# Three Essays on Urban Economics

**Lyndsey A. Rolheiser**

B.Sc. Mathematics
University of Alberta, 2005
M.A. Economics
Simon Fraser University, 2008

Submitted to the Department of Urban Studies and Planning
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Urban Economics

at the

Massachusetts Institute of Technology
June 2017



Signature of Author............................ Signature redacted
Department of Urban Studies and Planning
May 24, 2017

Certified by.................................. Signature redacted .....
Associate Professor Albert Saiz
Department of Urban Studies and Planning
Dissertation Supervisor

Accepted by.................................... Signature redacted
Associate Professor Brent Ryan
Chair, PhD Committee
Department of Urban Studies and Planning

# Three Essays on Urban Economics

by
Lyndsey A. Rolheiser

Submitted to the Department of Urban Studies and Planning
on May 24, 2017 in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Urban Economics

## Abstract

The three chapters contained in this dissertation represent a body of work concerned with ubiquitous municipal issues that affect the economic health, vibrancy, and stability of municipalities. These issues are generated through the interaction between agents within the municipality and the built environment of the municipality. The first chapter investigates the role of postwar housing characteristics in neighborhood decline. Extant literature hypothesizes that postwar vintage specific housing characteristics are contributing more to observations of decline than general housing age as the postwar home is no longer aligned with current consumer demand. I address this hypothesis by empirically separating aging and postwar vintage effects at the neighborhood level. Findings indicate previous empirical results linking postwar housing to decline confounded the age and vintage effect. Once separated, the postwar vintage effect is not a significant source of neighborhood decline as housing age is the driving factor.

In the second chapter, I explore the relationship between development patterns and municipal expenditures. Measures that capture the multidimensional aspects of land use patterns exist within the planning and landscape ecology literature but have not been applied to the 'Cost of Sprawl' discourse until now. Using a unique GIS data set covering all of Massachusetts, I construct measures of separation, continuity, centrality, integration, and concentration of residential and commercial land uses within municipalities. Findings suggest some aspects of land use patterns championed by Smart Growth and New Urbanism advocates produce lower levels of municipal expenditures per capita as compared to more sprawling development patterns.

The final chapter focuses on the issue of property tax incidence. With increasing reliance upon commercial property tax revenue, it is important that municipalities fully understand the implications of such reliance especially when it comes to attracting and retaining local business. Existing literature on commercial property tax is limited and only a small handful of studies focus on the issue of commercial property tax incidence. I contribute to this slim literature by asking one question in particular: who does the commercial property tax burden fall upon? Based on data from 96 Massachusetts municipalities over 26 years, I find nearly 100% of the burden is passed through to the renter.

Dissertation Supervisor: Albert Saiz
Title: Associate Professor

# Acknowledgements

This dissertation is the physical representation of the emotional, monetary, and intellectual support I have been blessed to receive over the years. As such, the work in these pages belongs to these support givers. To my parents, Daryl and Leslie, who never flinched when I announced I wanted to switch majors or wanted to pursue yet another level of education. You selflessly moved me across the country and across countries so that I could pursue my studies. To my siblings, Dayna, Lauren, and Derek. Thanks for putting up with me over the years. To Tom, who had the pleasure of meeting me during my doctoral studies and still stuck around in spite of it. Your support during this last year has been remarkable and unconditional. Without exaggeration, I'm not sure I would be in a position to write these acknowledgements in absence of the many meals and pep talks you provided me. To my partners in doctoral crime, Lily, Anthony, and Faizan, who made this challenging experience more humourous and less lonely. To my advisers and mentors, Albert, Bill, David, and Ingrid, who's guidance was exemplar and inspirational. Special thanks to the support team in the CRE as well as DUSP's Sandy and Duncan for all the small, potentially annoying questions you answered over the years. Lastly, to my animals, Mela and Deb, your nonjudgmental silent support was always perfectly timed.

# Contents

# List of Tables

# List of Figures

# Introduction

Municipalities are in a constant struggle over how to manage scare resources in ways that best serve their residents. This becomes especially difficult during times of economic and socioeconomic instability. How does a municipality support their vulnerable populations, promote economic growth, and efficiently manage their resources? The three chapters contained in this dissertation represent a body of work concerned with ubiquitous municipal issues that affect the economic health, vibrancy, and stability of municipalities. These issues are generated through the interaction between agents within the municipality and the built environment of the municipality. Specifically, I investigate topics of neighborhood decline, inefficient land use and municipal expenditures, and commercial property tax incidence.

As housing ages, households make decisions regarding housing maintenance and neighborhood migration. The aggregate result of these decisions may result in neighborhood decline. The influence of aging housing structures on neighborhood socioeconomic decline is well documented in both economics and planning literature. While planning literature recognizes the role of aging, it has also singled out postwar housing and its arguably undesirable characteristics (style, size, quality, and uniformity) as an additional source of decline. It is hypothesized that these postwar vintage specific characteristics are contributing more to observations of decline than housing age as the postwar home is no longer aligned with current consumer demand. I address this hypothesis by empirically separating aging and postwar vintage effects at the neighborhood level.

In the second chapter, I explore the relationship between development patterns and municipal expenditures. Extant literature has difficulty accurately capturing the cost of inefficient land use (or sprawl). Coarse measures of density are typically employed as proxies for inefficient land use as they are assumed to capture economies of scale. However, these measures mask the complex land use patterns that exist across various land use types and scales that may greatly influence the cost of municipal services. Measures that capture the multidimensional aspects of land use patterns exist within the planning and landscape ecology literature but have not been applied within the 'Cost of Sprawl' discourse until now. Using a unique GIS data set covering all of Massachusetts across four decades I construct measures of separation, continuity, centrality, integration, and concentration of residential and commercial land uses within municipalities. Hypotheses regarding what constitutes an efficient land use pattern put forth by Smart Growth and New Urbanism advocates are tested.

In the final chapter, I turn to the issue of property tax incidence. The past few decades have seen enormous amounts of public push-back against increasing residential property tax burdens. In their search for alternative sources of revenue, municipalities have turned to commercial property. With increasing reliance upon commercial tax revenue, it is important that municipalities fully understand the implications of such reliance especially when it comes to attracting and retaining local business. Unfortunately, the literature on commercial property tax is limited and only a small handful of studies focus on the issue of commercial property tax incidence. I contribute to this slim literature by asking one question in particular: who does the commercial property tax burden fall upon? In other words, do owners of capital internalize increases in tax payments or are these increases passed onto the renters of commercial space?

# Chapter 1

# Postwar Housing and Neighborhood Decline

Many postwar neighborhoods across the United States are exhibiting signs of decline previously only observed in central cities. This observation has led to hypotheses regarding the role of postwar housing and neighborhood characteristics in neighborhood decline. While extant empirical literature identifies a relationship between postwar vintage housing and neighborhood decline, the observed relationship confounds vintage with age. I contribute to the literature on postwar housing and neighborhood decline by separately identifying neighborhood-level vintage and age effects while controlling for city-level time effects. A large literature on micro level age and vintage effects exists, but such a distinction between age and vintage has not been operationalized at the neighborhood level. To do so, I develop two novel instrumental variables that model the geographic expansion of cities and the national evolution of housing stock over time. Using a fixed effects panel model covering three decades of geographically-consistent census tracts across American metropolitan areas, I show that postwar neighborhood decline is predominantly associated with the aging of housing structures and cannot be attributed to either postwar neighborhood characteristics or to the differential trends of cities that experienced more development during the postwar era.

## 1.1 Introduction

Postwar housing is aging en masse, much like the baby boomer generation it once sheltered. This housing, built between 1945 and 1969, began entering middle age[1] at the turn of the last century and now represents the largest volume of middle aged housing the United States has ever experienced at one time. The current state of many suburbs and neighbourhoods dominated by this vintage of housing is that of decline; as compared to outer-ring suburbs, 63% of metropolitan postwar suburbs are in decline in terms of income, population, and poverty (Hanlon, 2008). Economic growth within postwar suburbs (or inner-ring suburbs) is dwarfed by that of outer-rings and industry and population loss plagues many of these communities (Hanlon, 2012).

One particular corner of planning literature hypothesizes the decreased demand for the physical characteristics of housing and the typical subdivision layouts that dominated the postwar era as well as general aging of these structures to be one of the main factors contributing to the current observations of decline (Lucy & Phillips, 2000, 2006; Short, Hanlon, & Vicino, 2007; Hanlon, 2007, 2008, 2012). Hanlon (2007) states, "...style, size, and uniformity of postwar housing may be more important to suburban decline than merely the issue of housing age." Lucy and Phillips (2000) are more critical in their characterization of postwar homes:

> ...these [postwar homes] may be starter homes for first-time buyers, but the neighborhood - lacking many children and a school - has some liabilities for young families. Even for childless couples, small unit size makes space for a home office problematic. Since these structures are 40 or more years old, many building elements and operating systems are worn out. They need to be replaced, or they already have been replaced once or twice and are ready for a new replacement round. While the location is not as inconvenient as those of many other neighborhoods, this neighborhood was built at densities low enough that public transportation is not available within walking distance, commuting to work requires a car, and no major institutions are nearby. The architectural quality is mundane or worse. Except for mature trees, aesthetic attributes are scarce.

[1] Colloquially defined as 45 to 65 years of age.

While typological considerations and externalities may be important, many aspects of neighborhood decline can be sourced back to household behavior when faced with depreciating housing at the micro level. Specifically, homeowners face the decision to invest in the maintenance/renovation of their homes or to sell and migrate out of the neighborhood often to newer housing stock built at the urban fringes. This decision may be influenced by the characteristics of a housing structure (Gyourko & Saiz, 2004). Owners of middle-aged homes are in an especially precarious position as their homes are not yet old enough to warrant demolition or wholesale refurbishment, but nevertheless require significant costly maintenance (Rosenthal, 2008).

Factors internal and external to the housing structure influence the decision to reinvest through different channels. Internal factors are tied directly to the housing structure; they can be summarized as age and vintage effects which combine to represent the physical deterioration and functional obsolescence of a housing unit. I will refer to the combination of these effects as the composite structural effect henceforth. Age effects refer to physical wear and tear issues associated with aging whereas vintage effects refer to vintage-specific physical characteristics that may contribute to wear and tear but may also contribute to general desirability of the structure. External factors encompass neighborhood and regional externalities that are beyond the control of the homeowner but which significantly impact decision making. Often, larger scale regional externalities are characterized as time effects that capture shocks common across a geographic region that are experienced over time. These time effects summarize variations in economic and demographic evolutions across geographies.

The hypothesis regarding the negative macro level influence postwar housing has on neighborhood decline has not been explored empirically by the literature; however, the role of housing age in neighborhood decline and housing stock depreciation in general is well documented (Choldin, Hanson, & Bohrer, 1980; Malpezzi, Ozanne, & Thibodeau, 1987; Randolph, 1988; Bier, 2001; A. Goodman & Thibodeau, 1995; Rosenthal, 2008; Brueckner & Rosenthal, 2009; Rosenthal & Ross, 2014).

Here, I investigate whether or not the physical commonalities of postwar housing are contributing to neighborhood decline. More specifically: are we observing a true postwar vintage effect resulting from the characteristics of the postwar built environment?

This paper contributes to the literature on postwar housing and neighborhood decline by identifying vintage versus age effects at the neighborhood level while also accounting for metropolitan dynamics. A large literature on micro level age and vintage effects exists, but to the best of my knowledge, such a distinction between age and vintage has not been operationalized at the neighborhood level. To do so, I develop two novel instrumental variables that model the geographic expansion of cities and the national evolution of housing stock over time. Using a panel dataset covering three decades of geographically-consistent census tracts across American metropolitan areas, and saturating the model with year-and-metro area fixed effects, I show that postwar neighborhood decline is predominantly associated with the aging of housing structures. Thus, decline cannot be generally attributed to either postwar neighborhood characteristics or to the differing growth trends of cities during the 1950s and 1960s.

The paper proceeds with an overview of the history of postwar housing, the origins of neighbourhood decline as well as a discussion of age, vintage, and time effects. Sections 1.4 and 1.5 outline the methodology and data employed. Results are presented in Section 1.6.

## 1.2 Postwar Housing History and Characteristics

Why single out postwar housing characteristics when discussing recent and current observations of neighborhood decline? Some scholars lay blame on the humble postwar house due to the scale of development and homogeneity of this vintage coupled with a number of arguably undesirable structural characteristics. Here, I present an overview of the unique beginnings of this housing vintage as well as a discussion regarding common characteristics.

### 1.2.1 Supply, Demand, and Technological Change

Prior to the end of World War II, new housing starts dropped nearly 80% from 1941's high of 706,000 to 142,000 in 1944. Household formation and birth rates were on an upswing but new construction was limited and existing stock rarely benefited from consistent maintenance as capital and labor were in short supply (Jackson, 1985). By the end of the war, the need for new housing was so substantial the federal government was prompted to immediately underwrite a massive construction program. Supply was further bolstered when the Federal Housing Administration increased mortgage insurance authorization by $1.5 billion (Checkoway, 1980). On the demand side, returning veterans benefited from 4% mortgage rates with little to no down payment required (Jackson, 1985; Hayden, 2006).

Construction innovations and changes in land development also contributed to the increase in housing production. Prior the the late 1940s, the typical real estate developer was small and locally-based. This changed as demand for housing increased dramatically towards the end of the war. In order to address housing shortages, builders adopted mass production techniques utilized in automobile and military equipment production. The effective use of these mass production techniques required a new form of industrial organization as small builders did not have the capacity to implement such techniques. As a result, large regional and even national builders came to dominate residential development. Their foothold was supported and legitimized through the introduction of federal government programs directing aid specifically to large developers. By 1949, 70% of new homes were constructed by only 10% of existing firms (Checkoway, 1980).

The combination of cheap greenfield land, quick production, policy induced supply, and high demand eventually led to a staggering 15 million housing starts across the United States in the 1950s - twice the number of starts as compared to the 1940s and six times the number in the 1930s. As illustrative as the housing start statistic is, it masks the relationship between household formation and housing starts. Figure 1.2 plots both housing starts in 1000s and starts per 1000 households. The relative demand and supply of postwar housing peaked during the postwar era; from 1944 to 1950, starts per 1000 households increased

Every part of the house came prefabricated, even the stairs and fencing (Library of Congress).

Figure 1.1: Levitt, Cape Cod, 1948.

by more than an order of magnitude beginning at 3.7 and jumping to 44.4. While housing starts do not exhibit a downward trend post 1950 (aside from the drop associated with the 2007 housing crash), starts per 1000 households peaked in 1950 and have declined in a cyclical manner ever since.

### 1.2.2 Common Characteristics

Original postwar housing structures and developments displayed a number of consistent physical and social characteristics; this is not surprising given the scale of development by a relatively small number of developers over a short period of time. Jackson (1985) outlines a number of common characteristics. First, developments were often located peripherally to the city. By 1950 national suburban growth rate was ten times that of central cities. Vacant inner city lots were not amenable to mass production techniques and, for many buyers, these locations were undesirable. Second, structures were built at low densities. Between 1946 and 1956, 97% of all new single-family dwellings were completely detached and surrounded

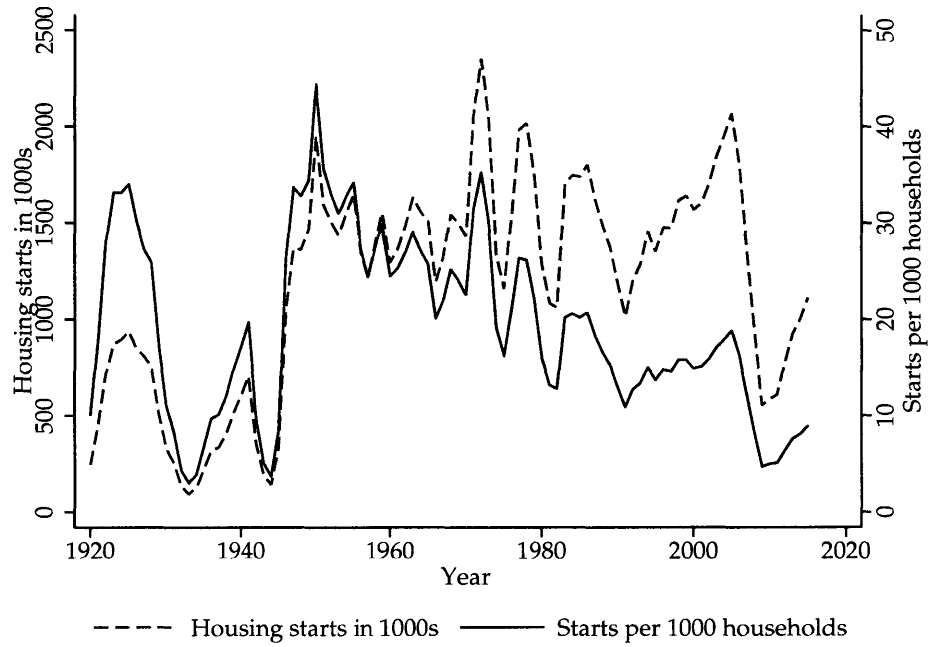



Figure 1.2: US housing starts from 1920 to 2015

on every side by their own plots. The typical uniform lot ranged from one-tenth to one-fifth of an acre. Nearly 80% of new development in the 1950s occurred within the suburbs (Checkoway, 1980). The third characteristic highlights the architectural similarity. Most large, national developers offered no more than a half-dozen basic house plans regardless of where their developments took place. Architectural styles tended to be variations on the simple ranch or cottage home. These one-storey or one and a half-storey homes typically boasted 8-foot ceilings, four to six rooms, two bedrooms, a front picture window, and no basement. Layouts were rectilinear or L-shaped often with an attached carport or single-car garage. This visual homogeneity, coupled with the easy availability of homes contributed to the fourth characteristic, the illusion of income homogeneity. However, Jackson asserts that true economic and racial homogeneity was created as families sorted themselves by income and race.

The simple postwar home and uniform development represented the pinnacle of construction technology and household desires during the postwar era. However, it is argued that postwar housing is no longer aligned with current consumer demand (Short et al, 2007). Preferences for housing have diverged from modest postwar structures to larger exurban compounds.[2]. The average size of a newly constructed single-family home in 1950 was 1,065 square feet. By 2015, the average jumped 155% to 2,721 square feet (American Community Survey). Two-thirds of new homes built in 1950 contained two or fewer bedrooms and only 4% had two or more bathrooms. In contrast, 47% of new single-family homes in 2015 contained four or more bedrooms, with only 10% containing two bedrooms or less. Just as extreme, only 4% of 2015 new-builds offered one and a half bathrooms or less while 38% contained three or more bathrooms.

As argued by Lucy and Phillips (2000, 2006), Short et al. (2007), and Hanlon (2007, 2008, 2012) these changes in housing and neighborhood preferences and, in general, the age of postwar structures may be contributing to neighborhood decline. However, the gap between current household tastes regarding housing amenities and characteristics alone cannot account for the observations of decline in these neighborhoods. If that were true, households would simply renovate and redevelop the housing structures as demand for the land would still exist. It is more likely that this gap, along with socio-economic and market forces influencing neighborhood level trends in housing maintenance, are combining to produce environments of decline. The interaction of these mechanisms is discussed in the following section.

## 1.3 Housing Depreciation and Neighborhood Decline

Neighborhood decline originates from the intersection of social and economic forces affecting housing inventory and neighborhoods. Grigsby et al. (1984) describe the process of decline as changes in social and economic variables causing households to act directly or indirectly through a system of housing suppliers and market intermediaries to make different decisions regarding housing maintenance and upgrading, whether to move, or where to move. These

[2]Data from homeowners who sold than purchased another home in Cleveland, Cincinnati, and Columbus, OH show a preference for buying larger, more expensive homes farther from city centers (Bier & Howe, 1998).

decisions produce changes in housing and neighborhood characteristics that initiate or reinforce existing trends of decline. Thus, the mechanisms associated with decline and the observation of decline are highly endogenous. In this section, I provide theoretical detail and empirical evidence of the process of neighborhood decline with specific focus on the role of housing age and vintage.

### 1.3.1 Housing Depreciation

As the separate life cycles of households and housing units evolve over time, the value of the house to its inhabitants depreciates as the bundle of services comprising the housing unit age and/or change. This bundle of services is not limited to the specific characteristics of the actual housing unit, it is also associated with external features such as neighborhood services and the socioeconomic status of neighbors. As the value of the bundle of services declines, households adjust their housing through repairing, maintaining, and improving current housing, or migrating out of current housing and possibly the current neighborhood in order to slow, stall, or avoid housing depreciation.

Factors affecting depreciation of a home that are internal to the housing unit include physical deterioration and functional obsolescence. Physical deterioration is closely associated with aging but can also be attributed to vintage and maintenance levels as physical deterioration depends on the original quality of construction and subsequent maintenance of the structure (Lusht, 1997; Wilhelmsson, 2008). Without maintenance, housing depreciates quickly given the deterioration of building elements and the wear and tear inflicted on operating systems over time (Chinloy, 1980; Shilling, Sirmans, & Dombrow, 1991; Weiss, 1994; Knight & Sirmans, 1996; Francke & van de Minne, 2016). Thus, physical deterioration is related to the composite structural effect and cannot be singularly associated with a general aging process or the characteristics specific to a housing vintage.

Functional obsolescence occurs when a "structural component is outmoded or inefficient when judged by current market standards" (Lusht, 1997). This would imply that the structural qualities of postwar homes are related not only to physical obsolescence given their age but perhaps also to functional obsolescence. The postwar house is often defined by its outmoded-ness, and thus one would expect functional depreciation to be associated with this vintage. A common example of functional obsolescence is housing size. Current square footage or number of bedrooms may no longer be adequate for a growing household.

Conditions and characteristics external to the housing structure that contribute to depreciation fall under the category of external obsolescence. Specifically, location and economic market conditions can affect how housing depreciates in value. A household may have a hard time justifying repair and maintenance costs when negative neighborhood externalities put downward pressure on the return to such an action (Ahlbrandt Jr & Brophy, 1975; Baer & Williamson, 1988; Gyourko & Saiz, 2004; Hanlon, 2007). Furthermore, as economic market conditions change, so too does the demand and supply of housing. These changes influence the value of housing capital over time and are akin to a time effect. External obsolescence of housing can thus be associated with local neighborhood externalities and broader municipal or regional externalities.

### 1.3.2 Neighborhood Decline

Two competing theories on neighborhood decline are derived from the internal and external forces influencing housing depreciation. The filtering model is related to the internal forces associated with the housing unit, specifically age of the housing stock. The filtering process sees the out-migration of upwardly mobile households away from depreciating housing stock towards newer neighborhoods and housing. Declining housing units are occupied by members of successively lower income groups and minority populations (Grigsby, 1963; Muth, 1973; Sweeney, 1974; Ohls, 1975; Bourne, 1981; Braid, 1984; Varady, 1986; Baer & Williamson, 1988; Weicher & Thibodeau, 1988; Arnott & Braid, 1997). A spatial aspect to filtering arises when forces influencing housing depreciation are common across a particular geography and begin to influence maintenance and out-migration decisions in aggregate. As the majority of stock within a neighborhood ages and deteriorates beyond reasonable repair, it becomes an attractive target for large-scale redevelopment.

External forces associated with the social and economic externalities of the neighborhood and the broader region drive the alternative theory of neighborhood change. Rosenthal (2008) outlines two possible mechanisms: (i) certain types of families may behave in ways that generate social capital and costs for the neighborhood influencing demand for that location, and (ii) families may choose to migrate into or out of a neighborhood based on the socioeconomic characteristics of their prospective neighbors.

Figure 1.3 summarizes the endogenous internal and external forces contributing and enforcing observations of neighborhood decline. Beginning with the actors within a neighborhood, decisions are made by households through a system of housing suppliers and market intermediaries to invest in the maintenance of their home; move out or stay in their home; and if they choose to move, where to move to. These decisions are based on the prospect of housing depreciation and the current socioeconomic characteristics/social capital of the neighborhood. When reinvestment in housing does not occur, depreciation takes hold and prices and rents fall. Large scale neighborhood level depreciation influences the decline in the physical and social environment along with changes in race/ethnicity of the neighborhood.[3] The collection of these changes in housing and neighborhood characteristics influence overall social and economic conditions: population and household income decline and the neighborhood becomes less attractive for business investment resulting in a decline in the consumer and product service mix. After observing these changes, a household must once again make decisions with respect to reinvestment and out-migration.

[3]Once a racial "tipping point" is reached, a neighborhood may become dominated by minority households (Schelling, 1971).

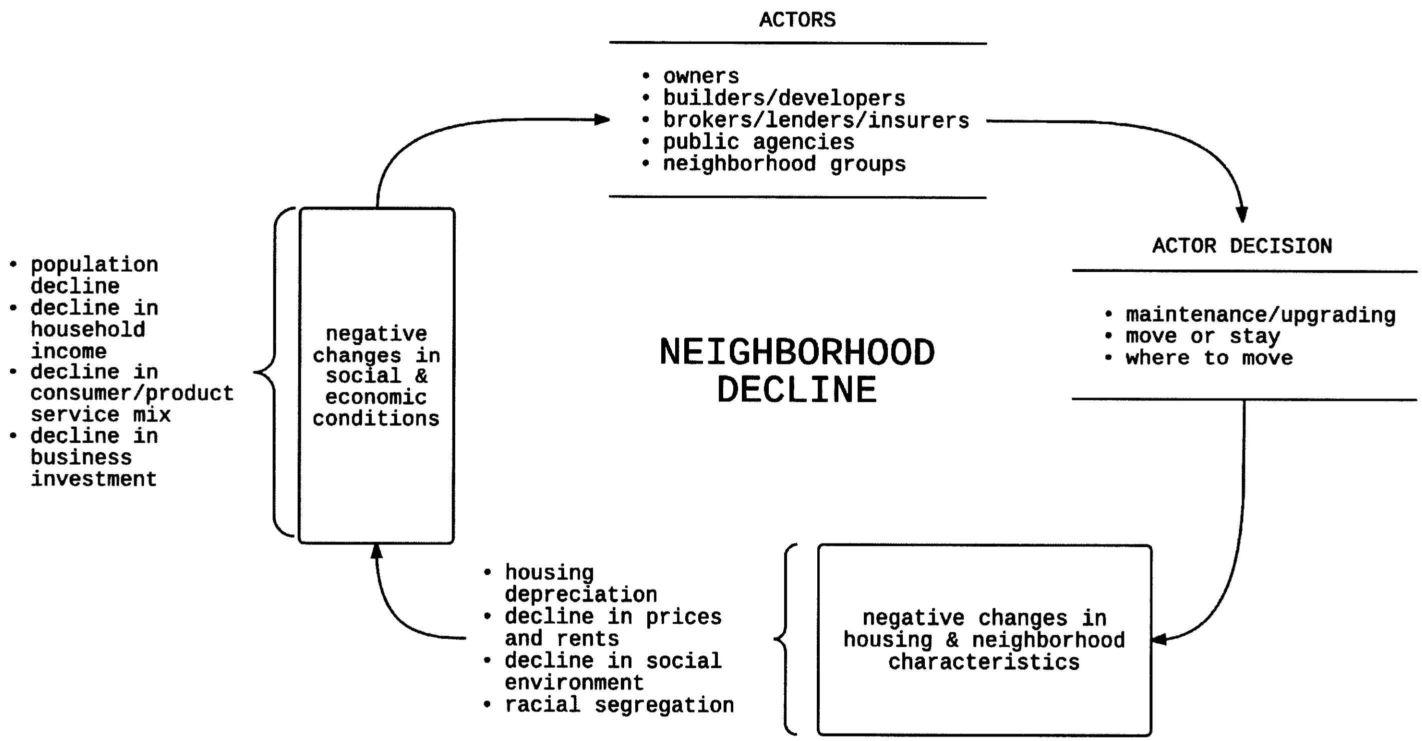



Figure 1.3: Model of neighborhood decline. Adapted from Grigsby et al. (1984)

Rosenthal (2008) provides empirical evidence for both internal and external forces influencing neighborhood decline across 35 American metropolitan areas. Specifically, short run changes in neighborhood economic status are influenced by changes in neighborhood socio-demographic composition and these changes strongly predict the rate at which neighborhoods transition up or down the income distribution. However, it is the housing age distribution that is the persistent source of status change over time. Rosenthal finds a U-shaped relationship between housing age and neighborhood decline; housing aged 10 to 19 years in 1980 has the most negative impact on the change in neighborhood economic status during the 1990s whereas housing aged 40 years or older in 1980 has the most positive impact. The author's favored explanation for this finding leans heavily on filtering and redevelopment theories; specifically, the presence of middle aged housing in 1980 implies that further depreciated middle aged housing will be present in the 1990s as they are not yet old enough for demolition. Lower income households will subsequently occupy the depreciated homes. However, there is a higher probability of redevelopment for the older cohort of housing during the 1990s. The mix of newly rebuilt and refurbished homes attracts higher income families and gentrification ensues.

In a related paper using the same dataset, Brueckner and Rosenthal (2009) focus specifically on the role of housing in the location of neighborhood level gentrification. Existing theories and empirical evidence on the location of gentrification cite proximity to amenities and public transit as main drivers (Brueckner, Thisse, & Zenou, 1999; LeRoy & Sonstelie, 1983; Nechyba & Walsh, 2004; Carlino & Saiz, 2008). Brueckner and Rosenthal's theoretical and empirical findings suggest that amenities and public transit cannot explain systematic changes in location of poor neighborhoods overtime; "gentrification is ultimately driven by the passage of time and the associated aging and obsolescence of the housing stock."

Neither Rosenthal (2008) nor Brueckner and Rosenthal (2009) address the possibility of vintage effects. The resulting coefficients on housing age bins included in their regression specification represent composite structural effect of age and vintage. Both papers do control for time-invariant metropolitan level variation through the use of fixed effects but not for metropolitan specific variation over time.

### 1.3.3 Identifying Age versus Vintage Effects

The identification of the separate effects associated with depreciation is not a straightforward task as the three main effects of interest - age net of maintenance, vintage, and time - are perfectly collinear as time = age + vintage (Sirmans, Macpherson, & Zietz, 2005; Coulson & McMillen, 2008). However, micro level empirical work is somewhat successful in disentangling the collinear effects. The majority of micro level studies focus on constructing housing price indexes or estimating the rate of housing depreciation within a hedonic price model setting.[4] The most common approach to the collinearity problem at the micro level is to ignore it by not accounting for both age and vintage effects within the model (Sirmans et al., 2005). When both effects are considered, specification restrictions are required to avoid collinearity. These restrictions generally take the form of testable non-parametric and polynomial specifications.

Early efforts to estimate depreciation rates employ housing age as a proxy for depreciation (see Malpezzi et al. (1987) for a review). This coarse proxy masks the effects of housing vintage and maintenance on depreciation as these effects tend to be correlated with age. When maintenance is considered, the rate of depreciation is much faster for poorly maintained homes (Chinloy, 1980; Knight & Sirmans, 1996). Furthermore, the omission of maintenance information increases the likelihood of age-related heteroskedasticity as the variance of price increases with age given the variation in maintenance (A. Goodman & Thibodeau, 1995).

Using a restrictive nonlinear transformation on age, Harding, Rosenthal, and Sirmans (2007) account for the effects of age net of maintenance and year of sale. Coulson and McMillen (2008) use a more flexible specification to simultaneously estimate age, vintage, and time effects. Maintenance is not included. A U-shaped age effect is found. Interestingly,

[4] In this setting, time is equivalent to time of sale and absorbs factors related to changes in construction costs and general market conditions.

Coulson and McMillen (2008) find a sharp downward shift in the vintage effect during the 1940s which they state may imply a certain level of cachet for pre-1940s homes. They caution against concluding that these result show "...the cumulative effects of the Great Depression and/or the exigencies of war placed a burden on the supply of new housing, lowering its quality relative to previous cohorts" as the effect they find continues for later vintages perhaps also implying that surviving pre-WWII units are higher quality or newer structures are of lower quality.

Francke and van de Minne (2016) also implement a simultaneous estimation of the various effects associated with housing depreciation. The authors are able to identify the effects of physical deterioration, functional obsolescence and vintage, and external obsolescence by separating land and structure depreciation. To do so, they make the following assumptions: "physical deterioration (age and maintenance), vintage and functional obsolescence (construction year) only affect the structure value and external obsolescence (time of sale) only the land value". Using data from the Netherlands, they find distinct vintage effects and significant influence of maintenance on housing depreciation.

### 1.3.4 Postwar Vintage Effect Evidence

An attempt has not been made to similarly empirically disentangle age and vintage effects with respect to neighborhood decline. This is a worthwhile endeavour as the factors influencing neighborhood level decline may be different from the factors influencing individual housing maintenance decisions.

Different factors influence neighborhood level trends in reinvestment. However, a number of descriptive studies exist within the planning literature that succeed in describing the state of decline at the neighborhood level and argue that the presence of postwar housing influences these observations of decline.

Lucy and Phillips (1997) present the first empirical evidence of a possible link between household income decline and postwar housing. Using census tract data from metropolitan areas in Virginia, they find that census tracts with large amounts of housing units constructed during the 1950s and 1960s display declines in median family income relative to the metro level between 1980 and 1990. Conversely, 60% of census tracts where 40% or more of the housing was built prior to 1940 increased in median family income over the same time period.

Lucy and Phillips (2006) expand their 1997 study of decline in two follow-up studies contained within "Tomorrow's Cities, Tomorrow's Suburbs". The first uses the suburbs of 35 metropolitan areas as a unit of analysis. They classify suburbs as "specializing in housing of a given time period" by comparing the age distributions within the suburb to the distribution of the metropolitan area as a whole. Groupings include six periods of construction: 1990s, 1980s, 1970s, 1960s, 1940s through 1950s, and pre-1940. Within each construction period, suburbs are classified as having 1.5, 2.0, or 2.5 or greater times the amount of housing of that particular vintage as compared to the aggregate metropolitan level. The authors find suburbs with 1.5 times more pre-1940s housing are more likely to increase in median household income relative to the metropolitan area as compared to all other vintages except the 1990s. Suburbs with 1.5 times or more 1960s housing were least likely to increase in relative income with the 1940 to 1959 vintage next least likely.

In the second study, Lucy and Phillips employ geographically consistent census tract data for the decades 1970, 1980, 1990, and 2000 within the metropolitan areas of Atlanta, Chicago, Los Angeles, Philadelphia, and Washington, D.C.. Census tracts are categorized as being old (substantial shares of pre-1940s housing), middle aged (1950s and 1960s), or new (1980s and 1990s). Census tracts with substantial housing from the 1940s are considered ambiguous but are believed to have more in common with middle aged census tracts than old census tracts given that the majority of construction in the 1940s occurred in the latter half of the decade. Census tracts with substantial housing from the 1970s are considered transitional. The authors find that the number of census tracts with relative incomes above the metropolitan average declined over the 1970-2000 period for census tracts specializing in housing built during the 1940s, 1950s, and 1960s. Conversely, census tracts dominated by pre-1940s housing were more likely to have relative incomes higher than the metropolitan average.

Leigh and Lee's (2005) descriptive study of Philadelphia explores differences in decline within various rings of development. Specifically, they address the following questions. Is intra-metropolitan disparity and divergence between the regions increasing over time? Are there significant differences in demographic, socioeconomic, and housing changes found between downtown, inner-city, inner-ring suburb, and outer-ring suburb regions? Geographically consistent census tract data for the years 1970, 1980, 1990, and 2000 are employed to construct the rings of development based on the year housing is built and a GIS process is used to construct the contour set of census tracts. Inner-city regions are defined as census tracts dominated by housing built prior to 1950. Inner-suburbs consist of census tracts dominated by housing built between 1950 and 1969. Outer-ring suburbs are dominated by housing built after 1970. Income inequality is found to be increasing in inner-ring suburbs relative to outer-ring suburbs. Furthermore, they find a decrease in white populations, increase in minority and low-income households, and rising poverty levels within inner-ring suburbs. Downtowns showed increases in white population and decreases in poverty. The authors conclude by asserting that postwar dominated neighborhoods are vulnerable to economic decline.

Leigh and Lee's (2007) follow-up study expands the Philadelphia case study to Atlanta, Cleveland, and Portland. Again, they construct vintage-based rings of development using the Neighborhood Change Database census tract data for 1970, 1980, 1990, and 2000. Implementing separate random effects generalized least squares regressions on each metropolitan region with a factor-analysis constructed measure of decline as the dependent variable, the authors find increasing levels of distress within the inner-ring suburbs of the four MSAs, decreasing levels of distress in outer-ring suburbs, and a mix of decline and improvement in downtowns and inner cities. The level of distress is larger in inner cities than in inner-ring suburbs; however, inner cities within fast growing regions are seeing improvements in their ability to attract upper and middle class households as compared to inner-ring suburbs.

Using 3,428 census designated places for the 100 most populated US urban areas as the unit of analysis, Hanlon (2008, 2012) constructs an index of suburban decline based on population, income, and poverty over the period of 1980 to 2000. Suburbs are ranked via the index of decline; two thirds of suburbs in crisis (first decile) are within older, inner suburbs. Moreover, half of the housing stock in these suburbs was built between 1950 and 1969. Hanlon also finds improvements over the 1980 to 2000 time period in suburbs dominated by housing built prior to 1940. Similar to Lucy and Phillips (1997, 2000, 2006), Hanlon concludes postwar housing lacks the cachet or exclusivity of older vintages and calls for policy to consider housing characteristics as a major factor in suburban decline.

Consistent across the literature summarized here is the confounding of age and vintage. All fail to disentangle the role of age and vintage in the observation of decline as only the relationship between decade of construction and decline is analyzed. It is of no surprise that many neighborhoods dominated by postwar housing are experiencing decline similar to that experienced by older inner city neighborhoods decades prior. This result may be consistent with the life cycle of housing depreciation and the filtering process that commonly ensues. Thus, it is difficult to confidently lay blame on the housing vintage specifically. Furthermore, additional confounding occurs when metropolitan level time trends are not taken into consideration. Metropolitan level exogenous factors like industry decline and housing supply shocks may also be influencing observations of decline within postwar neighborhoods. For example, rust belt cities like Cleveland and Detroit grew tremendously during the postwar era given strong manufacturing industries and have experienced subsequent economic decline. Again, observations of current neighborhood level decline in census tracts dominated by postwar housing are likely the result of a combination of effects outside of housing vintage.

Beyond the issue of age and vintage confounding, failure to account for the endogenous nature of neighborhood decline (see Figure 1.3) may also bias findings. Neighborhoods that were once dominated by postwar housing may experience redevelopment as they age given positive neighborhood shocks and thus the amount of postwar housing may endogenously decrease. The postwar neighborhoods that do not experience redevelopment will likely display endogenously lower levels of socioeconomic indicators. The observed influence of postwar housing on neighborhood socioeconomic status will be downward bias in this case.

## 1.4 Methodology

### Measuring Neighborhood Decline

There exists a wide variety of ways to measure neighborhood decline. Relative or standardized mean or median household income is by far the most common measure used (Lucy & Phillips, 2000, 2006; Hanlon, 2008; Orfield, 2011; Rosenthal, 2008; Brueckner & Rosenthal, 2009; Ellen & O'Regan, 2008, 2011). Other measures include population change, percent of female-headed households as a percentage of all households with children, percentage of children under five years of age in poverty, tax base per household (Hanlon, 2008; Orfield, 2011). I rely on a standardized income[5] variable as the dependent variable associated with neighborhood decline.[6]

### Age, Vintage, and Time Effects at the Neighborhood Level

Some of the techniques used and data available at the micro level that aid in the identification of separate age, vintage, and time effects are not amenable to the neighborhood level. In order to separately identify postwar vintage and age effects at the neighborhood level, I restrict both vintage and age effects to be stationary relative to each other over the time period considered while still controlling for metropolitan-area specific time effects in order to account for city trends. This is a reasonable assumption given Rosenthal's (2008) finding that housing age influences income decline over long time periods. Identification

[5]Using the census tracts across Core Based Statistical Areas (CBSAs) within the United States to represent a metropolitan neighborhood, I calculate the z-score of the census tract with respect to the CBSA average for each of the dependent variables considered.

[6]While median income is arguably a more desirable measure of income given its insensitivity to outliers, the use of normalized census tracts within the dataset I employ, prohibits the calculation of consistent CBSA level medians from which the standardized median would be calculated. The possible biases of mean household income will be kept in mind during the analysis of results.

takes advantage of a panel framework; estimation of the vintage effect relies on the separate movement of age and vintage over three time periods considered. Specifically, the age of the postwar vintage changes over the time; thus, there exists variation in age that is separate from the census tract level variation in postwar vintage. Both OLS and 2SLS specifications are considered.

### 1.4.1 Regression Models

#### Naïve Specifications

As a baseline, the first specification considered regresses standardized income (from census tract $i$, Core Based Statistical Area (CBSA) $m$, time $t$) on the proportion of postwar housing units in a census tract. Standard errors are clustered at the CBSA level. Next, I add geospatial and time controls. Specifically, controls for distance to the downtown, density of housing units (1000 units/sq km), a dummy variable for whether or not the census tract is in the central city within the larger metropolitan area, time and CBSA fixed effects along with the interaction of these fixed effects. Linear and squared distance to downtown variables capture variation in the dependent variables associated with distance to the central business district. The squared term allows for the possibility that the relationship may exhibit diminishing returns to distance. Time invariant CBSA fixed effects control for the influence metropolitan level characteristics may have on socioeconomic status across all census tracts within a given CBSA. These characteristics may include geography, weather, and industry among other omitted and unobserved variables. Metropolitan trends and economic shocks are captured by the interacted CBSA and time fixed effects.

The corresponding naïve OLS specification is as follows:

$$y_{imt} = \alpha + \beta \cdot Postwar_{it} + \lambda D_{it} + \gamma_m \phi_t + \epsilon_{imt} \tag{1.1}$$

where $y_{imt}$ is standardized mean household income; $Postwar_{it}$ is the proportion of postwar housing in a census tract for a given year, $D_{it}$ represents the vector of geospatial controls which includes distance to the CBD, squared distance to the CBD, housing density, and a central city dummy; $\gamma_m$ is a CBSA fixed effect; $\phi_t$ is a year fixed effect; and $\gamma_m\phi_t$ is the interaction of CBSA and year fixed effects. The subscripts are $i$ for the census tract, $m$ for CBSA, and $t$ for the year. $\widehat{\beta}$ is biased in this setting given the omission of housing age variables, i.e. *Postwar* is correlated with the omitted housing age variables and thus is correlated with the error term. A negative $\widehat{\beta}$ is expected given that *Postwar* is absorbing both vintage and housing age effects.

Next, I expand this specification to highlight vintage versus age effects. Specifically, five variables for the proportion of housing units that are aged 10-19 years, 20-29 years, 30-39 years, 40-49 years, and 50 years or older are included (age 0-9 is the omitted baseline variable):

$$y_{imt} = \alpha + \beta \cdot Postwar_{it} + \delta Age_{it} + \lambda D_{it} + \gamma_m \phi_t + \epsilon_{imt} \tag{1.2}$$

where $Age_{it}$ is the vector of housing age variables. Within the panel setting, the housing age distribution and the proportion of postwar homes are identifiable given that the age variation is observed for postwar homes. In other words, cross-sectional variation is exploited to identify the postwar effects whereas longitudinal variation identifies the age effect.

### Endogeneity

In identifying the post war vintage effect, the issue of endogenous redevelopment arises. Census tracts that were once dominated by postwar housing may have experienced redevelopment as they aged and thus the amount of postwar housing may endogenously decrease or the quality of postwar housing may endogenously improve. Reasons for redevelopment are not necessarily observed by the researcher; thus, there is a possibility that the postwar and housing age variables may be correlated with the error term. Postwar homes in relatively less desirable neighborhoods are likely to have a higher survival rate than those in desirable neighborhoods. This implies that the survival of postwar homes in "worse" neighborhoods has the potential to exaggerate the negative vintage effect.

To reduce the negative endogeneity resulting from the inclusion of contemporaneous housing age variables, existing literature (Rosenthal, 2008; Brueckner & Rosenthal, 2009) uses housing age distributions from a previous census year as a proxy. This proxy is cross-sectional and does not address issues of endogenous demolition of stock prior to year of the proxied variable and does not allow for age versus vintage identification. To better address the dynamics associated with redevelopment, I expand upon the proxy by constructing two new instruments based on the evolution of the housing age distribution over time and the spatial growth of cities.

*Housing Age Evolution*

The first instrument considered draws inspiration from Bartik's (1991) Shift-Share measure.[7] The "Age Evolution" instrument provides an expected distribution of housing age at time $t+1$ given a distribution at time $t$. The prediction is based on how the distribution of housing age evolves over time, on average, across the nation. A national average is used to predict census tract level evolution. Issues of endogenous redevelopment are avoided as the housing age distribution is only allowed to evolve according to the national average pattern of evolution of building age during the relevant time period and is not affected by neighborhood or city level factors.

I construct the national pattern of evolution by observing the survival rates of each housing vintage across all census tracts, i.e. what percentage of homes of a particular construction decade remain in the following census year. Average survival rates are recovered from the following regression:

[7]The Shift-Share measure is widely used in labor economics literature to create an exogenous measure of local labor demand. An estimate of local labor demand is produced by applying local industry employment weights to national industry employment growth. The resulting measure is correlated with local labor demand but not with changes in local labor supply.

$$H_{ik}^{t} = \theta_{k}^{t} \cdot H_{k}^{t-1} + \eta_{ik}^{t} \tag{1.3}$$

where $H_{ik}^{t}$ represents the census year $t$ count of housing units in census tract $i$ built in decade $k$. $\theta_{k}^{t}$ represents the survival coefficient from decade $t-1$ to decade $t$ for housing from constuction decade $k$. Construction decades considered include pre-1940, 1940-1949, 1950-1959, 1960-1969, 1970-1979, 1980-1989, 1990-1999, 2000-2009. $t$ covers the census years 1990, 2000, and 2010.

Next, the housing distribution (expected number of units within each construction decade $k$) for each census tract in 1990 is estimated based on $\widehat{\theta}_{k}^{1990}$ from equation 1.3:

$$\widehat{H}_{ik}^{1990} = \widehat{\theta}_{k}^{1990} \cdot H_{k}^{1980} \tag{1.4}$$

Survival coefficients for 2000 are applied to $\widehat{H}_{ik}^{1990}$ to construct the estimated housing distribution for 2000 and likewise for 2010:

$$\widehat{H}_{ik}^{2000} = \widehat{\theta}_{k}^{2000} \cdot \widehat{H}_{ik}^{1990} \tag{1.5}$$

$$\widehat{H}_{ik}^{2010} = \widehat{\theta}_{k}^{2010} \cdot \widehat{H}_{ik}^{2000} \tag{1.6}$$

Future vintage housing unit counts cannot be predicted as a function of past counts. For example, $\theta_{k}^{t}$ for $k = 1980-1989$ and $t = 1990$ does not exist. Likewise for $k = 1990-1999$ and $t = 2000$ as well as $k = 2000-2009$ and $t = 2010$. In these instances the counts are taken as given.

The preceding steps result in an estimated count of housing for each construction decade and each census year based on the national average evolution of housing. From these counts, estimated housing age distributions (0-9, 10-19 years, 20-29 years, 30-39 years, 40-49 years, and 50 years or older) are calculated for each census tract for each census year.

*Historical Growth*

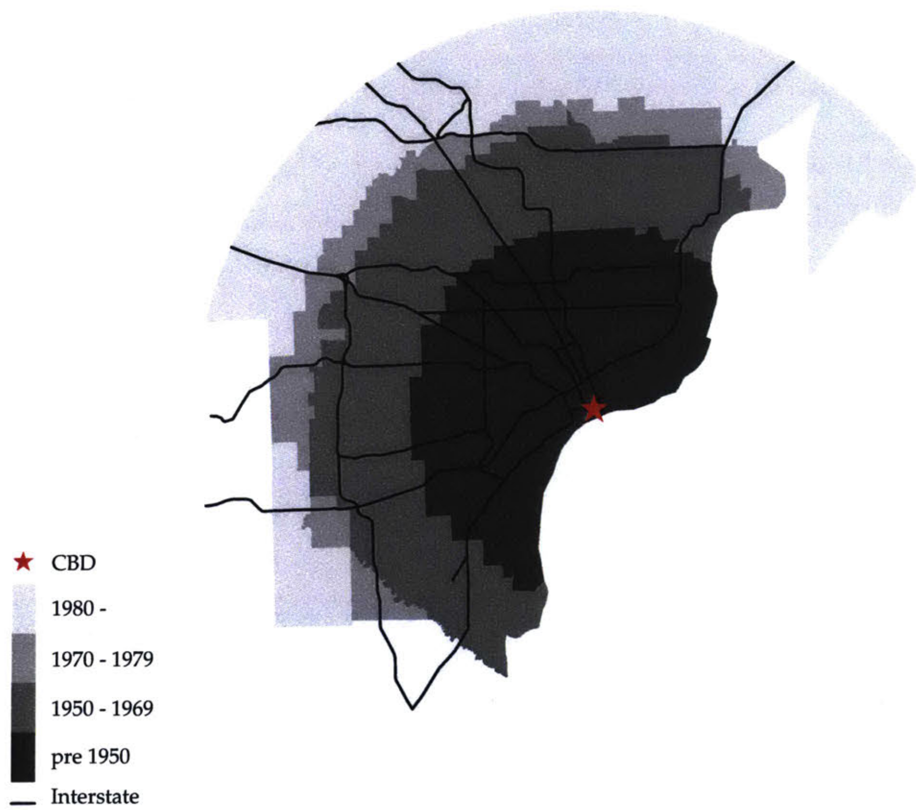
A second instrument focuses on identifying whether or not a census tract was primarily developed during the postwar. Using historical census data from 1950-1980, I observe the urbanized area of 144 major American cities. Urbanized area data consists of both the urban fringe and the central area. Monocentricity of the urbanized area is assumed with increases in urbanized areas over each decade adding another ring to the circle.[8] In other words, I treat urbanized areas as circles centered on downtown coordinates. The growth in the circular boundary from 1950 to 1980 produces rings of housing development. Each new ring is expected to be dominated by a particular vintage of housing. For example, the ring produced by the growth in the circular boundary from 1950 to 1960 is expected to be dominated by 1950-1959 vintage housing.

I make the assumption that as a city's population grows over time, the city expands in an outward fashion. This assumption is troublesome when considering cities adjacent to large bodies of water or cities with development restricting slopes of land. Furthermore, it poses problems for cities that are truly polycentric, e.g. Dallas-Fort Worth. I omit polycentric cities but adjust for land which is not developable due to extreme slopes and cities next to large bodies of water.

It is reasonable to assume that where we expect postwar housing to be the dominant housing stock based on how a city grew over time is uncorrelated with current socioeconomic status and thus redevelopment potential. A visualization of this instrument for Detroit, Atlanta, and Washington DC can be seen in Figures 1.4, 1.5, and 1.6 respectively.

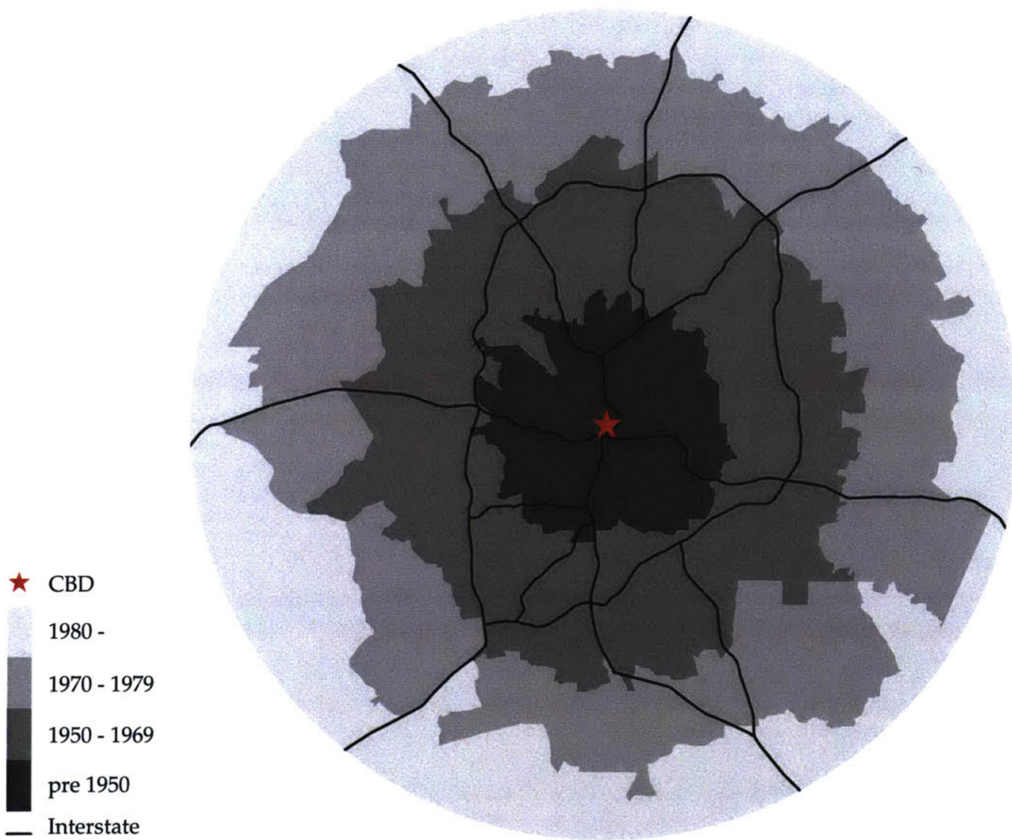
The steps taken to construct the historical growth instrument are as follows. Beginning with the 2010 Core Based Statistical Area (CBSA) shapefile, shapefiles for coastlines and major bodies of water are overlaid to remove water features from the CBSA shapefile. This is necessary as CBSA boundaries do not perfectly adhere to the boundaries of coastlines and major bodies of water. Using elevation data (USGS DEM), land with slope of 15% or

[8] Recent empirical work shows that outward expansion of cities is ongoing and pervasive (Romem, 2016).



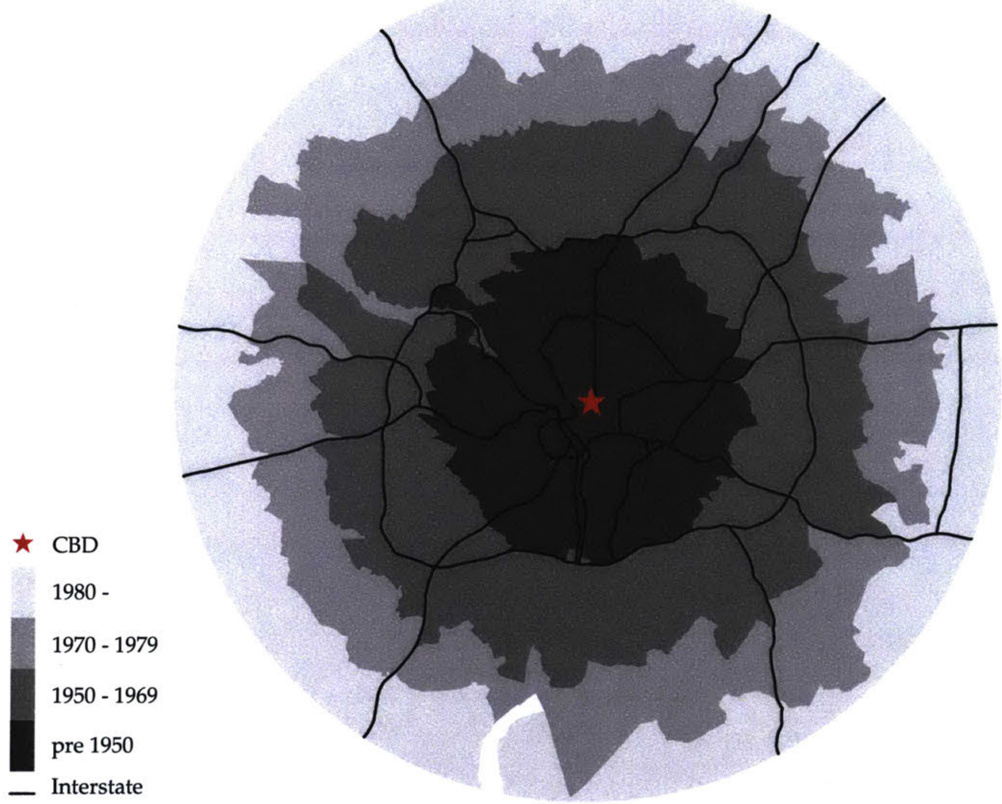

Detroit experienced large growth between 1950 and 1970; however, the larger population boom occurred in the first half of the century. Urbanized area grew from 1,023 square kms in 1950 to 1,281 by 1970. An increase of 121%. Population increased 49% from 2,659,398 to 3,970,584 over the two decades with the majority of both the population and area increase occurring during the 1950s. Growth stagnated after 1970 with population declining by 161,257 and urbanized area increasing 20% to 2,704 square kms.

Figure 1.4: Estimated Development Rings: Detroit




Unlike Detroit, Atlanta's population growth was more consistent across the decades and even more extreme. Population, just 507,887 in 1950, grew by 51% and 53% during the 1950s and 1960s respectively. A total population increase during the postwar period of 131% resulting in a population of 1,172,778 in 1970. Urbanized area growth was just as dramatic with area increasing from 273 square kms to 1,127 square kms during the 1950 to 1970 period. An increase of 312%.

Figure 1.5: Estimated Development Rings: Atlanta




In 1950, Washington D.C.'s population density was 2,786 people per square km. By 1980, density was more than halved to 1,322. Between 1950 and 1970, population nearly doubled from 1,287,333 to 2,481,489 and urbanized area increased from 462 square kms to 1,281 square kms. An increase of 177%.

Figure 1.6: Estimated Development Rings: Washington DC

higher is removed from the CBSA shapefile. This results in a dataset of developable land within each CBSA.

Next, a table of radii of 1 km increments is produced. For each increment, I calculate the amount of developable land area that exists within a circle of the given radius centered on a CBSA's downtown coordinates (see Table 1.1 for the developed areas associated with the first ten 1 km radii increments for the Detroit, MI). To produce development rings corresponding with particular decades of development, historic urbanized land area amounts from the 1950, 1960, 1970, and 1980 census are compared to the developed land areas in the constructed table. The radius that corresponds to the developable area that is the closest match to the historic urbanized area is used to delineate the vintage rings.

Table 1.1: Developable Land for Detroit, MI

| Radius (km) | Ring Area (km2) |
|---|---|
| 1 | 3.14 |
| 2 | 11.1 |
| 3 | 21.7 |
| 4 | 35.7 |
| 5 | 52.8 |
| 6 | 73.6 |
| 7 | 98.3 |
| 8 | 127 |
| 9 | 159 |
| 10 | 195 |

The predicted housing vintage of each census tract is based on the location of the census tract's centroid. Centroids within the pre-1950 radius are labeled as pre-1950. Centroids within the 1950-1959 radius but outside the pre-1950 radius are labeled as 1950-1959, and so on. A dummy variable is constructed that corresponds to the classification of the census tract as being developed between 1950 and 1969. This variable acts as coarse binary instrument for the proportion of postwar housing in a given census tract.

## 1.5 Data & Descriptives

### 1.5.1 Data

Observing consistent neighborhood units over time is imperative to this study; thus, I employ the Neighborhood Change Database (NCDB) as this dataset provides boundary consistent census tract level data for the years 1970, 1980, 1990, 2000, and 2010. To allow for consistent housing age variables (0-9, 10-19, 20-29, 30-39, 40-49, 50 or more years old) I consider only 1990, 2000, and 2010 in the analysis for comparison purposes and data reliability.[9]

Construction of the historical growth instrument requires historical census data as well as a number of spatial datasets. Historic Urbanized Area data is recovered from the Housing and Population Census for the years 1950, 1960, 1970, and 1980. Urbanized Areas are defined differently for census years preceding 1950 and following 1990 and are thus not included for consistency reasons. Downtown longitude and latitude coordinates are obtained from The Census of Retail Trade. Census tract shapefiles are provided by the Census Bureau. Shapefiles for coastline and major bodies of water are provided by ESRI. Digital Elevation Model (DEM) data with a spatial frequency of 30m is retrieved from the United States Geological Survey (USGS).

A balanced panel of census tracts is constructed after dropping census tracts with less than 83 people per square kilometer as this density level is associated with rural developments. This results in 49,522 census tracts representing 273 CBSAs across the years 1990, 2000, 2010. The construction of the historical growth instrument requires the use of historical urbanized area data which is limited to 144 cities. Furthermore, truly polycentric cities like Dallas-Fort Worth are not amenable to the monocentric instrument. With these restrictions, total census tracts drop to 31,356 representing 117 CBSAs across the same time period.

[9]The oldest housing age bin reported by the census is pre-1940; thus, consistent age bins can only be calculated for the years 1990, 2000, and 2010. In 1980, housing built before 1930 would be classified as 50 years or older but this housing vintage cannot be identified within the data.

### Descriptive Statistics

Descriptive statistics are provided in Table 1.2. The mean census tract level income across all years is $72,931 in 2010 dollars. On average, nearly 30% of homes within a given census tract are of the postwar vintage. The housing age distribution is dominated by either very new housing or housing that is 50 or more years old. Housing density averages around 805 homes per square km. The mean homeownership rate is 65% and the mean distance from a census tract centroid to the CBD is just over 20 km.

Table 1.2: Summary Statistics

| Variable | Mean | Std. Dev. | Min. | Max. |
|---|---|---|---|---|
| Mean income | 72,931 | 35,591 | 0 | 550,156 |
| Standardized Income | -0.017 | 1.04 | -4.103 | 11.323 |
| Prop. Black | 0.162 | 0.26 | 0 | 1 |
| Prop. Hispanic | 0.125 | 0.203 | 0 | 1 |
| Prop. College graduates | 0.256 | 0.179 | 0 | 1 |
| Prop. Aged 0-4 | 0.075 | 0.029 | 0 | 0.498 |
| Prop. Aged 5-17 | 0.175 | 0.055 | 0 | 0.553 |
| Prop. Aged 18-64 | 0.628 | 0.073 | 0 | 1 |
| Prop. Aged 65+ | 0.123 | 0.078 | 0 | 1 |
| Prop. Postwar homes | 0.288 | 0.201 | 0 | 1 |
| Prop. Homes aged 0-9 yrs | 0.167 | 0.197 | 0 | 1 |
| Prop. Homes aged 10-19 yrs | 0.16 | 0.147 | 0 | 1 |
| Prop. Homes aged 20-29 yrs | 0.156 | 0.129 | 0 | 1 |
| Prop. Homes aged 30-39 yrs | 0.152 | 0.128 | 0 | 1 |
| Prop. Homes aged 40-49 yrs | 0.118 | 0.116 | 0 | 1 |
| Prop. Homes aged 50+ yrs | 0.247 | 0.262 | 0 | 1 |
| Homes per square km (1000s) | 0.805 | 1.195 | 0 | 41.336 |
| Homeownership rate | 0.649 | 0.232 | 0 | 1 |
| Distance in km to downtown | 22.804 | 19.253 | 0.003 | 162.689 |
| N | | 97,395 | | |

### 1.5.2 Postwar Housing in Detail

Prior to analyzing regression results, I explore the relationship between postwar housing and socioeconomic variables endogenously related to income.

The overall distribution of the proportion of postwar housing across all census tracts is bimodal (see Figure 1.7): many tracts have little to no postwar housing; however, 50% of census tracts have proportions ranging from 0.14 to 0.42.

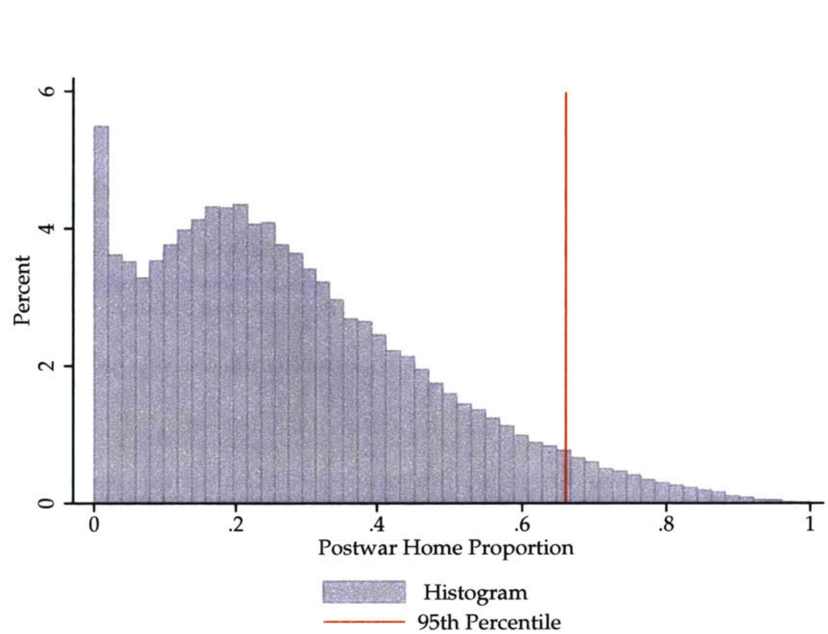



Figure 1.7: Histogram of Postwar Home Proportion within Census Tract

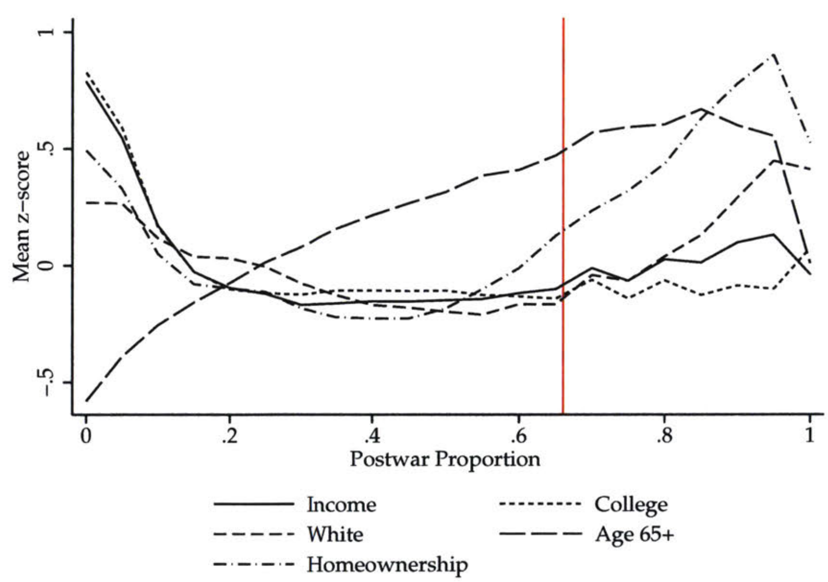



Figure 1.8: Neighborhood Socioeconomic Status and Proportion of Postwar Homes

As a first step in examining the relationship between neighborhood status and postwar housing, I plot the within city mean z-scores of the socioeconomic variables of interest for incremental levels of postwar proportion. In other words, for each postwar proportion interval of length 0.05, the mean z-score across all census tracts with postwar housing within this interval is calculated (Figure 1.8). The z-score of a census tract provides a relative measure of the variable of interest with respect to the city as a whole. Unsurprisingly, there is an initial drop in income, proportion white, proportion college educated, and homeownership rate as the proportion of postwar housing increases from 0. It is unclear if this drop is attributable to postwar housing, housing age distribution, or a number of spatial characteristics not controlled for in this simple plot; however, it is likely strongly correlated with the presence of new housing and the lack of older housing. What is surprising is the dramatic rise in proportion people aged 65 and over as well as the homeownership rate as the proportion of postwar homes increases. Proportion college educated remains somewhat flat for higher values of postwar proportion with income increasing marginally.

It should be noted that only 5% of the data contain postwar proportions greater than 0.66. The trends in the socioeconomic variables diverge substantially from patterns exhibited in the lower 95% of the data. Housing in this percentile is predominately owner-occupied and residents are whiter, older, and of moderately higher income as compared to the city as a whole. Given the high proportion of postwar housing in these census tracts, it is likely they were originally part of various large-scale postwar developments (LPWD) that sprung up during the postwar era. The suburbanization of urban populations at this time was influenced by a number of factors including white flight. In an attempt to explain the survivorship bias and the current socio-economic characteristics of these LPWDs, I provide a cursory exploration of the relationship between metro-level diversity, racial interaction, and the existence of LPWDs following the main results in Section 1.6. A logit model is employed to analyze correlation between the probability a census tract is an LPWD and measures of metropolian diversity and racial interaction.

For the main analysis presented in the paper, I opt to winsorize the top 5% due to the strong non-random characteristics displayed in these tail census tracts.

### Other Housing Vintages

Similar figures to Figure 1.8 are presented for housing between 1990-2009, 1970-1989, and prior to 1950 in Figures 1.9, 1.10, and 1.11 respectively. Trends in the socioeconomic variables are drastically different across the housing vintages. As the proportion of homes built between 1990 and 2009 increases with respect to all other vintages, income, homeownership, and college education steadily increase. The proportion of people 65 years and over displays a similar slight increase then gradual decrease across both vintages. The mean z-score for proportion of white residents in a census tract flattens out as the proportion of newer housing increases but continues to gradually increase when the proportion of housing built between 1970 and 1989 increases. These patterns are in stark contrast to older home dynamics seen in Figure 1.11. Here, there is a steady decline of income, white proportion, and homeownership while college flattens out and the proportion of residents 65 years and over increases then decreases as the proportion of homes built prior to 1950 increases.

## 1.6 Results

Table 1.3 contains the results of the OLS baseline specification, OLS models from equations (1.1), (1.2), followed by the IV version of equation (1.2) using the full collection of instruments. First stage results (Table 1.4) are strong and the null hypothesis of weak instruments from the Sanderson-Windmeijer F-test for multiple endogenous variables is soundly rejected for all instruments.[10]

Recalling from Figure 1.8, the relationship between standardized income and postwar proportion was nonlinear but predominately downwards sloping. This is represented by the large negative coefficient of -0.938 in the baseline specification in equation 1.1. The addition

[10]Unlike the OLS estimator, the 2SLS estimator may be biased in small samples with this bias increasing when instruments are "weak". Instruments are considered weak when the correlation with endogenous regressors is low.

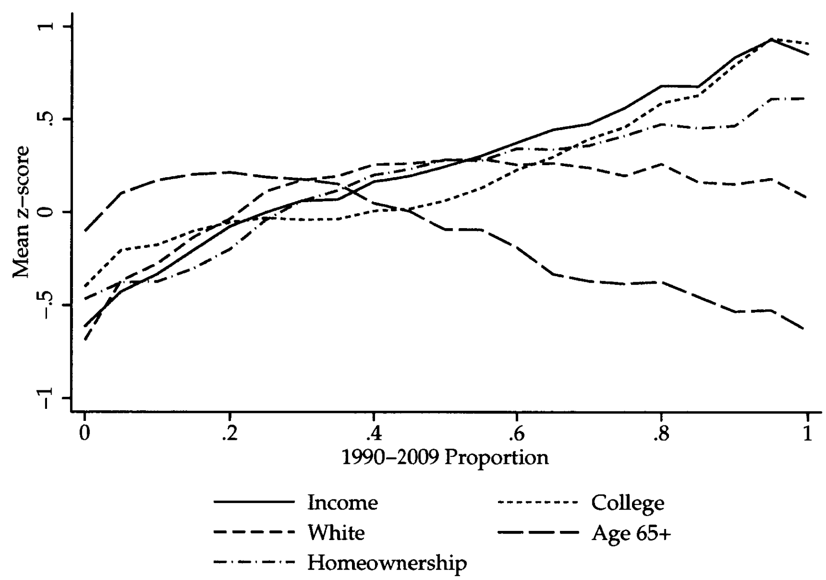



Figure 1.9: Neighborhood Socioeconomic Status and Proportion of 1990-2009 Homes

of spatial controls in equation 1.2 decreases the magnitude of the postwar coefficient slightly. Specifically, if the proportion of postwar housing increases by 0.10, income decreases by -0.0890 of a standard deviation. Using the mean income ($72,707) and the mean standard deviation of income within CBSAs ($35,785), this is equivalent to the mean income dropping by $3,185. The large significant negative coefficient on postwar housing is expected as the omission of housing age conflates the vintage and age effect. From this specification, it is unclear whether it is the age of the home or the vintage that is negatively impacting income levels.

Central city census tracts experience a significant overall decrease in income with respect to the mean income as compared to census tracts from the surrounding suburban cities. The coefficient of -0.224 translates to a drop in income of $8,016. Housing density displays the expected sign as increased housing density decreases the standardized household income by 0.211 of a standard deviation. In other words, increasing the number of housing units within a square kilometer by 100 decreases the mean income of $72,707 by $755. Distance from the central business district (CBD) has an overall small positive effect on income but

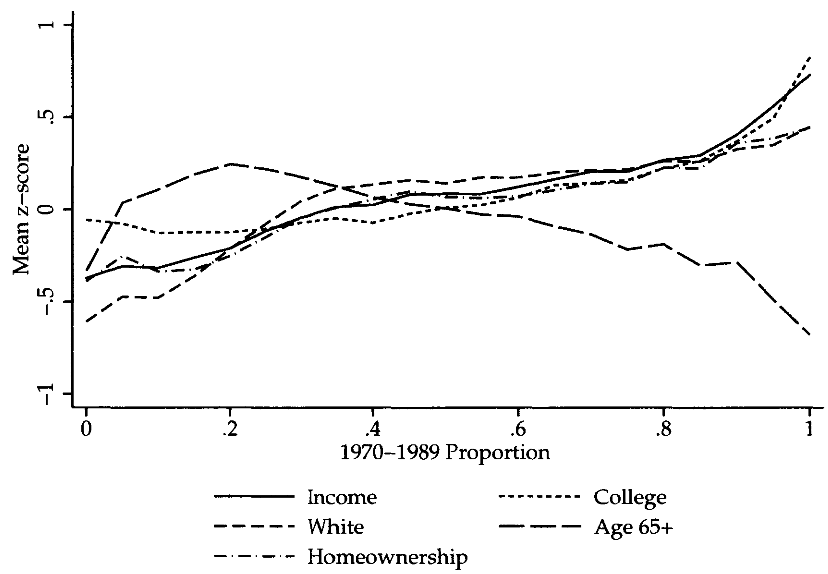



Figure 1.10: Neighborhood Socioeconomic Status and Proportion of 1970-1989 Homes

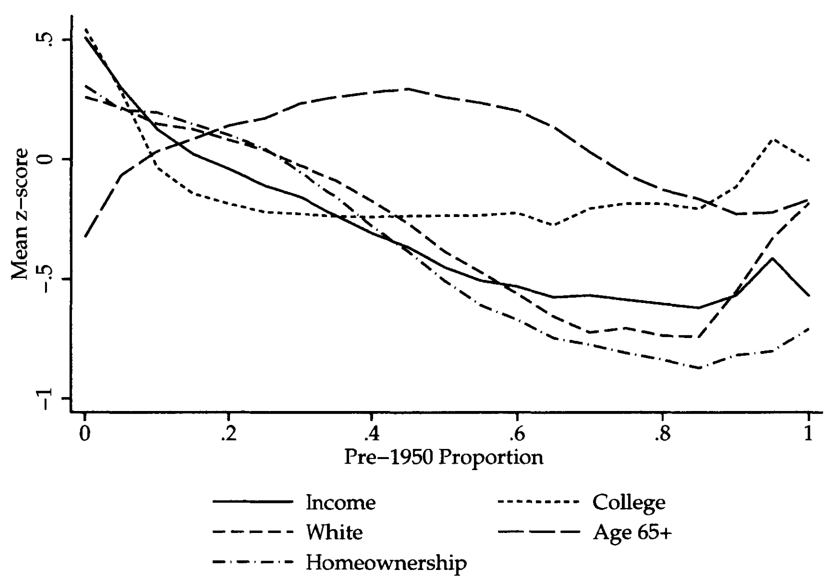



Figure 1.11: Neighborhood Socioeconomic Status and Proportion of Pre-1950 Homes

at large distances there are diminishing returns as represented by the negative coefficient on the squared distance to the CBD. Taking the mean distance of 23 km, an increase to 24 km increases the standardized income by 0.017 standard deviations which is equivalent to an increase of \$616 above the mean income of \$72,707.

Next, I include housing age distribution variables in the specification. The significant negative coefficient on percentage of postwar homes is now significantly positive whereas housing age variables are all significantly negative and follow a monotonically decreasing pattern. This pattern provides evidence of filtering taking place at the census tract level. As housing ages, occupying households decline in income relative to households occupying new housing.

The coefficients on central city and housing density remain negative and significant but are dampened slightly. This is expected as the housing age variables are likely picking up on spatial characteristics common across particular housing ages. The correlation between age and distance is enough to render the distance variable insignificant; however, there still exists a significant negative effect from very large distances albeit quite small. This simple OLS model is quite revealing as even without controlling for endogenous redevelopment, the inclusion of housing age reverses the sign on the postwar coefficient.

When endogenous redevelopment is controlled for within the IV specification, the postwar variable increases slightly but is rendered insignificant. Housing age distribution coefficients show a slight decrease in magnitude but the filtering effect is still present. There are no significant changes in the remaining control variables.

An investigation of the residuals from all four specifications in Table 1.3 allows for a visual interpretation of the fit of the models. Through the use of the full set of controls and the IV specification, I am able to partially explain the original nonlinear relationship displayed between household income and postwar housing proportion at the neighborhood level. Figure 1.12 graphs the median residuals from the four specifications against postwar proportion. Unsurprisingly, the model with no controls displays severe underestimation at low and high levels of postwar housing. Adding in spatial and temporal controls (central

Table 1.3: OLS & IV models with standardized income dependent variable

| | (1)<br>OLS 1 | (2)<br>OLS 2 | (3)<br>OLS 3 | (4)<br>IV |
|---|---|---|---|---|
| Postwar homes | -0.938*** | -0.890*** | 0.112*** | 0.115 |
| | (0.032) | (0.034) | (0.037) | (0.344) |
| Homes aged 10-19 yrs | | | -0.470*** | -0.260*** |
| | | | (0.029) | (0.035) |
| Homes aged 20-29 yrs | | | -1.093*** | -1.018*** |
| | | | (0.033) | (0.121) |
| Homes aged 30-39 yrs | | | -1.409*** | -1.105*** |
| | | | (0.037) | (0.177) |
| Homes aged 40-49 yrs | | | -1.700*** | -1.656*** |
| | | | (0.039) | (0.294) |
| Homes aged 50+ yrs | | | -1.952*** | -1.919*** |
| | | | (0.039) | (0.059) |
| Central city | | -0.224*** | -0.149*** | -0.148*** |
| | | (0.018) | (0.017) | (0.017) |
| Homes/$km^2$ (1000s) | | -0.211*** | -0.152*** | -0.150*** |
| | | (0.014) | (0.011) | (0.012) |
| Distance to CBD (km) | | 0.0172*** | 0.000397 | -0.000294 |
| | | (0.001) | (0.001) | (0.001) |
| $Distance^2$ to CBD | | -0.000217*** | -6.75e-05*** | -6.05e-05*** |
| | | (0.000) | (0.000) | (0.000) |
| Constant | 0.225*** | 0.203** | 1.201*** | 1.090*** |
| | (0.010) | (0.079) | (0.070) | (0.076) |
| CBSA X Year FEs | | X | X | X |
| Observations | 91,948 | 91,948 | 91,948 | 91,948 |
| R-squared | 0.023 | 0.124 | 0.205 | 0.204 |

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

city, density, distance to CBD, interacted year and CBSA fixed effects) has a marginal effect; however, the addition of housing age controls greatly reduces the magnitude of the residuals for all levels of postwar proportion. The IV model slightly decreases the magnitude of residuals for postwar proportions lower than 0.33 but increases the magnitude for proportions larger than 0.33.

Table 1.4: First Stage Regression Results

| | (1)<br>Postwar Homes | (2)<br>Homes 10-19 yrs | (3)<br>Homes 20-29 yrs | (4)<br>Homes 30-39 yrs | (5)<br>Homes 40-49 yrs | (6)<br>Homes 50+ yrs |
|---|---|---|---|---|---|---|
| Postwar Ring | 0.0571*** | -0.00295*** | -0.00108* | 0.00659*** | 0.00811*** | -0.000451 |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Inst Age 10-19 | 0.0137*** | 0.969*** | 0.0664*** | -0.0508*** | -0.0338*** | -0.0597*** |
| | (0.004) | (0.005) | (0.004) | (0.002) | (0.002) | (0.002) |
| Inst Age 20-29 | 0.307*** | -0.106*** | 0.975*** | 0.0467*** | -0.0341*** | -0.0383*** |
| | (0.006) | (0.003) | (0.004) | (0.003) | (0.002) | (0.002) |
| Inst Age 30-39 | 0.526*** | -0.156*** | 0.0170*** | 0.981*** | 0.0559*** | -0.0684*** |
| | (0.006) | (0.003) | (0.003) | (0.004) | (0.003) | (0.003) |
| Inst Age 40-49 | 0.842*** | -0.146*** | -0.0364*** | 0.0613*** | 0.990*** | -0.0307*** |
| | (0.008) | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| Inst Age 50+ | 0.104*** | -0.152*** | -0.0744*** | -0.0510*** | 0.00332*** | 1.066*** |
| | (0.004) | (0.002) | (0.001) | (0.001) | (0.001) | (0.003) |
| Central city | -0.0104*** | 0.000448 | -0.00126** | -0.00184*** | -0.00221*** | -0.00394*** |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Homes/km$^2$ (1000s) | -0.00910*** | 0.00103*** | 0.00334*** | 0.00402*** | 0.000321 | -0.00160*** |
| | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Distance to CBD (km) | -0.000613*** | 3.79e-05 | 5.67e-05 | 0.000444*** | 0.000178*** | -0.000305*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Distance$^2$ to CBD | 4.06e-06*** | -1.40e-06*** | 1.24e-06*** | -1.58e-06*** | -5.27e-07 | 6.09e-06*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Constant | 0.104*** | 0.111*** | 0.0269*** | 0.0225*** | 0.00659** | 0.0284*** |
| | (0.011) | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| CBSA X Year FEs | X | X | X | X | X | X |
| F-stat | 984.78 | 22,205.14 | 1,572.02 | 1,495.04 | 1,096.19 | 3,635.25 |
| Observations | 91,948 | 91,948 | 91,948 | 91,948 | 91,948 | 91,948 |
| R-squared | 0.511 | 0.801 | 0.748 | 0.738 | 0.781 | 0.9152 |

All models include year, CBSA, and year-by-CBSA fixed effects
Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1.13 plots the median residuals from the IV specification along with the 25th and 75th percentiles. The interquartile range is decently situated around 0 indicating insignificance; a slight nonlinearity for high and low values of postwar proportion is present. The likely cause of this nonlinearity is the omission of neighborhood socioeconomic controls that are highly correlated with observed levels of income and certain levels of postwar housing. These controls are endogenous within the current model. However, it is useful to separately analyze their relationship with postwar housing to help unpack the results contained in Table 1.3.

Household income is known to be highly correlated with race, education, and homeownership; thus, it is unclear what is driving the nonlinearities in the tails. I cannot adequately control for these variables within the current specification and do not favor the inclusion of lagged socioeconomic variables as proxies for contemporaneous variables given the likely complex endogenous channels that exist between income, housing age, housing vintage, and socioeconomic variables. Instead, I regress standardized versions of each socioeconomic variable on the same spatial and housing age controls used in the income regression and I allow postwar proportion to enter nonparametrically in order to capture the nonlinearities present in the residual plot. Postwar proportion is represented by eight bins with a baseline bin of 0-0.05. The regression is as follows:

$$y_{imt} = \alpha + \rho PW_{it} + \delta Age_{it} + \lambda D_{it} + \gamma_m \phi_t + \epsilon_{imt} \tag{1.7}$$

where $PW_{it}$ is the vector of dummy variables associated with the proportion bins [.05-.1), [.1-.2), [.2-.3), [.3-.4), [.4-.5), [.5-.6), and [.6-.66). Again, $D_{it}$ represents the vector of geospatial controls and $Age_{it}$ is the vector of housing age variables. Unfortunately, the instrument for postwar proportion cannot be utilized within the nonparametric specification given its binary nature.

Tables 1.5 and 1.6 present the OLS and IV specifications of standardized dependent variables for proportion of white residents, homeownership rate, proportion college edu-

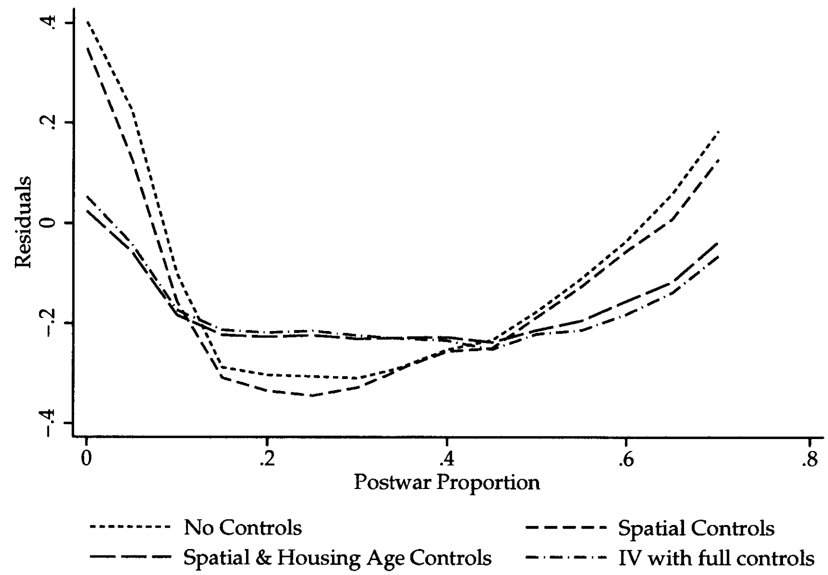



Figure 1.12: Median Residuals for OLS and IV Models

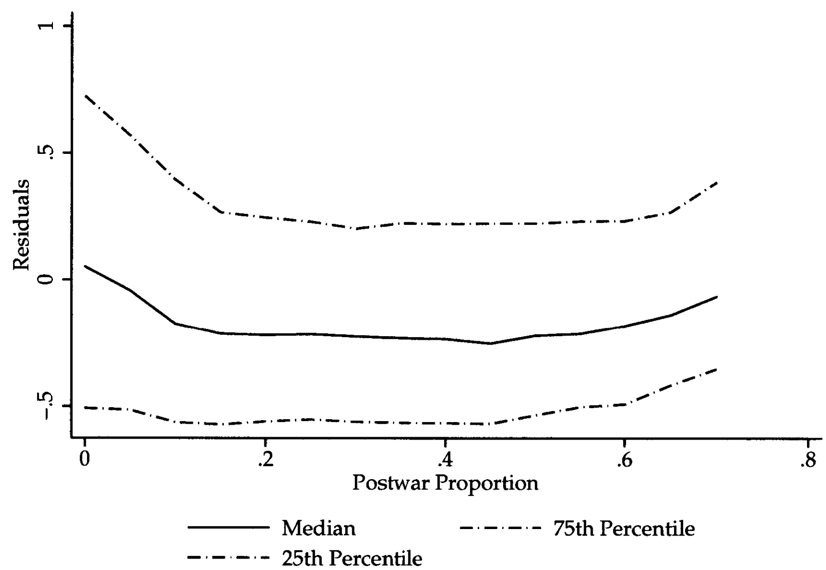



Figure 1.13: Median, 25th, and 75th Percentile Residuals for IV Model

cated, and proportion of residents 65 years or older. The IV specification instruments only the housing age variables using the housing age evolution instrument.

The IV regression with the standardized white resident proportion as the dependent variable shows a strong increasing relationship between postwar proportion and proportion white. As the proportion increases beyond 0-0.05, census tracts exhibit higher proportions in white population than the CBSA average. The last bin has the largest jump with a coefficient of 0.191. The homeownership regression displays a flat effect after an initial decrease with respect to the postwar bins. Homeownership spikes for census tracts containing 60 to 66% postwar housing as compared to census tracts with 0 to 5% postwar housing.

The relationship between postwar housing and college education is quite flat. Census tracts with more than 5% postwar housing display consistent negative coefficients indicating that there is a significant difference between census tracts with virtually zero postwar housing and those with positive amounts but very little difference as the proportion of postwar housing increases. In contrast, the proportion of residents aged 65 years and old displays a strong increasing relationship with respect to postwar housing. For census tracts with postwar proportions between 0.05 and 0.2, the proportion of residents age 65 or older is not significantly different from the baseline proportion.

Taking these regressions together, the nonlinearities observed in the residual plot in 1.13 can be linked to the variables analyzed here. The initial drop in income may be linked to the decline in college educated residents along with low levels of homeownership. The increase in income for census tracts with higher proportions of postwar housing may be associated with the increase in white population and homeownership in the tail. The effect of increases in older populations on income is likely ambiguous.

## 1.7 Large-Scale Postwar Developments and Segregation

Section 1.5 demonstrates that census tracts with extremely high proportions of postwar housing tend to have higher levels of owner-occupied housing than compared to the city on average. These census tracts also have higher household incomes, more residents over the age of 65, lower proportions of college graduates and higher proportions of white residents

Table 1.5: Proportion white & homeownership

| VARIABLES | (1) White OLS | (2) White IV | (3) Own OLS | (4) Own IV |
|---|---|---|---|---|
| Postwar [.05-.1) | 0.0146 | 0.0346** | -0.111*** | -0.110*** |
| | (0.014) | (0.014) | (0.016) | (0.016) |
| Postwar [.1-.2) | -0.00965 | 0.0321** | -0.241*** | -0.239*** |
| | (0.014) | (0.014) | (0.016) | (0.016) |
| Postwar [.2-.3) | -0.0117 | 0.0392** | -0.257*** | -0.258*** |
| | (0.016) | (0.016) | (0.017) | (0.018) |
| Postwar [.3-.4) | 0.00985 | 0.0588*** | -0.235*** | -0.242*** |
| | (0.018) | (0.018) | (0.019) | (0.019) |
| Postwar [.4-.5) | 0.0454** | 0.0934*** | -0.176*** | -0.187*** |
| | (0.020) | (0.021) | (0.021) | (0.021) |
| Postwar [.5-.6) | 0.0642*** | 0.112*** | -0.0314 | -0.0463** |
| | (0.023) | (0.023) | (0.022) | (0.022) |
| Postwar [.6-.66) | 0.141*** | 0.191*** | 0.158*** | 0.141*** |
| | (0.029) | (0.029) | (0.026) | (0.026) |
| Homes aged 10-19 yrs | 0.0376* | 0.120*** | -0.171*** | -0.0256 |
| | (0.022) | (0.024) | (0.027) | (0.028) |
| Homes aged 20-29 yrs | -0.273*** | -0.331*** | -0.654*** | -0.717*** |
| | (0.027) | (0.032) | (0.032) | (0.036) |
| Homes aged 30-39 yrs | -0.363*** | -0.240*** | -0.521*** | -0.355*** |
| | (0.035) | (0.039) | (0.036) | (0.039) |
| Homes aged 40-49 yrs | -0.707*** | -0.954*** | -0.0736* | -0.0210 |
| | (0.043) | (0.055) | (0.040) | (0.047) |
| Homes aged 50+ yrs | -0.723*** | -0.920*** | -0.799*** | -0.820*** |
| | (0.033) | (0.036) | (0.038) | (0.040) |
| Central city | -0.442*** | -0.431*** | -0.452*** | -0.449*** |
| | (0.017) | (0.017) | (0.016) | (0.016) |
| Homes/km$^2$ (1000s) | -0.0304*** | -0.0239*** | -0.264*** | -0.262*** |
| | (0.007) | (0.007) | (0.013) | (0.013) |
| Distance to CBD (km) | 0.0345*** | 0.0317*** | 0.0292*** | 0.0284*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Distance$^2$ | -0.000251*** | -0.000227*** | -0.000240*** | -0.000233*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Constant | -0.0940 | -0.0637 | 0.358*** | 0.327*** |
| | (0.076) | (0.076) | (0.072) | (0.072) |
| CBSA X Year FEs | X | X | X | X |
| Observations | 91,948 | 91,948 | 91,948 | 91,948 |
| R-squared | 0.246 | 0.244 | 0.350 | 0.349 |

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.6: College educated proportion and proportion aged 65+

| VARIABLES | (1)<br>College OLS | (2)<br>College IV | (3)<br>Aged 65+ OLS | (4)<br>Aged 65+ IV |
|---|---|---|---|---|
| Postwar [.05-.1) | -0.305*** | -0.309*** | -0.0303 | -0.00843 |
| | (0.017) | (0.017) | (0.019) | (0.019) |
| Postwar [.1-.2) | -0.411*** | -0.412*** | -0.0121 | 0.0281 |
| | (0.017) | (0.018) | (0.019) | (0.019) |
| Postwar [.2-.3) | -0.414*** | -0.424*** | 0.0443** | 0.0942*** |
| | (0.019) | (0.019) | (0.020) | (0.021) |
| Postwar [.3-.4) | -0.388*** | -0.413*** | 0.0875*** | 0.139*** |
| | (0.021) | (0.021) | (0.022) | (0.022) |
| Postwar [.4-.5) | -0.388*** | -0.426*** | 0.0987*** | 0.151*** |
| | (0.022) | (0.023) | (0.023) | (0.024) |
| Postwar [.5-.6) | -0.379*** | -0.427*** | 0.139*** | 0.193*** |
| | (0.024) | (0.025) | (0.026) | (0.027) |
| Postwar [.6-.66) | -0.339*** | -0.393*** | 0.181*** | 0.235*** |
| | (0.029) | (0.030) | (0.031) | (0.032) |
| Homes aged 10-19 yrs | -0.421*** | -0.272*** | 0.779*** | 0.388*** |
| | (0.026) | (0.028) | (0.033) | (0.037) |
| Homes aged 20-29 yrs | -0.719*** | -0.622*** | 1.176*** | 0.970*** |
| | (0.030) | (0.035) | (0.037) | (0.044) |
| Homes aged 30-39 yrs | -1.138*** | -0.898*** | 1.690*** | 1.507*** |
| | (0.036) | (0.040) | (0.043) | (0.050) |
| Homes aged 40-49 yrs | -1.459*** | -1.332*** | 2.521*** | 2.227*** |
| | (0.041) | (0.051) | (0.047) | (0.057) |
| Homes aged 50+ yrs | -1.686*** | -1.763*** | 1.152*** | 0.854*** |
| | (0.039) | (0.041) | (0.034) | (0.038) |
| Central city | 0.166*** | 0.171*** | -0.119*** | -0.110*** |
| | (0.018) | (0.018) | (0.017) | (0.017) |
| Homes/km$^2$ (1000s) | 0.0324*** | 0.0360*** | -0.00199 | 0.00292 |
| | (0.007) | (0.007) | (0.006) | (0.006) |
| Distance to CBD (km) | -0.0109*** | -0.0127*** | 0.00370*** | 0.00255** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Distance$^2$ | 2.95e-05*** | 4.56e-05*** | 2.87e-05** | 3.86e-05*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Constant | 1.441*** | 1.404*** | -1.361*** | -1.157*** |
| | (0.072) | (0.072) | (0.074) | (0.075) |
| CBSA X Year FEs | X | X | X | X |
| Observations | 91,948 | 91,948 | 91,948 | 91,948 |
| R-squared | 0.135 | 0.134 | 0.121 | 0.120 |

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

than compared to their respective city averages. In this section, I explore city level characteristics that are correlated with the existence of these highly postwar vintage census tracts: what relationships may help explain the observations of socioeconomic status as well as the persistence of the high proportion of postwar housing over many decades?

I hypothesize the observations of socioeconomic status within these census tracts may be linked to spatial patterns, prevalence, and persistence of segregation within cities. Current racial homogeneity within large-scale postwar developments (LPWDs) may be linked to the high original levels of homeownership and white enclave origin of LPWDs. High levels of homeownership are known to positively impact home maintenance, increase length of tenure, and improve neighborhood stability (Rohe & Stewart, 1996). Growing minority populations and persistent patterns of segregation at the city level may further entrench established racial homogeneity and other socioeconomic characteristics within the LPWD census tracts.

### Methodology

To examine the relationship between the magnitude of minority populations and the spatial patterns of black and white households and the prevalence of LPWDs, I employ a simple logit model. I calculate the odds of a census tract having a proportion of postwar housing greater than 0.66 based on metrics for segregation and diversity along with general measures of CBSA level socioeconomic status. Year 2000 census data is used.

Racial diversity is measured at the city level using the following metric:

$$Diversity_j = 1 - \sum_j \alpha_j^2 \tag{1.8}$$

where $\alpha_j$ is the population share of racial group $j$ in a given CBSA. A value of zero implies the entire population is of the same race. There are five racial groups considered (black, white, Hispanic, Asian, and other), thus the largest diversity score possible is 0.8. The exposure metric (Massey and Denton, 1988) is calculated as follows:

$$Exposure_{jk} = \sum_{i}(\frac{j_i}{J})(\frac{k_i}{t_i}) \tag{1.9}$$

where $j_i$ is the total population of group $j$ in census tract $i$, $J$ is the total population of group $j$ within a given CBSA, $k_i$ is the total population of group $k$ in census tract $i$, and $t_i$ is the total population of census tract $i$. The product is summed over all census tracts within a CBSA. The metric is bounded by 0 and 1 and can be interpreted as the probability that a member of group $i$ and a member of group $j$ exist within the same census tract. When $j = k$, the metric represents an isolation measure.

I also consider a spatial measure of centrality to control for the likelihood a city would have developed large-scale postwar neighborhoods to begin with. Rapidly suburbanizing cities in the 1950s and 1960s likely developed into decentralized metropolitan areas. In contrast, cities not experiencing the same postwar boom are likely to exhibit higher levels of centrality. Centrality is measured as the population weighted distance between a census tract centroid and the CBD and is adjusted for the spatial size of a given city:

$$Centrality = \frac{\frac{1}{N}\sum_{i}^{N} dist_i}{\sum_{i}^{N} p_i \cdot dist_i} - 1 \tag{1.10}$$

where $dist_i$ is the distance between census tract $i$ and the CBD, $p_i$ is the population weight for census tract $i$, and $N$ is the total number of census tracts in a given CBSA.

The logit regression includes the diversity and exposure metrics. A white isolation metric is not included as it is highly collinear with the diversity index. Census tract level spatial controls for housing density and distance to downtown are included. The regression specification is as follows:

$$\begin{aligned} logit(PW_{im}) = \alpha + \delta \cdot Diversity_m + \phi \cdot Exposure_m + \gamma \cdot Income_m \\ + \rho \cdot College_m + \theta \cdot Centrality_m + \zeta X_{im} + \epsilon_{im} \end{aligned} \tag{1.11}$$

where $PW_{im}$ is a binary variable that is 1 if the census tract contains more than 66% postwar housing, 0 otherwise and $X$ is the vector of census tract level controls. Standard errors are clustered at the CBSA level. I focus on year 2000 data.

Table 1.7: Logit Regression, dependent variable

| | (1)<br>Logit Coeff. | (2)<br>Odds Ratio |
|---|---|---|
| Diversity Index | 3.737*** | 41.95*** |
| | (0.810) | (33.971) |
| Black/White exposure | -4.837*** | 0.00793*** |
| | (1.455) | (0.012) |
| Income | 1.56e-05 | 1.000 |
| | (0.000) | (0.000) |
| College | -1.065 | 0.345 |
| | (2.854) | (0.984) |
| Centrality | -2.100*** | 0.122*** |
| | (0.808) | (0.099) |
| Central city census tract | 0.0278 | 1.028 |
| | (0.108) | (0.111) |
| Homes per square km (1000s) | -0.282*** | 0.754*** |
| | (0.093) | (0.070) |
| Distance in km to downtown | 0.0274 | 1.028 |
| | (0.029) | (0.030) |
| Squared distance in km to downtown | -0.00120** | 0.999** |
| | (0.001) | (0.001) |
| Constant | -4.886*** | 0.00755*** |
| | (0.410) | (0.003) |
| Observations | 49,512 | 49,512 |
| Psuedo R-squared | 0.056 | 0.056 |

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### Results

Referring to Table 1.7, overwhelmingly the more diverse the city is, the more likely the census tract will be a LPWD. Increases in exposure of black and white residents leads to lower odds a census tract is a LPWD. LPWD census tracts tend to be found in more decentralized cities. City average levels of income and college education are not significant.

I cannot claim a causal relationships from these exploratory observations. At most, the existence of LPWDs are correlated with high levels of racial diversity and black-white segregation at the city level. To unpack these causality issues, knowledge of the true boundaries and historical demographics of LPWDs along with historical city level diversity and segregation measures are required. I hope to explore these observations further in future research.

## 1.8 Conclusion

Literature has singled out characteristics common to postwar homes and housing developments as sources of present day neighborhood decline. However, the true role of postwar structures and developments cannot be fully ascertained without considering the confounding effect of housing age. The effect of aging housing structures on micro-level housing decline and macro-level neighborhood decline is well documented. However, previous literature identifies a composite structural effect and does not disentangle vintage and age effects at the neighborhood level. This paper contributes to both planning and economics literature by presenting the first attempt at neighborhood level disentanglement with a focus on postwar housing.

A further contribution is made through the development of two novel instruments that aid in controlling for endogeneity associated with redevelopment with respect to housing age and vintage. Namely, observations of housing age distributions with census tracts may be biased. Census tracts that were once dominated by postwar housing may have experienced redevelopment as they aged resulting in an endogenous decrease in postwar housing stock. Using historic urbanized area data, I observe whether or not a census tract lies within the estimated 1950 to 1969 growth ring. The second instrument provides an expected distribution of housing age based on national housing age distribution trends thus avoiding local endogenous trends of redevelopment. I employ an OLS and IV strategy across the census years 1990, 2000, and 2010. The panel framework of these data allow for the inclusion of year and CBSA interacted fixed effects which control for metropolitan specific time trends. Previous literature does not account for these trends adding bias to their results.

When age is not controlled for, the collective effect of the proportion of postwar homes and the housing age distribution on neighborhood mean household income is fully attributed to a postwar vintage effect. I find that the composite effect has a significant negative relationship with neighborhood mean income. Specifically, a neighborhood with a mean household income of $72,707 would see a drop of $3,185 for every 0.10 increase in the proportion of postwar housing. When both vintage and age are separately controlled for along with geospatial controls and city level variation, the negative vintage effect is eliminated. This finding is consistent across both the OLS specification and IV specification that controls for endogenous redevelopment in both vintage and age variables and implies postwar neighborhoods are not filtering at rates larger than would be predicted by age.

In investigating the levels of postwar housing in census tracts within and across cities, I find interesting correlations amongst socio-economic variables. Specifically, in census tracts with extreme proportions of postwar housing, 0.66 and greater, households are far more likely to be homeowners, whiter, older, and have higher incomes relative to other census tracts within the same city. I explore this observation further by hypothesizing the observed socio-economic status of these census tracts may be persisting due to, or in correlation with, city-wide segregation dynamics and historical racial tensions associated with large-scale housing developments of the 1950s and 1960s. A strong relationship is found between the odds a census tract contains a very high proportion of postwar housing and the city exhibiting high amounts of racial diversity. Furthermore, increases in black-white segregation across a city also increase the likelihood the city contains census tracts with high amounts of postwar housing. City-wide levels of income and education have no effect. I caution that these are preliminary findings and do not represent a full investigation.

The effect of aging is putting many postwar neighborhoods at risk as this vintage reaches a crucial age in terms of the cost of maintenance. Policies targeting aging neighborhoods, and postwar neighborhoods specifically, are few in number and arguably have limited impact. Furthermore, postwar era suburbs often land in "policy blindspots" as they are often not separate political jurisdictions and do not qualify for state or federal funding (Puentes & Orfield, 2002). The findings presented here provide evidence against the role of the postwar housing and neighborhood characteristics in the filtering of these neighborhoods down the income distribution. Thus, it is necessary that future policy not assume that the collective undesirability of these housing subdivisions are contributing to decline as existing literature has hypothesized. Instead, policies targeting aging housing in downtown regions may be readily adaptable to the surrounding areas. Such policies need not be vintage specific.

# Chapter 2

# Inefficient Land Use Patterns & Municipal Expenditures

Environmental, social, and economic ills are often attributed to inefficient land use patterns, most notably sprawl. Low density development patterns are accused of increasing traffic congestion and pollution, decreasing employment accessibility, increasing social isolation and racial segregation, and increasing the cost of providing municipal services. The literature that seeks to quantify the relationship between development patterns and municipal expenditures in particular, typically models inefficient development patterns through the use of a density measure. While density provides an intuitive link to economies of scale and congestion costs, it fails to capture other spatial characteristics that may greatly influence municipal expenditures. In this study, Massachusetts orthoimagery data is used to construct multidimensional measures of residential and commercial land use patterns. I regress municipal expenditures associated with public works, fire, and police on the multidimensional measures of land use as well as a number of demand controls. Not only do these metrics provide insight into the complex relationship between expenditures and the shape and location of development, they also lend support to some of the development patterns hypothesized to require lower levels of municipal expenditures by 'Smart Growth' advocates.

## 2.1 Introduction

Since the release of the infamous "The Costs of Sprawl" (COS) report in 1974, the relationship between land use patterns and municipal costs has gained modest academic attention. The COS report argues the lack of salience with respect to the relationship between inefficient land use patterns and their cost of development and maintenance leads to increased municipal expenditures (water, sewer, road infrastructure, public transit, etc.). Although this report suffers from methodological and measurement flaws (Altshuler, 1993; Burchell et al., 1998), it highlights an important relationship between the built environment and local public finance that may be costly if ignored.

The literature following COS finds middling success in addressing the complex relationship between development patterns and municipal expenditures. This may be due to three separate issues: first, sprawl is exceedingly hard to define; second, municipal costs are difficult to measure; and third, the channels through which inefficient development influences municipal costs may be opaque. Sprawl has an complex combination of qualities that renders single definitions inadequate although a single measure of density is often used to quantify it. Using density as the main unit of measurement focuses analysis on the expanding low-density quality of sprawl typically described through population, employment, and housing. While density captures important average vertical aspects of development, this single measure provides less and less information the larger the geographic unit is that one considers. Furthermore, there is some debate as to what the appropriate geographic unit should be to begin with. For example, a county level unit of analysis may include far too much rural land ("overbound") whereas Census Urbanized areas may omit legitimate urban fringe ("underbound") (Wolman et al., 2005).

Sprawl is not limited to describing density patterns of fringe suburban developments belonging to metropolitan areas. It is a catch-all term for inefficient land use, as defined by Blais (2011):

> Sprawl can be defined as an inefficient land-use pattern. It embodies a missallocation of resources that is wasteful, consuming resources that could have been put to more productive uses.

The shape and location of various forms of development and not just the overall density of development, have the potential to create inefficiencies with respect to the provision of municipal services. Advances in GIS techniques within planning and landscape ecology have helped to quantify the shape of developed land and spatial relationships between land uses using a rich set of multidimensional metrics (Torrens & Alberti, 2000; Galster et al., 2001; Allen, 2001; Ewing, Pendall, & Chen, 2002; Cutsinger, Galster, Wolman, Hanson, & Towns, 2005; Cutsinger & Galster, 2006; Clifton, Ewing, Knaap, & Song, 2008; Frenkel & Ashkenazi, 2008; Sarzynski, Galster, & Stack, 2014a, 2014b). These metrics quantitatively describe patterns of leapfrog development (discontinous/disperse development); ribbon sprawl/retailscape; the diversity, evenness, and clustering of the distribution of land use types; and the accessibility/proximity between uses. The relevance of using multidimensional metrics to describe development patterns is bolstered by findings showing the empirical independence of these metrics amongst themselves as well as with respect to density (Cutsinger et al., 2005; Frenkel & Ashkenazi, 2008; Jaret, Ghadge, Reid, & Adelman, 2009).

The literature outside of planning and landscape ecology is somewhat slow in adopting more detailed measures of development patterns most likely due to limited data availability and time-consuming construction. The extant literature on the topic of municipal expenditures and land use patterns offers methodologies that are variations on the same theme with density being the dominant measure used to describe land use patterns. Although density proves to be an intuitive measure given its connection to economies of scale and urban congestion, it is unlikely to adequately capture all of the possible economies (or diseconomies) associated the shape and location of land uses. To my knowledge, there does not exist analysis of the relationship between municipal expenditures and land use patterns that utilizes multidimensional measures.

The research presented here contributes to the literature by introducing multidimensional land use measures into the analysis of inefficient land use and municipal expenditures. Within a simple cross-sectional framework, I explore the relationship between these metrics, total municipal expenditures, as well as three disaggregated expenditure categories within Massachusetts municipalities. Metrics describing concentration, centrality, isolation, interaction, and fragmentation of residential and commercial uses are considered. I find large and significant effects for a number of metrics confirming the role of the shape and location of land use in the efficiency of municipal service provision. Furthermore, the collection of multidimensional measures considered provide more explanatory power with respect to municipal expenditures than does a simple density measure.

The paper proceeds as follows: in Section 2.2 I summarize the municipal expenditure and sprawl literature, Section 2.3 provides details on the construction of the land use metrics and the regression specifications, data and descriptives follow in Section 2.4, and results are presented in Section 2.5.

## 2.2 Municipal Expenditures and Land Use

In what ways are municipal expenditures influenced by development patterns? "Smart Growth" and "New Urbanism" development plans advocate for compactness in the form of dense, central, mixed, and connected land use patterns (Blais, 2011). These development patterns are thought to limit the negative impact of environmental, social, and economic externalities associated with more diffuse patterns. With respect to economic externalities, dense development patterns are assumed to exploit economies of scale associated with municipal services. When land use is sprawling (measured via density), these economies are underutilized (Corporation, (U.S.), Research, , & Management, 1974; Carruthers & Ulfarsson, 2002, 2003, 2008; Hortas-Rico & Sole-Olle, 2010; Hortas-Rico, 2014; C. B. Goodman, 2015). This underutilization is assumed to translate into larger expenditures per capita for more sprawling municipalities.

Opposing the benefits associated with intensity of development is the "harshness" accompanying high levels of density (Bradbury, Ladd, Perrault, Reschovsky, & Yinger, 1984; Ladd & Yinger, 1989; Ladd, 1992, 1994). It is posited by the local public finance literature that denser land use patterns may require larger expenditures for traffic safety and flow, waste management, and police services due to a greater intensity of use.

The work cited in support of the economies of scale and harshness arguments does not provide detail regarding the potential channels through which land use patterns may cause efficiencies or inefficiencies with respect to providing municipal services. While not an exhaustive review of engineering systems or fire and police operations, below I summarize potential built environment related cost sources with respect to select municipal services within the state of Massachusetts. As well, the current state of the empirical evidence investigating the connection between development patterns and municipal expenditures is presented.

### Water/Sewer

The majority of the expenditures associated with water and sewer infrastructure stem from the high cost of water distribution and sewer collection mains. These outlays dominate transmission main and pump costs (Blais, 2011). Larger distribution and collector mains are required for larger lots. Additionally, pumping costs increase with lot size as larger lengths of pipe tend to exhibit pressure losses. Larger lots also tend to have higher levels of water usage (Blais, 2011). This relationship between lot size and water and sewer costs is confirmed empirically by Speir and Stephenson (2002); doubling lot size increases costs by 30%.

While the connection between density and water/sewer provision is fairly clear, it is unclear how water and sewer expenditures may be related to other land use patterns championed by smart growth. Specifically, mixed use may require inefficient levels of water or sewer service as they must accommodate the highest capacity requirement, i.e. larger distribution and collection infrastructure than needed for lower capacity uses within a mixed development that contain buildings with high capacity requirements. Efficiencies may exist with respect to supplying services to residential land that is not fragmented, leapfrogging, or winding as less infrastructure would be required to service more contiguous development.

### Roadways

Unlike many other states, Massachusetts legally does not allow municipalities to charge property developers impact fees to cover costs associated with new road construction (MassDot, 2017). However, under Chapter 90, Massachusetts municipalities are reimbursed by the state for a portion of expenditures relating to road construction and maintenance (MassDot, 2017).

As such, the impact of road construction and maintenance on expenditures may be slightly dampened. Without access to state aid, a positive relationship between fragmented, low density development would be expected as more infrastructure per person would be required. However, the infrastructure requirements for many low density residential developments are minimal, e.g. no sidewalks, no curbs, gravel roads in some instances.

More mixed use development patterns may also produce ambiguous effects with respect to construction and maintenance expenditures. Less infrastructure may be required to link residential uses to commercial uses; however, the wear and tear on roads may increase if roads are used consistently throughout the day as opposed to in the morning and evening (residential) or during business hours (commercial).

### Fire

Massachusetts communities create service zone plans for emergency services based on the following: available resources, geography, population density, and community expectations (*Massachusetts Department of Fire Services*, 2011). In addition to the requirements laid out in service zone plans, the National Fire Protection Association (NFPA) stipulates that every ladder and engine be staffed by a minimum of four firefighters with more services to be made available depending on the hazard level associated with the buildings within a community.

High-hazard occupancies consist of schools, hospitals, nursing homes. Medium-hazard occupancies include apartments, offices, and industrial uses. Low-hazard occupancies include one to three family residential units as well as scattered small business and industrial buildings. Response capacities are based on the number and location of these various occupancies (NPA). Thus, density increases the number of structures accessible within a service zone; however, density may increase the fire hazard especially with high densities of high-hazard occupancies. More fragmented, irregular development patterns may hinder accessibility within a service zone, but it is unclear if these development patterns have a measurable impact on service costs.

### Police

Labor costs are the most substantial outlay associated with police expenditures (Spence, Webster, & Conners, 2016). In general, larger populations require larger police forces as economies of scale with respect to population are likely nonexistent or negative given the intense labor requirements and associated costs (Walzer, 1972; Gyimah-Brempong, 1987).

An additional influence on police expenditures is crime where increasing crime levels lead to increases in police services (Marvell & Moody, 1996). Following Jane Jacobs' (1961) emphasis on land use diversity and high levels of pedestrian activity as methods to reduce potential criminal activities, a substantial literature linking the built environment to crime levels has developed.[1] Jacobs posited that dense, mixed use neighborhoods increased pedestrian traffic throughout the day (Browning et al., 2010). The increased number of "eyes on the street" are thought to provide passive monitoring of potential criminal activity. However, the effectiveness of increased monitoring seems limited to the crime type. Browning et al. (2010) find reductions in violent crimes such as homicide and aggravated assault but increases in robberies which may rely on a densities of people and traffic. It remains to be seen whether the increase in non-violent crime in mixed use areas translates into increases in police expenditures per capita.

[1] See Sohn (2016) for a summary.

### 2.2.1 Empirical Evidence

Given the complex relationship between land use patterns and municipal services, the question of *if* or *how* various notions of compactness or seemingly inefficient land use influence municipal expenditures remains an empirical one. A small group of studies attempt to identify support for economies of scale and harshness in particular, but fail to address more complex patterns of land use.

Ladd (1992) is one of the first articles to apply a rigorous analysis of land use patterns and municipal costs using U.S. county level data. Ladd considers the effect of population growth and density on the cost of providing public services. Using a piece-wise regression model, she finds an initial decline in current spending as density increases at very low levels but an overall increase in the cost as population density increases. With respect to public safety costs, she finds that doubling a county's density from 500 to 1000 people per square mile increases per capita costs by about 30%. Results must be taken with a grain of salt given the county unit of analysis does not reflect the true spatial level of municipal expenditures.

Carruthers and Ulfarsson (2008) also perform an analysis at the U.S. county level. They employ a 'spillover model' approach that takes into consideration the municipal expenditures of neighboring counties. Two measures of sprawl are used: density of developed land and the percentage of county land area that is developed. A variety of expenditure variables are considered including per capita education, fire, housing, libraries, parks and recreation, police, roadways, sewerage, and solid waste. Density of developed land is found to have a negative effect on education, parks and recreation, police protection, and roadways. Percentage of county land area that is developed has an overall significant and positive effect on all expenditure categories expect for housing and solid waste where the effect is positive but not significant.

Hortas-Rico and Sole-Olle's 2010 paper is one of the few papers to use municipal level expenditure data. Data from 2500 Spanish municipalities for the year 2003 serve as the dependent variables within piece-wise linear regression models similar to Ladd (1992). A public spending equation is estimated for aggregate expenditures as well as six disaggregated categories that are grouped as community facilities, basic infrastructure and transport, local police, culture and sports, housing and community development, and general administration. Four different variables are used to measure sprawl with the main variable being urbanized land per capita. Residential houses, percentage of scattered population, and number of population centers all measured in per capita terms are also considered. These measures of sprawl are an improvement over the typical density measure but still do not fully describe the land use patterns. In general, the authors find that low-density development patterns lead to greater local expenditures and consequently higher levels of taxation for public services with the exception of housing and basic infrastructure and transport.

A panel approach is taken by Goodman (2015) using the same expenditure categories as Carruthers and Ulfarsson (2008) with the addition of natural resources. U.S. county level data from 1982 to 2012 are employed within a county and year fixed effects regression framework that allows for the land use variables of residents per acre and percentage of developed land within the county to enter cubically. Density is significant only within the education and natural resources specifications. In both cases, low levels of density are associated with lower expenditure levels. Increasing density then increases expenditures but at very high density levels the relationship becomes negative again. The percentage of developed land is oppositely related to all expenditures except housing, roadways, and sewerage. The relationship is initially positive for low levels of developed land, then negative, and eventually positive again for high levels of developed land. A positive linear effect is found for roadways, a negative linear effect for waste management, and no effect for housing. Goodman's findings suggest the percentage of developed land is a far more influential factor than density.

The literature summarized here does not provide an entirely clear depiction of the relationship between land use patterns and municipal expenditures. The argument in favor of the economies of scale theory does have some support; however, Goodman's findings show virtually no impact of density on numerous expenditure categories. Allowing density to enter nonlinearly improves the ability to assess the complexity of the relationship but it does not ameliorate the intrinsic problem with density, it is a blackbox. An entire urbanized area's built environment (or county's) is reduced to a one-dimensional average measure that omits possibly influential patterns of clustering, centrality, and accessibility between uses. On average, residential density within an urbanized area may be high but the shape and location of residential developments may be inefficient with respect to providing particular municipal services. In effect, any economies associated with shape and location not correlated with density are empirically unaccounted for.

## 2.3 Methodology

### 2.3.1 Land Use Metrics Construction

The collection of metrics described in detail within this section aim to provide a holistic representation of the complex land use patterns that exist within municipalities. Measures of concentration, centrality, and local proximity between uses are adapted from Galster et al. (2001). To pin down the fragmented characteristics of diffuse land use, two common landscape ecology metrics are used: mean patch area and mean perimeter to area ratio (Pickett & Cadenasso, 1995; O'Neill, Riitters, Wickham, & Jones, 1999; Wiens, Crawford, & Gosz, 1985).

The construction of multidimensional measures of land use patterns requires detailed data. For this study, I rely on a land cover/land use GIS datalayer produced from a high resolution orthoimagery dataset spanning the entire state of Massachusetts.[2] Land cover and land use within the datalayer are categorized into one of 33 land use categories that represent developed, potentially developable, and undevelopable land. Developed land contains residential use at four density levels, commercial, industrial, recreational land uses, and institutional uses.[3] Potentially developable land consists of agricultural, forest

[2]Further data details in following section.

[3]Institutional uses include schools, churches, colleges, hospitals, museums, prisons, town halls or court houses, police and fire stations, including parking lots, dormitories, and university housing. Also may include public open green spaces like town commons.

and shrub cover. Undevelopable land includes wetlands, beaches, cemeteries, land used for waste disposal infrastructure, mining, and transportation infrastructure. In addition to these categories of undevelopable land, I utilize elevation data to reclassify potentially developable land as undevelopable if the land is sloped greater than 15%. A datalayer consisting of protected lands within Massachusetts is further used to reclassify potentially developable land as undevelopable. Figure 2.1 depicts the various land uses within the town of Needham. A distinct commercial corridor is seen straddling the main roadway through the center of the town. The majority of residential land use is medium to high density residential with a small cluster of multi-family. Industrial land use is isolated away from the town center and is separated from residential uses by the interstate.

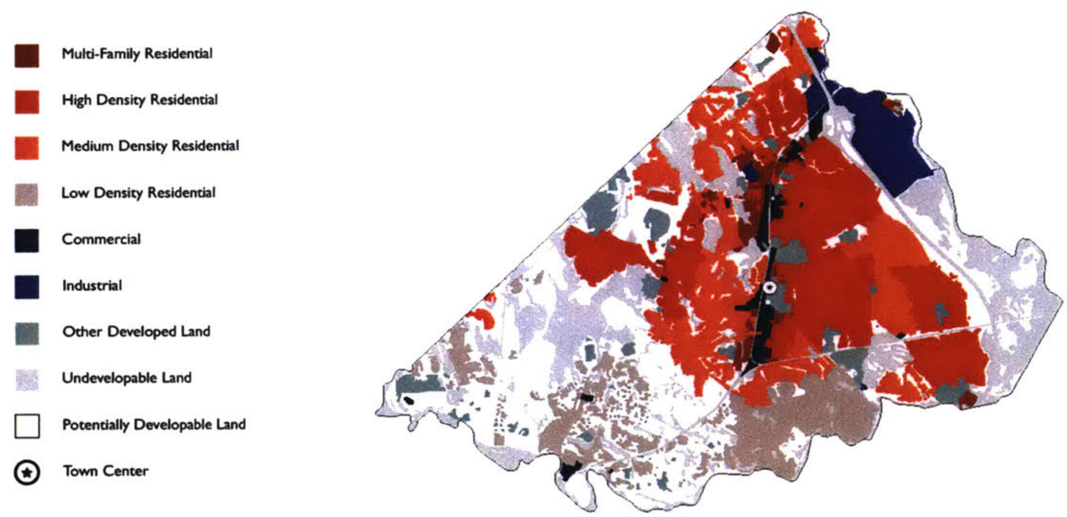



Figure 2.1: MassGIS Land Use Datalayer; Town of Needham

For analysis, low and medium density residential land use are grouped, as are high density residential and multi-family. Commercial land use is also considered but industrial land use is omitted as many municipalities do not contain this use type.

To construct the metrics pertaining to concentration, centrality, and local proximity, a uniform unit of analysis within each municipality is required. Similar to Galster et al. (2001), I make use of a grid comprised of 1.29 km square (1/2 mile square) cells overlaying the entire Massachusetts land use datalayer. The total area of the two residential groupings, commercial use along with totals for all developable, potentially developable, and undevelopable land within each cell within each municipality are calculated (see Figure 2.2). A benefit to working with Massachusetts data is that municipal boundaries are contiguous, meaning unincorporated land between municipalities does not exist. This allows for accurate identification of land use within municipal boundaries.

The disaggregated expenditures considered are limited to public works[4], fire, and police in the subsequent regression analysis in order to focus on municipal services that may have a more intuitive connection to economies or diseconomies of shape and location. Specifications using total expenditures are also constructed. Total expenditures include general government[5], police, fire, other public safety[6], education, public works, human services[7], culture and recreation[8], intergovernmental assessments, fixed cost[9], debt service[10], other expenditures[11], transfers to other funds, and other financing uses.

### Density and Proportion Developed

The density metric used is based on developed land within a given municipality's borders. Specifically, it is calculated as the number of residents per kilometer square of developed land. Alternative forms of density introduce bias based on the geography considered,

[4]Highways/streets snow and ice, highway/streets other, waste collection and disposal, sewerage collection and disposal, water distribution, parking garage, street lighting and other.

[5]Legislative, executive, accountant/auditor, collector, treasurer, law department, town/city counsel, public buildings/properties maintenance, assessors, operation support, license and registration, land uses, conservation commission and other.

[6]Emergency medical services, inspection and other.

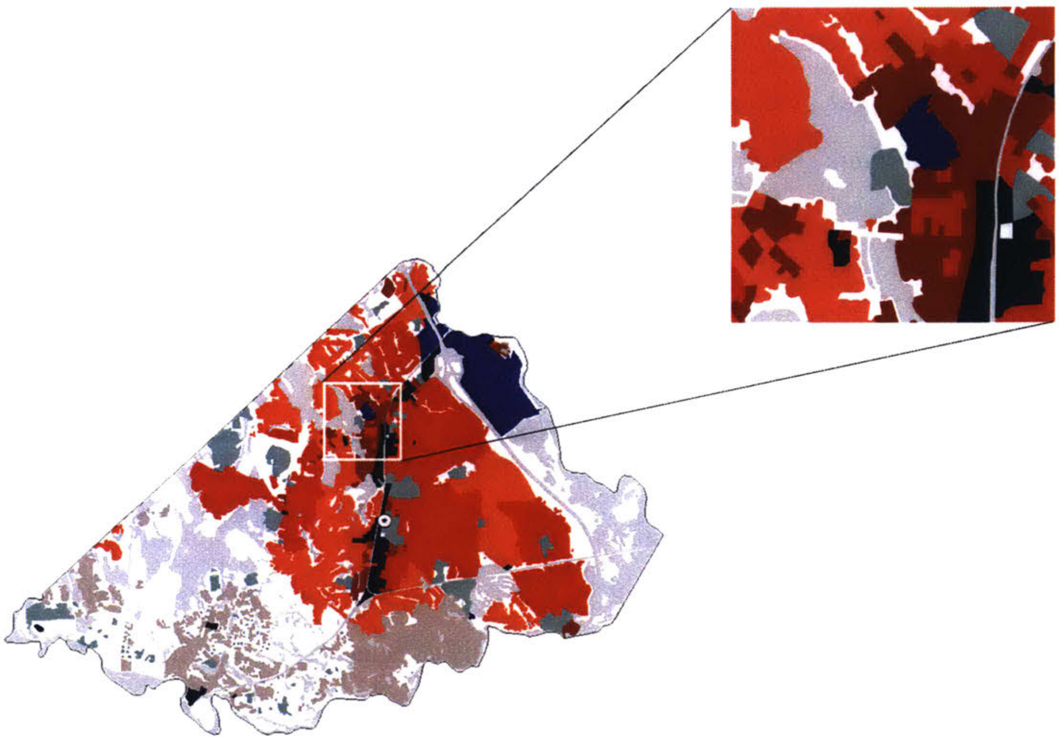
[7]Health services, clinical services, special program, veteran's services and other.

[8]Library, recreation, parks, historical commission, celebrations and other.

[9]Workers' compensation, unemployment, health insurance, other employee benefits, other insurance and retirement.

[10]Retirement of debt principal, interest on long term debt, interest on short term debt and other interest.

[11]Court judgments and other unclassified expenditures.



**Note**: Land use within a grid cell is highlighted by the zoomed in cell. Total areas of land use categories within each cell are used to construct the concentration, centrality and proximity metrics.

Figure 2.2: MassGIS Land Use Datalayer

whereas using the actual developed area within a municipality ameliorates issues that arise when utilizing urbanized area or county boundaries.

The proportion of land that is developed is often considered alongside a density measure. This metric captures a horizontal aspect of development (Goodman, 2015) with lower proportions associated with more compact development (Carruthers & Ulfarsson, 2003, 2008). This metric is calculated as the ratio of developed land to all developed and potentially developable land.

### Concentration

Concentration measures the degree to which a land use type is disproportionately located in relatively few areas or spread evenly across a municipality. It is akin to Massey and Denton (1988) dissimilarity index for racial segregation. Here, the value indicates the share of land use that would need to be redistributed in order to achieve a uniform distribution across the developable land within a municipality. Larger values indicate higher levels of concentration. Concentration of land use type type $i$ is calculated as follows:

$$Conc_i = \frac{1}{2}\sum_{m=1}^{N}(\frac{i_m}{I} - \frac{t_m}{T}) \tag{2.1}$$

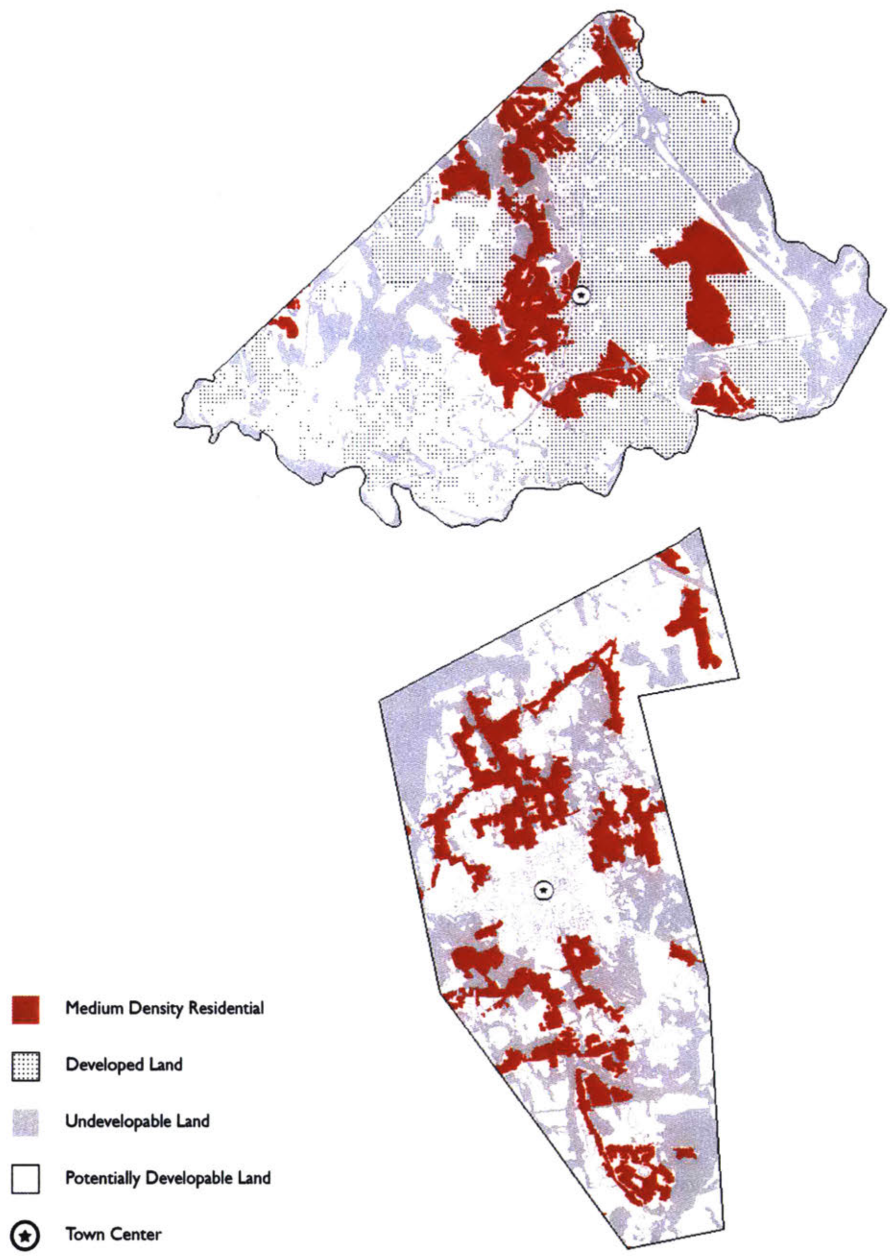
where $m$ refers to a cell within the municipality with $N$ being the total number of cells, $i_m$ is the area of land use type $i$ within cell $m$, $I$ is total area of land use type $i$ within a given municipality, $t_m$ is total developed and potentially developable area within cell $m$, $T$ is total developed and potentially developable area within a given municipality. Galster et al. (2001) associate lower levels of concentration with more sprawling development patterns.

### Centrality

Centrality measures the weighted average distance of a given land use type to the town or city center measured by the town hall location. Given the variation in municipality total area, I standardize the weighted average distance with respect to the average distance from all cells to the center:

$$Ctr_i = \frac{\sum_{m=1}^{N} d(m, center)/N}{\sum_{m=1}^{N} \frac{i_m}{I} d(m, center)} - 1 \tag{2.2}$$

where $d(m, center)$ is the distance from the centroid of cell $m$ to the town/city center and $\frac{i_m}{I}$ is the weight assigned to cell $m$ based on the proportion of land use type $i$ with respect to the total area of land use type $i$ within the municipality. A number larger than zero indicates a land use type is more centrally located than would be expected on average.




**Note:** The high concentration of medium density residential housing in the top figure highlights the limited coverage of this use type across the municipality. The municipality in the bottom figure (lower concentration) shows a more even distribution of medium density residential use.

Figure 2.3: High vs. Low Concentration

Centrality represents a very general notion of compactness where decentralization implies increasing sprawl. A more centralized residential pattern may allow for quicker access by fire and police services. As a municipality becomes less centralized additional stations are likely required in order to service residents within an required amount of time. Based on the notion that centrality measures an aspect of compactness, smart growth literature would anticipate a negative relationship between public works expenditures and low and medium density residential centrality. A greater amount of road, water, and sewer infrastructure is assumed to be needed to access less central locations. Centrality may also signal accessibility with respect to fire and police services.

### Local Proximity

Local proximity measures the extent to which land use type $i$ is exposed to land use type $j$ (interaction index) or the extent to which type $i$ is exposed to only type $i$ (isolation index). These indexes have a long history within the racial segregation literature (Massey & Denton, 1988; White, 1986; Sørensen, Taeuber, & Jr, 1975) and were adapted to be used within the land use context by Galster et al. (2001).

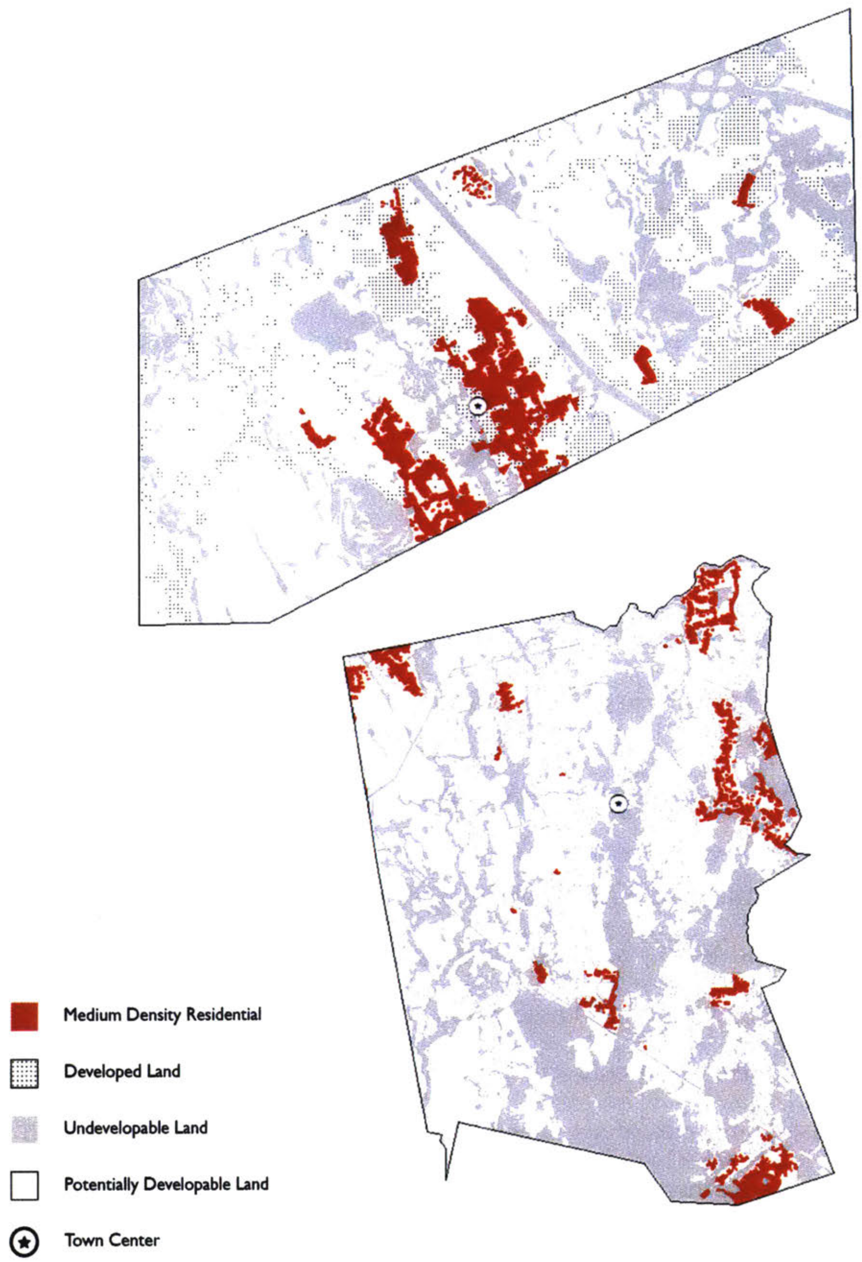
In general, both indexes can be calculated as follows:

$$p_{ij} = \sum_{m=1}^{N} \frac{i_m}{I} \frac{j_m}{J} \tag{2.3}$$

where the isolation index sets use type $i$ equal to use type $j$. However, this calculation does not account for the overall composition of the municipality and asymmetry exists between $p_{ij}$ and $p_{ji}$. White (1986) suggests the following adjustments to (2.3) for the isolation and interaction index respectively:

$$p^*_{ii} = \frac{p_{ii} - I/T}{1 - I/T} \tag{2.4}$$

$$p^*_{ij} = \frac{p_{ij}}{J/(I+J)} \tag{2.5}$$




**Note:** The higher centrality of medium density residential is clearly demonstrated by the top municipality. This use type is much more decentralized in the bottom municipality.

Figure 2.4: High vs. Low Centrality

where $T$ is the total area of developed and developable land in a given municipality. Thus, $I/T$ and $J/T$ are the proportions of land use $i$ and $j$ in given municipality. This adjusted metric is symmetric and is bounded by 0 and 1. It is analogous to the probability that two use types exist within the same cell or a use type exists on its own within a cell.

Isolation of residential uses from commercial uses may result from restrictive zoning practices such as minimum lot size zoning within different zoning categories (Galster et, al., 2001). As the distance between residences and employment/retail areas increase, so to does travel time. Larger amounts of infrastructure may also be required to connect the uses.

## Mean Patch Area

The landscape ecology term "patch" is analogous to contiguous land use shapes observed within the land use data. Mean patch area of land use type $i$ within a given municipality is calculated as follows:

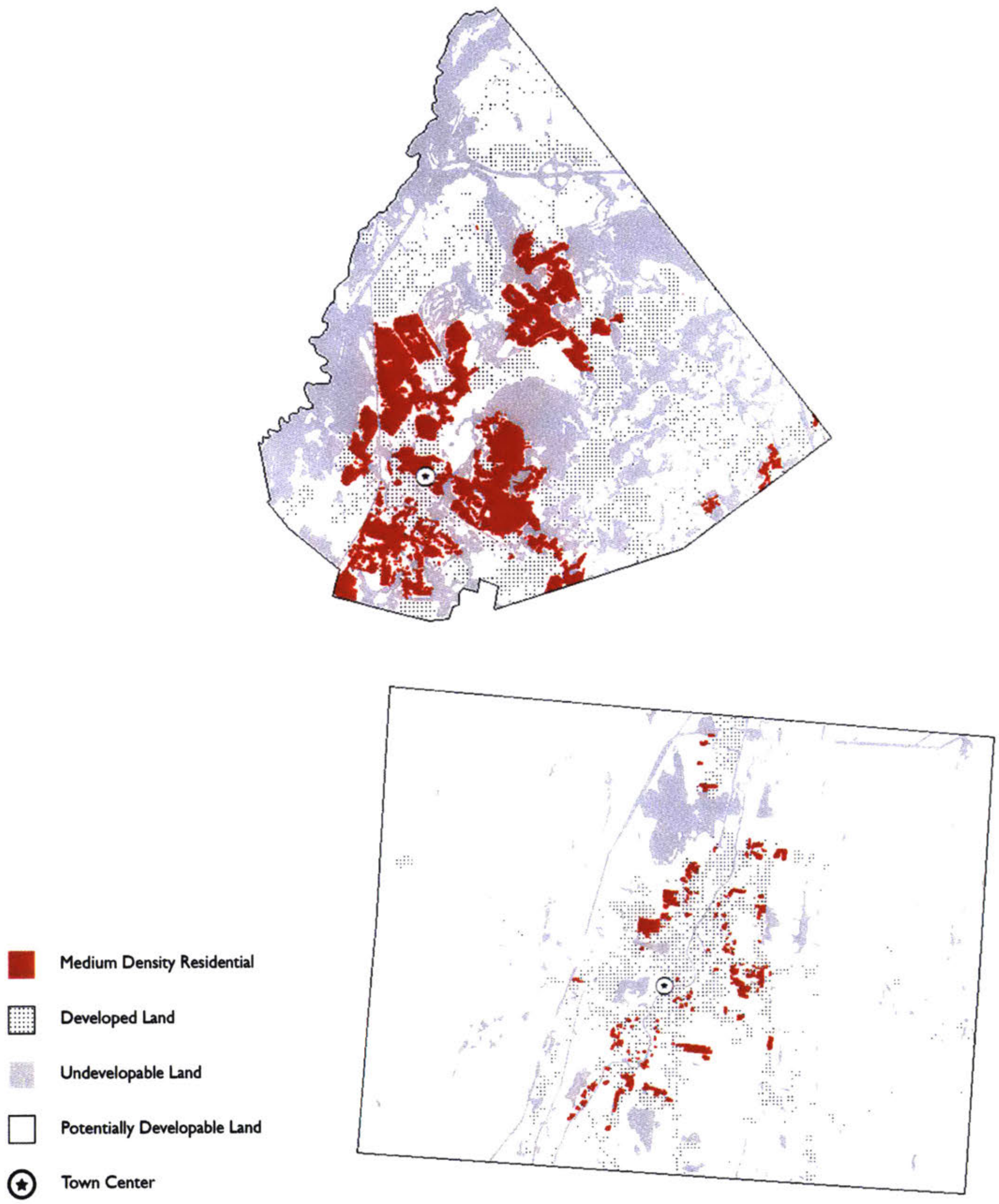
$$MPA_i = \frac{\sum_{l=1}^{n_i} a_{il}}{n_i} \tag{2.6}$$

where $n_i$ is the total number of patches of land use $i$ within a given municipality, $a_{il}$ is the area of patch $l$ with land use $i$. Smaller mean patch size values indicate higher degrees of fragmentation where fragmentation is associated with sprawling, leap frogging development patterns. This metric is constructed for residential and commercial use along with developed land in general. Based on issues of accessibility and connectedness, smart growth advocates expect fragmented development to be more costly to service.

## Mean Perimeter to Area Ratio

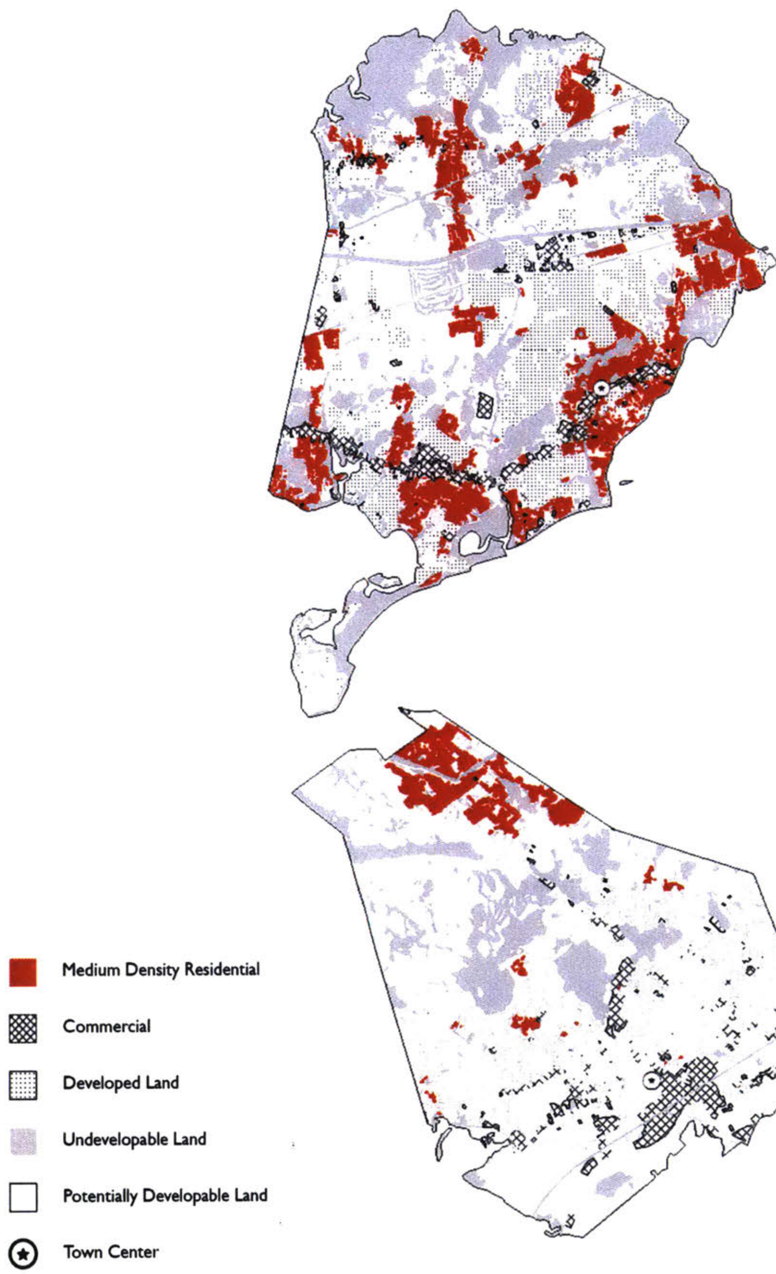
This metric captures the degree to which a patch is compact versus long, winding, or generally irregularly shaped.

$$MPAR_i = \frac{\sum_{l=1}^{n_i} r_{il}/a_{il}}{n_i} \tag{2.7}$$



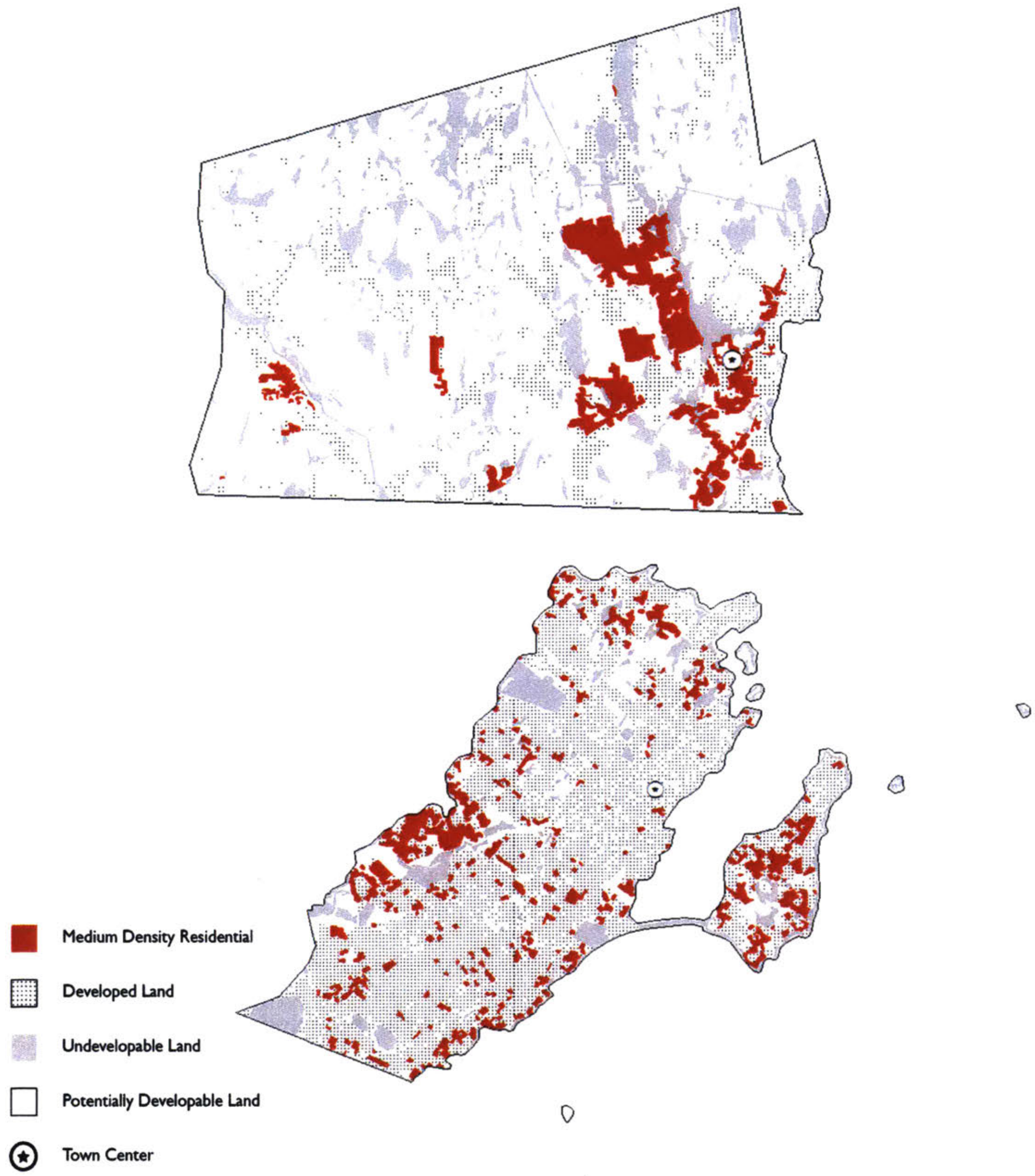

**Note:** Higher levels of medium density residential isolation are demonstrated by the top figure. Here, larger areas of unbroken land are covered by this land use type. The bottom figure shows a more sparse and isolated pattern of medium density residential land use.

Figure 2.5: Isolation




**Note:** High levels of interaction between medium density residential and commercial land use are seen in the top figure where the commercial land use is distributed in close proximity to the residential use. A nearly complete separation of the two uses is demonstrated in the bottom figure.

Figure 2.6: Interaction




**Note:** Larger medium density residential patch areas are seen clearly in the top figure as compared to the small patches in the bottom image.

Figure 2.7: Mean Patch Area

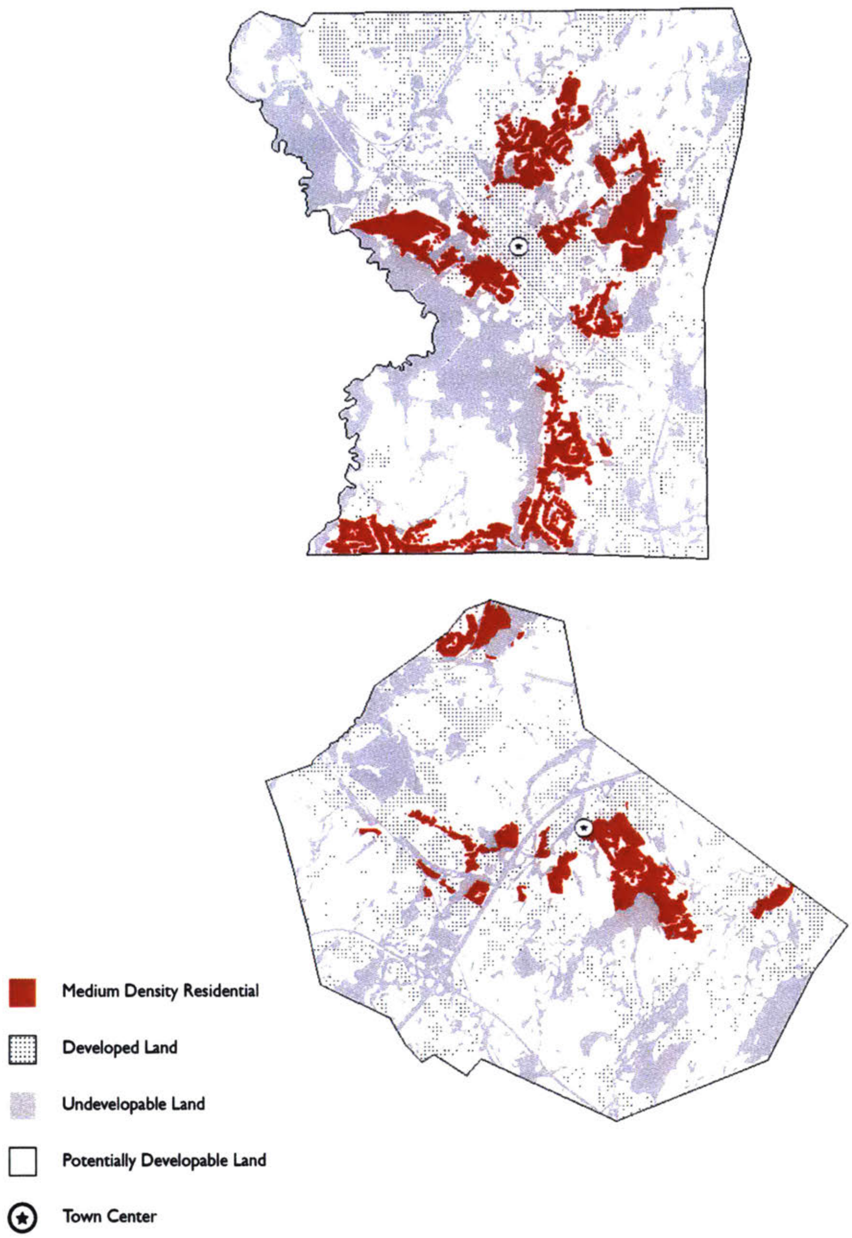
where $r_{il}$ is the perimeter of patch $l$ of land use $i$. A larger mean perimeter to area ratio indicates a more complex patch shape. In general, more winding, nonconvex developments are associated with inefficient delivery of municipal services. More infrastructure materials may be required to service a similar amount of people on a more compact patch. Accessibility is also diminished which in turn may affect police and fire response times as well as garbage pick-up and snow removal.

### 2.3.2 Regression Model

In order to identify the relationship between land use patterns and municipal expenditures, factors contributing to the demand and costs of municipal services must be considered. The costs associated with delivering municipal services include environmental as well as input costs (Bradford et al., 1969; Ladd, 1992, 1994; Hortas-Rico & Solé-Ollé, 2010). Land use patterns can be viewed as a source of environmental costs through the potential inefficiencies they impose on the delivery of municipal services. Other environmental costs include factors that affect the cost of providing municipal services but are beyond the control of local public officials (Ladd, 1992). Demand for municipal services cover demographic and socioeconomic characteristics that closely follow those outlined by Bergstrom and Goodman (1973) and Borcherding and Deacon (1972). Fiscal and political characteristics capture cross-sectional variation in sources of revenue and political ideology. With these factors in mind, I propose the following general form for per capita municipal expenditures:

$$\begin{aligned} Exp = f(&\textit{land use patterns, other environmental costs, demand for municipal services,} \\ &\textit{fiscal characteristics, unobserved factors}) \end{aligned}$$

Within other environmental costs I control for total employment per capita and proportion of homes built before 1940. Municipalities with a large workforce relative to residential population (commuters) may have different demands for municipal services which are likely




**Note:** The medium density residential land use in the top figure shows a long a winding pattern especially in the developments toward the lower border. The lower figure on average contains less complex patches of medium density residential.

Figure 2.8: Mean Perimeter to Area Ratio

to increase the total required level of services. With a large proportion of older homes, a municipality is likely to have aging infrastructure requiring higher levels of expenditures related to maintenance. Demand for municipal services may be influenced by total population, population growth, per capita income, age distribution of the population, proportion of individuals that are white, proportion of individuals with a Bachelor's degree or higher, the homeownership rate, and political ideology. These demand factors are each likely to have unique municipal service requirements and preferences. For example, by the homevoter hypothesis (Fischel, 2009), homeowners may demand higher levels of services in order to retain or increase the value of their home. I also control for fiscal characteristics through state revenue received by the municipality and the ratio between property tax revenue and total revenue. External sources of revenue may crowd-out or crowd-in local expenditures on municipal services and if a municipality is heavily reliant on its residents for revenue, expenditures may be attenuated.

Analysis begins with two baseline regressions: one consisting of no land use measures and another that includes only density.[12] The OLS model is cross-sectional with the municipality as the unit of analysis. Logged per capita municipal expenditures are regressed on demographic, socioeconomic, and fiscal controls.[13] The specification including density is as follows:

$$LogExp_i = \alpha + \beta \cdot D_i + \gamma X_i + \epsilon_i \tag{2.8}$$

where $LogExp_i$ is the log per capita expenditure for municipality $i$, $D_i$ is density, $X_i$ represents the vector of controls which include the additional environmental costs, demand for municipal services, and fiscal/political characteristics.

The two baseline regressions are compared to identify the strength of the relationship between density and the four municipal expenditure categories. Next, regressions are run

[12]Proportion of land developed was initially considered within this baseline regression but was omitted as it was not significant in any regression.

[13]Employment per capita, income per capita, and state revenue enter as natural logs.

separately for each multidimensional metric in place of density in order to view the influence of each metric in isolation prior to combining them in one regression:

$$LogExp_i = \alpha + \sum_j \lambda_j \cdot Z_{ij} + \gamma X_i + \epsilon_i \tag{2.9}$$

where $j$ = {med and low density residential, multi family and high density residential, commercial} and $Z_{ij}$ is a land use metric for land use type $j$.

Lastly, I consider a full specification consisting of all metrics and all controls for each expenditure category:

$$LogExp_i = \alpha + \sum_j \sum_k \phi_{jk} Z_{ijk} + \gamma X_i + \epsilon_i \tag{2.10}$$

where $k$ = {concentration, centrality, isolation, interaction, mean patch area, mean perimeter to area ratio}. $Z_{ijk}$ is the land use metric $k$ for land use type $j$ in municipality $i$.

Reverse causality may be a concern in the above regressions: are municipal expenditures influenced by land use patterns or are land use patterns the result of the cost associated with servicing particular forms of development. For example, a municipality may choose particular development patterns based on perceived costs to service. To minimize this issue, land use metrics are constructed using 2005 data whereas expenditure and control data is from 2010. The lag in expenditure and control data may be enough to break possible reverse causality; however, instrumental variable techniques were also initially considered but instruments proved to be weak likely a result of the low sample size.[14]

## 2.4 Data

Land use data is collected from MassGIS, a state government office tasked with providing a "comprehensive, statewide database of spatial information for mapping and analysis supporting emergency response, environmental planning and management, transportation

[14] The instrument was based on similar 1971 land use data.

planning, economic development, and transparency in state government operation" (*Office of Geographic Information (MassGIS)*, 2011). The Land Use (2005) datalayer is a digital dataset of land cover/land use and was constructed using semi-automated methods based on 0.5 meter resolution digital orthoimagery. The Protected Lands datalayer used to identify undevelopable land is also retrieved from MassGIS. Digital Elevation Model (DEM) data is provided by the United States Geological Survey (USGS) with a resolution of 30 meters (*National Elevation Dataset (NED) | The Long Term Archive*, 2016).

The Massachusetts Department of Revenue provides detailed annual expenditure and revenue data for all 351 municipalities. Annual population estimates are also provided. The 2010 Census provides demographic and socioeconomic data at the municipality level.

### 2.4.1 Descriptive Statistics

Not all land use types considered in the analysis exist within each of the 351 Massachusetts municipalities. However, a decent sample size remains (298) when selecting municipalities that contain the low and high density residential groupings along with commercial use. The municipality of Nantucket (which is an island) is removed from the analysis as it proves to be an extreme outlier for a number of the variables considered. Fire expenditure data is missing for 10 municipalities which lowers the the n to 288 within the fire expenditure regressions. Summary statistics are presented in Table 2.1.

Of the municipal expenditure categories, mean expenditures per capita on public works are the largest at $169. Total expenditures range across Massachusetts municipalites from a low of $1,029 per capita to a high of $6,824 per capita. The average number of residents per km$^2$ is 1,264 with a maximum density of 7,091 residents per km$^2$ within the city of Somerville.

Proportion of land within a municipality that is developed ranges from 0.02 to 0.83 with a mean of 0.27. Recalling that the concentration metric can be interpreted as the share of land use that would need to be redistributed in order to achieve a uniform distribution, on average med./low density residential is much less concentrated than both multi/high density residential and commercial. Multi/high density residential is more central then both med./low density residential and commercial use on average. Med./low density residential is the most isolated use type with an average isolation value of 0.277. Commercial use tends to be less isolated than the two residential groupings on average. Interestingly, multi/high density residential displays a lower interaction value with respect to commercial as compared to med./low density residential and commercial. The mean patch area for multi/high density residential land use is the largest at 0.047 $km^2$ with the mean patch for commercial use much smaller at 0.017. Multi/high density residential also displays the largest mean perimeter to area ratio at 0.194, although the mean variation across the uses is not large.

Massachusetts municipalities range in population from 327 in Tyringham to 617,594 in Boston. Population growth over the 2000 to 2010 time period was 9.0% on average. The range for proportion of population under 5 years old is quite small, 0.018 to 0.084, whereas the proportion over 65 ranges from 0.074 to a high of 0.398 in the town of Orleans. Orleans also claims the maximum value for proportion white residents at 0.988. There is large variation in the proportion of population with a Bachelor's degree or higher, 0.11 to 0.793 with a mean of 0.388. The mean Republican vote share for the 2010 Gubernatorial election was 0.388. Homeownership within Massachusetts municipalities appears to be quite high with a maximum value of 0.975 and a mean of 0.779. On average, a small portion of homes within municipalities were built before 1940, 0.026. The number of jobs per resident ranges from 0.053 to 1.611. Income per capita also has a large range from $5,440 to $289,184. The mean level of state revenue per capita is $505. The average proportion of revenue coming from property tax is quite high at 0.737 but does display a large range from 0.214 to 0.943.

### The multidimensional Metrics and Density in Detail

How are the multidimensional metrics related to one another and how are they related to density and proportion of developed land? Table 2.12 in Section 2.7 presents correlations between the traditional measures of density and proportion developed with the multidimensional metrics.

Table 2.1: Summary statistics

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Public works exp. per Capita | 169 | 117 | 28 | 1047 | 298 |
| Fire exp. per capita | 108 | 71 | 0.025 | 392 | 288 |
| Police exp. per capita | 167 | 83 | 3.216 | 721 | 298 |
| Total exp. per capita | 2,667 | 878 | 1029 | 6,824 | 298 |
| Residents per km$^2$ of developed land | 1,294 | 1,009 | 269 | 7,091 | 298 |
| Proportion developed | 0.267 | 0.179 | 0.019 | 0.826 | 298 |
| Med./Low. density resid. concentration | 0.124 | 0.105 | 0.000 | 0.444 | 298 |
| Multi/High density resid. concentration | 0.287 | 0.138 | 0.000 | 0.494 | 298 |
| Commercial concentration | 0.304 | 0.109 | 0.000 | 0.489 | 298 |
| Med./Low. density resid. centrality | 0.156 | 0.224 | -0.435 | 1.18 | 298 |
| Multi/High density resid. centrality | 0.704 | 1.205 | -0.499 | 10.993 | 298 |
| Commercial centrality | 0.552 | 0.745 | -0.398 | 7.321 | 298 |
| Med./Low. density resid. isolation | 0.277 | 0.074 | 0.099 | 0.497 | 298 |
| Multi/High density resid. isolation | 0.219 | 0.123 | 0.005 | 0.549 | 298 |
| Commercial isolation | 0.120 | 0.069 | 0.004 | 0.376 | 298 |
| Commercial to Med./Low. density resid. interaction | 0.246 | 0.113 | 0.014 | 0.595 | 298 |
| Commercial to Multi/High density resid. interaction | 0.177 | 0.140 | 0.000 | 0.690 | 298 |
| Med./Low. to Multi/High density resid. interaction | 0.305 | 0.111 | 0.009 | 0.629 | 298 |
| Med./Low. density resid. patch area | 0.033 | 0.029 | 0.004 | 0.193 | 298 |
| Multi/High density resid. patch area | 0.047 | 0.087 | 0.000 | 0.972 | 298 |
| Commercial patch area | 0.017 | 0.010 | 0.002 | 0.076 | 298 |
| Med./Low. density resid. perimeter to area | 0.11 | 0.277 | 0.049 | 4.614 | 298 |
| Multi/High density resid. perimeter to area | 0.194 | 1.267 | 0.015 | 20.791 | 298 |
| Commercial perimeter to area | 0.148 | 0.697 | 0.033 | 11.086 | 298 |
| Population (1000s) | 20.708 | 41.285 | 0.327 | 617.594 | 298 |
| 2000 to 2010 population growth | 0.090 | 0.112 | -0.515 | 0.642 | 298 |
| Proportion under 5 years old | 0.051 | 0.010 | 0.018 | 0.084 | 298 |
| Proportion over 65 years old | 0.151 | 0.046 | 0.074 | 0.398 | 298 |
| Proportion white | 0.910 | 0.090 | 0.416 | 0.988 | 298 |
| Proportion with Bachelor's or higher | 0.388 | 0.159 | 0.110 | 0.793 | 298 |
| Proportion Republican Vote | 0.442 | 0.123 | 0.106 | 0.661 | 298 |
| Proportion Independent Vote | 0.100 | 0.035 | 0.029 | 0.258 | 298 |
| Homeownership rate | 0.779 | 0.125 | 0.362 | 0.975 | 298 |
| Proportion of homes built before 1940 | 0.026 | 0.012 | 0.005 | 0.077 | 298 |
| Number of Jobs in Town per Capita | 0.380 | 0.240 | 0.053 | 1.611 | 298 |
| Income per capita | 35,387 | 25,534 | 5,440 | 289,184 | 298 |
| State revenue per capita | 505 | 381 | 43 | 2,103 | 298 |
| Proportion property tax | 0.735 | 0.142 | 0.214 | 0.943 | 298 |

For the most part, the multidimensional metrics show weak correlations with one another. This is encouraging given similar findings in previous studies (Cutsinger et al., 2005; Frenkel & Ashkenzai, 2008; Jaret et al., 2009). However, focusing on the traditional metrics of density and proportion of land developed, strong correlations exist. Specifically, residential density is highly positively correlated with proportion of land developed (0.72). This is unsurprising given that municipalities with large proportions of their land developed tend to be part of a metro area or sub-regional urban center where density is higher. Commercial concentration is also strongly related to density but in the opposite direction. As density increases, commercial use tend to spread out more evenly across the municipality. Proportion of land developed exhibits high negative correlations with both multi/high density residential and commercial concentration. This implies the concentration metric with respect to these uses is highly dependent on the proportion of land that is developed. This is kept in mind when interpreting regression results.

## 2.5 Results

### 2.5.1 Density

Table 2.2 presents the results from OLS regressions containing no land use measures and regressions with the addition of density. Beginning with the public works regression, the density is negative and somewhat significant: a one standard deviation increase in the number of residents per square km decreases expenditures by 6.84%. This negative result is in line with previous literature although it is not a large effect. The addition of density contributes marginally to the adjusted $R^2$ with the value increasing from 0.348 to 0.351. With respect to fire expenditures, a one standard deviation increase in density increases expenditures per capita by 20.75%. This signals support for the congestion costs/harshness argument and particularly that denser municipalities may have larger fire hazard levels. The addition of density into the regression adds just under 0.01 to the adjusted $R^2$. The density coefficient within the police expenditure regression is positive but insignificant. It is also insignificant but negative within the total expenditures regression.

Table 2.2: Density

| | Public Works | | Fire | | Police | | Total | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | No Land Use | Density | No Land Use | Density | No Land Use | Density | No Land Use | Density |
| Residents per km$^2$ of developed land | | -6.560e-5** | | 18.690e-5*** | | 1.050e-5 | | -1.150e-5 |
| | | (3.320e-5) | | (6.750e-5) | | (4.090e-5) | | (2.280e-5) |
| Population (1000s) | 6.007e-4 | 9.903e-4 | -0.430e-4 | -11.437e-4* | 2.799e-4 | 2.174e-4 | 0.871e-4 | 1.552e-4 |
| | (6.497e-4) | (6.236e-4) | (6.874e-4) | (6.383e-4) | (3.652e-4) | (3.913e-4) | (3.171e-4) | (3.307e-4) |
| 2000 to 2010 population growth | 0.351 | 0.314 | 0.127 | 0.227 | 0.284 | 0.289 | 0.534*** | 0.528*** |
| | (0.271) | (0.261) | (0.321) | (0.309) | (0.283) | (0.279) | (0.155) | (0.153) |
| Proportion under 5 years old | 8.851** | 8.885** | 2.906 | 2.806 | -7.128* | -7.133* | 0.642 | 0.647 |
| | (4.453) | (4.502) | (5.698) | (5.591) | (4.009) | (4.010) | (1.928) | (1.945) |
| Proportion over 65 years old | 4.173*** | 4.133*** | 4.739*** | 4.845*** | 2.588*** | 2.594*** | 2.753*** | 2.746*** |
| | (0.901) | (0.900) | (1.177) | (1.205) | (0.630) | (0.626) | (0.361) | (0.361) |
| Proportion white | -0.008 | -0.163 | -1.770*** | -1.336** | -0.860*** | -0.835** | -0.527*** | -0.554*** |
| | (0.466) | (0.501) | (0.659) | (0.623) | (0.318) | (0.353) | (0.192) | (0.198) |
| Proportion with Bachelor's or higher | 0.273 | 0.289 | 0.513 | 0.468 | -0.967 | -0.970 | 0.458** | 0.461** |
| | (0.409) | (0.410) | (0.930) | (0.927) | (0.642) | (0.642) | (0.188) | (0.188) |
| Proportion Republican Vote | -1.353*** | -1.290*** | 2.160*** | 1.981*** | 1.150*** | 1.139*** | -0.284* | -0.273 |
| | (0.336) | (0.337) | (0.591) | (0.597) | (0.347) | (0.348) | (0.166) | (0.166) |
| Proportion Independent Vote | 0.605 | 0.534 | 0.593 | 0.773 | -1.866 | -1.854 | -0.286 | -0.299 |
| | (0.988) | (0.987) | (2.740) | (2.746) | (1.300) | (1.294) | (0.403) | (0.407) |
| Homeownership rate | 1.253*** | 0.956** | -1.384* | -0.524 | -1.916*** | -1.869*** | 0.317** | 0.265 |
| | (0.340) | (0.401) | (0.718) | (0.857) | (0.525) | (0.616) | (0.156) | (0.195) |
| Proportion of homes built before 1940 | 10.333*** | 9.738*** | -26.772*** | -24.989*** | -2.750 | -2.655 | 3.798*** | 3.694*** |
| | (2.934) | (2.941) | (7.672) | (7.704) | (2.500) | (2.584) | (1.204) | (1.255) |
| Log per capita number of jobs | 0.102* | 0.076 | 0.216** | 0.291*** | 0.108 | 0.112 | 0.094*** | 0.090*** |
| | (0.062) | (0.066) | (0.103) | (0.107) | (0.078) | (0.086) | (0.027) | (0.030) |
| Log per capita income | 0.017 | 0.025 | 0.287 | 0.266 | 0.577* | 0.576* | 0.228*** | 0.229*** |
| | (0.114) | (0.112) | (0.245) | (0.245) | (0.297) | (0.299) | (0.055) | (0.055) |
| Log per capita state revenue | -0.007 | 0.005 | 0.019 | -0.014 | 0.046 | 0.044 | 0.177*** | 0.179*** |
| | (0.075) | (0.075) | (0.098) | (0.095) | (0.073) | (0.076) | (0.037) | (0.038) |
| Proportion property tax | 0.626 | 0.697 | 0.150 | -0.061 | 0.646 | 0.634 | 0.130 | 0.143 |
| | (0.488) | (0.489) | (0.686) | (0.690) | (0.511) | (0.530) | (0.240) | (0.246) |
| Constant | 2.556* | 2.769** | 2.710 | 2.098 | 0.745 | 0.711 | 4.060*** | 4.097*** |
| | (1.309) | (1.335) | (1.933) | (1.910) | (2.861) | (2.825) | (0.620) | (0.623) |
| Observations | 298 | 298 | 288 | 288 | 298 | 298 | 298 | 298 |
| R-squared | 0.379 | 0.384 | 0.436 | 0.447 | 0.429 | 0.429 | 0.631 | 0.631 |
| Adjusted R-squared | 0.348 | 0.351 | 0.407 | 0.416 | 0.401 | 0.399 | 0.612 | 0.612 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Signs and significance on control variables are as expected and exhibit little variation when density is added. The proportion of residents over 65 is positive and significant across all models. Proportion white exerts a negative effect on all expenditures although is not significant for public works expenditures. Higher education is only significant for total expenditures, where the relationship is positive. The Republican vote share from the 2010 Gubernatorial election has a strong negative effect on public works expenditures but positively affects fire and police expenditures. An increase in homeownership rates significantly increase public works expenditures and total expenditures, but decreases spending on fire and police. A larger amount of homes built before 1940 is associated with larger public works expenditures and total expenditures, but significantly lower fire services expenditures. Employment per capita positively influences all expenditures but is only significant for fire and total expenditures. Per capita income is positive and is weakly significant for police expenditures and strongly significant for total expenditures. Lastly, proportion of revenue coming from property tax is insignificant for all expenditure regressions.

### 2.5.2 Multidimensional Metrics

Beginning with concentration, Table 2.3 presents results for all expenditure categories with the inclusion of controls.[15] Med./low density residential concentration is positive and significant at the 5% level with respect to public works expenditures; increasing med./low density residential concentration by a standard deviation increases public works expenditures per capita by 9.92%. The concentration coefficient within the fire expenditure specification is positive but not significant. However, the coefficient is very large and highly significant with respect to police expenditures. Specifically, a one standard deviation increase in the concentration of low density residential land use decreases police expenditures by 14.49%. This is in contrast to the positive and significant coefficient on multi/high density residential concentration for police expenditures. Here, a one standard deviation increase in multi/high density residential concentration increases police expenditures by 12.04%.

[15]Control coefficients change very little and are not presented.

Commercial concentration produces strong negative coefficients for both fire and police expenditures; 31.30% and 20.10% increase in fire and police expenditures respectively. The centrality metric (Table 2.4) does not display a significant relationship with respect to any of the expenditure categories.

The majority of the isolation coefficients across the four regressions are negative (Table 2.5). Multi/high density residential and commercial isolation are significant at the 10% and 5% level with respect to public works. These coefficients translate into 6.47% and 6.73% decreases in public expenditures for one standard deviation increase in multi/high density residential and commercial isolation respectively. These metrics are also significant at the 5% level within the total expenditures regression but with smaller effects, 2.84% and 2.69% decrease in total expenditures for one standard deviation increase in multi/high density residential and commercial isolation respectively.

The interaction between commercial and med./low density residential uses produces the strongest coefficient. Specifically, a one standard deviation increase in the interaction of these uses increases police expenditures by 10.36%.

Mean patch area is not a strong predictor of the disaggregate expenditure categories, although commercial mean patch area does produce a somewhat significant negative coefficient for total expenditures. Here, a one standard deviation increase in the commercial mean patch area decreases total expenditures by 2.83%. The mean perimeter to area ratio proves to be a more statistically significant measure of fragmentation than patch area but its influence is minimal. With respect to public works expenditures, a one standard deviation increase in med./low and multi/high density residential perimeter to area ratio increases expenditures by 3.15% and 2.83%. Whereas a one standard deviation increase in the commercial ratio decreases public works expenditures by 2.18%. Fire expenditures increase by 7.95% and 4.85% when med./low density residential and commercial perimeter to area ratio increases by one standard deviation. Police expenditures are significantly but minimally impacted by a one standard deviation increase in the commercial ratio. The multi/high density residential ratio has a similarly significant but minimal impact on total expenditures.

Taking these results together, there are clear relationships between the various metrics and municipal expenditures, although some metrics prove to be inconsequential. A more thorough analysis is presented alongside the full regression specifications.

Table 2.3: Concentration

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Med./Low. density resid. concentration | 0.901** | 1.226 | -1.289*** | 0.233 |
| | (0.350) | (0.808) | (0.404) | (0.170) |
| Multi/High density resid. concentration | 0.095 | -0.241 | 0.824** | 0.263 |
| | (0.388) | (0.520) | (0.322) | (0.165) |
| Commercial concentration | 0.172 | -2.498*** | -1.680*** | -0.562* |
| | (0.565) | (0.666) | (0.573) | (0.286) |
| Constant | 1.763 | 1.955 | 1.876 | 3.883*** |
| | (1.267) | (2.303) | (2.568) | (0.633) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.398 | 0.466 | 0.484 | 0.645 |
| Adjusted R-squared | 0.361 | 0.433 | 0.453 | 0.623 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### 2.5.3 Full Model

Turning to the full regressions presented in Table 2.9, I include all multidimensional metrics as well as municipal controls.[16] For the most part, signs and magnitudes of the coefficients remain consistent in comparison to the individual metric regressions. To ease comparison and interpretation, Table 2.10 contains the expected percent changes in expenditures for a one standard deviation increase in the metric. Overall, I find lower expenditures per capita for some but not all of the development patterns associated with the smart growth movement.

[16] Control coefficients for all single multidimensional metric specifications can be found in Table 2.11 in the Appendix.

Table 2.4: Centrality

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Med./Low. density resid. centrality | 0.092 | -0.052 | 0.118 | -0.048 |
| | (0.145) | (0.486) | (0.204) | (0.087) |
| Multi/High density resid. centrality | -0.007 | -0.207 | -0.033 | -0.014 |
| | (0.021) | (0.126) | (0.032) | (0.010) |
| Commercial centrality | 0.056 | 0.121 | -0.004 | 0.008 |
| | (0.039) | (0.100) | (0.046) | (0.020) |
| Constant | 2.762** | 3.153* | 0.766 | 4.089*** |
| | (1.307) | (1.803) | (2.807) | (0.625) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.385 | 0.475 | 0.434 | 0.634 |
| Adjusted R-squared | 0.347 | 0.442 | 0.400 | 0.612 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

### Public Works

Beginning with public works expenditures, it is hypothesized that more compact land use patterns utilize less infrastructure and that mixing of residential and commercial uses increases accessibility and thus limits the amount of road infrastructure required. More complex, winding patterns of residential use are assumed to require more road, water, and sewer infrastructure per capita. Findings are mixed with respect to these hypotheses. Notions of general compactness (concentration and centrality) do not produce strong support of efficiency gains. And if anything, there is a somewhat positive relationship between the concentration of med./low density residential use and public works expenditures per capita. A one standard deviation increase, increases expenditures by 8.09%. Commercial centrality also appears to increase expenditures by 6.86% but is only marginally significant.

The effect of mixed use was hypothesized to be ambiguous, but there does appear to be some consistency across the coefficients. Signs on all interaction metrics are negative, but only the interaction between commercial use and med./low density residential produces a slightly significant negative coefficient implying a 6.26% decrease in expenditures for a one

Table 2.5: Isolation

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Med./Low. density resid. isolation | 0.100 | -1.104 | -0.319 | -0.088 |
| | (0.358) | (0.708) | (0.484) | (0.193) |
| Multi/High density resid. isolation | -0.510* | -0.957 | -0.255 | -0.288** |
| | (0.263) | (0.870) | (0.336) | (0.135) |
| Commercial isolation | -0.944** | 0.308 | -0.036 | -0.384** |
| | (0.476) | (0.808) | (0.477) | (0.186) |
| Constant | 2.596** | 3.180* | 0.871 | 4.122*** |
| | (1.304) | (1.866) | (2.791) | (0.628) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.395 | 0.452 | 0.434 | 0.645 |
| Adjusted R-squared | 0.358 | 0.418 | 0.400 | 0.624 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.6: Interaction

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Commercial to Med./Low. density resid. proximity | -0.424 | 0.611 | 0.872*** | 0.038 |
| | (0.325) | (0.542) | (0.264) | (0.140) |
| Commercial to Multi/High density resid. proximity | -0.573* | 0.516 | 0.178 | -0.199 |
| | (0.310) | (0.450) | (0.286) | (0.150) |
| Med./Low. to Multi/High density resid. proximity | 0.018 | -1.113* | 0.085 | -0.063 |
| | (0.228) | (0.589) | (0.234) | (0.100) |
| Constant | 2.465* | 2.499 | 1.035 | 4.065*** |
| | (1.316) | (2.001) | (2.760) | (0.636) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.389 | 0.445 | 0.455 | 0.635 |
| Adjusted R-squared | 0.352 | 0.410 | 0.422 | 0.613 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.7: Mean Patch Area

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Med./Low. density resid. patch area | 0.020 | 0.885 | 0.259 | -0.769 |
| | (1.731) | (1.089) | (0.718) | (0.467) |
| Multi/High density resid. patch area | -0.392 | 0.015 | -0.180 | -0.163 |
| | (0.400) | (0.369) | (0.205) | (0.107) |
| Commercial patch area | -2.689 | -3.644 | -4.171 | -2.788** |
| | (3.316) | (4.085) | (2.956) | (1.387) |
| Constant | 2.494* | 2.676 | 0.658 | 3.913*** |
| | (1.352) | (1.951) | (2.871) | (0.619) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.384 | 0.437 | 0.434 | 0.644 |
| Adjusted R-squared | 0.346 | 0.401 | 0.399 | 0.623 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8: Mean Perimeter to Area Ratio

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Med./Low. density resid. perimeter to area | 0.112** | 0.276*** | 0.046 | 0.006 |
| | (0.046) | (0.083) | (0.035) | (0.018) |
| Multi/High density resid. perimeter to area | 0.022*** | 0.006 | 0.002 | 0.016*** |
| | (0.006) | (0.008) | (0.006) | (0.003) |
| Commercial perimeter to area | -0.031*** | 0.068** | 0.023** | -0.004 |
| | (0.012) | (0.027) | (0.010) | (0.008) |
| Constant | 2.454* | 2.736 | 0.765 | 4.014*** |
| | (1.314) | (1.942) | (2.874) | (0.623) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.386 | 0.443 | 0.431 | 0.635 |
| Adjusted R-squared | 0.348 | 0.408 | 0.396 | 0.613 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

standard deviation increase. Isolation of use also produces consistent negative coefficients with multi/high residential density and commercial use significant at the 10% level. The magnitude of the effects with respect to a standard deviation increase is similar to the interaction coefficient, with standardized effects of -6.26% and -8.33%. Thus, there is suggestive evidence that servicing isolated uses (large residential or commercial developments) may provide efficiencies; however, savings also occur when med./low density residential and commercial uses in particular are proximate.

More fragmented use, or smaller patch areas were hypothesized to be more costly to service based on issues of accessibility. This relationship holds for lower density residential and commercial use although it is not significant.

Both residential perimeter to area ratios are significant and positively related to public works expenditures. This signifies some support of the hypothesis that expenses are greater when servicing complex winding land use patterns. However, the effect is small as a one standard deviation increase in the med./low and multi/high density residential ratios increases public expenditures by only 3.24% and 2.96% respectively.

Overall, the relationships between the multidimensional metrics and public works expenditures are suggestive but somewhat small in magnitude. This is likely due to two issues. First, public works contains a number of different expenditures that may not all be affected by development patterns in similar ways although it is likely there are some commonalities. Second, there are large variations in service levels across the municipalities considered. Of the 351 Massachusetts municipalities, 290 do not have water or sewer service, 14 have only sewer, 16 have only water, and 31 have both water and sewer (Massachusetts Water Resources Authority, 2017). With respect to waste management, low density residential areas in smaller municipalities are serviced by transfer stations, not curbside pickup.

### Fire

Previously, density was shown to increase fire expenditures substantially, lending strong support for the harshness hypothesis. A priori, it is unclear how residential concentration may impact fire expenditures although issues of accessibility may be related. Complex, winding land uses are hypothesized to increase expenditures given decreases in accessibility.

More concentrated (less evenly spread throughout the municipality) med./low density residential use increases fire expenditures by 20.10% for a one standard deviation increase. This is somewhat counterintuitive as accessibility is expected to increase with concentration but this notion may not be applicable when considering low density residential use. Multi/high density residential concentration produces the expected sign but is not significant. However, commercial concentration is large, negative, and highly significant. Fire expenditures per capita decrease by 23.48% for a one standard deviation increase in concentration.

Interaction between commercial use and the two residential groups produces positive coefficients with a large, slightly significant effect for the interaction between commercial and med./low density residential. Specifically, a one standard deviation increase in interaction increases expenditures by 17.30%. This increase is in contrast to the decrease associated with the interaction of the residential uses across densities. Here a one standard deviation increase in interaction, decreases expenditures by 19.40%. Given that residential use is considered low-hazard versus the medium-hazard of commercial uses, this result is not surprising. Furthermore, commercial isolation is positive but not significant and is positive and somewhat significant when interacted with residential use. This may represent increased service requirements for residential uses when they are located near higher hazard uses.

Coefficients for the perimeter to area ratio support the hypothesis of larger expenditures being required to service more complex, winding land use patterns. All coefficients are positive with med./low density residential perimeter to area ratio having the largest effect on expenditures; a one standard deviation increase in the ratio increases expenditures by 6.55%. The ratio is not significant for commercial uses.

### Police

Minimal hypotheses were made regarding the role of development patterns on police expenditures aside from the harshness hypothesis and the role of mixed use in crime prevention or attraction. The results indicate more concentrated patterns of commercial and lower density residential uses reduce police expenditures; a one standard deviation increase in the concentration of med./low density residential and commercial use decreases police expenditures by 9.84 and 14.18% respectively. However, increasing multi/high density residential concentration increases police expenditures by 9.49%. Higher crime rates in more dense residential neighborhoods may be driving these results and suggest support of the harshness hypothesis.

The isolation of residential use produces negative but insignificant coefficients. When commercial use are more proximate to other commercial use, police expenditures increase. Specifically, a one standard deviation increase in the isolation of commercial use increases expenditures by 7.63% and is significant at the 5% level. This may signal higher levels of crime in predominately commercial areas. Commercial proximity to residential use also increases police expenditures although the effect is only significant with respect to med./low density residential. A one standard deviation increase in the interaction between commercial and med./low density residential use increases police expenditures by 11.89%. This result may indicate more police services are required in municipalities with more mixed use areas and lends suggestive support to Browning et al.'s (2010) findings of increased nonviolent crime in mixed use areas.

The relationship between patch size and police expenditures is theoretically unclear although a significant relationships exist within the data. Specifically, larger contiguous developments of med./low density residential and commercial land use decrease expenditures. Perimeter to area ratio is not significant but does produce positive coefficients which are consistent with the results from public works and fire regressions.

### Total

Given the various expenditures included within the aggregated category, hypotheses regarding the effect of development patterns on the total expenditures are general at best. In the previous regressions with the inclusion of density, neither the economies of scale nor the harshness theory found support. With respect to the multidimensional metrics, I highlight consistencies across expenditure categories below.

All three disaggregate regressions display a negative relationship with commercial concentration. Total expenditures are also negatively related; a one standard deviation increase produces a 6.92% decrease in total expenditures. There is also consistency with respect to multi/high density residential isolation. All three disaggregate regressions and total expenditures produce negative coefficients. Patch area coefficients are negative for all use types and somewhat significant for the residential uses although the magnitude is small. Multi/high density residential perimeter to area is consistently positive and significant across all models with the exception of the police regression where it is insignificant but remains positive. With respect to total expenditures, a one standard deviation increase in the ratio increases expenditures by 2.44%. Not a large value, but consistent with other expenditures.

As compared to the regressions containing only density, explanatory power with the inclusion of the multidimensional metrics is larger across all four expenditure models. Specifically, adjusted $R^2$ increases by 0.024, 0.060, 0.051, and 0.031 for public works, fire, police, and total expenditures respectively. These results imply that the collection of multidimensional metrics are capturing significant relationships between land use patterns and municipal expenditures not captured by density.

Table 2.9: Full Regressions

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Med./Low. density resid. concentration | 0.741* | 1.744** | -0.987** | 0.344 |
| | (0.401) | (0.831) | (0.434) | (0.218) |
| Multi/High density resid. concentration | -0.071 | -0.645 | 0.657* | 0.138 |
| | (0.471) | (0.710) | (0.362) | (0.192) |
| Commercial concentration | -0.178 | -2.455*** | -1.403** | -0.658** |
| | (0.621) | (0.897) | (0.614) | (0.322) |
| Med./Low. density resid. centrality | 0.100 | 0.128 | 0.148 | -0.101 |
| | (0.159) | (0.472) | (0.199) | (0.089) |
| Multi/High density resid. centrality | 0.007 | -0.148 | -0.025 | -0.008 |
| | (0.022) | (0.100) | (0.028) | (0.010) |
| Commercial centrality | 0.089* | 0.110 | -0.021 | 0.010 |
| | (0.047) | (0.087) | (0.044) | (0.023) |
| Med./Low. density resid. isolation | -0.113 | -0.032 | -0.130 | 0.086 |
| | (0.423) | (0.714) | (0.557) | (0.203) |
| Multi/High density resid. isolation | -0.526* | -1.718 | -0.254 | -0.266** |
| | (0.294) | (1.064) | (0.362) | (0.134) |
| Commercial isolation | -1.261* | 0.929 | 1.066** | -0.107 |
| | (0.674) | (1.121) | (0.534) | (0.267) |
| Commercial to Med./Low. density resid. interaction | -0.681* | 1.412* | 0.994*** | 0.222 |
| | (0.390) | (0.807) | (0.352) | (0.185) |
| Commercial to Multi/High density resid. interaction | -0.535 | 0.598 | 0.313 | -0.063 |
| | (0.442) | (0.737) | (0.398) | (0.217) |
| Med./Low. to Multi/High density resid. interaction | -0.163 | -1.943** | -0.194 | -0.179 |
| | (0.288) | (0.840) | (0.298) | (0.124) |
| Med./Low. density resid. patch area | 1.732 | 0.164 | -1.903** | -0.950* |
| | (1.806) | (1.604) | (0.919) | (0.528) |
| Multi/High density resid. patch area | -0.375 | -0.547 | 0.073 | -0.221* |
| | (0.436) | (0.394) | (0.236) | (0.126) |
| Commercial patch area | 1.818 | -1.171 | -6.626* | -2.003 |
| | (4.721) | (5.296) | (3.423) | (1.859) |
| Med./Low. density resid. perimeter to area | 0.115** | 0.229*** | 0.043 | -0.005 |
| | (0.046) | (0.083) | (0.037) | (0.020) |
| Multi/High density resid. perimeter to area | 0.023*** | 0.018* | 0.008 | 0.019*** |
| | (0.006) | (0.011) | (0.006) | (0.003) |
| Commercial perimeter to area | -0.010 | 0.036 | -0.012 | -0.010 |
| | (0.018) | (0.029) | (0.015) | (0.007) |
| Constant | 2.249* | 2.389 | 1.600 | 3.717*** |
| | (1.339) | (2.029) | (2.463) | (0.675) |
| Controls | X | X | X | X |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.442 | 0.534 | 0.510 | 0.681 |
| Adjusted R-squared | 0.375 | 0.476 | 0.450 | 0.643 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.10: Standardized Coefficients

| | Public Works | Fire | Police | Total |
|---|---|---|---|---|
| Med./Low. density resid. concentration | 8.09* | 20.10** | -9.84** | 3.68 |
| Multi/High density resid. concentration | -0.98 | -8.52 | 9.49* | 1.92 |
| Commercial concentration | -1.92 | -23.48*** | -14.18** | -6.92** |
| Med./Low. density resid. centrality | 2.27 | 2.91 | 3.37 | -2.24 |
| Multi/High density resid. centrality | 0.85 | -16.33 | -2.97 | -0.96 |
| Commercial centrality | 6.86* | 8.54 | -1.55 | 0.75 |
| Med./Low. density resid. isolation | -0.83 | -0.24 | -0.96 | 0.64 |
| Multi/High density resid. isolation | -6.26* | -19.05 | -3.08 | -3.22** |
| Commercial isolation | -8.33* | 6.62 | 7.63** | -0.74 |
| Commercial to Med./Low. density resid. interaction | -7.41* | 17.30* | 11.89*** | 2.54 |
| Commercial to Multi/High density resid. interaction | -7.22 | 8.73 | 4.48 | -0.88 |
| Med./Low. to Multi/High density resid. interaction | -1.79 | -19.40** | -2.13 | -1.97 |
| Med./Low. density resid. patch area | 5.15 | 0.48 | -5.37** | -2.72* |
| Multi/High density resid. patch area | -3.21 | -4.65 | 0.64 | -1.90* |
| Commercial patch area | 1.83 | -1.16 | -6.41* | -1.98 |
| Med./Low. density resid. perimeter to area | 3.24** | 6.55*** | 1.20 | -0.14 |
| Multi/High density resid. perimeter to area | 2.96*** | 2.31* | 1.02 | 2.44*** |
| Commercial perimeter to area | -0.69 | 2.54 | -0.83 | -0.69 |

## 2.6 Conclusion

The results of this paper highlight the complex relationship between land use patterns and municipal expenditures while providing both support for and against development patterns advocated by smart growth and new urbanism literatures. Density provides an intuitive link between land use intensity, economies of scale, and congestion costs; however, it is ill-equipped in identifying more micro-level patterns and interactions of uses within a municipality. Using multidimensional metrics developed within the planning and landscape ecology literature, I show that municipal expenditures are very much influenced by complex aspects of land use and that these metrics are better predictors of expenditures than a simple density measure.

I find support for some aspects of compact, mixed, and connected development patterns with respect to municipal expenditures. Concentrated, or more clustered commercial use leads to decreases in fire, police, and total expenditures where higher levels of concentration are associated with less diffuse land use patterns. However, public works and fire expenditures increase as med./low density residential use becomes more concentrated. Although smart growth and new urbanism emphasize the importance of centrality, there appears to be no benefit to more centralized uses with respect to municipal expenditures. In opposition to what is prescribed by new urbanism and smart growth, I find that more isolated residential use is associated with lower public works expenditures and that proximity between residential and commercial use does not produce lower levels of fire, police, or total expenditures. In particular, fire and police expenditures per capita are significantly larger when proximity between lower density residential and commercial use increases. In support of smart growth hypotheses, I find more complex, winding residential land use patterns significantly increase public works, fire, and total expenditures.

Implications of these findings are both academic and policy related. When possible, researchers should consider additional measures of land use patterns in conjunction with or instead of density. A density measure may be masking the true channels through which relationships of interest are operating. On the policy side, new urbanists and smart growth advocates should reconsider the emphasis put on density for two reasons. First, increasing density does not necessarily translate into lower municipal expenditures per capita as the relationship is likely nonlinear. Second, attention should be paid to different land use patterns that show a more clear connection to municipal expenditures. More research is warranted on these metrics as the results presented here indicate some aspects of compact developments (aside from density) may in fact increase municipal expenditures.

## 2.7 Appendix

Table 2.11: Controls for Full Regression

| VARIABLES | (1) Public Works | (2) Fire | (3) Police | (4) Total |
|---|---|---|---|---|
| Population (1000s) | 4.978e-4 | -5.667e-4 | 6.917e-4 | -1.792e-4 |
| | (6.094e-4) | (6.516e-4) | (4.325e-4) | (3.070e-4) |
| 2000 to 2010 population growth | 0.244 | 0.274 | 0.279 | 0.382*** |
| | (0.275) | (0.381) | (0.296) | (0.143) |
| Proportion under 5 years old | 11.277** | 4.945 | -5.561 | 2.398 |
| | (4.513) | (5.748) | (3.913) | (1.781) |
| Proportion over 65 years old | 5.073*** | 3.526*** | 1.922** | 2.786*** |
| | (0.976) | (1.231) | (0.770) | (0.380) |
| Proportion white | -0.119 | -0.743 | -0.487 | -0.417** |
| | (0.520) | (0.607) | (0.350) | (0.189) |
| Proportion with Bachelor's or higher | 0.348 | 0.902 | -0.883 | 0.487*** |
| | (0.388) | (0.720) | (0.601) | (0.185) |
| Proportion Republican Vote | -0.714** | 2.330*** | 0.583* | -0.216 |
| | (0.350) | (0.705) | (0.345) | (0.170) |
| Proportion Independent Vote | 0.667 | 1.478 | -2.596** | -0.199 |
| | (0.890) | (2.422) | (1.281) | (0.420) |
| Homeownership rate | 0.837* | -1.003 | -1.697*** | 0.336* |
| | (0.442) | (0.779) | (0.564) | (0.196) |
| Proportion of homes built before 1940 | 5.386 | -22.429** | -1.104 | 3.099** |
| | (3.374) | (8.959) | (3.227) | (1.341) |
| Log per capita number of jobs | 0.133* | 0.206* | 0.067 | 0.111*** |
| | (0.071) | (0.111) | (0.096) | (0.032) |
| Log per capita income | 0.049 | 0.374 | 0.495* | 0.255*** |
| | (0.110) | (0.236) | (0.256) | (0.056) |
| Log per capita state revenue | 0.047 | -0.031 | 0.048 | 0.196*** |
| | (0.078) | (0.105) | (0.074) | (0.041) |
| Proportion property tax | 0.786 | -0.836 | 0.571 | 0.141 |
| | (0.528) | (0.850) | (0.519) | (0.276) |
| Constant | 2.249* | 2.389 | 1.600 | 3.717*** |
| | (1.339) | (2.029) | (2.463) | (0.675) |
| Observations | 298 | 288 | 298 | 298 |
| R-squared | 0.442 | 0.534 | 0.510 | 0.681 |
| Adjusted R-squared | 0.375 | 0.476 | 0.450 | 0.643 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.12: Correlation Matrix of Land Use Measures

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (2) | 0.72 | | | | | | | | | | | | | | | | | | |
| (3) | 0.04 | -0.38 | | | | | | | | | | | | | | | | | |
| (4) | -0.65 | -0.85 | 0.38 | | | | | | | | | | | | | | | | |
| (5) | -0.71 | -0.83 | 0.25 | 0.81 | | | | | | | | | | | | | | | |
| (6) | -0.38 | -0.50 | 0.31 | 0.47 | 0.39 | | | | | | | | | | | | | | |
| (7) | -0.12 | -0.27 | 0.20 | 0.28 | 0.26 | 0.41 | | | | | | | | | | | | | |
| (8) | 0.07 | -0.21 | 0.20 | 0.18 | 0.22 | 0.35 | 0.54 | | | | | | | | | | | | |
| (9) | 0.18 | 0.57 | -0.42 | -0.39 | -0.40 | -0.17 | -0.14 | -0.18 | | | | | | | | | | | |
| (10) | 0.51 | 0.49 | 0.12 | -0.36 | -0.40 | -0.25 | -0.06 | -0.01 | 0.15 | | | | | | | | | | |
| (11) | 0.05 | 0.36 | -0.33 | -0.17 | -0.21 | -0.20 | -0.10 | -0.17 | 0.45 | 0.19 | | | | | | | | | |
| (12) | -0.03 | -0.04 | 0.02 | 0.03 | 0.02 | -0.06 | -0.05 | -0.02 | -0.03 | -0.03 | -0.04 | | | | | | | | |
| (13) | 0.03 | -0.01 | 0.00 | -0.01 | 0.00 | 0.01 | 0.02 | 0.05 | -0.02 | 0.02 | 0.00 | -0.01 | | | | | | | |
| (14) | 0.05 | 0.08 | -0.05 | -0.08 | -0.13 | -0.01 | -0.01 | -0.05 | -0.01 | 0.15 | -0.02 | -0.01 | -0.01 | | | | | | |
| (15) | -0.10 | 0.05 | -0.03 | 0.07 | 0.05 | 0.19 | 0.21 | 0.10 | 0.29 | 0.08 | 0.24 | -0.11 | 0.07 | -0.06 | | | | | |
| (16) | 0.40 | 0.42 | -0.09 | -0.43 | -0.42 | -0.16 | 0.02 | 0.04 | 0.29 | 0.34 | 0.22 | -0.08 | 0.10 | 0.00 | 0.29 | | | | |
| (17) | 0.20 | 0.45 | -0.30 | -0.42 | -0.39 | -0.28 | -0.18 | -0.19 | 0.31 | 0.13 | 0.67 | -0.06 | 0.05 | -0.03 | 0.19 | 0.35 | | | |
| (18) | -0.30 | -0.13 | -0.36 | 0.23 | 0.15 | 0.25 | 0.16 | 0.15 | 0.29 | -0.27 | 0.03 | -0.07 | -0.07 | -0.01 | 0.21 | -0.22 | -0.27 | | |
| (19) | 0.72 | 0.70 | -0.04 | -0.69 | -0.67 | -0.28 | -0.01 | 0.10 | 0.22 | 0.57 | 0.14 | -0.05 | 0.11 | 0.11 | 0.03 | 0.59 | 0.28 | -0.39 | |
| (20) | 0.25 | 0.34 | -0.25 | -0.26 | -0.29 | 0.10 | 0.18 | 0.11 | 0.23 | 0.09 | 0.02 | -0.05 | -0.05 | 0.09 | 0.08 | -0.15 | -0.03 | 0.40 | 0.18 |

(1) Residents per $km^2$ dev. Land (2) Proportion of land developed (3) Low & Med. density resid. concentration (4) Multi & High density resid. concentration (5) Commercial concentration (6) Low & Med. density resid. centrality (7) Multi & High density resid. centrality (8) Commercial centrality (9) Low & Med. density resid. patch area (10) Multi & High density resid. patch area (11) Commercial patch area (12) Low & Med. density resid. perimeter to area (13) Multi & High density resid. perimeter to area (14) Commercial perimeter to area (15) Low & Med. density resid. isolation (16) Multi & High density resid. isolation (17) Commercial isolation (18) Commercial to Med./Low. density resid. interaction (19) Commercial to Multi/High density resid. interaction (20) Med./Low. to Multi/High density resid. interaction

# Chapter 3

# Commercial Property Tax Incidence: Evidence from the Rental Market

Within states that allow tax classification, many municipalities increase their reliance on commercial property tax revenue as an alternative to increasing the burden on residential property. However, empirical evidence on the implications of commercial property tax incidence in the rental market is scant. Specifically, a clear consensus does not exist regarding whether owners of capital internalize increases in tax payments or pass these increases onto the renters of commercial space. This is a difficult question to address given multiple sources of endogeneity. I make progress on this issue through the use of a rich panel of asking office rents in a nonlinear control function setting. Once the endogeneity issues are addressed, robust findings indicate that a \$1 increase in the commercial property tax payment per square foot increases rent per square foot by more than \$0.90. This transfer of tax burden to the renters of commercial space may in turn put upward pressure on the price of consumer goods and downward pressure on local wages.

## 3.1 Introduction

Property tax incidence is a well-trodden area of research within the local public finance literature. However, little attention has been paid to the incidence associated with commercial property taxation. Municipalities are increasingly reliant upon commercial tax revenue; thus, it is imperative that we fully understand the implications of such reliance. The paper presented here seeks to add to this understanding by investigating a particular aspect of commercial property taxation. Namely, do owners of capital internalize increases in tax payments or are these increases passed onto the renters of commercial space?

Anecdotally, it is assumed a landlord passes through 100% of tax increases onto tenants, although the limited extant literature does not support this hypothesis. A 100% pass through may be anticipated under specific forms of commercial rent contracts as some contracts stipulate that tenants are liable for tax increases. There are two main classifications of commercial leases: gross and net. Under gross leases, landlords will typically cover all or the majority of expenses associated with the property. This may include property taxes, insurance, and common area maintenance. A modified gross lease may require the tenant to cover a subset of the costs associated with the property. For example, a contract may stipulate the rent will increase based on the differential between the current tax rate and the baseline tax rate. This is known as a tax escalation clause. The modified gross lease is a hybrid between the gross lease and the net lease. A net lease will have a lower base rent but costs associated with property taxes, insurance, maintenance, utilities, property management, and so on are included net of the baseline rent. Even if a contract dictates that tax rate increases are to be covered by the tenant, the landlord may choose to exercise this right to varying degrees as increases in rent may negatively affect vacancy rates given the market faced by the landlord. Thus, it is not necessarily true that changes in the commercial tax rate are fully and automatically reflected in gross lease rates.

The question of commercial property tax incidence is not straightforward. There is an inherent simultaneity between rents, property valuations, and tax rates making causality hard to pin down. Consider a positive shock to rents. The effect on the tax payment is difficult to disentangle as two forces may be in play: the positive rent shock puts upward pressure on value which in turn dictates the tax payment and the tax rate may also be influenced by the rent shock as higher valuations will increase a town's tax base. With a larger tax base, a town may choose to decrease the tax rate. With these multiple sources of confounding, it is necessary to isolate exogenous changes in the tax rate and to account for the endogeneity between rents and valuations. I address these issues in two ways. First, I employ a nonlinear regression specification that eliminates endogeneity between rents and valuations. Second, a control function approach is utilized to capture the exogenous movement in the tax rate not associated with movements in property value. I exploit a rich panel of office rent data representing asking rents for new gross lease contracts covering 96 Massachusetts municipalities over 27 years.

I find that failing to control for the reverse causality between tax rates and rent as well as the endogeneity between rents and valuations produces a downward biased incidence parameter. However, within the nonlinear control function setting, tax incidence is positive and significant implying a nearly 100% pass through of tax payments into rent. This result is robust across numerous instruments and represents the primary form of commercial property tax incidence. The secondary form is related to the behavior of commercial space users facing rising rents. Thus, it is possible the tax burden may be partially borne by consumers through higher consumer goods prices and labor through lower wages as tenants seek to minimize the effect of a rent increase. I do not investigate secondary incidence in the present research, but consider it an important avenue of future research.

The paper proceeds with a brief overview of the property tax incidence literature in Section 3.2 followed by a summary of recent property tax history in Massachusetts. Sections 3.4 and 3.5 outline the methods and data used in the analysis. Results are presented in Section 3.6.

## 3.2 Tax Incidence Theory

One of the longest on-going debates in local public finance literature centers on how to model property tax incidence. Two views have emerged: the capital view and the benefit view. The capital view, whose origins lie in Harberger's (1962) general equilibrium model, hypothesizes that the average burden of property tax in the nation is borne by capital owners. Mieszkowski (1972) describes this portion of incidence as the profits tax component. If capital owners are predominately high-income individuals then property taxation from the national perspective is progressive. However, at the local level, the indirect effects of capital migration also contribute to the overall incidence. The distortionary excise tax component of the incidence arises out of the change in productivity of local factors resulting from capital migration. For example, with an increase in property taxes, capital migrates out lowering the productivity of local factors of production and the competitive returns to these factors. Commodity prices rise while wages and land prices fall. Fuest, Peichl, and Siegloch (2015) provide robust empirical evidence of incidence of a local business tax with respect to wages in Germany. Specifically, a 1-euro increase in the tax bill decreases the wage bill by 56 euro-cents.

Although these excise tax effects are thought to cancel out in the aggregate leaving the profits tax component the primary driver of incidence, they may lead to distorted allocations of capital across jurisdictions. Under the capital view, tax competition is interpreted as harmful as the tax base may relocate to a jurisdiction with a lower rate.

When focusing only on commercial property tax incidence, the argument that local land prices will fall may not be an accurate expectation. Specifically, land may be perfectly substitutable between residential and commercial uses. Thus, a decrease in the demand for commercial land does not necessarily lead to an overall reduction in land values if the land can be redeveloped for residential use with limited effort. The benefit view recognizes this substitutability and asserts that the incidence lies only on labour and/or capital (Fischel, 1974). If this is true, then a municipality has the incentive to tax commercial property when local residents are not the capital owners, workers, or consumers of local production (N. J. Lee & Wheaton, 2010).

The benefit view further emphasizes the relationship between taxation and resulting municipal services. Given that firms may have different preferences for municipal services than homeowners, placing a greater burden on the commercial property class without linking that tax revenue to services valued by commercial property owners may result in inefficient sorting of capital and not the efficient sorting Tiebout (1956) argues occurs from the household perspective. Related to this hypothesis, Kang, Skidmore, and Reese (2015) find that commercial property values are more responsive to changes in tax rates than changes in school spending. The opposite is found for residential property.

Empirically distinguishing between these two models has proved difficult[1]; however, testable implications exist within the rental market setting. In particular, forward shifting to lessees occurs under both views but backward shifting to capital owners is only present under the new view. Under the new view, higher rents result when demand for commercial rental property is less elastic than supply. Thus the tax burden is passed backward to the landlord when supply is relatively more elastic. There is limited empirical work in the rental market setting; I summarize these findings here.

Wheaton (1984) investigates the incidence of commercial property rates in the Boston Metropolitan area and finds that rental rates are unaffected by interjurisdictional tax differentials. He concludes that capital owners bear the burden of the differential and thus finds no support for the benefit view. It should be noted, however, that this study is cross-sectional in nature and thus focused on the variation in rates across municipalities within the Boston Metropolitan area.

In another cross-sectional approach situated in the downtown Chicago commercial rental market, McDonald (1993) investigates the tax incidence associated with changes in a building owner's tax payment. Building level rent and assessment data is used to construct a simple OLS model which includes building and location characteristics. McDonald finds

[1]see Nechyba (1997) for a review.

that a $1 increase in the tax payment per square foot increases rents by $0.45. It is not entirely clear that assessed value is exogenous to rent levels; thus, McDonald instruments the assessed value with total square footage of rentable space and the proportion of rentable space devoted to offices. A Hausman test concludes that assessed value per square foot is exogenous. These results suggest the demand for commercial rental space within a downtown region are not perfectly price elastic.

In contrast to the United States, occupiers of commercial property in the UK are formally liable for the commercial property tax. The tax burden may then be shifted to consumers via higher product prices, to employees via real lower wages, or to property owners through lower rents. Bond, Denny, Hall, and McCluskey (1996) explore this incidence over a period of six years for a panel of nearly 3000 commercial properties. Their findings indicate that in the long run, above average commercial tax rate increases are associated with below average increases in rents.

Tsoodle and Turner (2008) present more recent work in the rental market setting but for residential tax incidence. The authors use data from the American Housing Survey and the National League of Cities to construct a two-stage hedonic model of residential tax incidence at the municipal level. An increase in the property tax by 0.34% from its sample mean of 1.4% increases average municipal rent by roughly $400.

## 3.3 Massachussetts Property Tax Overview

### 3.3.1 Tax Classification

Prior to 1981, the state of Massachusetts required one property tax rate for all property classes within a municipality. Under single rate taxation, assessing practices across municipalities varied greatly and public concern over assessment bias gave rise to two Supreme Judicial Court cases. In Bettigole v Assessors of Springfield (1967), the city was found in violation of the constitutional mandate of proportional and reasonable assessments as well as the statutory requirement for full and fair cash values. Springfield city assessors routinely assessed multi-family, commercial, industrial, and personal property at rates substantially higher than single-family (DOR, 2004). Town of Sudbury v the Commissioner of Corporations and Taxation (1974) arose out of concern over the variation in assessing properties at full and fair cash values across municipalities. Many municipalities across Massachusetts assessed below the statutory standard. This variation led to "discrimination against those cities and towns whose assessors act lawfully, in favor of those whose assessors engage in illegal practice of fractional valuation" (1974). With many towns in violation of the constitutional mandate, a state-wide implementation of full and fair cash value assessments would have triggered a large shift in tax burden to residential property owners.

Interestingly, however, a 1978 Amendment to the Massachusetts Constitution allowed for legal classification of real property into at least four categories with the ability to tax each class at a different rate. Classification was officially implemented in 1981 with 15 of 351 municipalities choosing to classify their tax code. For municipalities already practicing fractional valuation, the adoption of classification changed tax burdens across classes very little and had no effect on revenue. For other municipalities, adoption similarly did not raise or lower tax revenue but did redistribute the tax burden across property owners. The original legislation prohibited the tax burden on commercial, industrial, or personal property (CIP) to exceed 150% of the burden borne by those classes under a uniform tax rate. A lower bound for residential tax burden, known as the minimum residential factor (MRF), was set at 65%.

Rapid appreciation of residential property values as compared to CIP properties in the late 1980s substantially increased the residential tax burden in some communities. Without the ability to set a residential factor below the MRF of 65%, these communities faced annual increases in residential tax burden. In 1988, the legislature allowed for municipalities with increasing residential burdens to lower the MRF to 50% and increase the factor on CIP classes to 175%. Residential property value increases in the mid 2000s brought about an additional, but temporary, change to the legislation. The MRF dropped further to 45% with the factor on CIP properties increasing to 200% for qualifying municipalities. As of

January 2016, the MRF is 65% and the factor for CIP properties is 150%. If after following these guidelines the residential levy share is greater than last year, a municipality can adjust the factor for CIP properties to 175% and residential taxpayers must pay the greater of 50% of their full and fair cash value or the lowest percentage share of the levy they have paid since classification began (DOR, 2016).

Municipalities that choose to classify do not set differential tax rates for each class; rather, rates vary based on changes to components used to calculate the tax rate. The municipality may change the amount of tax levied by a particular class or they may apply an adjustment to the assessed valuation of the class. Under both circumstances, the total taxes levied remains the same as under a single rate system. City council or town selectmen set the percentages of tax levied (or the valuation adjustment) on each class by adopting a residential factor. A residential factor of 1 implies a single tax rate across all classes. A residential factor less than one shifts the share of taxes levied by the residential class towards all other classes. A factor greater than 1 may also be adopted and thus will shift the burden towards the residential class.

### 3.3.2 Residential versus Commercial Tax Rates and Levies 1981-2016

The mean difference in residential and commercial rates across Massachusetts has increased since classification was legalized in 1981 (see Figure 3.1). Unsurprisingly, many municipalities are choosing to implement classification in order to shift the tax burden away from the residential class. This is evidenced by the increasing difference between the rates over time. In 1981, 15 of the 351 Massachusetts towns had different residential and commercial rates. By 1988, 101 municipalities were choosing to tax commercial property at a different rate, and in 2016, 113 municipalities had differing rates. Of the municipalities that chose to classify over the 1981-2016 period, there were only between 1-4 municipalities each year that taxed commercial property at a rate lower than residential.

Focusing on the actual tax burden instead of the rate, the difference in tax burden between municipalities that classify and those who do not is quite stark (see Table 3.1). The

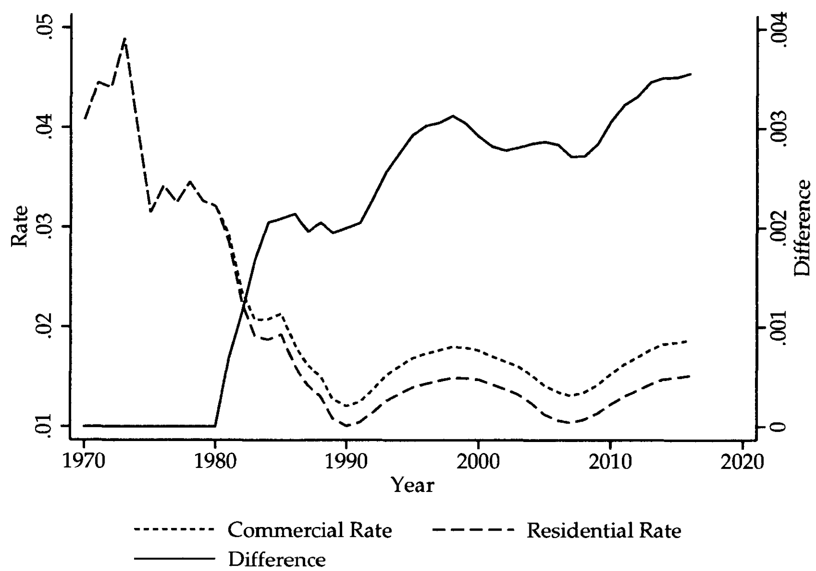



Figure 3.1: Average Residential and Commercial Rates

average share of the total levy borne by commercial properties in classified municipalities hovers around 0.17 over the years 1987 to 2015. This share is relatively stable over time with a minimum level of 0.16 and a maximum of 0.19. In single rate municipalities, this share averages 0.07 with a min and max of 0.05 and 0.09 respectively. In contrast, the residential share in classified municipalities is 0.66 with a min and max of 0.61 and 0.70. In single rate municipalities the share jumps to 0.88 with a min and max of 0.83 and 0.91.

## 3.4 Methodology

Consider the following simple model of property tax incidence:

$$R = R_0 + \beta \frac{l}{s} \tag{3.1}$$

where $R$ is average rent per square foot within a commercial building and $R_0$ is the base rent. The second term on the right-hand side represents the fraction of the property tax payment that is passed onto the tenant where $l/s$ is the building's levy per square foot and

Table 3.1: Average Levy Share of Residential, Commercial, and Industrial Property Across Municipalities

| | Classified | | | | Unclassified | | | |
|---|---|---|---|---|---|---|---|---|
| Year | Res. Share | Comm. Share | Ind. Share | Other Share | Res. Share | Comm. Share | Ind. Share | Other Share |
| 1988 | 0.633 | 0.181 | 0.123 | 0.063 | 0.835 | 0.084 | 0.036 | 0.045 |
| 1989 | 0.636 | 0.185 | 0.123 | 0.056 | 0.852 | 0.078 | 0.034 | 0.036 |
| 1990 | 0.641 | 0.186 | 0.120 | 0.053 | 0.860 | 0.076 | 0.032 | 0.032 |
| 1991 | 0.639 | 0.187 | 0.119 | 0.054 | 0.860 | 0.077 | 0.033 | 0.030 |
| 1992 | 0.639 | 0.185 | 0.117 | 0.060 | 0.860 | 0.077 | 0.033 | 0.030 |
| 1993 | 0.644 | 0.179 | 0.112 | 0.065 | 0.862 | 0.076 | 0.032 | 0.030 |
| 1994 | 0.647 | 0.178 | 0.110 | 0.065 | 0.863 | 0.073 | 0.032 | 0.032 |
| 1995 | 0.652 | 0.176 | 0.105 | 0.067 | 0.864 | 0.073 | 0.030 | 0.033 |
| 1996 | 0.656 | 0.175 | 0.110 | 0.059 | 0.867 | 0.072 | 0.029 | 0.033 |
| 1997 | 0.656 | 0.176 | 0.108 | 0.060 | 0.868 | 0.072 | 0.028 | 0.031 |
| 1998 | 0.651 | 0.179 | 0.109 | 0.061 | 0.869 | 0.072 | 0.028 | 0.031 |
| 1999 | 0.650 | 0.180 | 0.107 | 0.062 | 0.869 | 0.072 | 0.028 | 0.031 |
| 2000 | 0.646 | 0.181 | 0.114 | 0.059 | 0.869 | 0.071 | 0.028 | 0.032 |
| 2001 | 0.652 | 0.180 | 0.112 | 0.057 | 0.872 | 0.069 | 0.029 | 0.031 |
| 2002 | 0.654 | 0.180 | 0.110 | 0.057 | 0.876 | 0.067 | 0.029 | 0.028 |
| 2003 | 0.657 | 0.176 | 0.115 | 0.053 | 0.882 | 0.065 | 0.025 | 0.027 |
| 2004 | 0.671 | 0.167 | 0.109 | 0.053 | 0.889 | 0.061 | 0.024 | 0.027 |
| 2005 | 0.686 | 0.163 | 0.102 | 0.048 | 0.898 | 0.057 | 0.022 | 0.023 |
| 2006 | 0.699 | 0.158 | 0.095 | 0.048 | 0.902 | 0.056 | 0.021 | 0.022 |
| 2007 | 0.703 | 0.157 | 0.093 | 0.047 | 0.905 | 0.055 | 0.020 | 0.020 |
| 2008 | 0.702 | 0.159 | 0.092 | 0.048 | 0.902 | 0.057 | 0.021 | 0.020 |
| 2009 | 0.690 | 0.164 | 0.094 | 0.053 | 0.896 | 0.058 | 0.022 | 0.024 |
| 2010 | 0.681 | 0.167 | 0.096 | 0.056 | 0.893 | 0.059 | 0.022 | 0.025 |
| 2011 | 0.679 | 0.167 | 0.094 | 0.060 | 0.891 | 0.059 | 0.022 | 0.027 |
| 2012 | 0.682 | 0.166 | 0.093 | 0.059 | 0.891 | 0.058 | 0.022 | 0.029 |
| 2013 | 0.682 | 0.166 | 0.093 | 0.059 | 0.889 | 0.059 | 0.022 | 0.030 |
| 2014 | 0.681 | 0.167 | 0.093 | 0.059 | 0.888 | 0.059 | 0.022 | 0.031 |
| 2015 | 0.684 | 0.165 | 0.092 | 0.059 | 0.887 | 0.059 | 0.022 | 0.033 |

$\beta$ is the tax incidence parameter. A $\beta$ of 1 implies a complete transfer of the tax payment to the tenant. Unfortunately, I do not observe the actual tax payment (levy per square foot) made by a building owner, but (3.1) can be rewritten in terms of building value and commercial tax rate. Using the identity, $l = \tau \cdot v$ where $\tau$ is the commercial tax rate and $v$ is the assessed property valuation, (3.1) becomes:

$$R = R_0 + \beta\tau * \frac{v}{s} \tag{3.2}$$

I do not observe building level assessed value, but an estimate can be made using the present value discounted stream of rental payments:

$$v = \frac{s * R}{i} \tag{3.3}$$

where $i$ is the gross yield.

Substituting (3.3) into (3.2) highlights the endogeneity that exists between valuation and rent within the model for tax incidence, $R = R_0 + \beta\tau\frac{R}{i}$. Here rent is a function of value which is a function of rent in itself. However, collecting the rent terms on the left-hand side produces a nonlinear expression with respect to the tax rate:

$$R = \frac{R_0}{1 - \beta\frac{\tau}{i}} \tag{3.4}$$

The nonlinear reformulation in (3.4) internalizes the direct endogeneity between valuation and rent. However, endogeneity between rent and the commercial tax rate may still be a concern. In the presence of a positive rent shock that increases commercial valuations, a town may choose the lower the tax rate while still receiving previously targeted level of tax revenues. In this instance, causality is reversed with rent now influencing the tax rate. This endogeneity is addressed econometrically in the following section.

### 3.4.1 Econometric Specifications

#### Naïve Specification

Prior to creating a regression specification based on (3.4), I consider a naïve specification that does not control for either of the endogeneity issues associated with the tax payment. In this specification, log rent for building $j$ in municipality $m$ at time $t$ is regressed on the ratio of $\tau/i$, and year and building fixed effects, $\phi_j$ and $\gamma_t$ respectively.

$$R_{jmt} = \beta_0 + \beta_1 * x_{mt} + \phi_j + \gamma_t + \varepsilon_{jmt} \tag{3.5}$$

where $x = \tau/i$. In this setting the tax rate coefficient will provide an equilibrium effect of a change in the ratio of the tax rate to yield and will not reflect tax incidence independently. In an attempt to address the reverse causality associated with the tax rate, an IV specification is considered. Total levy per capita and the commercial property value to total property value ratio instrument $x$. These instruments control for changes in the tax rate associated with decisions to increase spending per capita and to increase or decrease the share of the tax burden on commercial property.

#### Constrained Least Squares

Equation (3.4) is estimable as a nonlinear model, and simplifications can be made to lessen the complexities that arise when controls are added. Taking logs linearizes the model:

$$logR = logR_0 - log(1 - \beta\frac{\tau}{i}) \tag{3.6}$$

Letting $x = \tau/i$, further simplifications are made by transforming the second term on the right-hand side of (3.6) into its Taylor series expansion about 0:

$$-log(1 - \beta x) \approx \beta x + \beta^2\frac{x^2}{2} + \beta^3\frac{x^3}{3} + ...$$

Given the above, (3.6) can be approximated as:

$$logR \approx logR_0 + \beta^2 x + \beta^3 \frac{x^2}{2} + \beta^4 \frac{x^3}{3} + ... \tag{3.7}$$

A constrained least squares specification is used to implement (3.7) as a regression specification in the following manner[2]:

$$\begin{aligned} logR_{jmt} &= \alpha + \beta x_{mt} + \beta^2 \frac{x_{mt}^2}{2} + \beta^3 \frac{x_{mt}^3}{3} \\ &+ \beta^4 \frac{x_{mt}^4}{4} + \beta^5 \frac{x_{mt}^5}{5} + \phi_j + \gamma_t + \varepsilon_{jmt} \end{aligned} \tag{3.8}$$

where $x_{mt} = \tau_{mt}/i_{mt}$, $\phi_j$ are building fixed effects, and $\gamma_t$ are year fixed effects.

There are over 1500 unique buildings in the data set; thus, to ease the computational burden I demean each variable using the within building average prior to running the regression rather than including dummy variables. Year fixed effects are also accounted for through a demeaning process. Standard errors are adjusted for the demeaning process.

The corresponding regression specification is as follows:

$$\begin{aligned} log\widetilde{R}_{jmt} &= \beta \widetilde{x}_{mt} + \beta^2 \frac{\widetilde{x}_{mt}^2}{2} + \beta^3 \frac{\widetilde{x}_{mt}^3}{3} \\ &+ \beta^4 \frac{\widetilde{x}_{mt}^4}{4} + \beta^5 \frac{\widetilde{x}_{mt}^5}{5} + \widetilde{\varepsilon}_{jmt} \end{aligned} \tag{3.9}$$

where $\widetilde{R}_{jmt}$ is the building and year demeaned rent. Likewise for $\widetilde{x}_{mt}$.

### Endogeneity Concerns

The first form of endogeneity associated with the tax payment is accounted for within the nonlinear framework. To address the second form of endogeneity, a similar approach to the one taken in the naïve regression is utilized. Within a control function setting, I consider a number of instruments that are considered to be possible sources of exogenous change

[2] The fifth order Taylor Series expansion is found to be an adequate approximation of (3.6).

in the tax rate. These instruments include state aid, employment, total levy, total levy per capita, estimated municipal expenditures based on regional trends, population, and commercial property value as a share of total property value. Table 3.2 summarizes the details and motivation behind using each instrument. Prediction for the expected direction of correlation with the commercial tax rate is also noted.

The first stage of the control function approach includes the exogenous instrument along with all regressors from the second stage.

$$x_{jmt} = \alpha + \theta \cdot X_{mt} + \phi_j + \gamma_t + v_{mt} \tag{3.10}$$

where $X_{mt}$ is an instrument, $\phi_j$ is the building fixed effect, and $\gamma_t$ is the year fixed effect.

A control function approach is favoured over the more widely used two-stage least squares method given that the endogenous variable enters into the second stage as a polynomial.[3] The second stage includes the residuals from equation (3.10):

$$logR_{jmt} = \alpha + \beta x_{mt} + \beta^2 \frac{x_{mt}^2}{2} + \beta^3 \frac{x_{mt}^3}{3} + \beta^4 \frac{x_{mt}^4}{4} + \beta^5 \frac{x_{mt}^5}{5} + \phi_j + \gamma_t + \rho \hat{v}_{mt} + e_{jmt} \tag{3.11}$$

The demeaned version of the regression in (3.11) is used to reduce computational burden of the building fixed effects.

Consistency of the control function estimator requires satisfying the two assumptions for valid instruments in the 2SLS setting along with two more assumptions. Valid instruments must be correlated with the endogenous variable and must be uncorrelated with the error term. Within the control function setting, sufficient but not necessary conditions for consistency are the independence of $v_{mt}$ and $\varepsilon_{mt}$ from the covariates in both equations (3.10) and (3.9). The second condition requires linearity of $E(\varepsilon|v)$. In particular, $E(\varepsilon|v) = \rho v$.

[3]Within a 2SLS model, each term of the polynomial would need to be separately instrumented.

Table 3.2: Control Function Instrument Details

| Instrument | Expected Sign | Details & Motivation |
|---|---|---|
| Total Levy | + | The total levy represents a municipality's required tax revenue. It can be argued that this is exogenously set by the municipality and thus not correlated with commercial rental market fluctuations. |
| Total Levy per Capita | + | Akin to spending per capita. All else constant, if spending per capita increases, tax rates must increase. Unlikely to be correlated with the commercial rental market. |
| Expenditures | +/- | Constructed by dividing the sum of municipal expenditures across all municipalities by the total population of all municipalities. This number is then multiplied by a given municipality's population and state aid is subtracted. Represents a regionally based estimate of local expenditures. Likely to be correlated with local tax rates but uncorrelated with the local commercial rental market. |
| Population | - | Larger populations may be able to exploit economies of scale in service provision allowing for lower property tax rates.[a] Unlikely to be correlated with the commercial rental market. |
| Employment | - | High levels of employment may indicate a strong commercial property tax base or may be an indicator of an attractive commercial tax rate level. Similar to population increase but may be somewhat correlated to fluctuations in the commercial rental market. |
| Commercial Value Share | +/- | As commercial value increases relative to total property value, municipalities may opt to tax it more heavily. However, if a municipality is already applying the maximum burden allowed, the commercial tax rate will decrease. Potential to be not fully exogenous with respect to commercial rental market if fluctuations in the share are dominated by changes in commercial value and not changes in residential value. |
| State Aid per Capita | + | As state aid increases, required tax revenue is likely to decrease. Unlikely to be correlated with the commercial rental market. |

[a] Bordignon et al. (2003) find strong evidence for this negative relationship

## 3.5 Data and Descriptives

### 3.5.1 Building Rent Data

Asking rent data is provided by Commercial Real Estate Services (CBRE) and represents 1539 individual office buildings. The rent value provided is the gross average asking rent per square foot per year based on all spaces for let in the building within the given quarter. Quarterly values are averaged to produce an average annual asking rent per square foot per year for a given building. Building level characteristics include building square footage, age of building, number of floors, and zip code. Summary statistics of these variables are presented in Table 3.3.

Table 3.3: Building Characteristic Summary Statistics

| Variable | Mean | Std. Dev. | Min. | Max. |
|---|---|---|---|---|
| Square Feet | 107886 | 150991 | 2800 | 1,755,398 |
| Number of Floors | 5.271 | 5.397 | 1 | 62 |
| Age | 40.15 | 38.42 | 0 | 231 |

126 municipalities are contained within this 27 year dataset (1988 - 2014). The eastern part of the state of Massachusetts is represented with good coverage for Boston and surrounding communities (see Figure 3.2). The vast majority of these communities lie within the 495 highway boundary and thus are within the Boston Metropolitan Statistical Area boundary. The dataset is reduced to 1512 individual buildings within 96 municipalities to ensure all buildings are observed at least twice and all municipalities are observed over at least two time periods.

Figure 3.3 presents the median and 25th and 75th percentiles of the rent variable over the years 1988 to 2014. The large boom and bust over the late 1990s to early 2000s coincides with the dot-com crash. Over the 1996 to 2001 time period, the median rent per square foot nearly doubled from \$18.62 to \$32.34. This correlation is unsurprising considering these data represent office buildings in and around the Greater Boston Area which had a large presence of dot-com related firms at the time. A smaller boom and bust cycle corresponds

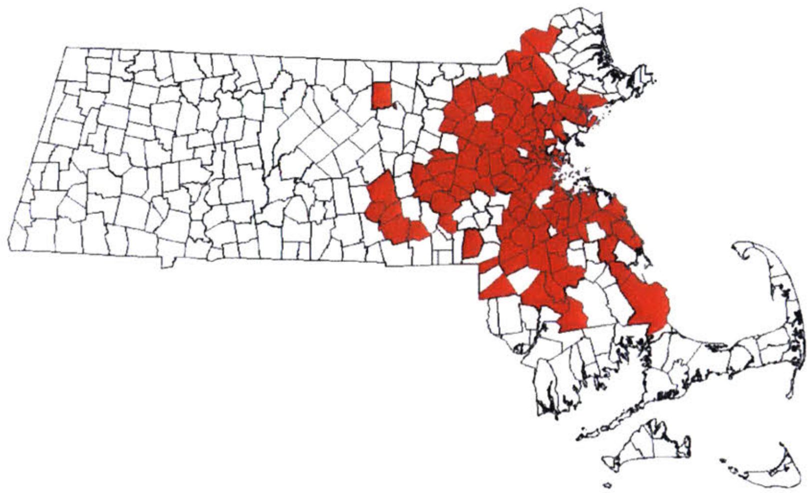


Figure 3.2: Rent Data Coverage

with the housing crisis of the late 2000s. The interquartile range indicates the presence of cross-sectional variation as well as a slight increase in variation over time. To further highlight the cross-sectional variation, Figure 3.4 presents the distribution of rents from 2014. The distribution is right-skewed with a large standard deviation of 12.56 and a mean of $27.32.

### 3.5.2 Tax Rate Variation

Section 3.3 highlights the average variation in the commercial rate overtime. These data also consist of a great deal of cross-sectional heterogeneity. From the descriptives presented in Table 3.4, we see that the mean and median are quite similar over time implying a lack of skewness. The interquartile range hovers around 0.01 for the years presented. In a given year, a typical municipality will adjust the rate by 0.001; thus, the interquartile range of 0.01 indicates an adequate amount of cross-sectional heterogeneity exists within these data. This variation is further illustrated in the histogram presented in Figure 3.5. The distribution for the commercial tax rate during 2014 highlights the non-normal and heterogeneous nature of the tax rate across the municipalities considered.


Figure 3.3: Median Rent per Sqft by Year

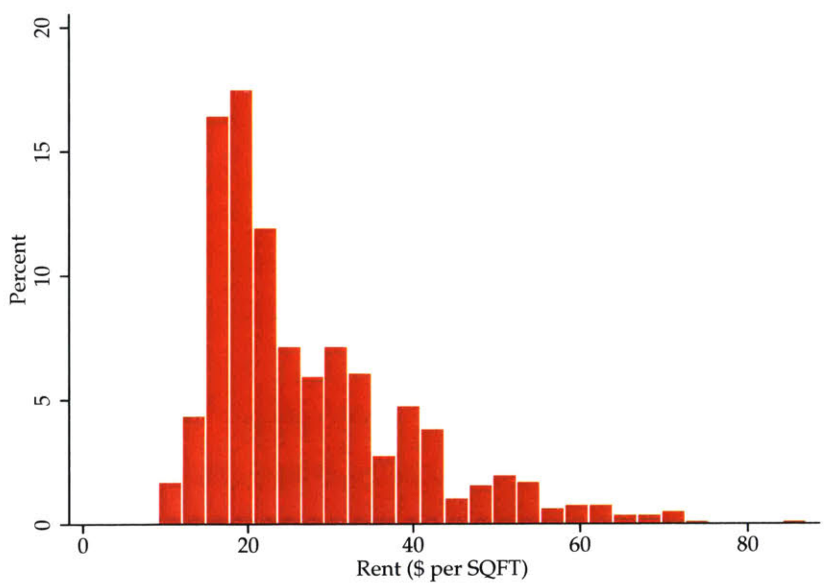



Figure 3.4: Rent Distribution 2014

### 3.5.3 Gross Yield

Gross yield is the ratio of effective gross income to asset value. Asset value for each property is not known, thus a property specific gross yield cannot be calculated. Furthermore, an

Table 3.4: Commercial Tax Rate Descriptives

| Year | $\mu$ | p25 | p50 | p75 | $\sigma$ | N |
|---|---|---|---|---|---|---|
| 1988 | 0.017 | 0.013 | 0.017 | 0.021 | 0.006 | 96 |
| 1989 | 0.014 | 0.011 | 0.014 | 0.017 | 0.005 | 96 |
| 1990 | 0.014 | 0.011 | 0.013 | 0.017 | 0.004 | 96 |
| 1991 | 0.015 | 0.011 | 0.014 | 0.019 | 0.004 | 96 |
| 1992 | 0.017 | 0.013 | 0.016 | 0.021 | 0.005 | 96 |
| 1993 | 0.019 | 0.014 | 0.018 | 0.023 | 0.006 | 96 |
| 1994 | 0.020 | 0.015 | 0.019 | 0.026 | 0.007 | 96 |
| 1995 | 0.022 | 0.016 | 0.021 | 0.027 | 0.007 | 96 |
| 1996 | 0.022 | 0.016 | 0.021 | 0.028 | 0.007 | 96 |
| 1997 | 0.022 | 0.016 | 0.021 | 0.028 | 0.007 | 96 |
| 1998 | 0.022 | 0.016 | 0.021 | 0.028 | 0.007 | 96 |
| 1999 | 0.022 | 0.016 | 0.022 | 0.027 | 0.007 | 96 |
| 2000 | 0.021 | 0.015 | 0.020 | 0.026 | 0.007 | 96 |
| 2001 | 0.020 | 0.015 | 0.020 | 0.024 | 0.006 | 96 |
| 2002 | 0.020 | 0.015 | 0.019 | 0.023 | 0.006 | 96 |
| 2003 | 0.018 | 0.013 | 0.018 | 0.022 | 0.006 | 96 |
| 2004 | 0.018 | 0.013 | 0.018 | 0.021 | 0.006 | 96 |
| 2005 | 0.017 | 0.011 | 0.017 | 0.021 | 0.006 | 96 |
| 2006 | 0.016 | 0.011 | 0.016 | 0.020 | 0.006 | 96 |
| 2007 | 0.016 | 0.011 | 0.017 | 0.020 | 0.005 | 96 |
| 2008 | 0.017 | 0.011 | 0.017 | 0.020 | 0.005 | 96 |
| 2009 | 0.018 | 0.012 | 0.018 | 0.022 | 0.006 | 96 |
| 2010 | 0.019 | 0.014 | 0.019 | 0.024 | 0.006 | 96 |
| 2011 | 0.021 | 0.015 | 0.020 | 0.025 | 0.007 | 96 |
| 2012 | 0.021 | 0.015 | 0.020 | 0.027 | 0.007 | 96 |
| 2013 | 0.022 | 0.016 | 0.021 | 0.027 | 0.007 | 96 |
| 2014 | 0.023 | 0.017 | 0.022 | 0.028 | 0.007 | 96 |

average gross yield series does not exist. Instead, I consider a cap rate series in its place where the cap rate is proportional to the gross yield given that gross yield is equivalent to (Effective Gross Income/Net Operating Income)*(Cap Rate).

A Boston specific cap rate series based on the Real Capital Analytics (RCA) commercial cap rate series was initially considered. This transaction based series includes office, retail, and industrial classes but is only available from 2002 onwards. A similar series for the same time period is also available for the national level. Alternatively, the National Council of Real Estate Investment Fiduciaries (NCREIF) provides appraisal based quarterly income returns that can be used to construct an annual cap rate series beginning in 1988. NCREIF's income return data is available for various real estate classes including office, retail, and

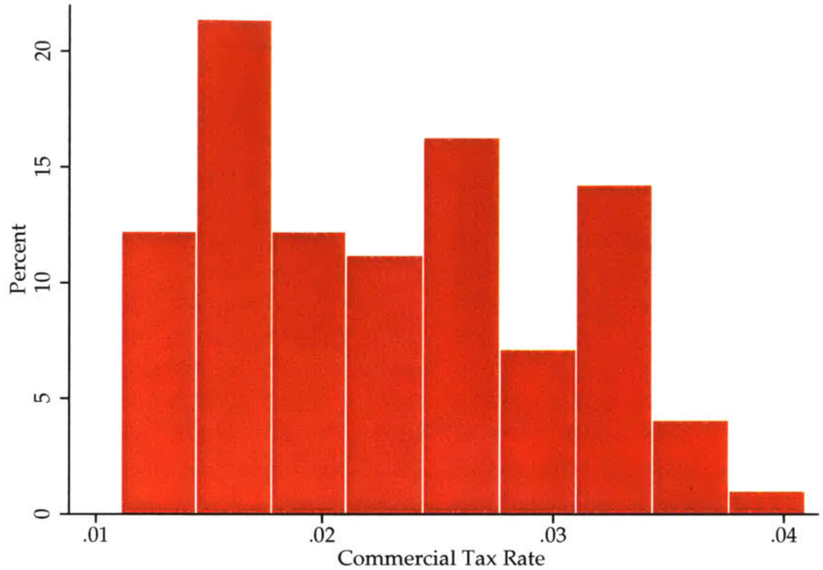



Figure 3.5: Commercial Tax Rate Distribution 2014

warehouse. I construct an annual series by multiplying the quarterly returns by four and averaging over the year. A further adjustment is made to account for the proportions of stock for each relevant class. I utilize Greater Boston Area total stock amounts of each class to construct weights. Specifically, I weight each return by their associated proportion as of Q2 2016. These proportions are 184.42 million, 242.56 million, and 218 million respectively.[4]

Figure 3.6 plots the Boston area and national RCA series against the weighted national NCREIF series. Stock followed by the NCREIF data tend to be large and institutional. In contrast, RCA data contains a wider range of building types; however, larger institutional buildings dominate the dataset in the early years of the series. This is evident in the closer tracking between the RCA and NCRIEF series for 2002 to 2005. Overall, the correlation is quite high over the 2002 to 2014 time period with the spread increasing over time as smaller buildings are captured within the RCA cap rate series.

[4]Greater Boston Area stock totals provided by CBRE.

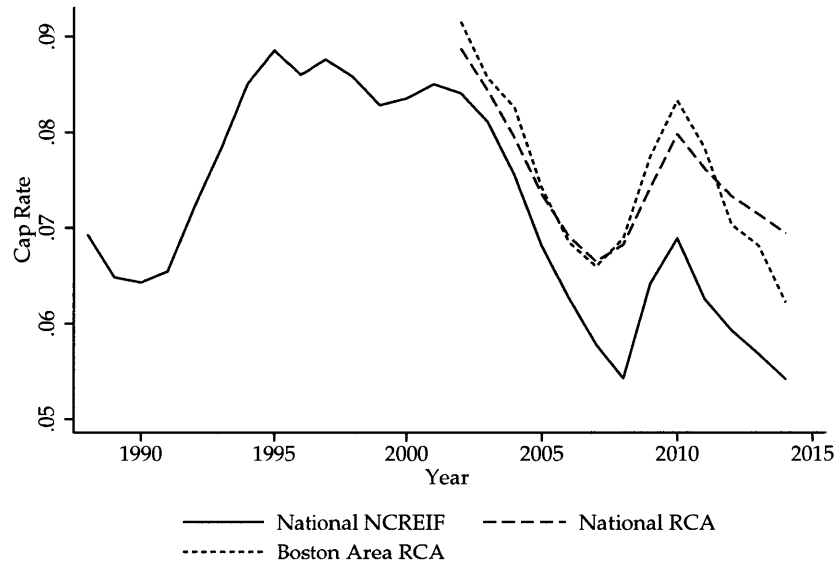



Figure 3.6: Cap Rate Comparison

## 3.6 Results

Beginning with the naïve specification from equation (3.5), the coefficient on $x$ is -0.800. Taking $x$ at the mean tax and cap rate values (0.0189 and 0.0729 respectively), a 100 bps increase in the tax rate holding the cap rate constant decreases rent by 11%. This effect represents the equilibrium outcome not accounting for the two sources of endogeneity. In an attempt to address the endogeneity associated with rent shocks influencing the tax rate, I instrument $x$ with both the levy per capita and the commercial property value to total property value ratio. The first stage of this IV regression is strong with a Sanderson-Windmeijer F statistic of 375. Regressing log rent on exogenous changes in the tax rate renders the coefficient on $x$ insignificant and positive. This result indicates a causality issue does exist between the tax rate and rent.

Turning to the nonlinear specification where both sources of endogeneity are accounted for, Table 3.6 presents the second stage results from the control function model. Beta represents the tax incidence parameter and Rho is the coefficient on the first stage residual term included in the second stage. The first column contains the results from the regres-

Table 3.5: Naïve OLS & IV Specification

| | (1) OLS | (2) IV |
|---|---|---|
| Tax Rate/Cap Rate | -0.800*** | 0.227 |
| | (0.034) | (0.167) |
| Constant | 3.196*** | 2.930*** |
| | (0.011) | (0.044) |
| Year FE | X | X |
| Building FE | X | X |
| Observations | 17,563 | 17,563 |
| R-squared | 0.830 | 0.175 |
| Root MSE | 0.150 | 0.154 |

*** p<0.01, ** p<0.05, * p<0.1

sion with no control function. The coefficient is negative, but small and not significantly different from zero as compared to the large negative and significant coefficient in the OLS setting. With the addition of the control function for the various instruments, the incidence parameter jumps to a range of 0.799 to 0.932. All coefficients are significant and there is a great deal of consistency across the instruments with the exception of log population. These results indicate the tax incidence parameter is positive and likely near 100%. A quick test of endogeneity of the tax rate can be performed by assessing the significance of Rho. All models have significant residual coefficients implying the tax rate is indeed endogenous. First stage results are presented in Table 3.7. Each instrument with the exception of log employment displays a strong F-statistic and all instruments are highly correlated with the tax rate. The direction of correlation with the instruments and the tax rate are in line with predictions.

There may be concerns that rents are spuriously trending with $x$ even after removing building and year fixed effects; however, fisher-type unit root tests demonstrate that demeaned rent and $x$ are stationary.

Table 3.6: Nonlinear Control Function

| | No Control | Log Levy | Levy/ Capita | Log Expend. |
|---|---|---|---|---|
| Beta | -0.149 | 0.927*** | 0.932*** | 0.899*** |
| | (0.129) | (0.165) | (0.156) | (0.152) |
| Rho | | -2.304*** | -2.299*** | -2.165*** |
| | | (0.658) | (0.625) | (0.559) |
| Observations | 17,563 | 17,563 | 17,563 | 17,563 |
| R-squared | 0.001 | 0.067 | 0.066 | 0.057 |
| Root MSE | 0.160 | 0.155 | 0.155 | 0.156 |

| | Log Pop. | Value Share | State Aid/ Capita | Levy/Capita & Value Share |
|---|---|---|---|---|
| Beta | 0.799*** | 0.913*** | 0.916*** | 0.903*** |
| | (0.163) | (0.127) | (0.175) | (0.124) |
| Rho | -1.930*** | -2.293*** | -2.231*** | -2.273*** |
| | (0.620) | (0.562) | (0.661) | (0.565) |
| Observations | 17,563 | 17,563 | 17,563 | 17,563 |
| R-squared | 0.048 | 0.070 | 0.061 | 0.067 |
| Root MSE | 0.157 | 0.155 | 0.155 | 0.155 |

Year and Building FE included in all specifications
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.7: Control Function First Stage

| | Log Levy | Levy/ Capita | Log Expend. | Log Pop. | Log Emp. | Value Share | Aid/ Capita | Levy/Capita & Value Share |
|---|---|---|---|---|---|---|---|---|
| Log Levy | 0.114*** | | | | | | | |
| | (0.006) | | | | | | | |
| Levy/Capita | | 2.75e-05*** | | | | | | 3.00e-05*** |
| | | (0.000) | | | | | | (0.000) |
| Log Exp. | | | -0.0208*** | | | | | |
| | | | (0.002) | | | | | |
| Log Pop. | | | | -0.307*** | | | | |
| | | | | (0.010) | | | | |
| Log Emp. | | | | | 0.00833 | | | |
| | | | | | (0.005) | | | |
| Value Share | | | | | | -0.342*** | | -0.356*** |
| | | | | | | (0.016) | | (0.016) |
| Aid/Capita | | | | | | | 5.18e-05*** | |
| | | | | | | | (0.000) | |
| Constant | -1.794*** | 0.236*** | 0.623*** | 3.748*** | 0.165*** | 0.348*** | 0.233*** | 0.328*** |
| | (0.112) | (0.002) | (0.037) | (0.110) | (0.058) | (0.004) | (0.002) | (0.005) |
| First Stage F stat | 344.47 | 199.66 | 99.40 | 1003.62 | 2.59 | 473.50 | 162.56 | 375.36[1] |
| Observations | 17,563 | 17,563 | 17,563 | 17,563 | 17,563 | 17,563 | 17,563 | 17,563 |
| R-squared | 0.910 | 0.910 | 0.909 | 0.914 | 0.908 | 0.911 | 0.909 | 0.912 |

[1]Sanderson-Windmeijer multivariate F test
Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

## 3.7 Conclusion

The tax burden on commercial property stands to increase as more municipalities adopt tax classification in an attempt to ameliorate residential tax burden. But what are the implications of such increases? In this paper I exploit municipal cross-sectional variation in tax rates as well as variation over time to investigate the degree to which commercial property tax payments are passed through to the renters of commercial office space within Massachusetts. Simultaneity between both rent and property value as well as rent and the tax rate are common issues within the commercial property tax incidence literature. I circumvent these issues through the use of a non-linear functional form that accounts for the endogeneity between rent and value. The functional form also makes use of instruments within a control function setting that address the endogeneity between rent and the tax rate within a control function setting.

A simple OLS and 2SLS setting illustrates the downward bias associated with not controlling for the possible reverse causality between rent and the tax rate. In this setting, a 100 bps increase in the tax rate decreases rent by 11%. Instrumenting for the tax rate erases this negative effect. Within the nonlinear control function specification, the tax incidence is strongly positive and highly consistent across various instruments. The majority of the coefficients are in the range of 0.90 to 0.93 indicating tax incidence may be nearly 100%. Furthermore, this result implies the commercial property rental market demand within 96 municipalities in the Greater Boston Area is highly inelastic as landlords are able to pass-through the majority of tax increases onto tenants.

Findings of a high tax incidence on renters of commercial space does not provide evidence in favor of the capital view nor the benefit view of property tax incidence as this finding is consistent with both. However, results strongly imply limited shifting of the incidence to landlords. Absent from the results presented here is an analysis of the secondary incidence on the prices of consumer goods and local wages. Future research on commercial property tax incidence will seek to provide a holistic view of the tax burden.

# Conclusion

The interaction between residents, property owners, and the built environment of a municipality generates numerous issues pertaining to the economic health, vibrancy, and stability of a community. Understanding the drivers of these issues is imperative in forming efficient local policy. This dissertation contributes to this understanding by utilizing novel specifications and instruments as well as high quality data to address important local issues of neighborhood decline, inefficient land use and municipal expenditures, and commercial property tax incidence.

Neighborhood decline is a complex dynamic influenced by many factors including neighborhood socioeconomic status, household maintenance and migration decisions, and housing age. Many of these factors are highly endogenous as they affect and are affected by one another. To address the question of whether or not characteristics of postwar homes and postwar developments are influencing household maintenance and migration decisions on aggregate and thus neighborhood decline, I account for possible endogenous redevelopment using two instruments. Furthermore, using a panel framework I am able to disentangle the affect of housing age and housing vintage. Previous empirical literature presented confounded evidence of the influence of postwar characteristics on neighborhood decline; whereas my findings indicate that it is housing age, not a postwar vintage that is associated with lower levels of neighborhood level household income.

Regarding future research on this topic, I plan to explore a number of empirical observations. Namely, neighborhoods that are overwhelmingly of a particular housing vintage tend to have higher levels of socioeconomic status. For example, census tracts where at least 65% of the housing is built between 1950 and 1969 have higher levels of income and homeownership rates than the city on average. This pattern is repeated for housing built between 1970 and 1989 as well as 1990 and 2009. I hypothesize these census tracts may have been part of large-scale master-planned developments.

Using historical residential development maps, I will explore the legacy of large-scale residential developments. In particular, how has the demand for this type of product evolved over time? What characteristics of large-scale developments might contribute to the consistent observations of income and homeownership levels across different vintages of developments? For the postwar vintage in particular, preliminary research shows that the postwar census tracts not only have higher incomes and higher rates of homeownership but residents are also whiter and tend to have larger populations over the age of 65 than the city on average. A closer look at these data reveal that more racially diverse cities are more likely to contain these predominately white, highly postwar census tracts. Is it possible one of the original sources of suburbanization in the postwar era, "White Flight", is contributing to these observations? In other words, is there a persistence to the segregation originally established in these large-scale developments that is not exhibited by smaller-scale developments?

The second chapter of this dissertation adds to the 'Cost of Sprawl' literature by employing more detailed measures of land use in order to better understand the complex relationship between municipal expenditures and development patterns. Using metrics that capture the separation, continuity, centrality, integration, and concentration of residential and commercial land uses I show that some development patterns championed by Smart Growth and New Urbanism advocates do produce lower levels of municipal expenditures as compared to more disconnected, sprawling patterns. I hope to expand this research to a national sample of municipalities in order to assess the robustness of these findings. This work has the potential to greatly influence how a municipality zones and develops land as well as the potential to inform future iterations of Smart Growth and New Urbanism doctrine.

The empirically difficult question of commercial property tax incidence is tackled in the final chapter. In this paper I exploit a novel instrumented structural specification to control for issues of endogeneity. Clear empirical evidence for whether or not owners of capital internalize increases in tax payments or pass these increases forward in the form of higher rents does not exist within the extant literature. Robust findings from this paper indicate that nearly 100% of a commercial property tax increase is passed through to the renter of the commercial space. With the burden clearly on commercial tenants, it is important to understand if and how this incidence gets further passed down to employees and consumers in the form of lower wages and higher consumer goods prices.

Together, this work succeeds in expanding the understanding of ubiquitous municipal issues. An emphasis is placed on developing empirically robust methodologies in order to account for serious issues of endogeneity and the use of measurements that adequately capture the phenomenon of interest.

## References


Ahlbrandt Jr, R. S., & Brophy, P. C. (1975). *Neighborhood revitalization: Theory and practice.* Lexington books.

Allen, E. (2001). INDEX: Software for community indicators. In R. K. Brail & R. E. Klosterman (Eds.), *Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools* (pp. 229–261). ESRI, Inc.

Altshuler, A. A. (1993). *Regulation for Revenue: The Political Economy of Land Use Exactions* (1St Edition edition ed.). Washington, D.C. : Cambridge, Mass: Brookings Institution Press and the Lincoln Institute of Land Policy.

Arnott, R. J., & Braid, R. M. (1997). A filtering model with steady-state housing. *Regional Science and Urban Economics*, *27*(4), 515–546.

Baer, W. C., & Williamson, C. B. (1988). The filtering of households and housing units. *Journal of Planning Literature*, *3*(2), 127–152.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

Bergstrom, T. C., & Goodman, R. P. (1973). Private demands for public goods. *The American Economic Review*, *63*(3), 280–296.

Bier, T. (2001). *Moving up and filtering down: Metropolitan housing dynamics and public policy* (Tech. Rep.). Washington, DC: Brookings Institute, Center on Urban and Metropolitan Policy.

Bier, T., & Howe, S. R. (1998). Dynamics of Suburbanization in OHIO Metropolitan Areas. *Urban Geography*, *19*(8), 695–713.

Blais, P. (2011). *Perverse Cities: Hidden Subsidies, Wonky Policy, and Urban Sprawl.* UBC Press.

Bond, S., Denny, K., Hall, J., & McCluskey, W. (1996). Effect of non-domestic rates on commercial property rents: an empirical study. *Journal of Property Valuation and Investment*, *14*(4), 18–30.

Borcherding, T. E., & Deacon, R. T. (1972). The demand for the services of non-federal governments. *The American economic review*, *62*(5), 891–901.

Bourne, L. S. (1981). *The geography of housing.* Halsted Press.

Bradbury, K., Ladd, H., Perrault, M., Reschovsky, A., & Yinger, J. (1984). State Aid to Offset Fiscal Disparities Across Communities. *National Tax Journal*, *37*(2), 151–170.

Braid, R. M. (1984). The effects of government housing policies in a vintage filtering model. *Journal of Urban Economics*, *16*(3), 272–296.

Browning, C. R., Byron, R. A., Calder, C. A., Krivo, L. J., Kwan, M.-P., Lee, J.-Y., & Peterson, R. D. (2010). Commercial Density, Residential Concentration, and Crime: Land Use Patterns and Violence in Neighborhood Context. *Journal of Research in Crime and Delinquency*, *47*(3), 329–357.

Brueckner, J. K., & Rosenthal, S. S. (2009). Gentrification and neighborhood housing cycles: will America's future downtowns be rich? *The Review of Economics and Statistics*, *91*(4), 725–743.

Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999). Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, *43*(1), 91–107.

Burchell, R. W., United States., Federal Transit Administration., National Research Council (U.S.)., Transportation Research Board., Transit Development Corporation., & Transit Cooperative Research Program. (1998). *The costs of sprawl–revisited.* Washington, D.C.: National Academy Press.

Carlino, G. A., & Saiz, A. (2008). *Beautiful city: Leisure amenities and urban growth* (Tech. Rep.). Philadelphia, PA: FRB of Philadelphia Working Paper.

Carruthers, J. I., & Ulfarsson, G. F. (2002). Fragmentation and Sprawl: Evidence from Interregional Analysis. *Growth & Change*, *33*(3), 312–340.

Carruthers, J. I., & Ulfarsson, G. F. (2003). Urban Sprawl and the Cost of Public Services. *Environment and Planning B: Planning and Design*, *30*(4), 503–522.

Carruthers, J. I., & Ulfarsson, G. F. (2008). Does 'Smart Growth' Matter to Public Finance? *Urban Studies*, *45*(9), 1791–1823.

Checkoway, B. (1980). Large builders, federal housing programmes, and postwar suburbanization. *International Journal of Urban and Regional Research*, *4*(1), 21–45.

Chinloy, P. (1980). The effect of maintenance expenditures on the measurement of depreciation in housing. *Journal of Urban Economics*, *8*(1), 86–107.

Choldin, H. M., Hanson, C., & Bohrer, R. (1980). Suburban Status Instability. *American Sociological Review*, *45*(6), 972.

Clifton, K., Ewing, R., Knaap, G., & Song, Y. (2008). Quantitative analysis of urban form: a multidisciplinary review. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *1*(1), 17–45.

Corporation, R. E. R., (U.S.), C. o. E. Q., Research, U. S. D. o. H. a. U. D. O. o. P. D., , & Management, U. S. E. P. A. O. o. P. a. (1974). *The costs of sprawl.* for sale by the Supt. of Docs., U.S. Govt. Print. Off.

Coulson, N. E., & McMillen, D. P. (2008). Estimating time, age and vintage effects in housing prices. *Journal of Housing Economics*, *17*(2), 138–151.

Cutsinger, J., & Galster, G. (2006). There is No Sprawl Syndrome: A New Typology of Metropolitan Land Use Patterns. *Urban Geography*, *27*(3), 228–252.

Cutsinger, J., Galster, G., Wolman, H., Hanson, R., & Towns, D. (2005). Verifying the Multi-Dimensional Nature of Metropolitan Land Use: Advancing the Understanding and Measurement of Sprawl. *Journal of Urban Affairs*, *27*(3), 235–259.

DOR. (2004). *Tax Classification Report* (Tech. Rep.). Boston, MA: Massachusetts Department of Revenue: Division of Local Services.

DOR. (2016). *Tax Classification Handbook* (Tech. Rep.). Boston, MA: Massachusetts Department of Revenue.

Ellen, I. G., & O'Regan, K. (2008). Reversal of fortunes? lower-income urban neighbourhoods in the us in the 1990s. *Urban Studies*, *45*(4), 845–869.

Ellen, I. G., & O'Regan, K. M. (2011). How low income neighborhoods change: Entry, exit, and enhancement. *Regional Science and Urban Economics*, *41*(2), 89–97.

Ewing, R., Pendall, R., & Chen, D. (2002). *Measuring sprawl and its impact: the character and consequences of metropolitan expansion* (Tech. Rep.). Washington, DC: Smart Growth America.

Fischel, W. A. (1974). Fiscal an environmental considerations in the location of firms in suburban communities. *Proceedings of the Annual Conference on Taxation Held under the Auspices of the National Tax Association-Tax Institute of America*, *67*, 632–656.

Fischel, W. A. (2009). *The Homevoter Hypothesis.* Harvard University Press.

Francke, M. K., & van de Minne, A. M. (2016). Land, Structure and Depreciation: Land, Structure and Depreciation. *Real Estate Economics.*

Frenkel, A., & Ashkenazi, M. (2008). Measuring Urban Sprawl: How Can We Deal with It? *Environment and Planning B: Planning and Design*, *35*(1), 56–79.

Fuest, C., Peichl, A., & Siegloch, S. (2015). *Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany* (SSRN Scholarly Paper No. ID 2739611). Rochester, NY: Social Science Research Network.

Galster, G., Hanson, R., Ratcliffe, M. R., Wolman, H., Coleman, S., & Freihage, J. (2001). Wrestling Sprawl to the Ground: Defining and measuring an elusive concept. *Housing Policy Debate*, *12*(4), 681–717.

Goodman, A., & Thibodeau, T. (1995). Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research*, *6*, 25–42.

Goodman, C. B. (2015). *The Fiscal Impacts of Urban Sprawl: Evidence from US County Areas.*

Grigsby, W. G. (1963). *Housing markets and public policy.* University of Pennsylvania Press.

Gyimah-Brempong, K. (1987). Economies of Scale in Municipal Police Departments: The Case of Florida. *The Review of Economics and Statistics*, *69*(2), 352-356.

Gyourko, J., & Saiz, A. (2004). Reinvestment in the housing stock: the role of construction costs and the supply side. *Journal of Urban Economics*, *55*(2), 238–256.

Hanlon, B. (2007). *The Decline of Older, Inner Suburbs: A New Reality in the U.S.*

Hanlon, B. (2008). The decline of older, inner suburbs in metropolitan America. *Housing Policy Debate*, *19*(3), 423–456.

Hanlon, B. (2012). *Once the American Dream: Inner-Ring Suburbs of the Metropolitan United States* (Reprint edition ed.). Philadelphia, Pa.: Temple University Press.

Harberger, A. C. (1962). The incidence of the corporation income tax. *The Journal of Political Economy*, 215–240.

Harding, J. P., Rosenthal, S. S., & Sirmans, C. (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. *Journal of urban Economics*, *61*(2), 193–217.

Hayden, D. (2006). Building the american way: public subsidy, private space. *The politics of public space*, *38*.

Hortas-Rico, M. (2014). Urban sprawl and municipal budgets in Spain: A dynamic panel data analysis. *Papers in Regional Science*, *93*(4), 843–864.

Hortas-Rico, M., & Sole-Olle, A. (2010). Does Urban Sprawl Increase the Costs of Providing Local Public Services? Evidence from Spanish Municipalities. *Urban Studies*, *47*(7).

Jackson, K. T. (1985). *Crabgrass Frontier: The Suburbanization of the United States.* Oxford University Press, USA.

Jaret, C., Ghadge, R., Reid, L. W., & Adelman, R. M. (2009). The Measurement of Suburban Sprawl: An Evaluation. *City & Community*, *8*(1), 65–84.

Kang, S. H., Skidmore, M., & Reese, L. (2015). The Effects of Changes in Property Tax Rates and School Spending on Residential and Business Property Value Growth. *Real Estate Economics*, *43*(2).

Knight, J. R., & Sirmans, C. (1996). Depreciation, maintenance, and housing prices. *Journal of Housing Economics*, *5*(4), 369–389.

Ladd, H. (1992). Population Growth, Density and the Costs of Providing Public Services. *Urban Studies*, *29*(2), 273–295.

Ladd, H. (1994). Fiscal impacts of local population growth: A conceptual and empirical analysis. *Regional Science and Urban Economics*, *24*(6), 661–686.

Ladd, H., & Yinger, J. (1989). *America's Ailing Cities: Fiscal Health and the Design of Urban Policy.* Baltimore: Johns Hopkins University Press.

Lee, N. J., & Wheaton, W. (2010). *Property Taxes Under" Classificiation:" Why Do Firms pay More?* (Tech. Rep.). Singapore: Institute of Real Estate Studies.

Lee, S., & Leigh, N. G. (2005). The role of inner ring suburbs in metropolitan smart growth strategies. *Journal of Planning Literature*, *19*(3), 330–346.

Lee, S., & Leigh, N. G. (2007). Intrametropolitan spatial differentiation and decline of inner-ring suburbs a comparison of four us metropolitan areas. *Journal of planning education and research*, *27*(2), 146–164.

LeRoy, S. F., & Sonstelie, J. (1983). Paradise lost and regained: Transportation innovation, income, and residential location. *Journal of Urban Economics*, *13*(1), 67–89.

Lucy, W. H., & Phillips, D. L. (1997). The post-suburban era comes to richmond: city decline, suburban transition, and exurban growth. *Landscape and urban planning*, *36*(4), 259–275.

Lucy, W. H., & Phillips, D. L. (2000). *Confronting suburban decline: strategic planning for metropolitan renewal.* Washington, D.C. ; Covelo, Calif: Island Press.

Lucy, W. H., & Phillips, D. L. (2006). *Tomorrow's cities, tomorrow's suburbs.* American Planning Association Chicago.

Lusht, K. M. (1997). *Real Estate Valuation: Principles and Applications.* Irwin.

Malpezzi, S., Ozanne, L., & Thibodeau, T. G. (1987). Microeconomic Estimates of Housing Depreciation. *Land Economics*, *63*(4), 372–385.

Marvell, T. B., & Moody, C. E. (1996). Specification problems, police levels, and crime rates. *Criminology*, *34*(4), 609–646.

*Massachusetts Department of Fire Services.* (2011). Retrieved from `http://www.mass.gov/eopss/agencies/dfs/index.html`

MassDot. (2017). *Chapter 90 Program - MassDOT Highway.* Retrieved from `http://www.massdot.state.ma.us/highway/DoingBusinessWithUs/LocalAidPrograms/Chapter90Program.aspx`

Massey, D. S., & Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, *67*(2), 281–315.

McDonald, J. F. (1993). Incidence of the property tax on commercial real estate: the case of downtown Chicago. *National Tax Journal*, 109–120.

Mieszkowski, P. (1972). The property tax: An excise tax or a profits tax? *Journal of Public Economics*, *1*(1), 73–96.

Muth, R. F. (1973). A vintage model of the housing stock. *Papers in Regional Science*, *30*(1), 141–156.

*National Elevation Dataset (NED) | The Long Term Archive.* (2016). Retrieved from `https://lta.cr.usgs.gov/NED`

Nechyba, T. J. (1997). Local property and sales income taxes: The role of interjurisdictional competition and collusion. *Journal of Political Economy*, *105*(2).

Nechyba, T. J., & Walsh, R. P. (2004). Urban sprawl. *The Journal of Economic Perspectives*, *18*(4), 177–200.

*Office of Geographic Information (MassGIS).* (2011). Retrieved from `http://www.mass.gov/anf/research-and-tech/it-serv-and-support/application-serv/office-of-geographic-information-massgis/index.html`

Ohls, J. C. (1975). Public policy toward low income housing and filtering in housing markets. *Journal of Urban Economics*, *2*(2), 144–171.

O'Neill, R. V., Riitters, K., Wickham, J., & Jones, K. B. (1999). Landscape Pattern Metrics and Regional Assessment. *Ecosystem Health*, *5*(4), 225–233.

Orfield, M. (2011). *American metropolitics: The new suburban reality.* Brookings Institution Press.

Pickett, S. T. A., & Cadenasso, M. L. (1995). Landscape Ecology: Spatial Heterogeneity in Ecological Systems. *Science*, *269*(5222), 331–334.

Randolph, W. C. (1988). Housing Depreciation and Aging Bias in the Consumer Price Index. *Journal of Business & Economic Statistics*, *6*(3), 359–371.

Rohe, W. M., & Stewart, L. S. (1996). Homeownership and neighborhood stability. *Housing Policy Debate*, *7*(1), 37–81.

Romem, I. (2016). *Has The Expansion of American Cities Slowed Down?* Retrieved from `https://www.buildzoom.com/blog/cities-expansion-slowing`

Rosenthal, S. S. (2008). Old homes, externalities, and poor neighborhoods. A model of urban decline and renewal. *Journal of Urban Economics*, *63*(3), 816–840.

Rosenthal, S. S., & Ross, S. L. (2014). Change and Persistence in the Economic Status of Neighborhoods and Cities. *Forthcoming in The Handbook of Regional and Urban Economics*, *5*.

Sarzynski, A., Galster, G., & Stack, L. (2014a). Evolving United States metropolitan land use patterns. *Urban Geography*, *35*(1), 25–47.

Sarzynski, A., Galster, G., & Stack, L. (2014b). Typologies of sprawl: investigating United States metropolitan land use patterns. *Urban Geography*, *35*(1), 48–70.

Schelling, T. C. (1971). Dynamic models of segregation†. *The Journal of Mathematical Sociology*, *1*(2), 143–186.

Shilling, J. D., Sirmans, C. F., & Dombrow, J. F. (1991). Measuring depreciation in single-family rental and owner-occupied housing. *Journal of Housing Economics*, *1*(4), 368–383.

Short, J. R., Hanlon, B., & Vicino, T. J. (2007). The Decline of Inner Suburbs: The New Suburban Gothic in the United States. *Geography Compass*, *1*(3), 641–656.

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of real estate literature*, *13*(1), 1–44.

Sohn, D.-W. (2016). Residential crimes and neighbourhood built environment: Assessing the effectiveness of crime prevention through environmental design (CPTED). *Cities*, *52*, 86–93.

Speir, C., & Stephenson, K. (2002). Does Sprawl Cost Us All?: *Isolating the Effects of Housing Patterns on Public Water and Sewer Costs. Journal of the American Planning Association*, *68*(1), 56–70.

Spence, D., Webster, B., & Conners, E. (2016). *Guidelines for starting and operating a new police department* (Tech. Rep.). Retrieved from `https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=236781`

Sweeney, J. L. (1974). A commodity hierarchy model of the rental housing market. *Journal of Urban Economics*, *1*(3), 288–323.

Sørensen, A., Taeuber, K. E., & Jr, L. J. H. (1975). Indexes of Racial Residential Segregation for 109 Cities in the United States, 1940 to 1970. *Sociological Focus*, *8*(2), 125–142.

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, *64*(5).

Torrens, P. M., & Alberti, M. (2000). *Measuring sprawl.* Retrieved 2016-04-12, from `http://discovery.ucl.ac.uk/1370/1/paper27.pdf`

Tsoodle, L. J., & Turner, T. M. (2008). Property taxes and residential rents. *Real Estate Economics*, *36*(1).

Varady, D. P. (1986). *Neighborhood upgrading: A realistic assessment.* SUNY Press.

Walzer, N. (1972). Economies of Scale and Municipal Police Services: The Illinois Experience. *The Review of Economics and Statistics*, *54*(4), 431.

Weicher, J. C., & Thibodeau, T. G. (1988). Filtering and housing markets: An empirical analysis. *Journal of Urban Economics*, *23*(1), 21–40.

Weiss, E. (1994). Depreciation of owner-occupied housing. *Federal National Mortgage Association.*

Wheaton, W. C. (1984). The Incidence of Inter-Jurisdictional Differences in Commercial Property Taxes. *National Tax Journal*, *37*(4).

White, M. J. (1986). Segregation and Diversity Measures in Population Distribution. *Population Index*, *52*(2), 198.

Wiens, J. A., Crawford, C. S., & Gosz, J. R. (1985). Boundary Dynamics: A conceptual framework for studying landscape ecosystems. *Oikos*, *45*(3), 421–427.

Wilhelmsson, M. (2008). House price depreciation rates and level of maintenance. *Journal of Housing Economics*, *17*(1), 88–101.

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.