanized area as "densely developed territory that contains 50,000 or more people." The land outside of designated urbanized areas is considered rural.

All of the data measuring sprawl are collected at the level of the urbanized area. But all other explanatory variables—such as unemployment rate, education levels, and local economic and demographic mix--are measured at the level of the central place.

As the literature review suggests, there are many possible approaches for measuring the extent of sprawl, or density, or centrality, in a geographic area. Some measures of sprawl are relatively simple to calculate, others are more sophisticated and complex indices. This study tests both kinds of measures. I use a straightforward measure of density within urbanized areas, drawn directly from Census data. I also employ the more complicated set of sprawl indices released by Ewing (2013). These include measures of density, centrality, connectivity, and mix of land uses.

Sprawl can be measured as a particular set of characteristics (density, centrality, mix, etc) at a particular time. But it can also be thought of as a process, in which communities become more or less sprawling over years. It is simple to generate measures of sprawl change or density change, using the sprawl measures above. However, this introduces the problem of endogeneity among variables. An independent variable is said to be endogenous if it is correlated with the error term—or the unexplained portion of the variation of a dependent variable—in a regression. In this case, if we observe a change in urban density and a change in urban poverty in the same time period, it will be difficult to determine if one change is causing the other, or if causation goes in both directions. One possible solution to this problem is to use the technique of two-stage least squares regression, which uses instrumental variables uncorrelated with the error term in place of the endogenous variable. This study uses data on growth control measures in U.S. cities and states to generate a set of instrumental variables in an effort to solve a potential problem with endogeneity of variables. These measures are adapted from Dawkins and Nelson (2003), with some changes suggested by Wassmer (2006). This dataset has been cited repeatedly by researchers (such as Paulsen 2013) seeking to discern the social impacts of such containment policies. The Dawkins and Nelson dataset shows states with a statewide set of growth controls, and also lists local governments using growth controls as of 2000. These are categorized according to whether the local growth control accommodates projected future growth ("accommodating" or "restrictive") and whether the growth measure makes provision for adequate land supply, affordable housing, infrastructure and land conservation ("strong" if so, "weak" if not).

The authors cite the well-known Portland Oregon model as an example of a "strong-accommodating" plan, characterized by strong housing and open-space policies and a plan for accommodating new growth within a certain planning horizon. At the opposite corner are "weak-restrictive" planning regimes, which lack effective containment and also lack goals for accommodating future regional growth. Appendix B shows the Dawkins and Nelson data for state and local growth controls.

Finally, differences in poverty rate may be due to regional factors, economic conditions and policies not controlled elsewhere in the model. For this reason, I am including dummy variables for each state in the regression analysis below. Following from the model above, I specified a complete set of explanatory variables.

Local economic factors$_i$ = f (Poverty rate in 2000$_i$, Unemployment rate$_i$, Percentage high school graduate$_i$, Percentage some college$_i$, Percentage who have obtained a bachelors degree$_i$, Percentage who have obtained a graduate degree$_i$, Percentage who have annual incomes above $200,000$_i$, Median household income for the central place$_i$, Median home value in the central place$_i$, Percentage employed in the construction sector$_i$, Percentage employed in the manufacturing sector$_i$,Percentage employed in the wholesale sector$_i$, Percentage employed in the retail sector $_i$, Percentage employed in the information sector$_i$, Percentage employed in the financial services sector $_i$, Percentage employed in other services$_i$

Local demographic factors$_i$ = f (Percentage age 19 or younger$_i$ Percentage age 20-44$_i$, Percentage age 45-64$_i$, Percentage African American$_i$, Percentage Hispanic$_i$, Percentage Asian$_i$, Percentage foreign born$_i$, Percentage single mothers$_i$.

Sprawl$_i$ = f (Urban Area density$_i$, Land use mix$_i$, Urban centering$_i$, Street connectivity$_i$)

Below I have listed each variable with a brief description where appropriate. In parentheses following each I have included a symbol "+", "-" or "?" to indicate the expected direction of influence on the dependent variable of change in poverty rates.

Local economic factors:

Percentage high school graduate (-)

Percentage who have obtained a bachelors degree or better (-)

Percentage who have annual incomes above $200,000 (-)

Median household income for the central place (-)

Median home value in the central place (-)

Percentage employed in the construction sector (+)

Percentage employed in the manufacturing sector (+)

Percentage employed in the wholesale sector (?)

Percentage employed in the retail sector (?)

Percentage employed in the information sector (-)

Percentage employed in the financial services sector (?)

Percentage employed in other services (?)

Variables among local demographic factors

Percentage age 19 or younger (?)

Percentage age 20-44 (?)

Percentage age 45-64 (older age group dropped for comparison) (?)

Percentage African American (White dropped for comparison) (?)

Percentage Hispanic (?)

Percentage Asian (?)

Percentage foreign born (+)

Percentage female head of household, no husband, with own children under 18 (single mothers) (+).

Variables among development pattern factors:

The following development pattern variables are derived from Census data:

Urban area population density in year 2000: calculated by dividing urbanized area population by the UA land area (-)

The percentage change in urban area population density between 2000 and 2010 (-)

The following development variables are derived from Ewing (2014):

Density factor, an index of factors measuring the overall population density of an urbanized area in the year 2000. Higher numbers are more dense (-)

Mix factor, an index of factor measure the mix of land uses in an urbanized area. Higher numbers have a more diverse mix of uses (-)

Centering factor, an index measuring the extent to which jobs and residents are centered in an urbanized area. Higher numbers are more centered (-)

Street factor, an index measuring street connectivity and block size in an urbanized area. Higher numbers are more well-connected (-)

Composite index, a composite of all the sprawl measures above. Authors define higher numbers as less sprawling (-).

The following growth control measures are adapted from Dawkins and Nelson (2003), and Wassmer (2006). All are dummy variables, designated “0” and “1”:

Presence of statewide growth control measures. (?)

Presence of a strong local policy that accommodates growth. (?)

Presence of strong local policy that restricts growth. (?)

Presence of a weak local policy that accommodates growth. (?)

Presence of a weak local policy that restricts growth in the urbanized area. (?)

Descriptive statistics

As noted above, all data in this study come from the U.S. Census Bureau. Half of the data comes from the 2000 decennial census, and half comes from the American Community Survey, five-year data from 2010. Most variables are measured in percentages. I measured urbanized area density by dividing the total population of the urbanized area by the size of the area in square miles.

Table 3-1. Descriptive statistics for all explanatory variables.

Variable	**Mean**	**Standard deviation**	**Min**	**Max**
Central place poverty change	21.62	28.15	-49.32	188.23
Urban Area density, year 2000	2937.38	1565.07	851.5667	7068.591
High school diploma (pct)	25.99	7.19	6	49.9
Bachelor degree (pct)	16.42	7.22	1.6	41.8
Graduate degree or higher (pct)	9.46	6.32	0.8	48.8
Median household income (dollars)	42534.48	13552.64	17206	100411
Percentage earning above $200,000	2.05	2.04	0.1	18

Poverty rate in 2000	13.81	7.20	1.6	46.9
Unemployment rate	3.98	1.51	.9	11
Median home value (dollars)	141680.9	86070.31	35900	708200
Percentage employed in construction	5.99	1.94	1.1	17.4
Percentage employed in manufacturing	13.58	7.17	1.5	47.3
Percentage employed in wholesale	3.63	1.35	0.4	10.2
Percentage employed in retail	12.25	1.96	5.6	20.8
Percentage employed in information	3.13	1.43	0	14.6
Percentage employed in financial services	6.81	2.43	1.3	24.9
Percentage employed in other services	4.96	0.99	2.1	10.3
Percentage age 19 and below	21.50	3.39	6.2	32.5
Percentage age 20 to 44	38.65	4.99	10.3	64.4
Percentage age 45 to 64	20.39	3.14	6.7	35.1
Percentage African American	13.55	15.79	0.1	92.4
Percentage Asian	4.86	7.59	0.1	61.8
Percentage Pacific Islander	0.18	0.59	0	11.4
Percentage Hispanic	15.58	18.40	0.5	97.7
Percentage foreign born	13.50	12.23	0.3	73.1
Percentage single mothers	8.24	3.11	1.8	30

Urban Area population density	2937.20	1565.99	851.57	7068.59
Percent change in urban area population density	-1.8	10.17	-41.97	81.92
Sprawl index	106.98	26.49	39.5	184.06
Density factor	130.05	43.33	49.14	219.66
Mix factor	113.18	18.78	55.21	162.91
Centering factor	106.71	24.06	50.78	170.57
Street factor	10155	27.64	39.5	189.55
Percent change in sprawl index	1.22	7.92	-27.73	32.33
Percent change in density	-1.99	5.11	-22.25	16.26
Percent change in mix	-5.29	9.69	-46.49	17.97
Percent change in centering	1.62	9.36	-44.42	49.36
Percent change in street	11.45	14.23	-12.13	94.12

Correlation analysis

In Appendix C, I have included a pairwise correlation matrix for all explanatory variables. Correlation analysis allows us test the strength, direction and statistical significance of the relationship between two continuous variables (Rogerson 2010). Pearson's correlation statistic produces an "R" value that ranges from -1 to 1. A value near zero represents no correlation between the variables. A value of 1 indicates a perfect positive or direct correlation, and -1 shows a perfect negative or indirect correlation. One reason for using correlation analysis before regression analysis is to turn up possible problems with multicollinearity among variables. A rule of thumb is that pairwise correlations of .80 or stronger may indicate multicollinearity, which can make it difficult to determine the magnitude of effect of one variable which is strongly correlated with another. (I will return to the subject of multicollinearity in the next chapter.)

Strong correlations in the dataset are found mostly where they are expected. There is a moderate and statistically significant correlation between median income and higher education (.56), and between income and home value (.70). The percentage of single mothers in a place has a moderately strong correlation with the baseline poverty rate in that place (.63). Denser places tend to be correlated with the percentage of foreign-born residents (.70). There is also a weak but statistically significant negative correlation between urban area density and the baseline percentage of people living in poverty. The correlation is -0.16 with a p-value below .01. Again, these correlations may offer some clues to the causes of poverty, or the social effects of urban density. But these processes are complicated, and simple correlation may mask the effects of one variable on another, or suggest a causal relationship where there is none. A regression analysis is needed to tease out the effects of these variables when taken together.

Conclusion

Above I described the broad theoretical model that I propose to explain how sprawl and spatial segregation may affect poverty rates in central places in the United States. Having assembled the data necessary and identified the appropriate dependent and independent variables, I am now prepared to move on to specifying and fitting the model and interpreting my initial regression results.

Chapter 4.
Findings

Explanation of regression models and techniques

This study uses linear regression as the main tool for understanding the possible effect of sprawl on central place poverty. Regression analysis allows us to control for confounding variables such as education level or the local employment base, or the baseline poverty and unemployment rates, which may be contributing to the local poverty rate. By controlling for these variables we are able to determine if sprawl truly accounts for a statistically and substantively significant difference in poverty rates between central places, in the presence of other factors.

Two regression techniques are used: ordinary least squares regression and two-stage least squares regression. In these models, both the dependent variable (change in poverty) and the key explanatory variable (change in sprawl) are measures of change over the same time period; this introduces the problem of simultaneity between the two variables. According to Studenmund (2006), simultaneity violates one of the "classical assumptions" of regression analysis, that all explanatory variables are uncorrelated with the error term because they are determined by factors outside the regression equation. But in this model poverty and sprawl are believed to be interdependent: flight from blight causes sprawl, but sprawl exacerbates poverty in central places. The two are endogenous, and so the coefficients for the change-in-sprawl measures used in this model will be biased. They will not truly measure the effect of changing sprawl upon changing poverty, but instead will reflect some mixture of the effect of sprawl on poverty and vice versa.

The two-stage least squares (2SLS) method can help to account for this endogeneity. Per Studenmund (2006), the 2SLS method works as follows: In the first stage, ordinary least squares (OLS) regressions are run on the reduced-form equations for each endogenous variable in the model. A reduced-form equation is an equation that explains the endogenous variables using only exogenous variables. In the second stage, the fitted values for the endogenous variables are then substituted back into the original structural equation and OLS is run. If the sample size is sufficiently large and the fit of the reduced form equations are good, then 2SLS can help eliminate bias in the coefficients for the endogenous variables.

There is added complexity because I am trying to test different measures of sprawl, suggested in the literature review above. One method is to use the simple population density of each urbanized area, calculated by dividing urbanized area population by the land area measured in square miles. Another method of measuring sprawl is borrowed from Ewing (2013), using Ewing's measures of the four components of sprawl—population and employment density, land use mix, population and job centering and street connectivity—along with a composite sprawl measure that combines all four. (See Chapter 3 for a more detailed explanation of Ewing's sprawl measures.)

Ewing only calculated sprawl index values for the 162 urbanized areas within the large and medium metropolitan statistical areas in the U.S. For this reason, I created two datasets, one which includes 835 central places within all 437 U.S. urbanized areas, and a smaller set of 498 central places within the urbanized areas for which Ewing's sprawl data is available. Regressions using the simple population density measure were run using both the small and large datasets. Regressions using the Ewing measures were performed only with the small dataset. This allows me to compare two different approaches to measuring sprawl using the same dataset, and also to test the population density measure on two slightly different datasets and compare the results.

After performing the two-stage least squares regressions, I moved on to a reduced form of the regression using ordinary least squares. Because of the different techniques and datasets used, I report results from more than a dozen regressions in the sections below. By learning which variables, if any, consistently appear to be statistically significant across these several models, I hope to better understand the relationship between urban form and poverty.

Early tests of these models showed that heteroskedasticity was present in the data (see Appendix D). Heteroskedasticity is the unequal variance in the error term among observations of a dependent variable, and it is common in datasets in which there is a large amount of variation in size (Studenmund 2006). Certainly a comparison of cities and other central places across the U.S. would fall into that category. Heteroskedasticity affects the standard errors of regression coefficients, and can lead to errors in determining the statistical significance of variables. This problem can be addressed here through the use of corrected standard errors that are "robust" to potential bias caused by heteroskedasticity. All of the models in this study are reported using robust standard errors.

Two stage least squares regression results

The two-stage least squares regression technique was calculated first. Sprawl is defined and measured in many different ways. Population density and centrality are major elements of sprawl, but sprawl is often also defined by how these elements or urban form change over time. The two-stage method is used to account for the simultaneity of changing central place poverty and changing urban area decentralization. In this model, I used available data (Dawkins and Nelson 2003, Wassmer 2006, Paulsen 2012) on state and local growth control measures to create instrumental variables—which can serve as proxies for the change in sprawl during the time period 2000-2010, but which should have no direct effect on the change in poverty rate in that time.

Table 4-1 shows the results of regressions testing the effect of changing urban area population density on change poverty rates in central places. Model 1 uses the large dataset; Model 2 uses the smaller dataset. Note that variables not designated with "change" are taken from 2000 Census data. The coefficients for state dummy variables are not reported here.

Table 4-1. Results of two-stage least squares regression, Models 1 and 2 (Dependent variable = Percentage change in poverty rate)

Variable	Model 1 (large dataset)	Model 2 (small dataset)
Constant	14.004 (45.623)	59.981 (61.860)
Change in urban area population density	-0.640 (1.085)	2.599* 1.342
Urban area population density in 2000, (1,000s)	-6.410** (3.179)	-1.582 (2.789)
Population (10,000s)	-0.006 (0.014)	-0.006 (0.013)
Percent African American	-0.039 (0.137)	-0.128 (0.201)
Percent Asian	0.047	0.228

	(0.169)	(0.225)
Percent Hispanic	-0.112 (0.109)	-0.255 (0.180)
Percent single mother	0.412 (0.763)	2.883*** (1.103)
Percent high school graduates	-0.469 (0.464)	-2.205*** (0.832)
Percent bachelor’ degrees and above	0.015 (0.405)	-0.680 (0.487)
Percent Foreign born	0.029 (0.268)	-0.260 (0.279)
Percent unemployed	0.733 0.871	0.268 (1.370)
Percent in poverty	-1.725*** (0.347)	-3.797*** 0.681
Percent Employed in Construction	0.739 (0.922)	1.874** 0.852
Percent Employed in Manufacturing	0.610*** (0.197)	1.002*** 0.329
Percent Employed in Wholesale	1.087 (0.923)	0.7293934 (1.368)
Percent Employed in Retail	0.742 (0.726)	0.479889 (1.272)
Percent Employed in Information	1.135 (0.899)	1.863425 (1.278)
Percent Employed in Finance	0.705 (0.558)	1.461961 (0.846)
Percent Employed in Other Services	-0.150 (1.201)	1.360388 (1.717)
Percent Household Income Above $200,000	-1.474 (1.117)	-0.574 (1.495)
Median household income ($1,000s)	-0.244 0.320)	-0.962** (0.407)
Median home value ($1,000s)	0.035 (0.028)	0.005 (0.041)
Percent Age 0 to 19	0.145 (0.485)	-0.043 (0.690)
Percent Age 20 to 44	0.359 (0.532)	0.374 (0.515)
Percent Age 45 to 64	0.275 (1.131)	1.393 (1.086)
R^2 **Number of Observations**	0.48 834	0.48 497

Each cell contains the regression coefficient above and the standard error of regression coefficient below in parentheses. “*” indicates p value .10 or below. “**” indicates p value .05 or below. “*” indicates p value .01 or below.

In this model there is no statistically significant relationship in the large dataset between change in urban area population density and change in poverty rates in a central place. There is however a statistically significant relationship in the large dataset between urbanized area population density in the year 2000 and the change in poverty rate over the following decade. This statistic shows that as we move from an urbanized area with a population density of 2000 people per square mile to a population density of $3000/mi^2$, we would expect to see a 6.4 percent decrease in the poverty rate, all other factors being held equal. The smaller dataset shows a statistically significant relationship between change in density and poverty rates, but no relationship with baseline density. In this case, the coefficient has an unexpected sign. As urban area density increases by one percent, the poverty rate changes by approximately 2.6 percent. The p-value for baseline urban area population density was .04 in the large dataset and .57 in that small dataset. Critical values of .01, .05 and .10 are commonly used to show levels of statistical significance, with values .01 or below indicating stronger statistical significance. In this study, statistical significance below .01, .05 and .10 are shown with ***, ** and *, respectively.

Model 2 shows a strong statistically significant negative relationship between the percentage of residents with a high school diploma and the change in poverty rate within a central place. For each one percent increase in the portion of the population with a high school diploma, we expect to see the poverty rate decrease by more than two percent, accounting for other factors. Conversely, the percentage of single mothers has a positive relationship with poverty rates—a single percentage point change in single mothers will boost poverty rates by nearly three percent. There is also a negative relationship between the baseline poverty rate and the change in the poverty rate. This is consistently true in later regressions and may be explained by reversion to the mean. And finally, the percentage of residents employed in construction and manufacturing has a direct and statistically significant relationship with the change in poverty rates.

The lack of statistical significance in the key explanatory variable may be because there is truly no relationship, but it may also be because the model is incorrectly specified. One specification problem could be with weak instrumental variables. The simplest test for weak instruments is to look at the pairwise correlations between the endogenous variable and the chosen instruments. Appendix C includes these pairwise correlation tables. None of the sprawl measures tested had a correlation of greater than 0.23 with any of the chosen instruments, thus indicating these instruments may be weak. Similarly, statistics generated from the first stage of each of the two-stage regressions showed low partial R^2 scores in both models: .0086 and .036, respectively. Likewise, F-statistics related to the instrumental variables were fairly low, 1.6 and 2.6 respectively. According to Cameron and Trivedi (2010), scores of 10 or above are the rule of thumb for determining strong instruments. These statistics further suggest the growth measures included in the model are weak instruments, and are not good proxies for change in urban area density. The weakness of the instruments may partly account for the lack of statistical significance in my key explanatory variables.

Two stage least squares with Ewing sprawl measures

I next moved on to a set of models using Ewing's sprawl measures as key explanatory variables, beginning with Ewing's composite sprawl measure, which was found to be statistically significant. In this index, higher scores indicate less-sprawling urban areas. As shown in Table 4.2, Model 3 suggests that for every one percent increase in an urban area's sprawl index score, we would expect to see about a 1.3 percent decrease in the poverty rate for central places within that urban area, all else being equal. (The median poverty rate for central places in the U.S. was 13.3 percent in 2010. So a 10 percent increase in the sprawl index would be expected to lower average central place poverty from 13.3 percent to about 11.5 percent.) But using a composite sprawl score like this obscures which aspects of sprawl might be correlated with higher or lower poverty rates. It could be density, as in the previous model. Or perhaps job centrality, or the mix of land uses. So I next tested the change in each the four parts of Ewing's sprawl index separately—development density, land use mix, centering of population and jobs, and street accessibility--to find what components of sprawl might be affecting poverty rates. The results are shown in Table 4-2.

Table 4-2. Two stage least squares regressions using Ewing sprawl measures.

Variable	**Model 3 Composite**	**Model 4 Density**	**Model 5 Land use m**	**Model 6 Centering**	**Model 7 Streets**
Constant	49.499 (54.289)	53.064 (53.372)	81.615 (73.003)	181.977 (118.783)	69.739 (52.986)
Sprawl change	-1.218 ** (0.564)				
Density change		-0.854 (0.638)			

Land use mix change			-2.609** (1.270)		
Centering change				2.310 (1.864)	
Streets change					-0.438** (0.242)
2000 UA pop. density/1000	-5.897*** (1.289)	-7.100 *** (1.419)	2.685 (4.919)	-6.80*** (1.588)	-8.04*** (1.646)
2000 Population	-0.005 (0.012)	0.001 (0 .012)	-0.011 (0.017)	0.012 (0.018)	-0.004 (0.012)
Percent African American	0.087 (0.184)	-0.009 (0.178)	0.191 (0.242)	-0.142 (0.221)	0.001 (0.185)
Percent Asian	0.184 (0.207)	0.206 (0.195)	0.175 (0.254)	0.423 (0.290)	0.230 (0.202)
Percent Hispanic/Latino	-0.164 (0.136)	-0.136 (0.126)	-0.243 (0.166)	-0.190 0.158)	-0.162 0.133
Single mother	2.443** (1.015)	1.984** (0.896)	1.725 (1.305)	1.010 (1.497)	2.607*** (0.989)
High school graduate	-1.734*** (0.615)	-1.617 ** (0.641)	-2.694*** (0.901)	-3.336** (1.624)	-1.869 *** (0.649)
Bachelors or above	-0.360 (0.398)	-0.492 (0.407)	-0.785 (0.508)	-1.503 (0.943)	-0.468 (0.395)
Foreign born	-0.033 (0.226)	-0.077 (0.221)	-0.261 (0.291)	-0.704 (0.559)	-0.095 (0.222)
Percent unemployed	1.435 (1.236)	1.751 (1.188)	1.069 (1.483)	2.216 (1.713)	2.064 (1.285)
Percent in poverty	-3.507*** (0.573)	-3.478*** (0.531)	-4.114*** (0.691)	-4.194*** (0.803)	-3.621*** (0.559)
Percent Employed in Construction	1.883** (0.929)	1.568** (0.774)	1.013 (1.241)	1.183 (0.987)	2.477*** (0.933)
Percent Employed in Manufacturing	0.902*** (0.276)	0.620** (0.252)	1.337** (0.531)	.874** (0.378)	.842*** (0.262)
Percent Employed in Wholesale	1.788* (1.055)	1.378 (1.024)	2.276 (1.445)	1.106 (1.237)	2.023* (1.133)
Percent Employed in Retail	1.585* (0.939)	1.265 (1.005)	1.189 (1.143)	0.680 (1.336)	1.268 (0.978)
Percent Employed in Information	1.299 (1.075)	1.298 (1.082)	0.698 (1.132)	-0.277 (1.626)	1.159 (1.056)
Percent Employed in Finance	-0.196 (0.684)	0.110 (0.657)	-0.140 (0.911)	1.2312 (0.924)	0.558 (0.630)
Percent Employed in Other Services	2.515 (1.610)	1.823 (1.469)	1.704 (1.828)	-1.800 (3.264)	1.110 (1.463)
Above $200,000	-0.308 (1.432)	-0.561 (1.411)	0.339 (1.550)	-0.017 (1.532)	-0.378 (1.401)
Median Income	-0.610 (0.442)	-0.627 (0.395)	-1.193** (0.521)	-1.457** (0.724)	-0.713* (0.413)
Median Home Value	0.039	0.041	0.035	0.057	0.046

	(0.034)	(0.035)	(0.040)	(0.045)	(0.035)
Age 0 to 19	-0.370 (0.700)	0.001 (0.641)	-0.052 (0.827)	0.370 (0.780)	-0.383 (0.688)
Age 20 to 44	0.172 (0.466)	0.290 (0.469)	-0.069 (0.604)	0.195 (0.576)	0.261 (0.448)
Age 45 to 64	0.772 (0.920)	0.842 (0.976)	1.424 (1.140)	1.884 (1.467)	0.802 (0.961)
R^2 N	.58 497	.59 497	.41 497	.40 497	.58 497

Each cell contains the regression coefficient above and the standard error of regression coefficient below in parentheses. "*" indicates p value .10 or below. "**" indicates p value .05 or below. "*" indicates p value .01 or below.

In Model 4, change in Ewing's density measure is not shown to be statistically significant, but in models 5 and 7, land use mix and street accessibility are statistically significant. In models 3, 4, 5 and 6, baseline urban area population density is shown to be statistically significant. And the effect is substantial. When urban area density increases by 1000 people, we expect to see a 6 to 8 percent reduction in the poverty rate, all other factors held equal. And as with Model 2, high school education, baseline poverty, construction and especially manufacturing employment also have a significant effect on poverty rates.

I again tested the instruments for Models 3 through 7. For some measures, the growth control instruments are stronger. Table 4-3 shows partial R-squared and F-statistics for the instrumental variables in each of those models.

Table 4-3. Partial R^2 and F statistics for instrumental variables in models 3-7.

Model	**Partial R^2**	**F-statistic**
3 (Composite)	.095	6.65
4 (Density)	.122	7.23

5 (Land use mix)	.031	2.026
6 (Centering)	.009	.736
7 (Street accessibility)	.133	15.3

These regressions suggest there is some evidence of a relationship between the change in sprawl and change in central place poverty rates. And there may be a difference in the strength of the Ewing sprawl measures compared to the simpler measure of population density. Diagnostic statistics still suggest the instrumental variables used in these regressions are weak. However, there are solid theoretical grounds to believe that sprawl is having an effect on poverty rates in these communities. These models consistently show that greater density is associated with lower poverty. The theory suggests there should be a relationship between change in density and poverty rates, and there is not much evidence of that in the regressions above. There is some evidence that communities which become less sprawling in terms of land use mix and street connectivity also saw lower increases in poverty. The concern about the weak instruments led me to simplify my approach.

Ordinary least squares regression results

I next moved on to a reduced-form equation, using the ordinary least squares method. Instead of using the *change* in sprawl as a key explanatory variable, I used measures of the 2000 baseline level of sprawl and density in each urban area and tested their effects on the change in central place poverty over the decade 2000 to 2010. As explained above, a reduced-form equation explains an endogenous variable from the structural equation using only exogenous variables. In this case the variables are exogenous because poverty in 2010 cannot possibly cause sprawl in 2000.

In half of the models, state and local growth measures from Dawkins and Nelson, described in Chapter 3, were also included as explanatory variables. The results of these regressions are shown in Table 4-4, and Table 4-5. The first set of regressions (Table 4-4) test the population density variable, in both the small and large datasets. The second set (Table 4-5) test the Ewing sprawl measures and use the smaller dataset of 497 central places.

In the first set of OLS regressions, the models consistently show that urban area population density has a strong indirect (negative) relationship with change in central place poverty between 2000 and 2010. And increase of urban area density of 1000 people per square mile was associated with a -4.5 to -6.3 change in poverty rate within a central place.

Table 4-4. OLS regressions using urban area population density as key explanatory variable

Variable	**Model 8 (Large dataset)**	**Model 9 (Small dataset)**	**Model 10 (Large, growth measures omitted)**	**Model 11 (Small, growth measures omitted)**
Constant	3.986 (36.977)	57.208 (58.597)	-2.583 (35.904)	54.567 (56.995)
Urban area population density in 2000 (by 1000)	-4.571*** (1.123)	-5.852*** (1.420)	-4.633*** (1.112)	-6.392*** (1.392)
Population (10,000)	-0.012 (0.011)	-0.001 (0.012)	-0.013 (0.010)	-0.0001 (0.011)
Percent African American	-0.009	-0.022	-0.006	-0.001

	(0.131)	(0.186)	(0.133)	(0.191)
Percent Asian	-0.026 (0.177)	0.090 (0.217)	0.045 (0.176)	0.241 (0.212)
Percent Hispanic	-0.113 (0.109)	-0.204 (0.133)	-0.099 (0.111)	-0.132 (0.134)
Single mother	.030 (0.682)	1.673* (0.960)	0.196 (0.694)	2.091** (0.961)
High school graduates	-0.690 (0.428)	-1.793*** (0.664)	-0.555 (0.430)	-1.669 ** (0.690)
Bachelor degrees and above	-0.199 (0.298)	-0.613 (0.430)	-0.125 (0.298)	-0.508 (0.434)
Foreign born	-0.084 0.206	-0.128 0.237	-0.064 0.206	-0.125 0.238
Percent unemployed	0.698 (0.896)	1.107 (1.189)	0.770 (0.859)	1.295 (1.171)
Percent in poverty	-1.784*** (0.379)	-3.600*** (0.578)	-1.707*** (0.373)	-3.535*** (0.582)
Percent Employed in Construction	0.256 (0.670)	1.201 (0.799)	0.384 (0.702)	1.669** (0.824)
Percent Employed in Manufacturing	0.636*** (0.180)	0.602** (0.277)	0.658*** (0.180)	0.666** (0.270)
Percent Employed in Wholesale	1.134 (0.866)	2.059* (1.153)	0.828 (0.852)	1.283 (1.128)
Percent Employed in Retail	0.604 (0.669)	0.994 (1.082)	0.607 (0.674)	1.142 (1.069)
Percent Employed in Information	0.904 (0.907)	1.154 (1.142)	1.097 (0.931)	1.245 (1.161)
Percent Employed in Finance	0.956** (0.474)	0.709 (0.657)	0.880* (0.456)	0.393 (0.642)
Percent Employed in Other services	-0.727 (1.220)	1.243 (1.663)	-0.454 (1.198)	1.642 (1.564)
Above $200,000	-0.958 (1.155)	-0.070 (1.469)	-1.329 (1.140)	-0.520 (1.506)
Median income ($1000)	-0.368 (0.324)	-0.932** (0.415)	-0.261 (0.328)	-0.717 (0.438)
Median value ($1000)	0.019 (0.031)	0.022 (0.038)	0.032 (0.029)	0.040 (0.038)
Age 0 to 19	0.384 (0.463)	0.430 (0.674)	0.238 (0.462)	-0.015 (0.681)
Age 20 to 44	0.611* (0.327)	0.383 (0.497)	0.579* (0.326)	0.319 (0.501)
Age 45 to 64	0.892 (0.740)	1.242 (1.052)	0.749 (0.736)	0.981 (1.063)
Statewide growth control	-13.714* (8.322)	-11.265 (10.999)		
Strong accommodating growth	0.989 (2.748)	2.231 (4.313)		

Strong restricting growth	11.801*** (4.430)	16.850*** (6.460)		
Weak accommodating growth	2.607 (3.223)	9.825** (4.268)		
Weak restricting growth	4.305 (3.948)	0.173 (6.780)		
R^2 N	.51 834	.61 497	.50 834	.60 497

Each cell contains the regression coefficient above and the standard error of regression coefficient below in parentheses. "*" indicates p value .10 or below. "**" indicates p value .05 or below. "*" indicates p value .01 or below.

Table 4-5. OLS regressions using Ewing sprawl measures as key explanatory variable.

Variable	Model 12 (Composite measure)	Model 13 (All sprawl measures)	Model 14 (Composite, growth controls omitted)	Model 15 (All sprawl measures, growth controls omitted)
Constant	69.152 (61.908)	54.844 (63.093)	62.074 (59.313)	47.684 60.087
Composite index	-0.164** (0.069)		-0.078 (0.064)	
Density		-0.260*** (.077)		-0.300*** (.072)
Land use mix		0.025 (0.096)		0.058 (0.097)
Urban centering		0.057 (0.089)		0.152** (0.077)
Street accessibility		-0.002 (0.080)		0.012 (0.076)
Population	-0.005 (0.012)	0.0001 (0.012)	-0.006 (0.011)	-0.001 (0.011)
Percent African American	-0.172 (0.176)	0.005 (0.190)	-0.186 (0.185)	0.031 (0.195)
Percent Asian	0.058 (0.217)	0.100 (0.220)	0.189 (0.217)	0.261 (0.215)
Percent Hispanic	-.278** (0.135)	-0.183 (0.138)	-0.241* (0.139)	-0.111 (0.139)
Single mother	2.065** (0.968)	1.531 (0.986)	2.672*** (0.965)	1.787* (0.984)
High school graduates	-1.998*** (0.689)	-1.714*** (0.664)	-1.864*** (0.716)	-1.656** (0.674)
Bachelor degrees and above	-0.647 (0.456)	-0.580 (0.439)	-0.485 (0.458)	-0.485 (0 .432)
Foreign born	-0.252 (0.239)	-0.142 (0.237)	-0.222 (0.245)	-0.162 (0.237)

Percent unemployed	0.853 (1.202)	1.040 (1.218)	0.971 (1.167)	1.082 (1.216)
Percent in poverty	-3.543*** (0.576)	-3.626*** (0.581)	-3.321*** (0.552)	-3.451*** (0.570)
Percent Employed in Construction	1.111 (0.838)	1.145 (0.818)	1.827** (0.862)	1.643* (0.859)
Percent Employed in Manufacturing	0.419 (0.286)	0.589** (0.293)	0.530* (0.275)	0.732*** (0.280)
Percent Employed in Wholesale	1.136 (1.161)	2.108* (1.174)	-.311 (1.118)	1.013 (1.124)
Percent Employed in Retail	1.481 (1.119)	0.910 (1.061)	1.776 (1.113)	1.301 (1.036)
Percent Employed in Information	0.886 (1.200)	1.277 (1.153)	1.081 (1.263)	1.470 (1.169)
Percent Employed in Finance	0.643 (0.669)	0.762 (0.667)	0.174 (0.642)	0.355 (0.656)
Percent Employed in Other services	0.337 (1.662)	1.438 (1.688)	0.341 (1.587)	1.712 (1.591)
Above $200,000	0.1399 (1.451)	-0.145 (1.471)	-0.286 (1.487)	-0.446 (1.501)
Median income	-1.053*** (0.403)	-0.956** (0.422)	-0.737* (0.423)	-0.733 (0.445)
Median home value	0.008 (0.037)	0.024 (0.038)	0.017 (0.037)	0.041 (0.038)
Age 0 to 19	0.598 (0.655)	.0548 (0.695)	-0.005 (0.659)	0.091 (0.696)
Age 20 to 44	0.287 (0.493)	0.407 (0.508)	0.139 (0.482)	0.267 (0.495)
Age 45 to 64	1.267 (1.060)	1.307 (1.056)	0.7642 (1.067)	0.875 (1.045)
Statewide growth control	-8.852 (13.018)	-12.300 (11.075)		
Strong accommodating growth	4.624 (4.207)	0.699 (4.897)		
Strong restricting growth	22.661*** (6.839)	20.809*** (6.722)		
Weak accommodating growth	12.366*** (4.449)	11.011*** (4.239)		
Weak restricting growth	-3.198 (6.589)	-1.927 (6.630)		
R^2 N	.60 497	.61 497	.57 497	.51 497

Each cell contains the regression coefficient above and the standard error of regression coefficient below in parentheses. “*” indicates p value .10 or below. “**” indicates p value .05 or below. “*” indicates p value .01 or below.

Models which used the Ewing sprawl measures also showed a statistically significant relationship—less sprawling areas saw smaller increases in central place poverty. But again, density appears to be the driving causal factor, and factors of land use mix, centrality and street connectivity had no statistically significant relationship. The exception is Model 15 in which urban centering appears to have a small statistically significant and positive effect on poverty change. This result first seems counter to the hypothesis, more centered communities are less sprawling, but in this regard are experiencing slightly higher increases in poverty. But the coefficient may also be capturing some underlying difference between more monocentric and more polycentric communities, such as age of the community or differences in the economic base.

Finally, these regressions showed evidence of a relationship between certain state and local policies on urban growth and change in poverty rates. Specifically, in the presence of statewide growth controls appear to have a large negative effect on the change in central place poverty (as much as 15 percent) while certain kinds of local growth controls may actually have a positive effect on poverty rates. The coefficient for statewide growth control could be capturing some underlying difference between states that actually has little to do with growth policy. But those differences should be largely accounted for by the inclusion of state dummy variables. So this may be evidence that states that aggressively manage growth at the start of a decade are seeing economic benefits in lower poverty rates within central places throughout the next decade. On the other hand, in Models 12 and 13, local growth controls appear to have a strong upward effect on poverty rates. In areas with “strong restricting” growth controls, we expect to see increases in poverty rates of anywhere from 12 to 23 percent, all else held equal. In "weak accommodating" jurisdictions, the effect was 10 to 12 percent. That is a very large effect, and begs for an explanation. Perhaps this is because the "strong restricting" growth controls are reducing opportunities for affordable housing in one area, leading to greater spatial mismatch between low skilled workers and new jobs areas—thus further concentrating poverty. On the other hand, "weak accommodating" growth controls could be contributing to a greater supply of low income housing, thus attracting more low income residents. Overall caution should be used in interpreting these results because these growth measures were not the intended subject of this analysis, and because the effect may be due to other variables that are not included in the models. But it may be evidence that growth controls can have unintended consequences if they are not carefully crafted.

As mentioned above**,** baseline poverty continues to have an indirect relationship with poverty change— the higher baseline poverty rate, the smaller the increase tends to be. Manufacturing and Construction seem to be positively associated with poverty change. Consistently, across almost all of the models, a percentage point increase in manufacturing employment is associated with a .6 to 1 percent increase in poverty rates. This may reflect a long-term trend of overall decline in quality manufacturing jobs in American cities. The construction effect may also be related to a decline in new home construction during the recession in that decade. The median percentage in manufacturing employment was 12.5 in the year 2000. The highest that year was Dalton, Georgia, which has long billed itself as, "The Carpet Capital of the World." In 2000, 47.3 percent of its jobs were in manufacturing. By 2010, that had dropped to 40 percent, and poverty had increased by more than 50 percent. The loss of American manufacturing jobs accelerated in the 2000-2010 decade. Dalton was hit especially hard because carpet and floor coverings are so closely tied to the boom and bust of the U.S. housing market (Severson 2012).

Multicollinearity and corrections

Multicollinearity presents another potential problem in regression analysis. Multicollinearity is often present when two or more explanatory variables are highly correlated with each other. The principal consequence of multicollinearity is to increase the standard errors of the coefficients in a regression equation, leading to t-statistics (and calculated likelihoods of statistical significance) that are unreliable. As mentioned in Chapter 3, correlation analysis showed no correlations among the explanatory variables above 0.80, and thus is one bit of evidence that multicollinearity may not be a problem. However, more sophisticated statistical tests are needed to be certain about presence or absence of multicollinearity. Appendix E shows the “variance inflation factor” scores for each of the OLS models above. The rule of thumb is that a VIF score of 5 or higher indicates a high degree of multicollinearity. And the VIF scores do indicate a high degree of multicollinearity in each of the models. But it is important to recognize that this is only a problem if the regression coefficient is determined to be statistically insignificant.

For example, bachelor degrees and high school degrees seem to be highly collinear, and regression coefficients on bachelor's degree never statistically significant. The statewide growth control measure also shows a high VIF score, as does the possibly related dummy variable for the state of Florida. One solution for multicollinearity is to remove variables with high VIF scores. However, omitting variables presents other problems, such as omitted variable bias. Another possible solution for multicollinearity is to transform collinear explanatory variables, often by combining one or more variables together. For example, age groups or education levels can be combined.

By and large these high VIF scores do not affect the variables that I am most interested in, specifically the key explanatory variables. And, mulitcollinearity is never considered to be a problem if the variables of interest remain statistically significant. In Table 4-6, I report results of regressions with corrections for multicollinearity. This included transforming the age variables, and substituting a simplified education variable showing percentage of all residents with a high school diploma or better. I also dropped the variables for percentage of the population that is foreign born (collinear with Latino) and for percentage of population earning above $200,000 (collinear with median income). The collinearity between the Florida dummy variable and the statewide growth control measures was more problematic. I could drop the Florida variable and substitute in a dummy variable for California, which had been the single state variable omitted earlier for comparison. However, the California variable and the statewide growth control measure are highly collinear as well. In the end, I decided to leave the Florida dummy variable in the equation, understanding that it may be causing the lack of statistical significance of the statewide growth control variable.

As shown in Table 4-6, these corrections had little overall effect on the statistical significance of the key explanatory variables compared to the earlier models. The Appendix E also contains tables with VIF scores for Models 16, 17 and 18.

Table 4-6. Model 16, 17 and 18 regressions corrected for multicollinearity

Variable	Model 16	Model 17	Model 18
Constant	-8.363 (31.964)	58.027 (47.160)	67.536 (51.627)
Urban area population density, by 1000	-4.777*** (1.075)	-6.619*** (1.353)	
Density factor			-0.282*** (0.077)
Land use mix			0.049 (0.092)
Urban centering factor			0.057 (0.097)
Street factor			-0.024 (0.078)
Population (10,000)	-0.013 (0.011)	-0.001 (0.013)	0.0009 (0.013)
Percent African American	-0.006 (0.137)	-0.023 (0.185)	-0.008 (0.183)
Percent Asian	-0.075 (0.138)	-0.017 (0.164)	-0.011 (0.166)
Percent Hispanic/Latino	-0.078 (0.108)	-0.228* (0.127)	-0.230* (0.123)
Single mother	0.316 (0.586)	1.703** (0.799)	1.742** (0.802)
Percent high school graduates (or better)	-0.080 (0.240)	-0.560* (0.314)	-0.610* (0 328)
Percent unemployed	0.543 (0.864)	0.766 (1.194)	0.650 (1.213)
Percent in poverty in 2000	-1.666*** (0.309)	-3.255*** (0.503)	-3.345*** (0.520)
Percent employed in construction	0.120 (0.644)	0.226 (0.786)	0.154 (0.796)
Percent employed in manufacturing	0.572*** 0.188	0.249 0.287	0.182 (0.304)
Percent employed in Wholesale	1.246 (0.850)	2.075* (1.135)	2.123* (1.142)
Percent employed in retail	0.520	0.562	0.493

	(0.698)	(1.169)	(1.137)
Percent employed in information	0.780 (0.942)	0.676 (1.234)	0.774 (1.235)
Percent employed in finance	0.780* (0.458)	0.656 (0.699)	0.627 (0.719)
Percent employed in other services	-0.823 (1.126)	1.240 (1.510)	1.383 (1.539)
Median income ($1000)	-0.211 (0.213)	-0.450* (0.238)	-0.449* (0.240)
Median home value ($1000)	-0.008 (0.026)	0.012 (0.034)	0.011 (0.035)
Percent who are working age	0.811*** (0.300)	0.649 (0.462)	0.655 (0.475)
Statewide growth control	-13.912 (8.477)	-14.137 (10.967)	-16.227 (11.141)
Strong accommodating growth	0.587 (2.670)	0.939 (4.240)	-0.225 (4.880)
Strong restricting growth	11.484*** (4.294)	14.511** (6.228)	18.805*** (6.482)
Weak accommodating growth	3.016 (3.379)	11.431 ** (4.927)	12.696*** (4.850)
Weak restricting growth	3.993 (3.905)	.879 (6.517)	-16.227 (11.141)
R^2 N	.51 834	.61 497	.61 497

Each cell contains the regression coefficient above and the standard error of regression coefficient below in parentheses. "*" indicates p value .10 or below. "**" indicates p value .05 or below. "*" indicates p value .01 or below.

Conclusion

This density effect found above appears to be consistent with what the literature suggests: more sprawling areas reduce economic opportunity, through segregation of the poor from jobs, social capital, quality schools and other amenities.

It is important to keep the magnitude of this effect in perspective. Fairly large differences in urban density may account for less than a percentage point difference in the poverty rate between two places. And when compared to other factors in the model, the population density effect is relatively large. In 2010, the median poverty rate for central places, according to the U.S. Census was 13.3 percent. The median population density for urbanized areas was 2506 people per square mile. Moving from a low density urban area like Atlanta, with a population of 1783/mi^2 to the density of Sacramento County (3776/mi^2), or going from Sacramento to Los Angeles (7068/mi^2), we would expect to see poverty rates move down as much as 1 to 1.5 percentage points—all else being held equal. Over a large urban population, a statistically significant relationship between sprawl and poverty could mean real differences in the quality of life for thousands of people.

But these regression results should be looked at with care. Even with the large number of explanatory variables included in each model, each falls far short of fully explaining the variation in the change in central place poverty. And while I have tried to account for confounding variables –such as land values (proxied by home prices) or education levels, the mix of industries in each central place—there are other factors not included in the model.

Additionally, there may be a causal relationship between some variables and only correlations between others. For example, population density may have a causal relationship with poverty, while urban centering is only correlated with poverty rates.

Finally, caution should be used in interpreting the meaning of the regression results related to the growth management measures used. This study was not designed to test the effect of such growth measures on poverty. But it is clear that there is a relationship between certain kinds of local growth controls and poverty rates, even if the exact mechanism is not clear. In the concluding chapter I will discuss the results this study in the context of the larger literature on urban form and income inequality, and make some suggestions for further research.

Chapter 5.

Conclusions, implications and future research

Place and policy

The results of this study suggest that population density has an effect on poverty rates similar in magnitude to other factors such as education, industrial base and portion of single mothers. Governments explicitly make policies regarding education or jobs, with the goal of reducing poverty. Why do we not make policies about urban density with the same goal in mind?

If one already believes that we should be curbing suburban development and directing more development into downtown, then this study adds supporting evidence. At the same time, and as shown in Chapter 1, people like living in suburbs, and there are plenty of academics and policy makers ready to fight against any proposals for “forced densification.” And in fact, there is evidence in this study that policies aimed at slowing suburban growth may actually cause poverty to worsen. However, policy makers should recognize that there is a relationship between urban form and poverty. Understanding the causal mechanisms underlying that relationship can only help to make better policy.

For example, if there is empirical evidence that low-density development drives up poverty rates because it undermines transit networks, then jurisdictions can make appropriate choices: promote higher densities to support transit, or subsidize automobile access for low-income workers (Grengs 2012) or pursue some other solution.

Likewise, if sprawl is affecting educational equity—through income segregation, disinvestment in urban school districts and negative "peer effects" from concentrated poverty-- then that presents difficult policy questions too. How can we provide students living in low-income zip codes with quality schools, without limiting a more affluent family's right to sort themselves into school districts with lower poverty rates and higher test scores? Understanding all the causal pathways through which urban population density might be affecting poverty rates means trying to understand the connections between place and several other policy areas, such as transportation, education, and economic development.

If urban population density is negatively associated with poverty, that would seem to support the argument for strong limits on outward growth. However, the results in Chapter 4 show evidence that some local growth measures actually may exacerbate poverty rates. Recall that in urban areas where "weak accommodating" growth controls were in place, we expect to see 10 to 12 percent increases in the poverty rate, accounting for other factors. In jurisdictions with "strong restricting" growth controls, the effect was even larger, increasing poverty rates 12 to 23 percent. Growth policies are complex, and in Chapter 4 I offered the possibility that growth controls may limit the production of affordable housing, or conversely they may promote the production of affordable housing. Either might plausibly explain a change in poverty rates. Growth controls are adopted for many different purposes—to protect property values, to preserve natural habitat, reduce greenhouse gas emissions from automobiles and to save money on infrastructure, among other reasons. Often they may not be intended to address income inequality at all, but simply to preserve the character, and property values, of a community. Still, Addison (2012) reviewed several studies that found on balance growth controls do reduce the availability of affordable housing to low and moderate income families, because they limit the amount of available land and increase costs. The findings in this study are an additional reminder that social outcomes are tangled up together, regardless of what a particular policy is intended to do. All the more reason to try and better understand how place affects poverty.

Review of the purpose and findings of this study

In Chapter 1, I showed that recent research on economic mobility and sprawl had generated considerable interest among academics, and in the popular media. And I showed that research was built on a foundation, built over many years, of economic theory on spatial mismatch, neighborhood choice and economic segregation.

In Chapter 2 of this study, I described the theoretical literature on economic segregation and spatial mismatch, which guided me in crafting my own hypothesis about the effects of sprawl on poverty rates. That literature review also gave me an understanding of the methods used to by other researchers to define sprawl, to measure sprawl and measure the effects of sprawl on social outcomes. This literature review also showed me where there are gaps in the empirical studies of urban form and economic inequality, and where I might contribute something new.

In Chapter 3, I presented my hypothesis—that sprawl affects central place poverty rates—along with a broad theoretical model supporting that hypothesis. And I presented the data I would use to test the model, and included descriptive statistics and correlation analysis to provide a context for the regression analysis in the next chapter.

In Chapter 4 I specified the regression model, fitted the model, and reported the initial results. I consider the initial two-stage least squares results only partly successful. Change in population density was not statistically significant as a factor in changing poverty rates, or it appeared to have a positive relationship—which was counter to my hypothesis. These results may be partly due to weak instrumental variables.

However, both the two-stage and ordinary least squares regression results consistently showed a statistically significant and substantial relationship between urban area population density in 2000 and change in poverty rates from 2000 to 2010, holding all other factors constant. The density effect is there even when we account for a wide range of potentially confounding variables, such as education levels, demographic mix, and differences in the economic base between places.

This finding adds evidence in support of theories of spatial mismatch and the role of agglomeration in creating economic opportunity (O'Sullivan 2009, Glaeser 2011). We must assume that density is working on poverty rates through any number of intermediary processes. Perhaps lower density communities are not generating the kind of entrepreneurial activity that reduces poverty. Economies of agglomeration may be playing a role. Agglomeration economies can increase productivity when firms cluster together—through the sharing of labor, information and production inputs (O'Sullivan 2006). Fallah (2011) found that U.S. metro areas with higher levels of sprawl had lower levels of labor productivity, an effect the authors credit to greater benefits of agglomeration for dense communities. Agglomeration is a plausible explanation for the density effect seen in this study. However, large cities are commonly what we think of when we think of agglomeration. But the regression model in this study takes population into account—meaning that density has a negative effect on poverty rates regardless of city size. This makes me doubt that agglomeration explains much of the density effect.

What about spatial mismatch? Fan's (2012) review of the literature on spatial mismatch identifies several causal mechanisms identified by researchers. "Modal mismatch," separates carless low-income workers from jobs because of auto-dominated land-use patterns. At the same time, lower density communities may not support an effective public transit system (O'Sullivan 2009) as an alternative to driving. The theory of "skill mismatch" holds that even if workers can get to job sites, employment is functionally inaccessible because of poor education. In this scenario, low density may not be causing skill mismatch, but there may be a correlation with land use patterns that cause greater income segregation between urban and suburban school districts. The density effect shown in this study is consistent with either of these types of spatial mismatch. However, because the effect is happening over a short time—densities in 2000 are having an effect on 2010 poverty rates—I suspect the spatial separation of the poor from jobs is the more likely explanation.

What this study does not show

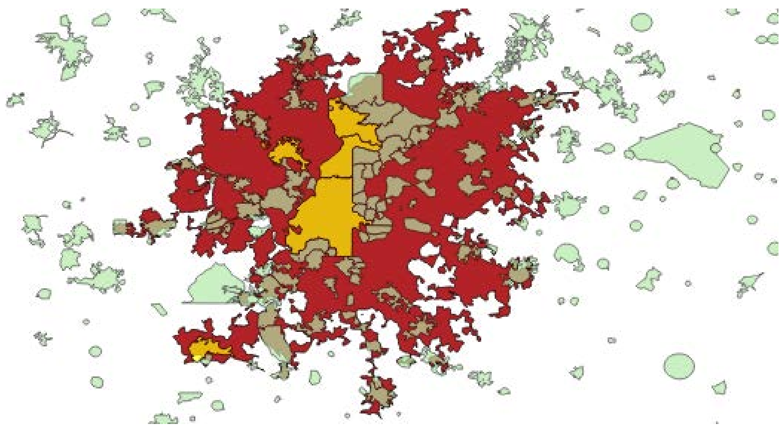
As discussed above, there are theoretical explanations for why density would have an effect on poverty rates, including the theory of spatial mismatch and urban agglomeration. But whatever the causal mechanisms at work, this study does not explicitly say. A logical next step would be to begin test those causal explanations. For example, one approach might be to compare density, transit share and change in poverty rates. It would be simple to repeat this study with the addition of data on transit share included in the model as an explanatory variable.

To test whether educational disparities are part of the mismatch, another approach would be to compare the portion of schools within a central place that fail to make "adequate yearly progress" according to federal education standards, and compare that measure to density and to poverty rates.

Also, because spatial mismatch may be playing a part, it makes sense to measure spatial mismatch. An approach would be to measure the location of jobs and compare this to location of low-income residents, and generate an index of spatial mismatch. This could then be compared to poverty rates or to the change in poverty rates.

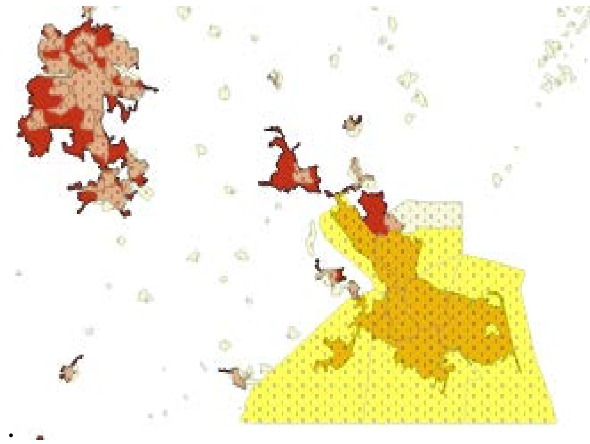
A potential flaw in the model has to do with the unit of analysis. This model assumes that urban areas look something like a fried egg—or perhaps multiple eggs in a pan. The dense central place are the yolks, surrounded by the smooth suburban white around the outside. In fact, many urbanized areas are like this, with the central place nested neatly inside the larger urbanized area. But in many cases the reality is messier. Sometimes the census place spills outside the boundary of the urbanized area, and sometimes the land area of the central place is actually larger than its associated urbanized area. Sometimes central places fall within two urbanized areas. Figure 5-1 shows the Atlanta urbanized area, the large dark area that looks something like a snowflake, and the central places are the lighter polygons inside the UA. This is the classic pattern.

Figure 5-1. Atlanta urbanized area and central places.



But now consider Figure 5-2, which shows the Virginia Beach central place and urbanized area. Here the central place is again the lighter polygon, but as you can see it surrounds the census-designated urbanized area, which is the opposite of the expected relationship.

Figure 5-2. Virginia Beach urbanized area and central place.



The model is still valid. More sprawling urbanized areas are less dense, regardless of the relationship between central place and urbanized area boundaries. But it is clear that not all of my observations follow the pattern of older, inner city and first ring suburbs within larger suburban areas. Also, there is tremendous variation among central places. Central places can include large inner city and older suburb neighborhoods. But they also include cities that have had a lot of new suburban growth. For example, one of the central places in the Sacramento urbanized area is the city of Sacramento, a large city with many older urban neighborhoods. However, there are other "central places" within the UA that are much newer and more suburban in character--such as the cities of Elk Grove and Folsom. So, the designation of central place may not adequately capture the dynamic between older urban communities and younger suburban ones that I am trying to model. A different approach might be to identify "rings" of development within a metropolitan statistical area, defined by the average age of the housing stock in that jurisdiction. Then factors like poverty, economic inequality or job mismatch could be measured to understand the differences between rings. Some of the methods described in the literature review—using GIS techniques and census tract level data (such as Lee, 2011) —could be useful in this regard, although it was beyond the scope of this study.

Other opportunities to build on this research would be much simpler. This study shows that urban density has a relationship to the change in poverty rate between 2000 and 2010. It would be fairly easy to replicate this work using different census decades. Perhaps the effect is different, or not present at all, when 1990 and 2000 data are compared, or 1980 to 1990. And with the full implementation of the American Community Survey in 2005, it is now possible to measure income inequality at the place level (and indeed down to the Census tract level), through the use of a Gini coefficient. Because the Gini index was not available for Census designated places before 2005, it was not possible in this study to measure the change in Gini scores for central places between 2000 and 2010. But the Gini measure could be a useful way of exploring the relationship between place and income inequality in future research.

Conclusion

I began this study with a description of the current policy environment, in which there is heightened interest in poverty and income inequality. But the title of this study, “Place and poverty,” also reflects my abiding interest in how the qualities of the built environment can affect our quality of life. I do not mean this just in the sense of street-level urban design; I am not (only) talking about mixed-use development projects, sidewalk cafes, and bike lanes and community gardens. Instead, I mean that the policies we make about land-use and development—where we decide to put people—help determine outcomes in education, law enforcement, transportation, and other policy areas. At least, that is what I suspected based on reading and casual observation. And that was the kind of question that I was interested in testing, empirically. I believe this study does contribute something to the understanding of the reasons for poverty and income inequality. But it also highlights the interplay between policy areas like transportation, education and land use, and the complicated ways that "quality of place" shapes our quality of life.

.

Appendix A.
Literature Tables.

Table A-1. Causes of sprawl.

Author	**Dependent variable**	**Key explanatory variables**	**Method**	**Findings**
Brueckner and Fansler (1983)	Land area of 40 Census designated Urbanized Areas, in square miles.	Population, Pct commuters using transit, pct commuters owning autos, median ag land value, household income	Linear regression	Elasticities: Population: 1.097, 1.086 Ag land value: -0.234, -0.231 Household income: 1.497, 1,496 Commuters using transit and commuters owning autos not significant.
Razin and Rosentraub (2000)	Residential sprawl measure, indexing pct single-unit homes, pop per sq km, and dwellings per sq km. Included U.S. and Canadian metropolitan areas	Index of political fragmentation. Based on number of local governments per 10,000 residents, general purpose governments per 10,000 residents, proportion of population in largest city, proportion of population in cities >100,000.	OLS regression	Metro areas with high land values, Canadian metropolitan areas, and larger metro areas were less sprawling. Fragmentation not significant.
McGrath (2005)	Urbanized land area in square miles for 33 largest metropolitan statistical areas.	Population, income, ag land values, transportation costs, decade.	OLS regression	Elasticities: Population 0.76 Income: 0.33 Transportation cost: 0.28

	Source is U.S. Census	Source: U.S. Census, Consumer Price Index, USDA ag land prices		Ag land value: 0.10
Burchfield (2006)	Sprawl index based on pct undeveloped land in a square kilometer around an average residential development built 1976-92. Source is satellite images from USGS Landsat data.	Central sector employment, streetcar passengers per capita in 1902, pop growth 1920-70, elevation in urban fringe, ruggedness in urban fringe, mean cooling degree days, mean heating degree days	Linear regression. Reported in standardized coefficients measuring absolute change in sprawl index per one Std Dev change in EV.	TK RESULTS
Song and Zenou (2006)	Land area of Census designated Urbanized Areas, in acres.	UA population, median household income, aggregated effective property tax rates of jurisdictions within UA.	OLS regression 2SLS regression Property tax rates are lagged three years GIS techniques used to aggregate tax rates within a UA.	Elasticities: Population: 0.519 Income: 0.724 Transportation expenditure: 0.288 Preperty tax mil rate: -0.401
Wassmer (2006)	Census Urban Area land area, square miles.	Presence of local urban containment policy. Length of time that urban containment has been in place.	Regression.	Presence of statewide local growth policy resulted in -.159 percent decrease in urban area size.

		Presence of statewide local growth management. Length of time that statewide growth management law has been in place.		Years of urban containment somewhere in UA: -0.004 Years of vertically integrated local growth management: -0.023
Wassmer (2007)	Census Urban Area land area, square miles. Urban Area population density.	Population, ag land price, per capita income, pct housholds married, pct households owning autos, pct central place poor population	OLS regression	Elasticities (Square miles): Population: 0.9 Ag land price: .05 Pct households married: .959 Income: 1.139 Autos:0.052 Pct households > $100k: -0.372 (Pop Density) Income: -1.010 Pct households married: -1.134 Pct households > $100k: 0.445
Woo (2011)	Population Density, and Employment Density, as described by Census Traffic Analysis Zones (TAZ) in 135 Metropolitan Statistical Areas.	Presence of state and local urban growth boundaries, presence of local urban service area.	Three stage least squares regression.	State urban growth boundaries reduced population density gradient by .014 percent, and employment density gradient by .013 percent. Statistically significant at 5 percent.
Fallah (2012)	Metropolitan area sprawl	Standard deviation from	OLS regression	Uncertainty is negatively

	index number (see above), between 0 and 1 with higher number being more sprawling. Looked at a panel with DV from 1980, 1990 and 2000.	past annual population change. This is assumed to be a proxy for uncertainty regarding future land rent.		correlated with sprawl.

Table A-2. Methods of measuring sprawl

Author	Method	Pros and cons
Razin and Rosentraub (2000)	Calculated a residential sprawl measure for Census Metropolitan Areas. Measure is an index of: percentage of dwellings that are single-unit detached houses, population per square km and, housing units per square km.	Pro: Easily applicable to all US MSAs. Con: Coarse, says little about distribution of population, and nothing about employment centers.
Wassmer (2000)	Compared centralization in 6 Census designated Metropolitan Statistical Areas. Centralization=Central City Pop/Region Pop. Also measured change in Central City retail sales and change in farm land percentages, and relationships between suburb and central city in poverty, unemployment, income and commute.	Pro: Can compare cities to suburbs. Can relate extent of sprawl to social outcomes. Con: Limits the definition of sprawl to centralization.
Fulton (2001)	Sprawl measure is metropolitan area population (per decennial census) divided by growth of urbanized land (per USDA's National Resources Inventory).	Pro: Simple and easy to apply to all U.S. metro areas. Con: Coarse, says nothing about distribution of population and jobs, no differentiation between central cities and suburbs.
Wassmer (2001)	Looked at Census designated Urbanized Areas and Metropolitan Statistical Areas in California and the Western U.S. Used a sprawl index of pct change in urban fringe (non central	Pro: Census data are relatively easy to acquire. Advantages in comparing California or Western UAs which may by more similar in some ways compared to older Eastern cities.

	place) land divided by pct change in urban population. Looked at other relationships such as change in farm land area compared to metro land area, and change in central place population to metro area population.	Con: Little discernment of local differences. UA may not capture leap frog development. Metro Areas mask local development patterns.
Galster (2001)	Divided 13 Census designated Urban Areas in one-mile grids, and measured data for each. Developed 8 dimensions of sprawl.	Pro: Takes into account many different attributes of sprawl. Con: Limited to a few metro areas.
Ewing et al. (2003)	Another multidimensional approach. Variables from census data combined into four factors: density, land use mix, centering, streets accessibility	Pro: Nuance, allows researchers to study types of sprawl. Con: Complexity, problem of weighting variables
Cutsinger (2005)	Extended Galster's 2001 index. Added Extended Urban Area to capture "leap frog" areas.	Pro: Considers many attributes of sprawl. Expanded to 50 UAs. Con: Unwieldy, still limited in the number of metro areas.
Wolman et al (2005)	Uses Extended Urban Area (EUA), a Census defined Urban Area, in addition to outlying 1sq mile cells that contain >= 60 dwellings and are in census tracts where >= 30 percent of worker commute to UA.	Pro: More complete measure of urbanized area than UA Con: More complicated Each EUA must be constructed by the researcher.
Burchfield (2006)	Developed a "scatteredness index" using satellite images and measuring pct undeveloped land in a square kilometer around an	Pros: Fine detail. Measures a type of newer development.

	average residential development built 1976-92.	Cons: Data is difficult to gather. Leaves out other types of development. Possible issues with different images resolution in each dataset.
Glaeser and Kahn (2001)	Used U.S. Department of Commerce's Zip Code Business Patterns to count number and size of businesses in zip codes around most population MSAs. Then measured distance between each zip code and CBD.	Pro: Relatively simple, can be applied to many metro areas. Con: Monocentric, ignores residential patterns.
Glaeser and Kahn (2003)	Used employment density above. Also used population density in zip codes and measure zip code distance from CBD.	Pro: Simple, includes some measure of residential patterns. Con: Lacks measures for other aspects of sprawl such as open space or connectedness
Lopez and Hynes (2003)	Used Census tract data to develop a sprawl index. For each of 330 Metropolitan Statistical Areas, tracts were sorted into high density (above 3500 persons per sq mi) and low density groups. Percentage of total population living in low density tracts was compared to high density tracts to derive a number between 0 and 100, with 100 being an MSA in which the population lives entirely in low density tracts.	Pro: Census data is available for all metro areas. Density is widely accepted measure of sprawl. Con:
Song (2004)	Using GIS and Census tract data, developed an index for Portland, OR, neighborhoods, measuring	Pro: Detailed, measures several facets of urban form to describe the extent of sprawl

	density, access to transit, mix of uses, pedestrian access and street interconnectivity.	Con: Difficult to gather data for more than one city or urban area.
Tsai (2005)	Measured density and distribution of employment and population based on Census Transit Analysis Zones (TAZ).	Pro: Data readily available. TAZ are built on census block data, so possible to measure clustering within a MSA or UA. Con: Time consuming compared to simpler indices.
Batisani (2009)	Used Landsat images and land use maps for Centre County, PA. 30,000 polygons were assigned land use values of 1 (urban) or (0) non-urban. Urbanization increased from 15.4 to 17.2 between 1993-2000	Pro: Fine description of urbanization in one location. Captures changes that might not be obvious in Census data. Con: Data is for one county, hard to compare several areas.
Jargowsky (2009)	Used four measures of suburbanization based on U.S. Census data. 1) A density gradient comparing portion of central city population to the suburban population, 2) The metropolitan area average population density, 3) The percent of the metro area population living in the central city, and 4) The average commuting time to work	Pro: Data is available for all U.S. metro areas. Variety of measures recognize multifaceted nature of sprawl. Con: Metro area is very broad, may not capture much of the local differences in land use and development.
Jia and Jiang (2010)	Used Open Street Maps database to measure street nodes as a proxy for urban sprawl.	Pro: Considers “natural cities” and naturally defined urban boundaries rather than more arbitrary political boundaries.

		Con: Non standard, reliability of data uncertain. Intensive computing necessary.
Kugler (2011)	Used Census Tract data to calculate 11 metrics of sprawl, including road connectivity, separation of uses and homogeneity of housing types, in each tract.	Pro: Fine detail, captures the "know it when you see it" characteristics of sprawl. Author says tract information can be aggregated to regional level. Con: 53,000 observations.
Fallah (2012)	Used Census block data to come up with density index. Within MSA, compared percentage of population living in blocks groups with pop density below US media to pct population in blocks with above median density. Index value is between 0 and 1 with higher number more sprawling	Pro: Census data are available to all. Con: Density is only one measure of sprawl.
Weitz (2012)	Used zip code business pattern from U.S. Census and GIS information to look at location of employment centers relative to residential areas.	Pro: Recognizes the polycentric nature of sprawl. Con: Still primarily a job sprawl measure, says little about residential sprawl
Yang (2012)	Developed a measure of polycentricity. Calculated regional population density based on Census Metropolitan Statistical Area. Calculated population density for census tracts. Divided tracts into four groups based on density. Compared distribution of high and moderate density tracts to calculate polycentricity. Compared	Pro: Captures the idea of a polycentric urban idea in a way other methods do not. Con: Only looks at population density, not other aspects of sprawl.

	this measure to commute behavior and travel times.	
Pereira (2013)	Created an urban centrality index (UCI), an employment density measure. Applied index to different datasets for U.S. and European cities. U.S. cities UCI based on data from the National Historic Geographic Information System website. Data is by census tract.	Pro: Data available for all U.S. areas. Centrality can be a better measure than just density. Con: Only a measure of economic/job centrality, not residential patterns

Appendix B.
Local governments employing growth control measures.

Table B-1. Local growth control measures by urban area and type of growth control (per Dawkins and Nelson 2003).

Urbanized area	Strong Accommodating	Strong Restricting	Weak Accommodating	Weak Restricting
Albuquerque, NM				■
Anchorage, AK	■			
Appleton, WI			■	
Atlantic City, NJ	■			
Balitmore, MD	■			
Beloit, WI-IL			■	
Bellingham, WA	■			
Bend, OR	■			
Bismarck, ND				■
Bloomington, IN				■
Bonita Springs, FL		■		
Boston, MA-NH-RI				■
Boulder, CO				■
Bremerton, WA	■			
Cape Coral, FL	■			
Charleston, SC			■	
Charlotte, NC-SC			■	
Charlottesville, VA			■	
Chico, CA		■		
Clarksville, TN-KY			■	
Corvallis, OR	■			
Davis, CA		■		
Dayton, OH			■	
Denver-Aurora, CO	■	■	■	■
Dover, DE			■	
Eau Claire, WI			■	
Eugene, OR	■			
Fayetteville, NC			■	
Flagstaff, AZ	■			
Fond du Lac, WI			■	
Fort Collins, CO	■	■		
Fresno, CA		■		

Gainesville, FL		■		
Greeley, CO				■
Green Bay, WI			■	
Hagerstown, PA				■
Honolulu, HI				
Iowa City, IA			■	
Jacksonville, FL	■			
Janesville, WI			■	
Joplin, MO			■	
Kalamazoo, MI			■	
Kennewick, WA	■			
Kenosha, WI			■	
Knoxville, TN		■		
La Crosse, WI-MN		■		
Lancaster, PA			■	
Lawton, OK			■	
Lexington, KY			■	
Lincoln, NE			■	
Little Rock, AR			■	
Madison, WI			■	
Medford, OR			■	
Merced, CA		■		
Miami, FL	■			
Minneapolis-St. Paul, MN			■	
Modesto, CA			■	
North Port-Punta Gorda, FL			■	
Ocala, FL	■			
Olympia-Lacey, WA	■			
Orlando, FL	■			
Oshkosh, WI	■			
Palm Bay-Melbourne, FL	■		■	
Portland, OR-WA	■		■	
Port St. Lucie, FL	■			■
Racine, WI			■	
Raleigh, NC			■	
Riverside-San Bernardino, CA	■			
Rochester, MN			■	
Sacramento, CA	■			
St. Cloud, MN			■	
Salem, OR			■	■

San Diego, CA	■			
San Francisco-Oakland, CA	■	■		
San Jose, CA	■	■		
San Luis Obispo, CA		■		
Santa Barbara, CA		■		
Santa Cruz, CA		■		
Santa Rosa, CA		■		
Sarasota-Bradenton, FL	■		■	
Seattle, WA	■			
Sheboygan, WI			■	
Sioux City, IA-NE-SD			■	
Sioux Falls, SD			■	
Spokane, WA-ID	■			
Springfield, IL			■	
Tallahassee, FL	■			
Tampa-St. Petersburg, FL	■			
Titusville, FL				■
Tucson, AZ	■		■	
Vallejo, CA		■		
Virginia Beach, VA			■	
Visalia, CA				■
Washington, DC-VA-MD				■
Wausau, WI			■	
Wichita, KS			■	
Wilmington, NC			■	
Yakima, WA	■			
York, PA			■	
Yuba City, CA	■			
Yuma, AZ-CA			■	

Appendix C.

Pairwise correlations for explanatory variables

	Urban area pop. Density (1000)	Pop. (10000)	Pct African American	Pct. Asian	Percent Hispanic	Single Mother	High school graduate
Urban area pop. Density (1000)	1						
Population (10000)	0.13*	1					
Pct African American	-0.18*	0.12*	1				
Percent Asian	0.56*	0.05	-0.19*	1			
Percent Hispanic	0.5*	0.09*	-0.18*	0.11*	1		
Single Mother	-0.15*	0.07*	0.7*	-0.26*	0.20*	1	
High school graduate	-0.37*	-0.08*	0.16*	-0.38*	-0.23*	0.36*	1
Bachelors degree above	0.06*	0.01	-0.20*	0.29*	-0.31*	-0.60*	-0.75*
Foreign born	0.71*	0.12*	-0.19*	0.54*	0.73*	-0.08*	-0.42*
Percent unemployed	-0.02	0.11*	0.47*	-0.15*	0.30*	0.69*	0.14*
Percent in poverty	-0.16*	0.10*	0.47*	-0.19*	0.22*	0.64*	0.10*
Percent in construction	-0.05	-0.03	-0.15*	-0.24*	0.18*	-0.03	0.22*
Pct in manufacturing	0.02	-0.08*	-0.06*	0.02	0.06*	0.15*	0.28*
Percent in wholesale	0.44*	-0.01	-0.24*	0.18*	0.44*	-0.04	-0.06*
Percent in retail	-0.30*	-0.16*	-0.17*	-0.28*	-0.12*	-0.01	0.42*
Percent in information	0.25*	0.13*	-0.10*	0.25*	-0.09*	-0.34*	-0.44*
Pct in finance	0.19*	0.12*	-0.11*	0.13*	-0.16*	-0.36*	-0.25*
Pct in other services	0.06*	0.07*	0.17*	-0.10*	0.19*	0.23*	0.12*
Pct. Above $200,000	0.22*	0.04	-0.15*	0.28*	-0.11*	-0.45*	-0.59*
Median income (1000)	0.39*	-0.04	-0.38*	0.43*	-0.04	-0.58*	-0.50*
Med. home val (1000)	0.53*	0.02	-0.33*	0.55*	0.10*	-0.50*	-0.60*
Pct. age 0 to 19	0.14*	0.01	0.12*	-0.06*	0.49*	0.52*	0.005
Pct. age 20 to 44	0.13*	0.09*	-0.01	0.14*	0.06*	-0.09*	-0.42*
Pct. age 45 to 64	0.004	-0.04	-0.12*	0.14*	-0.33*	-0.35*	0.02
Statewide growth cntrl	-0.06*	-0.05	-0.03	-0.04	-0.06*	-0.11*	-0.02
Strong/Accommodating	0.19*	0.01	-0.08*	0.19*	0.06*	-0.15*	-0.21*
Strong/Restricting	0.23*	0.002	-0.07*	0.33*	0.02	-0.20	-0.28*
Weak/Accommodating	-0.09*	0.0006	-0.04	-0.07*	-0.13*	-0.05	0.005
Weak/Restricting	-0.01	-0.01	-0.06*	0.04	-0.03	-0.11*	-0.10*

	Bachelor degree	Foreign born	Pct Unemp	Pct in Poverty	Pct in Construc.	Pct in Manuf.	Pct Wholes.
Bachelors degree above	1						
Foreign born	0.02	1					
Percent unemployed	-0.40*	0.07*	1				
Percent in poverty	-0.31*	-0.0017	0.75*	1			

Percent in construction	-0.39*	-0.03	-0.13*	-0.18*	1		
Pct in manufacturing	-0.33*	0.03	0.01	-0.05	-0.12*	1	
Percent in wholesale	-0.20*	0.42*	-0.06*	-0.19*	0.08*	0.23*	1
Percent in retail	-0.38*	-0.29*	-0.09*	-0.04	0.36*	-0.05	0.14*
Percent in information	0.56*	0.19*	-0.26*	-0.30*	-0.18*	-0.31*	-0.06*
Pct in finance	0.39*	0.08*	-0.37*	-0.49*	-0.02	-0.33*	0.16*
Pct in other services	-0.29*	0.14*	0.17*	0.16*	0.23*	-0.20	0.003
Pct. Above $200,000	0.70*	0.18*	-0.35*	-0.38*	-0.21*	-0.15*	0.02*
Median income (1000)	0.57*	0.25*	-0.59*	-0.79*	-0.03	0.003	0.21*
Med. home val (1000)	0.57*	0.46*	-0.34*	-0.42*	-0.14*	-0.09*	0.11*
Pct. age 0 to 19	-0.40*	0.14*	0.40*	0.24*	0.14*	0.23*	0.28*
Pct. age 20 to 44	0.39*	0.16*	0.10*	0.17*	-0.24*	-0.07*	-0.15*
Pct. age 45 to 64	0.22*	-0.10*	-0.48*	-0.55*	0.03	-0.10*	0.06*
Statewide growth cntrl	0.07*	0.10*	-0.10*	-0.09*	0.13*	-0.28*	-0.03
Strong/Accommodating	0.16*	0.26*	-0.11*	-0.15*	0.10*	-0.22*	0.004
Strong/Restricting	0.29*	0.21*	-0.16*	-0.15*	-0.01	-0.09*	-0.13*
Weak/Accommodating	0.06*	-0.15*	-0.10*	-0.11*	0.004	0.02	-0.06*
Weak/Restricting	0.21*	0.07*	-0.15*	-0.13*	-0.02	-0.16*	-0.16*

	Percent in retail	Pct in information	Pct in finance	Pct in other services	Pct. Above $200,000	Median income (1000)	Median home value (1000)
Percent in retail	1						
Percent in information	-0.27*	1					
Pct in finance	-0.09*	0.39*	1				
Pct in other services	0.03	-0.03	-0.06*	1			
Pct. Above $200,000	-0.34*	0.45*	0.47*	-0.15*	1		
Median income (1000)	-0.25*	0.46*	0.48*	-0.23*	0.69*	1	
Med. home val (1000)	-0.36*	0.48*	0.33*	-0.13*	0.72*	0.71*	1
Pct. age 0 to 19	0.01	-0.25*	-0.28*	0.04	-0.30*	-0.05	-0.26*
Pct. age 20 to 44	-0.30*	0.31*	-0.04	-0.20*	-0.01	0.03	0.18*
Pct. age 45 to 64	-0.01	0.14*	0.38*	0.04	0.43*	0.43*	0.28*
Statewide growth cntrl	0.11*	0.08*	0.09*	0.17*	0.04	0.02	-0.01
Strong/Accommodating	-0.05	0.23*	0.10*	0.10*	0.13*	0.19*	0.22*
Strong/Restricting	-0.21*	0.34*	0.03	0.07*	0.26*	0.31*	0.41*
Weak/Accommodating	0.04	-0.01	0.04	-0.08*	-0.01	0.03	-0.07*
Weak/Restricting	-0.17*	0.30*	0.08*	0.12*	0.11*	0.18*	0.14*

	Pct. age 0 to 19	Pct. age 20 to 44	Pct. age 45 to 64	State growth control	Strong Accomm -odating	Strong Restrict- ing	Weak Accomm -odating
Pct. age 0 to 19	1						
Pct. age 20 to 44	0.07*	1					
Pct. age 45 to 64	-0.52*	-0.60*	1				
Statewide growth cntrl	-0.28*	-0.14*	0.21*	1			
Strong/Accommodating	-0.15*	0.04	0.14*	0.55*	1		
Strong/Restricting	-0.15*	0.15*	0.10*	0.09*	0.39*	1	
Weak/Accommodating	-0.02	0.02	0.04	0.004	0.04	0.03	1
Weak/Restricting	-0.12*	0.13*	0.06*	0.16*	0.19*	0.37*	0.05

Appendix D

Tests for heteroskedasticity

Heteroskedasticity is the unequal variance in the error term among observations of a dependent variable. It affects the standard errors of regression coefficients, and can lead to errors in determining the statistical significance of variables. The Breusch-Pagan test is used to determine if heteroskedasticity is present in the data. Using STATA statistical analysis software, I performed the Breusch-Pagan test following three regressions, Model 8, Model 9 and Model 12. In each case, p-values below .01 indicate I must reject the null hypothesis of constant variance. Thus all regressions are reported in Chapter 4 using robust standard errors.

Model 8

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of CpPctPovChg

chi2(1) = 84.66
Prob > chi2 = 0.0000

Model 9

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of CpPctPovChg

chi2(1) = 60.92
Prob > chi2 = 0.0000

Model 12

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of CpPctPovChg

chi2(1) = 56.81
Prob > chi2 = 0.0000

Appendix E.

Tests for multicollinearity

Table E-1. VIF Scores for Models 8 and 9

Model 8 variable	VIF	Model 9 Variable	VIF
Bachelor degrees and above	25.58	Bachelors degrees and above	34.42
Median income	21.04	Statewide growth control	30.45
High school graduate	16.84	Median Income	25.5
Statewide growth control	16.52	FL	23.88
FL	12.73	High school graduate	20.89
Percent Hispanic	11.81	Percent Hispanic	15.61
Single mother	11.16	Single mother	14.63
Foreign born	11.04	Percent in poverty	14.55
Percent in poverty	10.5	Age 0 to 19	12.16
Median home value	9.37	Foreign born	11.89
Age 0 to 19	8.78	Median home value	10.13
Percent African American	7.32	Above $200,000	8.81
Above $200,000	6.76	Percent African American	8.44
Age 45 to 64	6.37	MD	7.66
WA	4.82	Age 45 to 64	6.94
Age 20 to 44	4.82	WA	6.67
MD	4.4	Percent Asian	5.33
Urban area population density (by 1000)	4.33	Urban area population density	5.12
Percent Asian	4.14	Percent unemployed	4.97
Percent Unemployed	3.88	Manufacturing	4.79
Manufacturing	3.65	OR	4.57
PA	3.26	Strong restricting growth control	4.12
TX	3.24	Age 20 to 44	4.11
OR	3.02	Strong accommodating growth control	3.72
MA	2.97	Finance	3.67
MI	2.94	VA	3.6
Construction	2.73	Weak restricting growth control	3.24
OH	2.7	Retail	3.21
Retail	2.58	MI	3.18

Finance	2.55	OH	3.17
GA	2.52	TX	3.13
VA	2.47	Wholesale	3.11
NC	2.43	Construction	2.98
Wholesale	2.42	PA	2.86
Strong accommodating growth control	2.37	NY	2.76
IL	2.36	CO	2.74
Information	2.35	NC	2.73
WI	2.27	IL	2.71
NY	2.27	Information	2.57
Weak restricting growth control	2.23	MN	2.55
Strong restricting growth control	2.18	NJ	2.47
IN	2.16	Other services	2.45
AL	2.01	CT	2.22
MN	2	GA	2.19
CT	1.99	Weak accommodating growth control	2.18
CO	1.96	WI	1.84
Other services	1.92	AL	1.81
NJ	1.87	IN	1.79
RI	1.86	AR	1.73
SC	1.86	MO	1.71
LA	1.77	LA	1.71
TN	1.77	NV	1.67
ME	1.74	ME	1.66
Weak accommodating growth control	1.68	DC	1.64
AR	1.67	SC	1.58
IA	1.63	AZ	1.57
MO	1.63	OK	1.54
HI	1.47	TN	1.49
WV	1.42	NM	1.46
NH	1.42	KS	1.42
NV	1.4	MS	1.41
OK	1.4	UT	1.41
AZ	1.38	IA	1.39
MS	1.37	DE	1.26
KS	1.37	NE	1.24
NM	1.3	KY	1.23
UT	1.29	Population	1.21
DC	1.27	WV	1.19
VT	1.26	ID	1.08
KY	1.24		
ID	1.22		
ND	1.18		
Population00	1.17		
MT	1.17		
DE	1.16		

SD	1.14		
NE	1.13		
AK	1.09		
WY	1.08		
Mean VIF	3.99	Mean VIF	5.58

Table E-2. VIF scores for Models 12 and 13

Model 12 variable	VIF	Model 13 variable	VIF
Bachelor degrees and above	35.02	Bachelor degrees and above	36
Statewide growth control	30.42	Statewide growth control	31.27
Median income	25.33	FL	26.02
FL	23.88	Median income	25.61
High school graduate	20.87	High school graduate	21.34
Percent Hispanic	15.44	Percent Hispanic	15.99
Percent in poverty	14.54	Single mother	14.89
Single mother	14.53	Percent in poverty	14.61
Age 0 to 19	12.14	Density factor	12.43
Foreign born	11.86	Age 0 to 19	12.25
Median home value	10.12	Foreign born	11.99
Above $200,000	8.8	Median home value	10.16
Percent African American	8.06	Above $200,000	8.88
MD	7.64	Percent African American	8.77
Age 45 to 64	6.96	MD	7.91
WA	6.66	WA	7.49
Percent Asian	5.34	Age 45 to 64	6.98
Manufactu~00	5.06	Street accessibility	5.67
Percent unemployed	4.97	Percent Asian	5.38
OR	4.63	Manufacturing	5.35
Strong restricting growth	4.35	Percent unemployed	5.16
Age 20 to 44	4.11	Urban centering	5.12
Strong accommodating growth	3.73	Strong accommodating growth	4.98
Finance	3.68	OR	4.88
VA	3.59	Land use mix	4.84
Weak restricting growth	3.38	Strong restricting growth	4.27
Composite sprawl measure	3.21	Age 20 to 44	4.13
Retail	3.15	MI	4.03
OH	3.07	IL	3.98
Construction	3.06	TX	3.82
MI	3.04	Finance	3.74
Wholesale	2.99	PA	3.7
TX	2.98	VA	3.67

PA	2.8	OH	3.64
CO	2.74	NY	3.57
NY	2.66	NJ	3.48
Information	2.57	Weak restricting growth	3.37
IL	2.54	MN	3.27
MN	2.54	Retail	3.27
NJ	2.53	Wholesale	3.18
NC	2.48	NC	3.11
OtherSvcs00	2.39	Construction	3.08
Weak accommodating growth	2.19	CT	3.03
GA	2.11	CO	2.95
CT	2.08	Information	2.59
WI	1.84	GA	2.49
IN	1.76	Other services	2.47
AL	1.7	NV	2.31
LA	1.69	WI	2.2
AR	1.69	Weak accommodating growth	2.19
NV	1.67	LA	2.09
ME	1.67	IN	2.08
DC	1.64	AL	2
MO	1.64	SC	1.93
AZ	1.59	ME	1.88
SC	1.51	MO	1.87
OK	1.51	AR	1.87
TN	1.49	AZ	1.81
NM	1.46	TN	1.73
UT	1.4	OK	1.65
IA	1.38	DC	1.65
KS	1.38	NM	1.57
MS	1.36	KS	1.56
NE	1.25	MS	1.55
DE	1.25	IA	1.48
KY	1.23	UT	1.43
Population	1.21	WV	1.37
WV	1.19	NE	1.34
ID	1.08	DE	1.29
		KY	1.27
		Population	1.22
		ID	1.1
Mean VIF	5.53	Mean VIF	5.99

Table E-3. VIF scores for Models 16 and 17

Model 16 variable	VIF	Model 17 variable	VIF
Median income	14.16	Median income	16.64
Percent in poverty	8.97	Percent Hispanic	11.77
High school graduates	6.38	Percent in poverty	11.37
Age 45 to 64	5.74	Density factor	9.12

Median home value	5.25	High school graduates	6.96
Age 0 to 19	4.52	Age 0 to 19	6.39
Percent Hispanic	4.19	Age 45 to 64	6.33
Age 20 to 44	4.04	Median home value	6.26
Urban area population density (by 1000)	3.9	Percent unemployed	4.93
Percent unemployed	3.6	Street accessibility	4.48
Percent African American	3.06	Percent African American	4.25
Manufacturing	2.73	Percent Asian	3.98
Construction	2.38	Land use mix	3.78
Retail	2.34	Urban centering	3.61
Wholesale	2.31	Manufacturing	3.6
Finance	2.27	Age 20 to 44	3.52
Informati~	2.21	Wholesale	2.96
Percent Asian	2.14	Finance	2.91
PA	1.94	Retail	2.81
MI	1.88	Construction	2.66
OH	1.68	Information	2.48
TX	1.67	NY	2.34
Other services	1.65	Other services	2.2
NC	1.57	CT	2.18
MA	1.56	OH	2.11
GA	1.53	MI	2.1
IL	1.49	PA	2.09
WI	1.49	NJ	2.01
NY	1.47	IL	1.91
IN	1.44	TX	1.83
VA	1.42	MD	1.78
CT	1.4	NC	1.77
AL	1.36	NV	1.69
MD	1.35	MN	1.62
SC	1.35	VA	1.59
WA	1.33	WA	1.58
TN	1.33	GA	1.5
MN	1.32	UT	1.43
NJ	1.3	WI	1.42
LA	1.27	CO	1.39
MO	1.26	OR	1.36
CO	1.24	AZ	1.34
NV	1.22	MO	1.33
AR	1.21	IN	1.32
IA	1.21	TN	1.31
UT	1.2	AL	1.29
WV	1.19	LA	1.28
AZ	1.17	SC	1.28
OK	1.16	MS	1.22
NH	1.16	OK	1.22
Population	1.16	Population	1.2
MS	1.15	ME	1.19
OR	1.15	WV	1.19

KS	1.13	AR	1.18
HI	1.12	IA	1.18
ID	1.12	KS	1.17
ME	1.09	DC	1.17
NM	1.09	DE	1.16
RI	1.09	NM	1.13
DC	1.09	NE	1.11
KY	1.09	KY	1.1
ND	1.08	ID	1.06
AK	1.07		
DE	1.07		
MT	1.07		
SD	1.06		
NE	1.04		
VT	1.04		
WY	1.04		
Mean VIF	2.11	Mean VIF	3.04

Table E-4. VIF scores for Model 18.

Model 18 Variable	**VIF**
Statewide growth control	30.76
FL	22.39
Percent high school graduate or better	18.53
Density factor	12.13
Percent in poverty	12.03
Percent Hispanic	9.84
Single mother	8.49
Median income (1000)	7.96
Percent African American	7.49
WA	7.22
MD	6.7
Median home value (1000)	5.95
Percent in Manufacturing	5.44
Street factor	5.33
Centering factor	5.18
Percent unemployes	4.91
Strong accommodating growth control	4.81
Mix factor	4.72
OR	4.64
Strong restrictive growth control	4.04

Weak restrictive growth control	3.33
VA	3.15
TX	3.15
Percent in wholesale	3.12
Percent in finance	3.1
IL	2.89
MN	2.85
MI	2.85
Percent in retail	2.7
CO	2.7
Percent working age	2.68
Percent in construction	2.66
Percent Asian	2.64
NC	2.49
Percent in information	2.41
OH	2.32
NY	2.29
NJ	2.28
Percent in other services	2.27
CT	2.17
Weak accommodating growth control	2.16
PA	2.1
NV	2.1
GA	1.97
LA	1.77
ME	1.75
AZ	1.73
WI	1.71
AL	1.7
SC	1.7
AR	1.63
MO	1.61
TN	1.61
IN	1.56
DC	1.53
OK	1.46
NM	1.46
MS	1.45
KS	1.37
WV	1.28
IA	1.25
UT	1.23
NE	1.23
Population (10000)	1.2
KY	1.19
DE	1.18
ID	1.07
Mean VIF	4.25

References

Addison, C., Zhang, S., & Coomes, B. (2013). Smart Growth and Housing Affordability A Review of Regulatory Mechanisms and Planning Practices. *Journal of Planning Literature*, *28*(3), 215-257.

Anacker, Katrin B., (2012) Shaky Palaces? Analyzing property values and their appreciation in minority first-suburbs. *International Journal of Urban and Regional Research* 36:4 (2012) 791-816

Atikison, Robert (2012), et al "Worse than the Great Depression: What experts are missing about American manufacturing decline." Information, Technology and Innovation Foundation. Retrieved from http://www2.itif.org/2012-american-manufacturing-decline.pdf

Batisani, N., & Yarnal, B. (2009). Urban expansion in Centre County, Pennsylvania: Spatial dynamics and landscape transformations. *Applied Geography*, *29*(2), 235-249.

Brueckner, Jan K., and Largey, Ann G. (2008) Social Interaction and Sprawl. *Journal of Urban Economics* 64 (2008) 18-34

Brueckner, Jan K., and Helsley, Robert, W., (2011) Sprawl and Blight. *Journal of Urban Economics*. 69 (2011) 205-213

Brueckner, J. K., & Fansler, D. A. (1983). The economics of urban sprawl: Theory and evidence on the spatial sizes of cities. *The Review of Economics and Statistics*, *65*(3), 479-482.

Burchell, R. W. (1997). Economic and fiscal costs (and benefits) of sprawl. *Urban Law.*, *29*, 159.

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

Cameron, A. Colin and Trivedi, Pravin K., (2009), *Microeconomics Using Stata*, College Station, Texas: Stata Press

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). *Where is the land of opportunity? the geography of intergenerational mobility in the united states* (No. w19843). National Bureau of Economic Research.

Cutsinger, Jackine, Galster, George, Wolman, Michael, Hanson, Royce, and Towns, Douglas, (2005) Verifying the multi-dimensional nature of metropolitan land use: Advancing and understanding the measurement of sprawl. *Journal of Urban Affairs* 27:3 (2005) 235-259

Datel, R., & Dingemans, D. (2008). Immigrant space and place in suburban Sacramento. *Twenty-First-Century Gateways: Immigrant Incorporation in Suburban America*, 171-199.

Dawkins, C. J., & Nelson, A. C. (2003). State growth management programs and central-city revitalization. *Journal of the American Planning Association*, *69*(4), 381-396.

Ewing, R., Pendall, R., & Chen, D. (2003). Measuring sprawl and its transportation impacts. *Transportation Research Record: Journal of the Transportation Research Board*, *1831*(1), 175-183.

Ewing, R., Meakins, G., & Hamidi, S. (2013). Measuring Urban Sprawl and Validating Sprawl Measures.

Fallah, B. N., Partridge, M. D., & Olfert, M. R. (2011). Urban sprawl and productivity: evidence from US metropolitan areas. *Papers in Regional Science*, *90*(3), 451-472.

Fallah, B., Partridge, M., & Olfert, M. R. (2012). Uncertain economic growth and sprawl: evidence from a stochastic growth approach. *The Annals of Regional Science*, *49*(3), 589-617.

Fan, Y. (2012). The Planners' War against Spatial Mismatch Lessons Learned and Ways Forward. *Journal of Planning Literature*, *27*(2), 153-169.

Fry, R., & Taylor, P. (2012). The rise of residential segregation by income. *Pew Research Center*.

Fulton, W. B., & Pendall, R. (2001). *Who sprawls most?: How growth patterns differ across the US*. Washington, DC: Brookings Institution, Center on Urban and Metropolitan Policy.

Fusch, R., & Ford, L. R. (1983). Architecture and the Geography of the American City. *Geographical Review*, 324-340.

Galster, George, Hanson, Royce, Ratcliffe, Michael R., Wolman, Harold, Coleman, Stephen, and Freihage, Jason (2001) Wrestling sprawl to the ground: Defining and measuring an elusive concept. *Housing Policy Debate* 12:4 (2001) 681-716

Glaeser, Edward and Kahn, Matthew (2001) Decentralized employment and the transformation of the American city. National Bureau of Economic Research working paper. (February 2001)

Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and urban growth. *Handbook of regional and urban economics*, *4*, 2481-2527.

Glaeser, Edward, (2011). Triumph of the City: How our greatest invention makes us richer, smarter, greener, healthier , and happier. Penguin Press, New York, NY

Grengs, J. (2010). Job accessibility and the modal mismatch in Detroit. *Journal of Transport Geography*, *18*(1), 42-54.

Hanlon, Bernadette (2009) City and Community 8:3 A typology of inner ring suburbs: Class, race and ethnicity in U.S. Suburbia (September 2009)

Holcombe, Randall G., and Williams, DeEdgra W., (2010) Urban sprawl and transportation externalities. *The Review of Regional Studies*. 40:3 (2010) 257-273

Jaret, Charles, Ghadge, Ravi, Williams Reid, Lesley, and Adelman, Robert, M., (2009) The measurement of suburban sprawl, an evaluation. *City and Community* 8:1 (March 2009)

Jargowsky, Paul A. (2009) Cause or consequence? Suburbanization and crime in U.S. Metropolitan Areas. *Crime and Delinquency* 55:28 (2009)

Jargowsky, Paul and Squires, G. D. (Ed.). (2002). *Urban sprawl: Causes, consequences, and policy responses*. The Urban Insitute.

Jia, T., & Jiang, B. (2010). Measuring urban sprawl based on massive street nodes and the novel concept of natural cities. *arXiv preprint arXiv:1010.0541*.

Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *The Quarterly Journal of Economics*, 175-197.

Kew, B., & Lee, B. D. (2013). Measuring Sprawl across the Urban Rural Continuum Using an Amalgamated Sprawl Index. *Sustainability*, *5*(5), 1806-1828.

Kotkin, Joel (2013) “How can we be so dense?” Forbes, website, August 8, 2013. Retrieved on July 15, 2014, from http://www.forbes.com/sites/joelkotkin/2013/08/08/how-can-we-be-so-dense-anti-sprawl-policies-threaten-americas-future/

Krugman, Paul (2013), "Density," New York Times website, April 16, 2013, retrieved from
http://krugman.blogs.nytimes.com/2013/04/16/density/?_php=true&_type=blogs&_r=0

Kugler, T. A. (2011). A multi-scale comparative study of shape and sprawl in metropolitan regions of the United States.

Lee, Sugie and Leigh, Nancy Green (2007) Intrametropolitan Spatial Differentiation and Decline of Inner-Ring Suburbs. *Journal of Planning Education and Research* 27 (2007) 146-164

Lee, Sugie (2011) Metropolitan Growth Patterns and Socio-Economic Disparity in Six U.S. Metropolitan Areas, 1970-2000 *International Journal of Urban and Regional Research* 35:5 (2011) 988-1011

Lopez, Russ and Hynes, H Patricia (2003) Sprawl in the 1990s: Measurement, Distribution and Trends, Urban Aff Rev Thousand Oaks Calif 2003 January 1; 38(3): 325–355.

Massey, D. S., & Rothwell, J. (2009). The effect of density zoning on racial segregation in US urban areas. *Urban Affairs Review*.

McGrath, D. T. (2005). More evidence on the spatial scale of cities. *Journal of Urban Economics*, *58*(1), 1-10.

O'Sullivan, Arthur (2009) *Urban Economics,* Seventh Edition, McGraw-Hill Irwin, New York, NY

Paulsen, K. (2012). Yet even more evidence on the spatial size of cities: Urban spatial expansion in the US, 1980–2000. *Regional Science and Urban Economics*, *42*(4), 561-568.

Paulsen, K. (2013) Geography, policy or market? New evidence on the measurement and causes of sprawl (and infill) in US metropolitan regions. *Urban Studies*, 0042098013512874.

Pereira, R. H. M., Nadalin, V., Monasterio, L., & Albuquerque, P. H. (2013). Urban Centrality: A Simple Index. *Geographical Analysis*, *45*(1), 77-89.

Puentes, Robert, and Warren, David (2006) One-fifth of America: A comprehensive guide to America's first suburbs. Brookings Institution Survey Series, February 2006

Razin, E., & Rosentraub, M. (2000). Are fragmentation and sprawl interlinked? North American evidence. *Urban Affairs Review*, *35*(6), 821-836.

Rico, Miriam Hortas, (2013) Sprawl, blight and the role of urban containment policies: Evidence from U.S. Cities. *Cities and Innovation* 2 (2013)

Rogerson, P. A. (2010). *Statistical methods for geography: a student's guide*. Sage Publications.

Rosenthal, Stuart S. (2008) Old homes, externalities, and poor neighborhoods. A model of urban decline and renewal. *Journal of Urban Economics* 63 (2008) 816-840

Sanburn, Josh (2014), "U.S. Cities are slowing but suburbs are growing." May 22, 2014, Time, retrieved July 15, 2014, from http://time.com/107808/census-suburbs-grow-city-growth-slows/

Severson, Kim (2012), "No end to housing bust in 'Carpet Capital of the World,.'" *New York Times*, August 20, 2012, retrieved from http://www.nytimes.com/2012/08/21/us/in-carpet-capital-of-the-world-housing-fallout-continues.html?_r=0

Sharkey, Patrick (2013) *Mobility and the Metropolis: How communities factor into economic mobility*. Pew Charitable Trusts.

Shihadeh, E. S., & Ousey, G. C. (1996). Metropolitan expansion and black social dislocation: The link between suburbanization and center-city crime. *Social Forces*, *75*(2), 649-666.

Silver, Nate, (2014) "'Inequality' Booms on MSNBC And Fox News (CNN's looking for Flight 370)", retrieved from, http://fivethirtyeight.com/datalab/inequality-booms-on-msnbc-and-fox-news-cnns-looking-for-flight-370/

Song, Yan, and Knapp, Gerrit-Jan (2004) Measuring Urban Form: Is Portland Winning the War on Sprawl? *Journal of the American Planning Association* 70: 2 (Spring 2004)

Song, Y., & Zenou, Y. (2006). Property tax and urban sprawl: Theory and implications for US cities. *Journal of Urban Economics*, *60*(3), 519-534.

Stoll, M. A. (2006). Job sprawl, spatial mismatch, and black employment disadvantage. *Journal of Policy Analysis and Management*, *25*(4), 827-854.

Studenmund, A. H. (2005). *Using econometrics: A practical guide* 5th Edition Addison Wesley, New York, NY

Tiebout, Charles, (1956). A pure theory of local expenditures'. *The Journal of Political Economy*, *64*(5), 416-424.

Tsai, Y. H. (2005). Quantifying urban form: compactness versus' sprawl'. *Urban Studies*, *42*(1), 141-161.

Wasmmer, Robert W., (2000) "Urban Sprawl in a U.S. Metropolitan Area: Ways to Measure and a Comparison of the Sacramento Area to Similar Metropolitan Areas in California and the U.S." *CSUS Public Policy and Administration Working Paper*, (2000-03).

Wassmer, R. W. (2001). An Economist's Perspective on Urban Sprawl: With an Application to Metropolitan Areas in the American West. *Sacramento, CA: California State University Sacramento Graduate Program in Public Policy and Administration*.

Wassmer, R. W. (2006). The Influence of Local Urban Containment Policies and Statewide Growth Management on the Size of United States Urban Areas*. *Journal of Regional Science*, *46*(1), 25-65.

Wassmer, R. W. (2008). Causes of urban sprawl in the United States: auto reliance as compared to natural evolution, flight from blight, and local revenue reliance. *Journal of Policy Analysis and Management*, *27*(3), 536-555.

Weaver, R.C., and Sharmistha, Bagchi-Sen (2013) Spatial analysis of urban decline: The geography of blight. *Applied Geography* 40 (2013) 61-70

Weitz, Jerry, and Crawford, Tom (2012) Where the Jobs are Going: Job Sprawl in U.S. Metropolitan Regions, 2001-2006 *Journal of the American Planning Association*. 78:1 (Winter 2012)

Wheelan, Charles (2013) *Naked Statistics*, W.W. Norton and Company, New York, NY

Wheeler, C. H. (2006). Urban decentralization and income inequality: Is sprawl associated with rising income segregation across neighborhoods?. *Federal Reserve Bank of St. Louis Working Paper Series*, (2006-037).

Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., & Sarzynski, A. (2005). The fundamental challenge in measuring sprawl: which land should be considered?. *The Professional Geographer*, *57*(1), 94-105.

Woo, M., & Guldmann, J. M. (2011). Impacts of urban containment policies on the spatial structure of US metropolitan areas. *Urban Studies*, *48*(16), 3511-3536.

Yang, Jiawen, et al. "Measuring the Structure of US Metropolitan Areas, 1970–2000: Spatial Statistical Metrics and an Application to Commuting Behavior." Journal of the American Planning Association 78.2 (2012): 197-209.

Yang, R., & Jargowsky, P. A. (2006). Suburban development and economic segregation in the 1990s. *Journal of Urban Affairs*, *28*(3), 253-273.

Zhao, Zhenxiang, and Kaestner, Robert (2010) Effects of Urban Sprawl on Obesity Journal of *Health Economics* 29 (2010) 779-797