Summer 8-2-2012

Understanding Land Use Grain: An Evaluation of Meaning and Measurement

Benjamin N. Williams
University of New Orleans, link0612@gmail.com

Understanding Land Use Grain: An Evaluation of Meaning and Measurement

A Thesis

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Master of Urban and Regional Planning

By

Benjamin N. Williams

B.S., Cornell University, 2010

August 2012

Acknowledgements

I would like to thank Michelle Thompson for her guidance and encouragement with this thesis and with my general studies, Karen Parsons for her valuable input on my thesis and for providing me with practical experience during my career at UNO, and John Renne for his input on my thesis. I would also like to thank my partner Joe Muller for his infinite support and feedback.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FUNCTIONS

LIST OF TABLES

ABSTRACT

Land use grain is a commonly-used measure of the mixture of land uses in the urban environment in transportation planning and public health, but there is no standard measurement practice in place. This thesis examines the meaning and common measurements of land use grain in these subfields. The entropy-based equation, the jobs-to-housing ratio, and the Herfindahl-Hirschman Index (HHI) are among the most common measures of land use grain, but results from these metrics differ depending upon how researchers choose a sample area and upon how land use categories are defined. All three metrics are performed, in a single context with varying assumptions, using the neighborhoods of Roxbury and Dorchester in Boston, MA. The entropy-based equation was deemed the most appropriate measure in a general context, with the HHI and the jobs-to-housing ratio potentially appropriate in specific contexts.

Keywords: active transportation; Boston; buffer; entropy; Herfindahl-Hirschman Index; land use; land use grain; land use mix; non-motorized transportation; public health; transportation demand

CHAPTER I: INTRODUCTION

A growing need for a variety of viable transportation modes for cities has been recognized in all facets of urban planning work (transportation planning, land use planning, environmental planning, equity-focused planning, and so on). This need is also recognized in the fields of transportation engineering and public health. A greater awareness of this need is filtering into the general political discourse in the United States, shaping new public policy, government spending habits, and planning goals.

Existing academic literature summarizes the impacts of particular modes of transportation on the urban form and demographics[1] in particular cities. This complex of the urban form and demographics will be referred to as the *urban fabric* for the rest of this work. There is a need for planners and policy makers to emphasize accessible, sustainable, and human-scaled transportation modes such as walking, bicycling, and public transit, referred to as *non-motorized transportation*. The goal is to effect healthier, more economically-successful, and ultimately more sustainable urban fabrics in cities. In sharp contrast to this recognized need, U.S. cities as a whole have been identified as relying on a single mode of transportation: the private automobile.[2]

The extant literature has begun to focus on the impacts of the urban fabric on the modal share[3] of non-motorized transportation. One intended goal of this body of research is to discern which urban fabrics are most conducive to non-motorized transportation, with the hope of then encouraging shifts in public policy towards an urban fabric which will encourage non-motorized transportation. This research faces the difficulty of deciding how best to measure the urban fabric, particularly the urban form. Broad characteristics of urban form (such as density, design, and diversity[4]) are difficult

[1] The impacts of particular modes of transportation on public policy and *vice versa* are tacitly being explored in academic planning literature as well. For examples, see e.g. Ruth Steiner's work or recent work from the Lincoln Institute of Land Policy (Steiner, Bejleri, Wheelock, Cahill, & Perez, 2008) (Ingram, Carbonell, Hong, & Flint, 2009).

[2] For a comprehensive overview of the justification for more sustainable transportation options, particularly in U.S. cities, see e.g. Schiller et al.'s Introduction to Sustainable Transportation (Schiller, Bruun, & Kenworthy, 2010).

[3] Modal share is the proportion of the trips taken by a particular mode of transportation to the total number of trips taken in a particular geography. This is a standard measure in transportation work.

[4] Robert Cervero defines these three dimensions of urban fabric in his work "Travel Demand and the 3 D's," which will be investigated more closely in Chapters II and III (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997).

to quantify and objectively compare. This work will focus on the measurement techniques for the consistently-considered but not-consistently-measured *land use grain*—a component of diversity, and a dimension of the urban form loosely defined as the qualities of the distribution of land use in a given urban environment. Land use grain is often referred to as *land use mix* in contemporary literature. This thesis uses the term *grain* instead in order to re-associate the concept with other aspects of the grain of the urban environment. Detailed definitions of the term, and its application and use, will be described in Chapter II.

Research Questions

The research questions addressed in this work are as follows:

Research Question 1

What is meant by the term "land use grain" in academic literature, and how does this meaning inform measurement of land use grain? (Chapter II)

Research Question 2

What standard measurement techniques quantify land use grain in the fields of transportation engineering, transportation planning, and public health? (Chapter III)

Research Question 3

How do standard measurement techniques and their results compare with one another when performed in context? (Chapter IV)

Statement of Purpose

The primary purpose of this work is to provide an evaluation of how to use existing measurement techniques of land use grain. Demonstration of the techniques will better inform researchers on which technique should be considered when conducting research on land use grain. This work summarizes existing techniques and prior applications of these techniques by previous researchers. The information gained from this analysis can be used to tailor specific research projects by providing a discussion of how various measurement techniques may be used, which inputs may be necessary for these measurement techniques, and which techniques may be appropriate in particular contexts.

In order to establish a similar definition/framework for analysis, the first part of this report focuses on defining land use grain (RQ1). A summary of commonly-used land use grain metrics will be described and analyzed, with a focus on identifying the key components of each metric. Potential measurement contingencies which may affect the results of measurements will also be discussed (RQ2). This research will use these metrics to perform or demonstrate each technique in a single urban context. The report will conclude with summarizing, how the techniques and results differ, as well as how they are impacted by potential contingencies (RQ3).

Rationale

Land use grain is a common measure of urban form because it has been repeatedly correlated with walking, bicycling, and transit usage in both North America and Europe.[5] However, measuring land use grain is a relatively new task, and researchers use several methods to make these measurements. The lack of clear standards in defining, measuring, and incorporating land use grain into studies of transportation, land use, and public health inhibits the transferability of study methodology and study results. Land use grain is often incorporated into comprehensive indices and more expansive models used to evaluate urban form's impact on particular behaviors or outcomes. As a result measurement techniques are at times obscured (and at best are footnoted), concealing potentially important differences in measurement practices. Therefore transferring these indices and models to other studies must be approached gingerly to ensure that comparisons are not made between incomparable land use grain measurements.

The lack of standard measurement practices can also create a steep learning curve for new researchers aiming to perform studies on the relationships between urban form and other aspects of the urban fabric. The plurality of land use grain literature has been written by a limited number of researchers. These researchers at times re-apply their own methods to subsequent studies without adequate attention to other measurement options used in other disciplines. For a new researcher entering this field, this creates a broad range of measurement standards, many of which have never been

[5] For a planning perspective, see e.g. (Frank, Schmid, Sallis, Chapman, & Saelens, 2005). For a public health perspective, see e.g. (Berrigan, Pickle, & Dill, 2010).

compared with one another. Deciding which measurement standard to apply in which contexts requires a good deal of effort and time on the researcher's part, due to the lack of adequate comparison among competing measurement options, eating into time which could be spent analyzing results.

Land use grain—as opposed to other common measures of the urban form such as street connectivity and density—requires study because it does not have one standard measurement technique. Measuring street connectivity is an established practice in public health, transportation engineering, and transportation planning, and measuring density is a standard practice in a variety of academic fields. While these measurements of the urban form are still much debated, the potential models are more mature and more standardized than any particular model of land use grain. This implies that land use grain requires some independent attention outside of these larger studies of the impacts of urban form on non-motorized transportation.

Transportation planning and engineering at the municipal, regional, and state levels generally rely upon transportation models which focus on travel demand. These models do not readily incorporate land use, bicycle and pedestrian modes, or trip degeneration as goals. This disconnect hinders the development of non-motorized transportation modes as a viable means of transportation and creates confusion and lack of communication between professionals working to mitigate traffic and professionals working to encourage alternative transportation modes.

Justification of Importance of the Study

This study will address the transferability issue, the multitude of measurement options that arise with the use of various land use grain measurement methods in current research, and the disconnect between non-motorized transportation advocates and transportation engineering. Existing metrics will be evaluated based upon the stability given variable inputs, intuitiveness, and conceptual vision of land use grain. As land use grain is a commonly-used indicator in indices of the built environment, gaining a better understanding of the concept and appropriate methods of measuring it will aid use of these indices individually and their ability to be compared to one another. Providing a consistent manner of measuring land use grain will also aid new researchers in choosing a single, consistent approach in measuring land use grain for their own work. Discussing an important factor contributing to non-motorized transportation demand in language accessible and meaningful to the transportation engineering community will help bridge gaps between these two subfields of transportation planning, and between the disciplines of transportation planning and public health.

Scope

This study is comprised of four parts. This introductory chapter outlines the scope of the study and the rationale behind studying land use grain.

Chapter II will evaluate common definitions of land use grain, using the most comprehensive and the most influential definitions in the literature to date. Understanding the term conceptually will better inform measurement approaches. This chapter will examine the relationship between land use grain and other important aspects of the urban fabric, as identified in contemporary research, in order to show why this concept is important to a variety of academic disciplines.

Chapter III will discuss measurement and analysis practices from the literature, cognizant of the concept of land use grain outlined in Chapter II. Common metrics of land use grain will be evaluated to ascertain different methods of defining areas of study/capture and different methods of defining and classifying land use by categories (the primary contingencies upon which land use grain measurements rely). This chapter will also evaluate minor issues that arise when using measures of land use grain in larger studies.

Chapter IV will use data from the Roxbury and Dorchester neighborhoods in Boston, MA to apply the techniques described in Chapter III. This demonstration will demonstrate strongest measurement practices in context, allowing for a visual representation of the results of a land use grain analysis, and allowing for a side-by-side comparison of measurement results given varying assumptions.

Chapter V will reassess the concept and measurement of land use grain given the information from the above chapters, and will use this to answer the three research questions identified above.

Limitations

As this thesis does not treat the use of the resulting land use grain measurement options in studies which use land use grain as a predictor of some other behavior or outcome, this study cannot stand alone as a user guide for land use grain measures. Researchers will have to use their best judgment in constructing models which include a land use grain value, based on their individual goals. As a corollary, this work will not provide judgment on the quality of the relationship between land use grain and participation in non-motorized transportation, as previous studies have consistently found a measure of significance in their studies. The relative importance or unimportance of land use grain in the transportation network is for other research to discern.

This thesis will not look in-depth at the rich background of land use grain (and other forms of grain in the urban fabric) as a metaphysical concept in urban design, architecture, or political philosophy. Certain works about the importance of land use grain will be included to inform a conceptual outline of the term, but the full history as a concept, its identification as an element of urban form, its relation to the human perception of urban form, its role in local political engagement, and its importance to the *goodness* of cities in general will not be analyzed in this work. As a result, the full importance of land use grain will not be treated here. Rather, this study assumes that there is some established importance to land use grain, based on previous work and interest in the planning field.

This thesis does not aim to identify the sole legitimate metric of land use grain. As this concept is used in a variety of fields for a variety of purposes, different measurements may in the end be required. However, this thesis will make apparent differences in the primary land use grain metrics, particularly when considering the varying inputs upon which final measurements are contingent.

Chapter II: Why Land Use Grain?

Land use grain is a powerful concept in urban planning and urban design. It has been established as a factor in the inherent *goodness* of a place, and has been identified strongly with the value of cities in general. Land use grain has also been consistently correlated with transportation choices, mobility, and accessibility. This chapter will evaluate the origin and existing definitions of the term using literature from planning and urban design. In order to claim that land use grain is important to urban form and to transportation systems, a precise definition with a historical reference is provided. A discussion on the intersections between land use grain and academic work surrounding urban transportation will follow. Finally, the intersections that land use grain has with other urban-related issues, including legibility, zoning law, and property values will be described. This work will not be exhaustive, but it will firmly establish the importance of land use grain as a measure of urban form.

Land Use Grain as a Concept

Due to the intangible nature of land use grain, an exact measure is difficult to undertake. The following holistic concepts of the grain of the urban fabric offer a clear vision of what grain is to mean at its most complete, and what land use grain means as a portion of the grain of urban fabric.[6] In an effort to better understand the concept as a whole, below are presented two of the foundational concepts of land use grain.

Kevin Lynch

Kevin Lynch wrote extensively on land use grain in his book *A Theory of Good City Form*, which is a discussion of which dimensions or virtues comprise a good city. In this book, Lynch identified five dimensions (and two sub-dimensions) of a good city: <u>vitality, sense, fit, access, and control</u> (with efficiency and justice as sub-dimensions which may apply to each dimension). Each of these virtues exists on a scale with extremes at each end, and the “good” somewhere in the middle. As cities are inherently political, the exact location of the “good” on each scale is to be agreed upon by the citizens themselves through the planning and political process.

[6] “Grain” as identified below encompasses a variety of aspects of the urban fabric, including the distribution of activity, land uses, demographics, commercial enterprise, building styles, and so on through physical space and through time.

Lynch applied this framework to different attributes of cities, one of which was the urban form. In discussion of urban form, Lynch outlines three crucial attributes of the city through which it may be judged: density, grain, and transportation. His loose definition of grain: "the way in which the various different elements of a settlement are mixed together in space." When applied to land uses, grain would then by the way in which different land uses in the city are mixed together.

Lynch's concept of grain is comprised of two fundamental characteristics: *fineness* and *sharpness*. By fineness, Lynch means how finely mixed different aspects of the urban environment are in a single context, with a fine land use grain showing "like elements … widely dispersed among unlike elements" and a blurred land use grain showing "extensive areas of one [element] separated from extensive areas of another." By sharpness, Lynch means how abrupt the transitions between different elements of the urban environment are from one another, with a sharp land use grain showing abrupt shifts between elements and a blurred land use grain showing more gradual transitions.[7]

Lynch writes during a period of coarsening land use grain in the U.S. (or activity grain, as he writes), and identifies benefits and drawbacks from this process. Lynch is not as vocal as Jane Jacobs on the merits of a fine grain of activity. The coarse grain of activity as a feature of heavily planned societies is identified, while stating that finely grained settlements are "more closely fitted to the varying activities of occupants." This sentiment is in line with contemporary planning efforts to return to human-scaled urban forms.

Lynch uses this concept in examples of demography, land use, time[8], and control. The wide variety of applications of the grain of the urban fabric is important to Lynch's identification of grain as a separate crucial element distinguishing cities from one

---

[7] The fineness of land use grain is generally what is being analyzed in quantitative studies of land use grain, as will be shown in Chapter III. The sharpness of land use grain is an important aspect of the transect theory of urban design, however, and plays a large role in the rising encouragement of form-based zoning codes over standard Euclidian zoning. (Duany A. , 2002) (Lynch, A Theory of Good City Form, 1981)

[8] Time is one feature of land use grain which has been touched upon in the literature on cities and urban form, but has not been researched closely.

another, and from the countryside. As an attribute of all dimensions of the urban form, grain is inherently important to any consideration of any dimension of urban form.[9]

Jane Jacobs

Jane Jacobs also discusses the concept of land use grain in her influential book *The Death and Life of Great American Cities*.[10] Jacobs, in contrast with Lynch, is not engaged in abstract academic theory of the city. Jacobs is engaged in a political project aiming to achieve particular political goals in a particular context (New York City). However, in pursuit of this goal, Jacobs makes an effort to define the loose concept of land use grain as a crucial element of a functional, living city.

For Jacobs, land use grain is one of four[11] generators of diversity: the basic positive component of cities. Diversity consists of a mixture of uses, streets, public spaces, and buildings at a high level of density. This mixture is an inherent part of any city, and it is to be encouraged in order to create a good city. In this scheme, land use grain runs on a spectrum from bad to good, with a good land use grain being one in which there is a very fine and blurred grain of different land uses in a single area.

Land uses in *Death and Life* differ from the generally-used categories of residential, commercial, and industrial. Jacobs, instead, discusses land uses as activity generators, with a primary focus on the variety (or lack thereof) in individuals attracted to use them. The focus for Jacobs is on the variety of people present in a given location, as opposed to what is being done at that particular location. In this construction, land use grain is a proxy for the diversity of human usage of urban space. This distinction means that two land uses which generally would be considered the same (say, two restaurants with similar hours) could be considered as adding to the diversity of a particular street if they have different consumer bases, as they would be attracting two separate groups of individuals to the street.[12]

---

9 (Lynch, A Theory of Good City Form, 1981)
10 (Jacobs, 1961)
11 The other three being street connectivity/short blocks, diversity in the age and economic yield of individual buildings, and density.
12 Jacobs, like Lynch, also includes a temporal aspect to her concept of land use grain. In fact, the timing of activity in urban space is key to her concept of land use grain. However, this thesis will not discuss metrics which include timing, as no such metrics are known to the author. Clearly, as more data become available, more research could be done on temporal distributions of activity in the urban form.

Lynch and Jacobs invest the largest effort in identifying and defining the concept of urban grain, including land use grain. However, as Lynch's and Jacobs's example-laden styles indicate, land use grain is best understood when placed in context of its impacts on the urban fabric. Below, the various impacts of land use grain on transportation and public health will be discussed, focusing on empirical literature.

LAND USE GRAIN AND TRANSPORTATION

Empirical approaches to land use grain and its impacts on transportation range widely between subfields of research and across time. The major transportation subfields which discuss the concept are transportation demand modeling, non-motorized transportation research, and transit demand research.

Transportation Demand Modeling

As early as the late 1970s, academics in transportation demand modeling (TDM) were looking at the relationship between land use grain and transit usage. In 1977, Pushkarev and Zupan found that there were a set of *land-use thresholds* which justified different transit investments for particular urban forms, as different densities and land use mixtures could not accommodate the same transit options.[13] Pushkarev et al. continued publishing books through the early 1980s, including *Urban Rail in America (1982)*, an exploration of which metropolitan areas in the U.S. could support rail systems of what intensities based on their respective land use patterns.[14] As public funding for transit stagnated in the mid-1980s through the expansion of highway construction, studies of this nature became less common.[15]

TDM research began looking at land use grain in earnest in the late 1980s with Robert Cervero's report *American Suburban Centers: A Study of the Land-Use Transportation Link*. This report proposed a relationship between land use and transportation trends in U.S. suburbs, by focusing on travel demand patterns in suburban employment centers. In this report, Cervero showed that mixed-use

[13] (Pushkarev & Zupan, Public transportation and land use policy, 1977)
[14] (Pushkarev, Zupan, & Cumella, Urban Rail in America: An Exploration of Criteria for Fixed-Guideway Transit, 1982)
[15] To see the expansion of the Highway Trust Fund (and the subsequent stagnation of transit funding) since 1982, see the FHWA report on the history of the HTF: (USDOT Federal Highway Administration, 2011).

developments have an impact on travel demand by internalizing trips within their bounds. This report does not include land use grain as a concept, but does demonstrate empirically clearly that employment centers with a mixture of land uses have less congestion, fewer trips, higher incidences of ridesharing, and higher transit ridership.[16]

TDM research did not begin to attempt close measurement of land use grain alongside transportation models until the 1990s, when Cervero and Kockelman published an article in 1997 testing how broad measures of density, diversity (their measure of land use grain and other forms of urban grain), and design played into VMT, trip rates, and mode choice using survey data from the San Francisco Bay Area. This study gathers together some of the many different methods to measure land use that were used at the time, and included several of the measures which will be looked at in Chapter III. Cervero and Kockelman found that diversity was a statistically significant predictor of trip degeneration, specifically decreases in trips in single-occupancy vehicles, decreases in personal vehicle usage, and increases in non-auto commuting.[17] This work signals the first major effort in TDM research to focus on mitigation of trip demand or on transportation mode choice.

In 1998, Boarnet and Sarmiento applied a different aspect of TDM, behavioral analysis, to land use attributes. Boarnet and Sarmiento did not find strong relationships between travel behavior and land use grain proxies. However, their analysis used employment densities and residential density to proxy for land use grains and long commutes and number of vehicles owned as proxy for travel behavior. This study does not describe the intersection of associations between land use grain and travel behavior, but it does suggest that appropriate measurement techniques must be used to find the statistical rue relationship between these concepts.[18]

Ewing et al. in 2003 used a variety of measures of land use grain to measure the impacts of sprawl on transportation mode choice. This study found that more or less finely-grained land use distributions are associated with fewer vehicles per household,

---

[16] (Cervero R. , America's Suburban Centers: A Study of the Land Use-Transportation Link, 1988)
[17] The models calculated in Cervero & Kockelman's study only explain a small portion of the variability in responses (R^2 ranges between 0.171 and 0.203), but in each of the cases above, land use grain measures were still statistically significant ($p < 0.05$) (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997).
[18] (Boarnet & Sarmiento, 1998)

higher percentages of commuters using public transportation or walking, daily vehicle miles traveled per capita, lower levels of traffic fatalities, and lower levels of air pollution in the form of ozone. This study did not find a significant relationship between the value of land use grain and average delay or average commute times, however. These associations have strong implications for the use of land use grain measures, particularly the association with traffic fatalities.[19]

The 2012 publication on measuring land use grain and travel demand is by Reid Ewing and Cervero. *Travel and the Built Environment* is an updated from the original work in 2001. Both studies are meta-analyses of planning studies analyzing the relationship between urban form and transportation demand. This study analyzed the elasticity of travel given changes in urban form variables across a large number of previous analyses. As would be expected, travel is generally inelastic with respect to any individual variable in the urban form[20], although Ewing and Cervero hypothesize that a change in several variables in the urban form would effect a greater change in the travel demand of a particular individual within that form. Notably, Ewing and Cervero hypothesize that—given their results—density may be a proxy for other measures of urban form, as places which are dense are more likely to have diverse land uses, more careful design standards, and shorter distances between destinations.[21]

While TDM research has acknowledged the value of trip degeneration and mode choice, and their relationship with land use patterns, this knowledge is only just beginning to enter practice. TDM is performed in the US by federally-designated Metropolitan Planning Organizations (MPOs) or by transit agencies. These agencies overwhelmingly use TransCAD, a transportation modeling software suite by Caliper. While this software suite is useful for forecasting travel demand, the travel demand models do not readily incorporate land use considerations or mode choice into travel demand forecasts. While some Smart Growth-based software packages are now available to MPOs and transit agencies to better account for effecting changes in mode

[19] (Ewing, Pendall, & Chen, Measuring Sprawl and Its Transportation Impacts, 2003)
[20] This is expected because travel is a derived demand. Individuals will travel at relatively steady rates given certain demographic variables such as income, and changes in any one aspect of the built environment are unlikely to effect a large absolute change in the number of trips necessary (Mannering, Washburn, & Kilareski, 2009).
[21] (Ewing & Cervero, Travel and the Built Environment, 2010)

choice by better incorporating land use and other considerations, most agencies engaging in TDM have yet to add these capabilities.[22] This disconnect between research and practice indicates there is still more work in studying and advocating for land use considerations to be included in transportation analysis.

Non-Motorized Transportation & Physical Activity

The rise of non-motorized transportation as a research topic in the late 20th century (among both planners and public health advocates) led to a burst of studies using transportation demand to measure the inclination of individuals to walk, bike, or take public transit. This is often referred to as mode choice in the transportation modeling literature. This literature shows associations between land use grain and non-motorized transportation which may be more meaningful than the relationships between non-motorized transportation and other measures of the urban form.

In a study measuring the association between urban form and physical activity in 2005, Lawrence Frank et al. found that walkability—defined as 6*(z-score of land use grain) + (z-score of net residential density) + (z-score of intersection density)[23]—had a statistically significant relationship between physical activity when demographic considerations are accounted for ($p < 0.001$). This study measured the land use grain of the neighborhoods immediately surrounding study participants, the association between land use grain (among other characteristics of urban form), and the amount of physical activity (as measured by an accelerometer worn by participants). The measure would assess whether the urban form made it more likely for an individual to reach 30 minutes of moderate physical activity a day (an important indicator in the public health realm).[24]

In a follow-up study in 2007, Frank et al. investigated more closely whether urban form had a role in effecting greater physical activity among individuals, or whether individuals were self-selecting into particular neighborhoods. This study found that self-selection did play some role in whether individuals engaged in physical activity (specifically walking for individual trips). Additionally, urban form made individuals

---

[22] (Caliper, 2012) (Lu & Nimbole, 2008) (New Orleans Regional Planning Commission, 2010)
[23] This walkability index was created to work around issues of multicollinearity across these variables.
[24] (Frank, Schmid, Sallis, Chapman, & Saelens, 2005)

approximately 1.62x more likely to walk for any trip than individuals in an "un-walkable" neighborhood, given particular self-selection factors.[25]

Oliver et al. studied physical activity for recreation and for utility among survey respondents in Vancouver, and found a difference between the impacts of urban form on physical activity (non-motorized transportation) for utility purposes and physical activity for non-work / leisure purposes. Specifically, utility trips or physical activity for exercise were not affected strongly by urban form (echoing results from transportation demand studies which showed that these trips were inelastic given urban form variables), but walking for errands or for leisure was strongly affected by urban form. Studying mode choice among recreational or non-work trips (or other more elastic reasons for transportation) may tell researchers more about the effects of urban form on the desire to walk than studying mode choice among work trips. This also suggests that a more mixed land use grain may increase the enjoyment associated with walking, encouraging individuals to walk for non-work trips.[26]

Sundquist et al. found in a 2011 study of non-motorized transportation in Sweden that neighborhoods with high walkability (defined similarly as the walkability index in Frank et al. 2005) was associated with higher levels of walking for transportation and for leisure, and was also associated with higher levels of physical activity in general.[27]

A 2011 study by Buehler and Pucher analyzed cycling in 90 American cities and the relationship between cycling rates and the presence of bicycling infrastructure and supportive policies.[28] This work found that land use grain measurements (part of the sprawl index used in this study) was statistically significant at the 5% level ($p < 0.05$) for all but one model tested.[29]

Ewing and Cervero's 2012 meta-analysis detailed in the previous section also discussed the implications of diversity (their measure of land use grain) on the likelihood of individuals to walk or to take public transit. Unlike overall travel demand,

[25] (Frank, Saelens, Powell, & Chapman, 2007)
[26] (Oliver L. , Schuurman, Hall, & Hayes, 2011)
[27] (Sundquist, et al., 2011)
[28] This study made use of the several land use grain metrics detailed in Ewing et al.'s 2003 work. (Ewing, Pendall, & Chen, Measuring Sprawl and Its Transportation Impacts, 2003)
[29] (Buehler & Pucher, 2012)

the likelihood of walking or taking public transit is highly associated with land use grain.[30]

Land use grain—and indices built which include land use grain as a dimension—has a relationship with transportation choices, both utilitarian and recreational. However, the magnitude of this relationship may be obscured by differing measurement techniques over time and across subfields.

OTHER IMPLICATIONS OF LAND USE GRAIN

Land use grain has other implications outside of its impacts on transportation. Below, three other dimensions of land use grain's impact on urban form and the urban fabric are discussed: *legibility and community, zoning codes*, and *property values*.

Legibility and Community

Since land use grain informs the legibility and efficiency of a space, Kevin Lynch considers this as an important component of a good city. Legibility is a key dimension of Lynch's conception of the good city. By legibility, Lynch refers to the ability of an individual within an urban space to understand where one is within a particular city, who one is likely to meet, and what others are likely to be doing. To be plain, a coarse, sharp land use grain could indicate a single land use, a single set of potential passersby, and a narrow set of likely activities. This contrasts with a region with a fine, blurred land use grain, in which passersby may be partaking in a variety of activities and may be from a variety of social and ethnic backgrounds. Land use grain, then, contributes a good deal to the legibility of a space by assisting individuals in recognition of the types of uses, persons, or activities which may surround them in a given environment, raising or lowering comfort levels appropriately.[31]

Politically, a more evenly distributed land use grain may have implications for the strength and quality of the association between citizens of a single urban fabric. Specifically, recent research has shown that the amount and type of contact between citizens of different social and ethnic backgrounds may have a significant effect on local politics and on individual political behavior. For one example, the 'racial threat' theory

---

[30] (Ewing & Cervero, Travel and the Built Environment, 2010)
[31] (Lynch, A Theory of Good City Form, 1981)

in political science looks at the impact of the presence of racial outgroups on the political behavior of other racial groups.[32]

Wood et al. found a complex relationship when studying participants' *sense of community* compared with neighborhood walkability and land use grain (2010). The relationship included a wide variety of land use grain and walkability thresholds above or below which walking habits and individuals' sense of community changes. In short, walking for leisure increases with walkability to a certain threshold, beyond which walking decreases. Sense of community follows a similar pattern of increasing to a certain point and then decreasing. Wood et al. hypothesize that beyond a certain level of density, the sheer volume of residents, visitors, and workers detracts from the sense of community and walkability that a more moderate density and grain may encourage.[33]

Rise of Form-Based Codes

As the present coarse distribution of land uses in U.S. cities has become more critically analyzed, planners have looked towards methods of increasing the mixture of land use. One influential idea that has arisen is form-based codes: zoning codes based around the design and bulk of the buildings as opposed to the uses within them. A primary goal of the proponents of form-based codes is to encourage a more fine and blurred land use grain than currently exists by breaking down barriers to the mixing of dissimilar land uses.

Standard zoning codes in the U.S. arise out of *Village of Euclid v. Ambler Realty Co (272 U.S. 365 (1926))*, which allowed that municipalities could regulate development by land use restrictions which demonstrably protect the health, safety, and welfare of the community. Euclidian zoning, or use-based zoning, has been adopted by every U.S. state and by most municipalities in the U.S..[34]

Use-based zoning encourages a coarse, sharp land use grain by regulating which uses can be present in a given location and at what densities. In recent decades, this has become considered a negative trait, and architects and planners have looked for means around use-based zoning. Form-based codes are one such means, and are growing

---

[32] Political science and philosophy have growing literatures on the relationship between contact and political behavior. For more on racial threat specifically, see e.g. (Key, 1949), (Enos).
[33] (Wood, Frank, & Giles-Corti, 2010)
[34] See, e.g. (Lawlor, 2011)

rapidly in the U.S.. Form-based codes regulate development by the size, shape, bulk, and design of buildings, but allow for flexible use of buildings. This is intended to encourage a more fine land use grain by allowing property owners to make use of their property how they see fit.[35, 36] As these new form-based codes are implemented across the U.S., land use grain patterns are expected to change. These codes have been concentrated in suburbs and rural towns, generally in the Southeast and the West, but these codes have increasingly been adopted by cities nationally. This includes larger cities, such as Miami and Denver.[37] A rising level of form-based codes, particularly as they emphasize changing land use grain patterns and land use development patterns, requires a more concrete knowledge of land use grain.

Relationship to Property Values

Property taxes tend to be a major source of income in U.S. cities, so maximizing property value is one way in which cities can maintain adequate budget programs. New research in planning and in real estate has just begun to study the approximate differences in property value in areas with different land use grain patterns.[38] This growing area of study contrasts with some transportation demand modeling strategies. Transportation demand modeling research has attempted to isolate the effect of urban form on transportation choices, in part by accounting for factors of residential self-selection. Literature on property values, however, focuses largely on residential self-selection itself (among other things), and the value this trait brings to cities.

Matthews and Turnbull in 2007 analyzed the relationship between a wide variety of land use and transportation variables to find their impacts on property values in the Seattle area. Proximity to retail, office space, industry, institutional uses, hotels, and apartments all had a significant ($p < 0.05$, in some cases $p < 0.01$) impact of the value of individual residential properties. Additionally, removing visible uses from the model did not decrease the effects (and in some cases enhanced the effects), so their model indicated that the presence of different land uses nearby (despite not being visible) had

---

[35] (Duany & Plater-Zyberk, Neighborhoods and Suburbs, 1995)
[36] (Duany & Talen, Making the Good Easy: The Smart Code Alternative, 2001-2002)
[37] (Lawlor, 2011)
[38] For the most recent example of popularizing work on this topic, see Emily Badger's short article in The Atlantic on Asheville's property and sales tax studies. (Badger, 2012)

impacts on property values.[39] This indicates that, contrary to intuition, proximity to a wide variety of land uses may improve residential property values and therefore residential property tax receipts.

[39] (Matthews & Turnbull, 2007)

CHAPTER III: MEASUREMENT CONSIDERATIONS

Having discussed the meaning and the importance of land use grain to subfields of transportation planning and public health in the previous chapter, the following section summarizes the standard land use grain measurement practices. Three metrics which are in use will be described and analyzed. Important questions which arise when employing any particular one of the three metrics will also be described, resulting in a summary of options for defining area of capture and defining land use categories to use in analysis.

METRICS

The metrics considered in this chapter are the entropy-based equation (used in transportation modeling and public health), the jobs-to-housing ratio (used in government and planning), and the Herfindahl-Hirschman Index (used in planning and public health).

Table 1: Three chosen metrics by academic discipline

	Entropy-based	Jobs-to-housing Ratio	HHI
Public Health			
Transportation Demand Modeling			
Urban Planning			

Entropy-Based

The most often-used measure of land use grain is the entropy-based equation, pioneered by Cervero and Frank. The equation is based on the physical property of entropy, or the measure of disorder in physics. Generally, entropy governs the distribution of atoms in a physical space. This measure derives from the entropy of mixing, which ranges from a perfectly disorganized space where atoms are uniformly spread out and mixed to a perfectly organized space where atoms are grouped by element and clearly structured.[40] Applied to land use grain, this metric approximates the amount of mixing among dissimilar land uses by measuring the amount of the total space is taken up by each of the component land use categories.

[40] (Bhadeshia)

Function 1: Entropy-based LUM Equation

$$-\sum_{i=1}^{n}(p_i^* \ln p_i)/\ln n$$

where p_i is the proportion of square footage attributed to land use i, and n is the number of land use categories.[41] This definition provides results "ranging between 0 (homogeneity, wherein all land uses are of a single type) and 1 (heterogeneity, wherein developed area is evenly distributed among all land use categories."[42] This form of the equation, published in 2007, is a generalization of an equation developed by Frank and Gary Pivo in Frank's doctoral dissertation at the University of Washington in 1994,[43] which was in turn based on preliminary work done by Robert Cervero in 1989.[44] Cervero gives the best explanation for the formula in his 1997 article linking travel demand with density, diversity, and urban design.[45] These three composite variables were used to model the built environment's impact on travel demand.[46] In this spirit, the entropy-based equation of land use grain is most often used in walkability indices made up of street connectivity, residential density, and land use grain.[47]

Notably, the entropy-based equation of land use grain has been strongly associated with utilitarian walking, but its relationship with recreational walking may be more tenuous. Christian et al. in particular encourage researchers to be careful when defining land use categories, as different land use category sets have different relationships with utilitarian vs. recreational walking. This issue will be examined in further detailed later in this chapter.[48]

The entropy-based equation of land use grain is most often taken at face value, but Wood et al. classified the results into a categorical variable (low, medium, and high land use mix) in order to account for potentially flawed and non-normal data. The

[41] (Frank, Saelens, Powell, & Chapman, 2007)
[42] (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997)
[43] (Frank & Pivo, 1994)
[44] (Cervero R. , 1989)
[45] (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997)
[46] "Diversity" in this article is land use grain, and Cervero uses seven indicators to make up the composite variable as a whole: dissimilarity, entropy, vertical mixture, per developed acre intensities, activity center mixture, commercial intensities, and proximities to commercial-retail uses.
[47] See, e.g. (Frank, Schmid, Sallis, Chapman, & Saelens, 2005), (Frank, Saelens, Powell, & Chapman, 2007), (Christian, et al., 2011), etc.
[48] (Christian, et al., 2011)

decision whether to treat land use grain as a scalar or a categorical variable in a particular study is up to the discretion of the researcher, but should be seriously considered in cases where land use categories are less than certain.[49]

Jobs-to-Housing Ratio

The *jobs-to-housing ratio* is a conceptually simpler measure of urban diversity, often used as a proxy for land use grain in transportation-related studies. The jobs-to-housing ratio is most often used when researchers are directly concerned with utilitarian transportation, especially commutes. The jobs-to-housing ratio is also notable for being the measure of diversity or land use grain in the U.S. E.P.A.'s Smart Growth Index (SGI), and in subsequent Smart Growth-related E.P.A. publications.[50]

The formula for a jobs-to-housing ratio is shown below in Function 2:

Function 2: Jobs-to-Housing Ratio

$$\frac{Total\ number\ of\ jobs\ in\ study\ area}{Total\ number\ of\ workers\ in\ study\ area}$$

This formula has many more measurement contingencies than the above entropy-based equation, which makes this metric more susceptible to manipulation on the part of the researcher, accidental or not. The greatest contingencies are the definitions of the numbers of jobs or workers. This data may not be readily available at the same geographies as the study—or may be available but preparing it for usage would take an undue amount of resources—these measures are sometimes approximated, using standard numbers of workers per household or jobs per workplace (or type of workplace). A minor change in the approximation scheme can have significant effects on the ratio, and in turn on the analysis. The third important contingency is how the study area is defined, which will be discussed later in this chapter.

Ewing & Cervero, in their 2012 meta-analysis of studies analyzing the relationship between travel and the built environment, find that studies which use a jobs-housing balance or jobs-to-housing ratio measure of land use mix show a greater elasticity in travel mode choice (particularly the likelihood of walking trips) than the

---

[49] (Wood, Frank, & Giles-Corti, 2010)
[50] (Ewing & Cervero, Travel and the Built Environment, 2010), (United States Environmental Protection Agency, 2012)

more-commonly-used entropy-measure detailed above. This same analysis, when done on housing density vs. jobs density, found greater elasticities associated with housing density. Additionally, Ewing & Cervero were focusing on non-work travel. This analysis indicates that land use grain itself is less integral than the location of work to the decision to walk, bicycle, or take public transit, countering some assumptions of the importance of land use grain.[51]

According to the E.P.A., the standard calculation for a jobs-to-housing ratio should be the ratio of total jobs to total "housed workers," assuming 1.4 workers per household. (The assumption of number of workers per household should be made using local knowledge in any study, as this figure is simply an average the E.P.A. uses for the nation as a whole.)[52]

Herfindahl-Hirschman Index (HHI)

One growing alternative to the entropy-based approach used by Frank et al. is to use a different diversity index: the Herfindahl-Hirschman Index (HHI). The formula for this index is shown in Function 3: Herfindahl-Hirschman Index (HHI) below:

Function 3: Herfindahl-Hirschman Index (HHI)

$$H = \sum_{i=1}^{K} s_i^2$$

where *s* is the share of the total space taken up by each individual land use category. The HHI is most commonly used in economics, and was crafted to measure market concentration. It is commonly used by the U.S. Department of Justice, where mergers are analyzed by their effect on the HHI of the industry of the individual companies merging. The HHI ranges from 0 to 10,000, where 10,000 indicates a single firm controlling the entire market share (or a monopoly) and 0 indicates a very large number of firms each with near 0% of the total market share. According to the guidelines governing antitrust laws in the U.S., a low HHI is below 1500, a moderate HHI is

[51] It is important to note that entropy-based measures of land use grain still were associated with a higher rate of walking, bicycling, and taking public transit. However, this association wasn't as powerful as the association between non-motorized transportation and a jobs-to-housing ratio. (Ewing & Cervero, Travel and the Built Environment, 2010)
[52] (United States Environmental Protection Agency, 2003)

between 1500 and 2500, and a high HHI is over 2500.[53] When applied to land use, an HHI score of 0 would indicate a very fine mix of land use grain, and an HHI score of 10,000 would indicate an environment wholly made up of a single land use category. While the U.S. Department of Justice uses the market share as an integer (20% is 20 in D.O.J. work), a researcher could use the market share as a proportion (20% is 0.20), resulting in a more intuitive scale of 0 to 1. This more intuitive scale will be used in this paper, as it parallels the results from the entropy-based equation.

This formula's usage in planning has been pioneered by Song and Rodriguez in unpublished work through the Active Living Research grant in 2005, and has been included in the NEAT-GIS manual compiled by the Design for Health team, a network of researchers and academics studying the intersection between urban design and public health.[54] The HHI is not as well-tested as the other two measures outlined above, but it is included because of its potential as a useful quantification and because of its relative simplicity.

Sundquist et al., in their 2011 study of neighborhood walkability, used the HHI in addition to the entropy-based equation. Sundquist et al. used the entropy-based equation for an absolute land use grain score as part of a walkability index, *a la* Frank et al., whereas they used the HHI to assess the level of mixture of land uses as a stand-alone concept.[55]

While there is not as much literature on the use of the HHI in a public health or planning context, this index has several strengths. For one, it is more straightforward than the entropy-based equation (and some other alternatives), both conceptually and mathematically. This makes the HHI more accessible to planning professionals. For another, the HHI still measures the grain of land use, whereas the jobs-to-housing ratio (and some other alternatives) is in actuality measuring a slightly different indicator of non-motorized transportation. This allows the HHI to be used in non-transportation contexts—including in public health, where it has been applied—which may be interested in better understanding land use grain and its impacts.

---

53 (United States Department of Justice)
54 (Design for Health, 2012)
55 (Sundquist, et al., 2011)

Other Metrics

In addition to the above three focus metrics, several other land use grain metrics have been employed with varying degrees of success. Some of the following may be useful in very particular contexts, though broad application of these may not be practical.

The dissimilarity index is the most complex alternative land use grain measure used by Cervero & Kockelman in their 1997 study. This index calculated the proportion of dissimilar land uses among grid cells within a single study area (in this case a census tract). The index itself is calculated with the function:

Function 4: Cervero's Dissimilarity Index

$$\left\{\left[\sum_{j}^{k}\sum_{l}^{8}(X_l/8)\right]/K\right\}$$

where K is the number of actively developed cells in the study area and $X_l = 1$ if the land-use category of neighboring cells differ from cell j (0 otherwise). (See Figure 1 below) This measure is a useful tool for calculating how much of the study area is directly adjacent to a differing land use, regardless of whether the differing land use is represented in any proportion across the area of study.

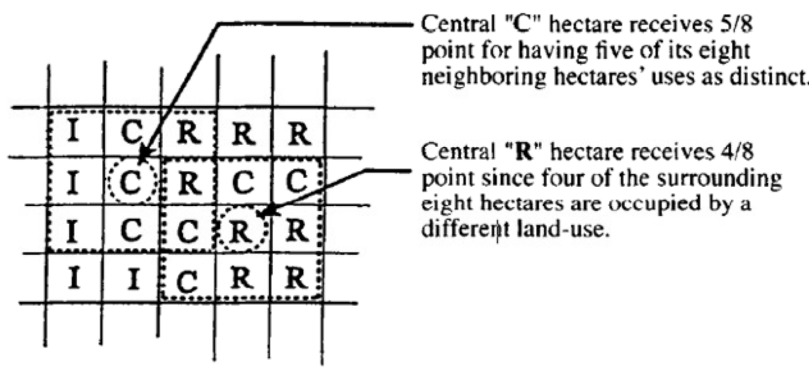

Figure 1: Computation of the dissimilarity index. (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997)

This method is distinctly different from the three focus metrics, as the dissimilarity index is measuring the distribution of land use adjacencies across the area of study, instead of the distribution of land uses themselves. The dissimilarity index may be a very useful tool for analyzing more closely *how* land uses are distributed within a region, as opposed to the very fact of whether land uses are distributed. For example, the dissimilarity index could be extremely low while the three focus measurements could all show a fine land use grain if the area of study has large blocks of differing land uses in equal proportions. As the dissimilarity index is measuring a somewhat different dimension of land use grain, it will not be treated fully in this thesis. However, the author recommends follow-up on this method in future research.[56]

Another potential metric used by Cervero & Kockelman is what they define as the Accessibility-to-Jobs indicator,[57] but is now called the gravity model.[58] This measurement is based on travel times between two origin-destination (O-D) pairs and the impedance between them. This function is at its core on the mathematical model of gravity, and approximates the nearness of distinct land uses and the compactness of land use categories. This model may warrant further exploration in the future, and has been used since in land use grain analysis. In the *NEAT-GIS* manual, the gravity model lists the potential of using parcels instead of O-D pairs, which may create more nuanced results. This model is not explored in this thesis primarily due to its complexity. The gravity model requires background in traffic engineering, data on impedances, an appropriate distance decay parameter, attractiveness scores, exact distances between involved units of study, and so on.[59] This complexity makes this particular model infeasible for practical planning use.[60]

Vertical mixture was a different method Cervero & Kockelman employed in their article described above. Vertical mixture is the proportion of commercial or retail parcels with more than one land-use category on the site. This measure may be useful in a particularly dense environment, an area with a large number of tall or large buildings, or in a heavily mixed use neighborhood. For most neighborhoods, however, this measure may be misleading, as parcels and buildings may not align, or there may be mega-parcels, etc.

---

[56] (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997)
[57] (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997)
[58] (Design for Health, 2012)
[59] For more information on some of these concepts, please see an introductory traffic engineering textbook, such as Mannering et al.'s *Principles of Highway Engineering and Traffic Analysis*. (Mannering, Washburn, & Kilareski, 2009)
[60] (Design for Health, 2012)

Cervero & Kockelman also tested a method to evaluate how diverse or fine-grained uses were at existing activity centers, by combining an entropy-based score measuring only commercial-retail sites and proportions of activity centers which have more than one category of commercial-retail sites available. This metric may be useful in an area with multiple small activity centers, such as a streetcar suburb. In regions with a single commercial center, or regions which rely on automobile-oriented shopping centers, this metric is likely to not provide much information as to the actual distribution of land.

Berrigan et al., in a study of the impact of land use and street connectivity measures on non-motorized transportation in 2010, included employment density as a proxy for land use grain.[61] This study was mainly focused on street connectivity as a measure, so using a proxy for land use grain was in the interests of time and ease. Because the importance of land use grain in transportation studies is explained by the ability to walk from one destination to another (generally one's home to one's work), employment density captures *some* aspects of the measure. However, this proxy is less informative than the jobs-to-housing ratio described above.

Outside of the listed measurements above, there are other alternative methods of measuring land use grain which have been used once or twice in the literature. However, the entropy-based equation and the jobs-to-housing ratio are still the most widely used, while the HHI has been used more frequently in recent literature than the other alternatives. This thesis will focus on these three metrics.

AREA OF CAPTURE

To use any metric of land use grain, one must decide on an area to measure within. Any analysis with a basis in geography is subject to the modifiable areal unit problem (MAUP)[62]. The MAUP is a statistical bias inherent in grouping data which is particularly dangerous for artificial boundaries such as census geography, urban boundaries, or in some places parcels. Statistical results change depending upon the geographic resolution of the data (municipality versus county, for instance) and upon where boundaries happen to be arbitrarily drawn (census tracts versus true

[61] (Berrigan, Pickle, & Dill, 2010)
[62] (Openshaw, 1984)

neighborhoods). This issue requires defining the study area in a meaningful and transparent manner with regards to the purpose of the study and to the unit of analysis.

Duncan et al., in their study of the relationship between land use grain and non-motorized transportation, found that pure measures of land use grain (as this study is working towards) do not have a significant relationship with walking or bicycling, but "refined" measures of land use grain (accounting for either the geographic scale of measurement or only the most relevant land uses) did have a significant relationship.[63] This work shows clearly that researchers can alter the results of statistical analysis using land use grain by altering the measurement area. While this particular study encourages the use of their area correction strategies (detailed later in this chapter), this study also shows that researchers must be careful when choosing a measurement area.

The most variable methods of calculating an area of capture are studied in Cervero & Kockelman's 1997 work, as this work also used the most measurement techniques[64]. In the interests of space, only the most repeated or most potentially useful techniques will be explained in detail here.

Area-Based

The first method of area capture is to simply define an area of study and perform measurements within that area. This method is not very well represented in the literature to date because of the architecture of existing studies, but has been used in isolated examples. The most common method of performing an area-based area of capture is to create a grid system across a city or metropolitan region and perform a calculation across the grid, giving each cell a land use grain measure. This method has been used in the first stages of larger research designs, with the goal of choosing areas in which to perform closer study.[65]

Duncan et al. in 2010 performed an analysis of land use grain across Adelaide, Australia, using Census Collection Districts (CCDs) as the unit of analysis instead of using individuals. This work primarily compared land use grain scores and census-collected data on transportation in CCDs across the urban area, in order to discern

---

[63] (Duncan, et al., 2010)
[64] (Cervero & Kockelman, Travel Demand and the 3Ds: Density, Diversity, and Design, 1997)
[65] See, e.g., (Frank, Schmid, Sallis, Chapman, & Saelens, 2005), (Oakes, Forsyth, & Schmitz, 2007)

whether there was an association between the two. Duncan et al. introduce several potentials methods to account for differences in the sizes of areas of study which may be helpful for future researchers. Duncan et al. area-corrected their entropy-based land use grain measurements by dividing the grain measurement by the ratio of the area of the CCD to the average area of all CCDs in the study. This area-correction solves issues of comparison across CCDs which are diverse in geographic area.[66]

Area-based methods are often not an option in studies which incorporate land use grain, as these studies generally use individuals as the unit of study, as opposed to geographies. For a study with individuals as the unit of study, defining the land use grain of the neighborhood in which they live or work is not as meaningful as defining the land use grain of the area immediately surrounding their exact location. The area being captured is often defined by time or distance from these particular locations—rather than as a signifier of distinct geographies—because these studies are often studying transportation mode choice.

An area-based method of capture could be useful in a comparative study of transportation mode choices, however, particularly if individual-level data is not being collected or aggregate data is preferred. An area-based method could also be useful in a pure treatment of land use grain itself. Land use grain could be used as a response variable in very specific studies, and an area-based method of capture may be more useful in this context.

Simple Buffers

Simple buffers are the least computationally difficult method of capturing an area of study based on individuals as a unit of study. In short, simple buffers are areas within a certain distance of a point (usually the home of an individual study participant), with that distance being defined by the purposes of the study at hand. Simple buffers are not the most descriptive of the individually-based methods available, but it is by far the quickest and the most simple conceptually. This simplicity allows for a very large sample to be buffered quickly, and may be most useful with extremely large sample sizes or for planning professionals with limited technical resources. Simplicity also allows these

[66] (Duncan, et al., 2010)

studies to filter into other doctrines or into the political mainstream more easily, as simple buffering is a more popular and understandable means of defining walkable areas.

The distance of the buffer ranges depending on the research being done, but generally spans from 200m – 2.5km, depending on the goals of the study. In the extensive post-hoc work on the Twin Cities Walking Study, Forsyth et al. make use of simple buffers of 200m, 400m, and 800m.[67] Similarly, Frank et al. use 1km in their 2005 walking-related study.[68] However, when related to bicycling, a researcher may want to use a larger buffer based upon the distance which a cyclist may be able to reasonably travel, ranging on 1.5mi or a bit less than 2.5km.[69]

Network Buffers

Network buffers, based on graph theory flow networks[70], are a more detailed implementation of buffering which allow a polygonal area to be defined by the network available to the individual being studied, rather than as a circle around the individual's location.

Frank et al. utilized a network buffer for their area of capture in the 2005 study. Frank et al. chose a sample of participants from their study area (Atlanta, GA) and followed their movements for a two-day period. The study then attempted to model the data using land use grain (among other factors) as an explanatory variable. As the homes of the individual study participants were known to the researchers, Frank et al. used GIS software to analyze transportation networks and land use considerations in the area immediately surrounding the participants' homes. The researchers created network buffers of 1km around each participant's home and analyzed land use grain in the resulting areas.[71]

---

[67] (Design for Health, 2012)
[68] (Frank, Schmid, Sallis, Chapman, & Saelens, 2005)
[69] The literature on standard walking and bicycling distances is somewhat contentious, and lies outside of the purview of this study. Therefore, this work will not take a stance on a standard distance for buffers, leaving that to the discretion of the individual researcher.
[70] Flow networks are the basis of network buffers in most GIS suites. (Ahuja, Magnanti, & Orlin, 1993)
[71] (Frank, Schmid, Sallis, Chapman, & Saelens, 2005)

The majority of studies following Frank et al. also used network buffers, either in part of in whole, and this method has become the standard means of selecting study areas in situations where the study relies upon a sample of individuals.

Line-Based Network Buffers

A line-based network buffer takes this a step further, by clipping the street network by the polygon-based network buffer and then creating a small simple buffer around the clipped street network (See Figure 2: Comparison of simple buffers, network buffers, and line-based network buffers. below). The intention of this buffer style is to ensure that only accessible parcels are selected as part of the analysis.

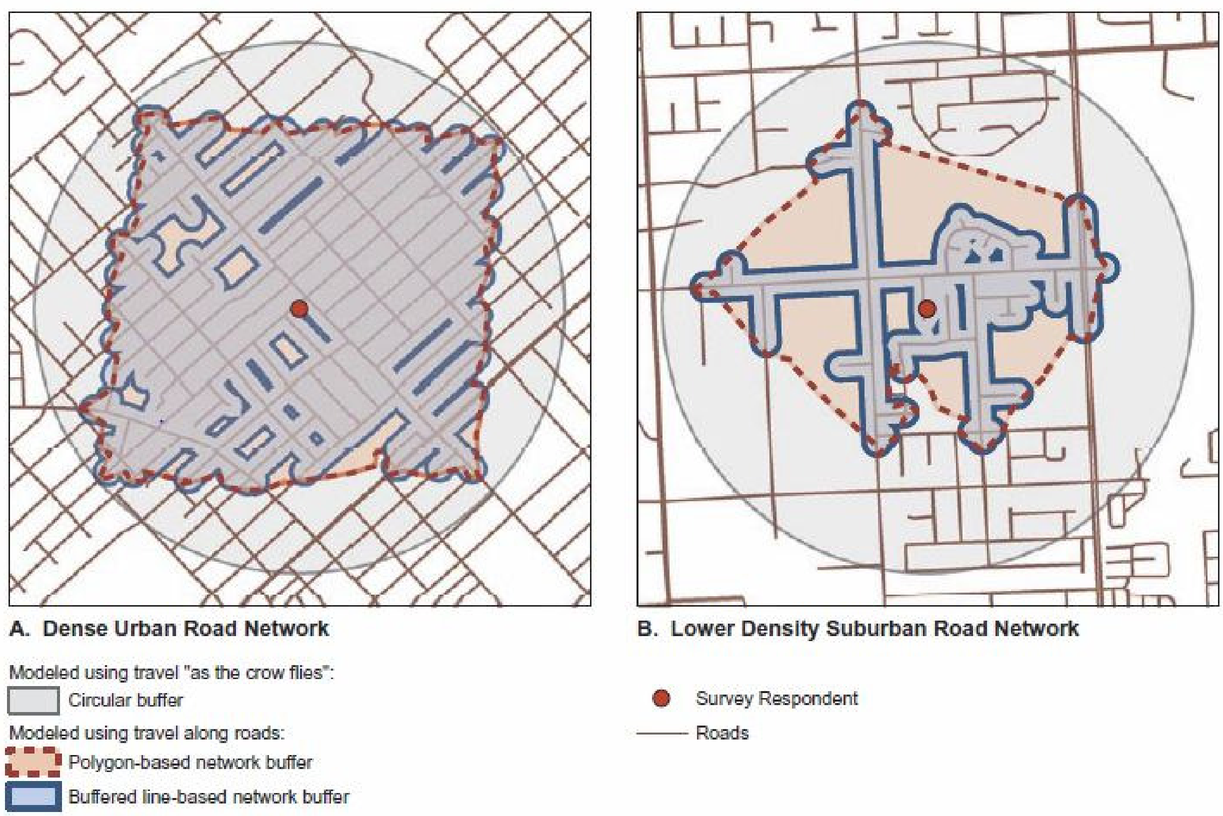

Figure 2: Comparison of simple buffers, network buffers, and line-based network buffers. (Oliver, Schuurman, & Hall, 2007)

The line-based network buffers above were constructed by Oliver et al. in order to create a more nuanced version of the polygon-based network buffer, which has a tendency to include lots not facing the streets available to pedestrians, for example.

Oliver et al. set a network-based buffer of 950m with an additional line-based buffer of 50m in order to obtain a 1km accessible area.[72] Once this buffer is made, the area of study is chosen by collecting all parcels which touch this buffer, rather than using the buffer itself.

This method does not filter out back parcels (parcels not in contact with the formal street network), and therefore may not adequately describe the population of accessible parcels if the street network data is incomplete. This deficiency is minor, and for smaller datasets can be solved by reviewing selected parcels. But the time required to carefully review selected parcels for accessibility may not be possible on a very large dataset.

There is potential for this method to be refined further using building footprints data. Because the purpose of the line-based network buffer is to more closely approximate where individuals can walk, looking at which buildings an individual could reach within a reasonable walking time would be more informative than looking at which parcels an individual could reach. This method may also resolve some issues with the MAUP by preventing reliance upon arbitrary boundaries (parcels, in this case). This refinement of the line-based network buffer is purely hypothetical, and will not be closely examined in this thesis.

Transit Network Area Captures

A final method of capturing a study area, transit network buffers and line-based transit network buffers, will not be treated closely in this work, as this method has no strong precedent in the literature. However, as the available of transit data increases, this may become a viable means of selecting an area of study.[73] This method would be more useful in studying transit as a potential transportation mode choice, or in studying the accessibility of metropolitan areas to neighborhoods or the demographics within neighborhoods.

---

[72] (Oliver, Schuurman, & Hall, 2007)

[73] For early implementations of this, see Mapnificent or Walk Score's Apartment Finder tool (Wehrmeyer), (Walk Score).

LAND USE CATEGORIES

Land use categories are another contingency upon which most land use grain measurements rely.[74] How a researcher defines and generalizes existing land use data into categories is of great importance both to the outcomes of the individual study and to the transferability of the study to differing contexts. Christian et al. found that the land use categories used in the entropy-based equation of land use grain, for instance, affects the strength and quality of the statistical relationship between the final land use grain measure and different types of walking behaviors. The study found using land use categories of 'Residential,' 'Retail,' 'Office,' 'Health, Welfare and Community,' and 'Entertainment, Culture and Recreation' have a strong association with utilitarian walking, whereas including 'Public Open Space,' 'Sporting Infrastructure,' 'and 'Primary and Rural' land use categories had a stronger association with recreational walking.[75]

These results are similar to the results from Duncan et al. discussed in the previous section. Duncan et al. found that "refining" land use categories to ones researchers deemed relevant changed the statistical relationship between land use grain and non-motorized transportation to be significant, when a more general measure was not significant.[76] This carries with it the implication that any change in the definition of land uses or land use categories for use in land use grain analysis can have a significant effect on the outcome of the research.

The number and types of categories used in the existing literature depends largely on the types of land use grain being studied, and the function land use grain is hypothesized to have in the context of the study.

Residential, Commercial/Retail, and Office

[74] Jobs-to-Housing ratios do not rely on land use categories, as they do not measure land use directly.
[75] (Christian, et al., 2011)
[76] Duncan et al. encourage the use of more specific models of land use grain, which only include land use categories which are deemed relevant to the study at hand. While using only land use categories which are relevant to the study at hand may more often give statistically significant results, it necessarily clouds land use grain itself. (Duncan, et al., 2010)

In their 2005 study, Frank et al. made use of three land use categories: residential, commercial, and office.[77] Frank et al. attempted to create a model linking physical activity levels with built environment indicators, one of which was land use grain.[78] Residential was chosen as it is the home is the starting point and end point for most citizens' days, and the majority of citizens' trips. Commercial and Office Space were chosen to signify the proximity of places to work and places to shop, and were differentiated because these two spaces fulfill different roles.

In their 2007 study, Frank et al. describes more fully the reason these three categories were chosen. For Frank et al., these three categories represent the 'walkable' destinations, with all other land use categories considered 'un-walkable.' To account for other land uses, the total square footage with which the square footage of chosen land uses was compared included all land in the study area, as opposed to only the proportion of these land uses to one another. This allows the entropy-based equation to distinguish between areas with very little residential, commercial, or office space and areas with a large amount of the above land use categories.[79]

This land use categorization scheme may be useful for discussing utilitarian (and some commute) non-motorized transportation in a particular urban environment (hence its use in transportation-related studies), but this categorization scheme fails to take into account recreational space and industrial work places. This categorization also fails to account for other implications of a mixture of land uses, particularly aesthetic-, health-, and safety-related implications. To account for other land use mixtures and their impacts on land use grain and on other aspects of the urban fabric, a more multivariate land use categorization scheme is necessary.

Recreational

[77] 'Commercial' often includes office space, but for Frank et al.'s analysis, commercial meant retail space, with offices being a separate category. The justification for this was not given. (Frank, Schmid, Sallis, Chapman, & Saelens, 2005)

[78] It was decided to use a holistic model of built environment indicators (or urban form measures, as they call them), because there tends to be a large amount of interaction between different built environment indicators such as density and land use grain. This is broader than the purposes this paper, but future research using the measure of land use grain should take this into consideration. (Frank, Schmid, Sallis, Chapman, & Saelens, 2005)

[79] (Frank, Saelens, Powell, & Chapman, 2007)

One land use category: recreational lands, appears to provide a greater insight to the impacts of land use grain on recreational transportation mode choice and on some other aspects of the urban fabric.

Oliver et al., in their study of available buffering techniques done in Vancouver, BC, used five land use categories: 'Recreational/Parks,' 'Residential,' 'Commercial,' 'Industrial,' and 'Institutional' (with mixed use properties' areas prorated as each use, depending on their split). An 'Other' category was developed for all other land uses, but was excluded from analysis, and lumped with empty land use parcels to help provide a proportion of relevant land use categories to total land area. These land use categories were composites built out of the Greater Vancouver Land Use Data set, which assigns detailed land use categories by parcel. [80]

Christian et al., in their analysis of the impact of differing land use categorization schemes on land use grain scores, use the Western Australia Ministry for Planning's land use categorization scheme, which breaks recreational land into 'Entertainment/Recreation,' 'Public Open Space,' and 'Sporting Facilities.' This differentiation proved to be important to Christian et al.'s results, as 'Entertainment/Recreation' lands were significantly related to utilitarian transportation while 'Public Open Space' and 'Sporting Facilities' were significantly related to recreational transportation.[81] This speaks for the potential to further classify recreational lands based upon how, when, and what scales they are used by the population in the neighborhood.

Adding recreational land as a category helps capture more detail about the urban environment, and is of particular importance to studies of physical activity and recreational transportation in public health. All referenced literature regarding non-utilitarian transportation included recreational land of some type in the analysis of land use grain.

Industrial

[80] Oliver et al. created simple logistic regression models of tendency to walk versus land use category distribution to find which land use categories had a strong statistical impact on the tendency to walk, rather than an entropy score, so the ability to add extra land use does not conflict with any discussed land use grain metrics. (Oliver, Schuurman, & Hall, 2007)

[81] (Christian, et al., 2011)

Including industrial lands can help public health research looking at proximity to noxious industries, and may capture some influence on commute transportation mode choice if industrial companies employ a significant portion of the population. Most studies which include industrial lands in land use grain measures also use a variety of other standard land use categories, lumped into an 'Other' category.[82]

Seven Land Use Categories

The *NEAT-GIS* manual proposes using seven land use categories: 'Residential,' 'Commercial,' 'Institutional,' 'Office,' 'Parks and Recreation,' 'Industrial,' and 'Vacant'[83] (with vacant properties removed from analysis). This categorization is used mostly by members of the Design for Health team, including Forsyth, Oakes, and Rodriguez.[84] [Public health, general land use mix work]

Ultimately, the land use categories used in analysis do depend on 1) the goal of the study and 2) the data or local knowledge available. For the most cohesive land use grain measurement, however, the more land use categories the better.

Data Requirements

The following table synthesizes the data requirements for the three focus metrics:

Table 2: Data requirements for the three focus metrics

	Entropy-based	Jobs-to-housing Ratio	HHI
Parcels shapefile			
Land Use classifications			
Street network shapefile*			
Employment Numbers			
Population			

**Street network data are required for network- or line-based buffers*

[82] See, e.g. (Duncan, et al., 2010)
[83] Vacant lands pose a particular issue for land use grain measures, as vacant land is by definition the absence of formal land use. Most studies have avoided this issue by excluding vacant land from the analysis, while noting vacant land if there is a significant amount. Studies on the impact of vacant land could look at the grain of vacant land within the urban fabric, but studies of this sort have not yet been performed.
[84] (Design for Health, 2012)

As shown in the Table 2, these analyses do not require much in the way of individual data sets. However, these data must be cleaned for analysis. The types of cleaning necessary for analysis will be discussed in Chapter IV, as part of the demonstration.

Of particular interest is the land use classification data. Land use classification data is generally available at the municipal level through the municipal planning agency, and may also be available through regional or state planning agencies. Unfortunately, national land use classification datasets are created either from satellite imagery or from an amalgamation of local land use data sets. Satellite imagery-based data often return vague results for urbanized areas, such as *Highly Developed* instead of residential or commercial categorization. Amalgamated datasets are more likely to contain classification errors due to the broad range of zoning regulations across the US. This thesis will not look at data complications arising from crossing state boundaries, as the research on land use grain has studied individual metropolitan areas.

Chapter IV: A Demonstration

The remainder of this thesis will consist of a demonstration of the land use grain measurement techniques, with consideration to the land use categories and area of capture discussions (as discussed in Chapter III). This chapter will provide the results of the visual and numerical assumptions which are used in land use grain metrics. This chapter will also provide the results of each metric given these assumptions. Once calculations have been performed, this chapter will discuss how each metric performed across assumptions and compared to one another.

Area of Study

This demonstration looks at the Dudley district, a portion of the Roxbury and North Dorchester neighborhoods in Boston, MA. This region was chosen as parcel-level data and street network data were readily available, and because parallel research in the Dudley neighborhood allows for this demonstration to align with existing work on the part of the author.

Data Used

This demonstration requires a street network shapefile[85], parcel-level land use data for the City of Boston[86], a chosen starting location[87], and access to geographic information systems (GIS) software. Street shapefiles were obtained from the U.S. Census Bureau's TIGERLine dataset, although local governments may have better street shapefiles available through their GIS department. The TIGERLine dataset is only updated sporadically, and has no attribute information for the street lines. A local government streets shapefile would be more useful for network analysis through GIS software and is more likely to have any new streets which may have been created since the previous census. Availability of parcel-level land use data is more subject to variability, but is likely available through the local or state government, or through a state land grant university.

---

[85] (US Census Bureau, 2010)
[86] (Harvard Geospatial Library, 2011)
[87] In this case, 504 Dudley Street, the offices of the Dudley Street Neighborhood Initiative, were chosen.

The parcel-level land use data will likely need to be cleaned before beginning analysis. Municipalities use a variety of land use categorization schemes, and regardless of the scheme used by the municipality there are likely to be parcels mis-categorized. If

The data used for this analysis is originally classified using the State of Massachusetts official land use codes, which divide land into fine categories using a three-digit coding system. These land use classifications do not neatly coincide with the land use categories used for the following demonstration, so some land use categories are changed or generalized. Additionally, some local knowledge was utilized to correct flaws in the base data. Because this demonstration is on a single point of interest, reclassifying land uses at the parcel level was not time-intensive. For more information on the State of Massachusetts official land use codes and the judgments for this analysis, see Appendix I.

Maps and Discussion

This section will look closely at sample maps using various methods of capturing study areas and of land use categorization. The land use grain will be calculated using the three described measures. The jobs-to-housing ratio will be calculated, but it is not dependent upon the land use categorizations shown in the following maps. The results will be summarized and more closely analyzed at the end of the chapter.

Simple Buffer

The simple buffer will be measured first, as it is the least complex method of capturing an area of study. This buffer style allows for a quick analysis, as it does not require the use of a streets shapefile or any local knowledge outside of the land use shapefile.[88] Below are two maps, the first showing the simple buffer performed with a 1km radius around a single point. The second map shows all of the parcels this selection process includes. The point used in all of these examples if the office of the Dudley Street Neighborhood Initiative in Boston, MA.

[88] No local knowledge is necessary *provided* the land use data is trustworthy and recent. Otherwise, land use data will most likely have to be reclassified to fit within the scope of the analysis.

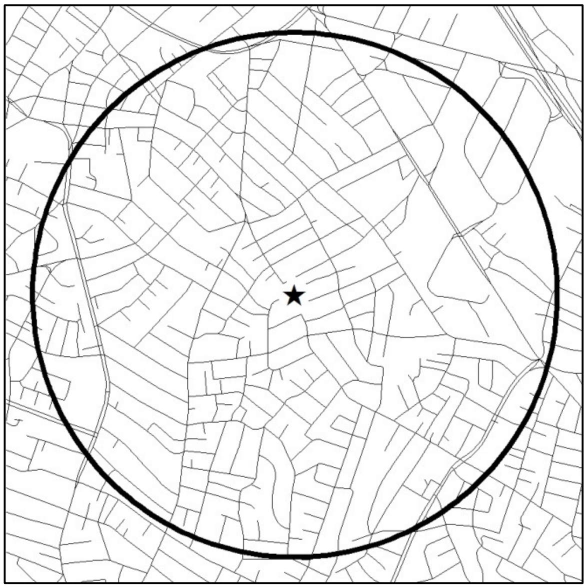

Figure 3: A simple 1km buffer.

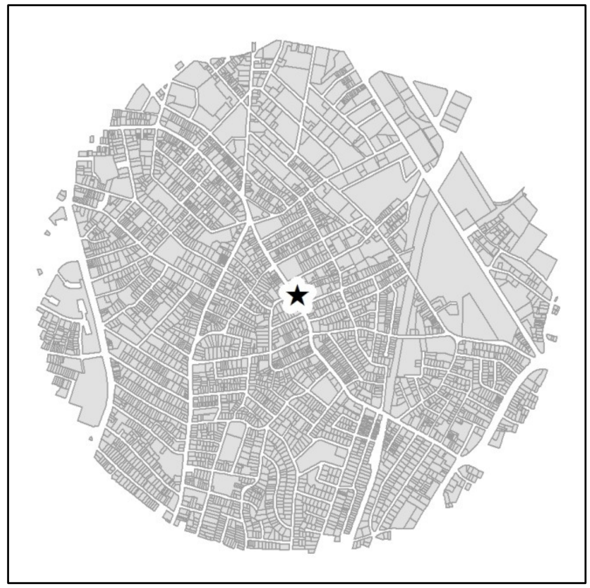

Figure 4: Parcels within 1km simple buffer, calculated by selecting and exporting parcels wholly within buffer.

Number of Categories

Now that parcels have been selected, it is possible to perform the analysis using the different land use categorization schemes outlined in Chapter III. Below are four maps showing the parcels chosen via the simple buffer with either 3, 4, 5, or 7 land use categories.[89]

[89] Note that the 7 land use category scheme includes vacant parcels, which are then excluded from the analysis. Therefore only 6 land use categories are actually being used in the calculations.

Figure 5: Selected 1km simple buffer parcels classified by 3 land use categories.

Figure 6: Selected 1km simple buffer parcels classified by 4 land use categories

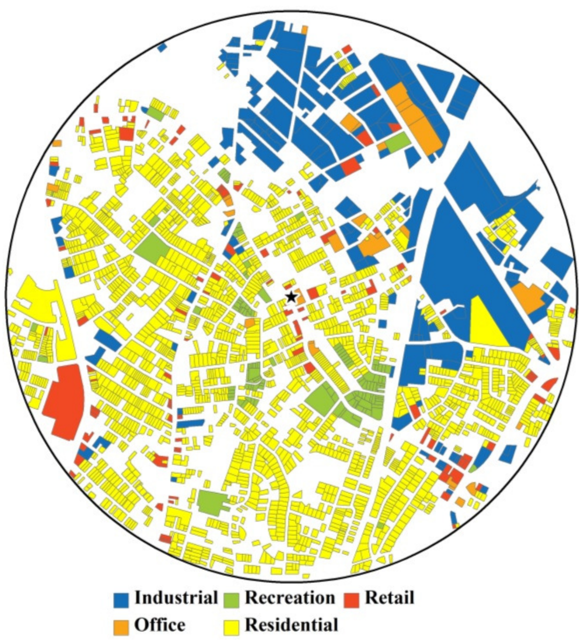

Figure 7: Selected 1km simple buffer parcels classified by 5 land use categories

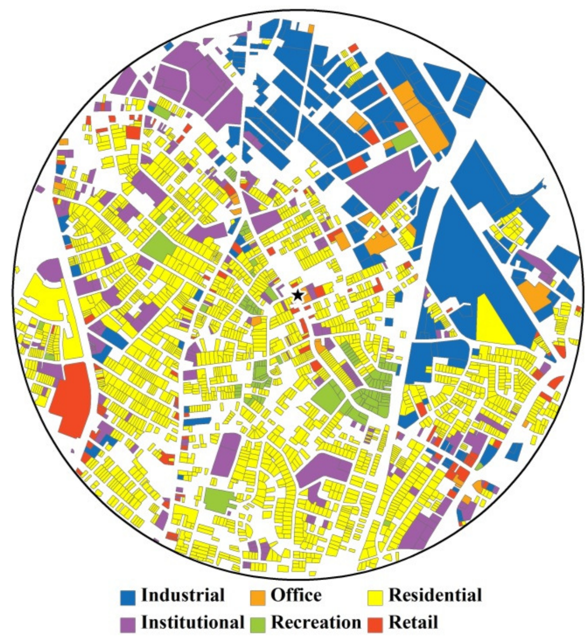

Figure 8: Selected 1km simple buffer parcels classified by 7 land use categories

Network Buffer

The network buffer, shown in the maps below, collects a significantly smaller number of the parcels near the point of interest. Using the network buffer tool in ArcGIS requires a street network shapefile, preferably as up-to-date as possible. The network buffer tool can also weigh particular streets as undesirable, allowing for a more nuanced result. This requires some local knowledge as to where major barriers to the subject of interest may lie.[90] For this demonstration, two high-traffic roads (both divided and greater than four lanes wide) and one elevated rail line were identified as potential barriers.

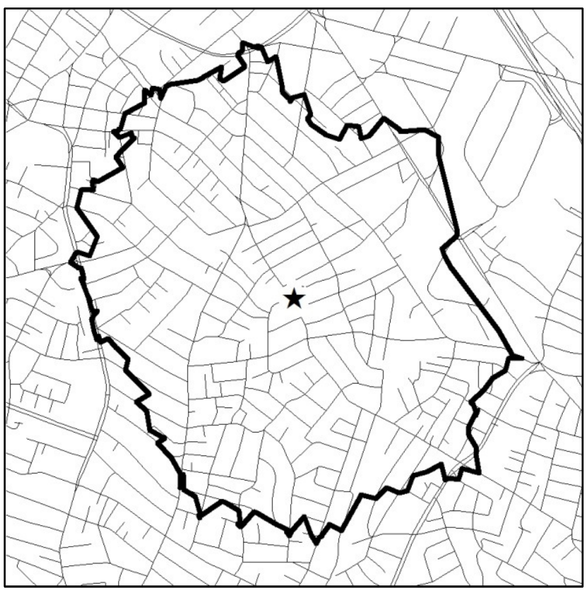

Figure 9: A 1km network-based buffer.

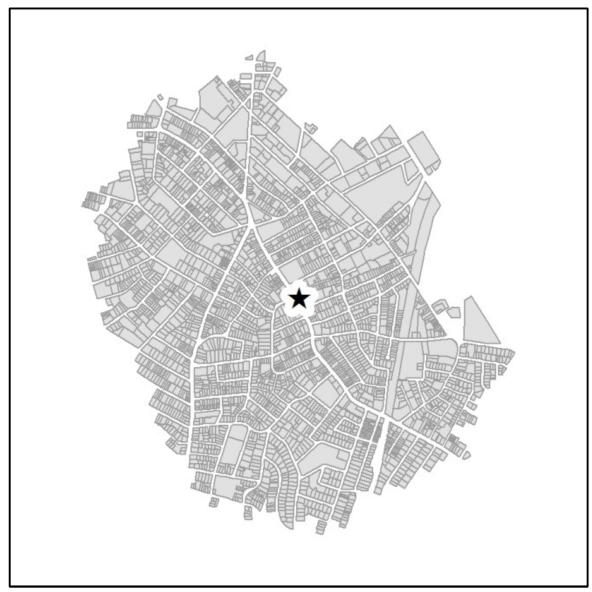

Figure 10: Parcels counted within 1km network-based buffer, calculated by selecting and exporting parcels wholly within buffer.

Number of Categories

Now that parcels have been selected, it is possible to perform the analysis using the different land use categorization schemes outlined in Chapter III.

---

[90] For example, if a study were looking at walking, knowing if a particular street was difficult to cross or whether there were an elevated highway that was particularly dangerous to walk under would be important to include in the analysis. These barriers are also important to note when not looking at transportation, as neighborhoods can be defined by their edges as much as by their centers. (Lynch, The Image of the City, 1960)

Figure 11: Selected 1km network buffer parcels classified by 3 land use categories.

Figure 12: Selected 1km network buffer parcels classified by 4 land use categories

Figure 13: Selected 1km network buffer parcels classified by 5 land use categories

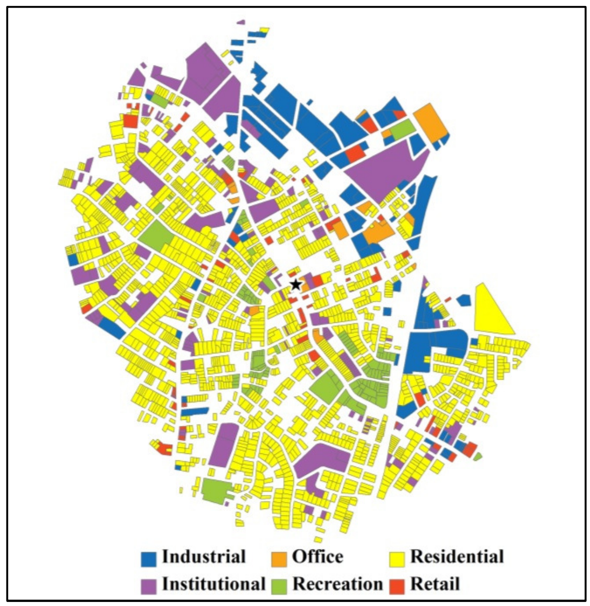

Figure 14: Selected 1km network buffer parcels classified by 7 land use categories

Line-Based Network Buffer

The line-based network buffer is, as detailed in Chapter III, a refinement of the general network buffer. This is evident as the selected parcels in this section are very similar to the selected parcels in the network-based buffer section above, with two exceptions: 1) the line-based network buffer does not include many infill lots which are inaccessible on the formal street network, and 2) the line-based network buffer does include some edge parcels that the network-based buffer does not, as those parcels may not fall wholly within the network-based buffer zone. This would indicate that these parcels are inaccessible to an individual walking along this street network, most likely due to being a back lot.

Excluding these physically inaccessible parcels is an important refinement. As land use grain is first and foremost a proxy for the *experience* of variety in the urban environment, the experience of the individual on the street is of utmost importance. If the individual does not experience these back lots, it is useful to exclude back lots from the analysis.

For regions wholly made up of small parcels and a well-connected street network (as most of this study area happens to be), the second difference (the inclusion of more edge parcels in certain locations) is negligible at worst and beneficial at best. If an individual can reasonably walk up to the front door of a business or institution, for instance, what does it matter to her if she cannot walk reasonably walk to the back of the lot? However, this difference is important for regions made up of large lots, big box stores, or other automobile-oriented environments. This issue is made apparent in the easternmost portion of the selected parcels below. Large shopping center lots have been included in this analysis because a portion of their parking lots were accessible by walking, despite the possibility that the door to the establishment may be much farther away or entirely inaccessible.[91]

[91] Out of the large lots in the eastern portion of the line-based network buffer parcels in this dataset, the closest a storefront gets to a street is approximately 250 meters. This would clearly exclude some of these parcels from a true look at the area within reasonable distance.

Figure 15: A 1km line-based network buffer.

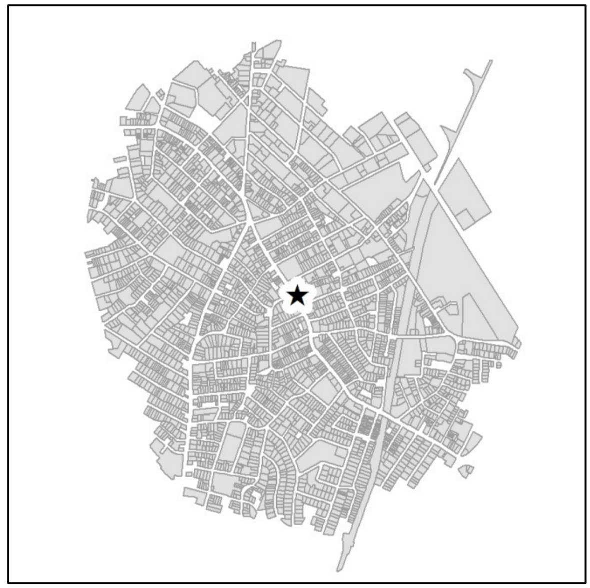

Figure 16: Parcels within 1km line-based network buffer, calculated by selecting and exporting parcels intersecting with buffer.

Number of Categories

It is useful to compare the results from this buffering technique with the results from the purely network-based buffer. Below are the two populations of selected parcels side-by-side for visual comparsion:

Figure 17: A side-by-side comparison of the network-based buffer and the line-based buffer results

Note that some large parcels in the eastern portion of the study area are included in this buffer which were dropped by the network-based buffer. This is because these parcels did not fall wholly within the network-based buffer, but did touch accessible streets in the line-based network buffer. Also note that the network-based buffer produced a more compact result, with fewer fringes and all internal parcels selected, while the line-based network buffer excluded some internal parcels and included extra edge parcels.

Now that parcels have been selected, it is possible to perform the analysis using the different land use categorization schemes outlined in Chapter III. Below are four maps showing the parcels chosen via the simple buffer with either 3, 4, 5, or 7 land use categories.

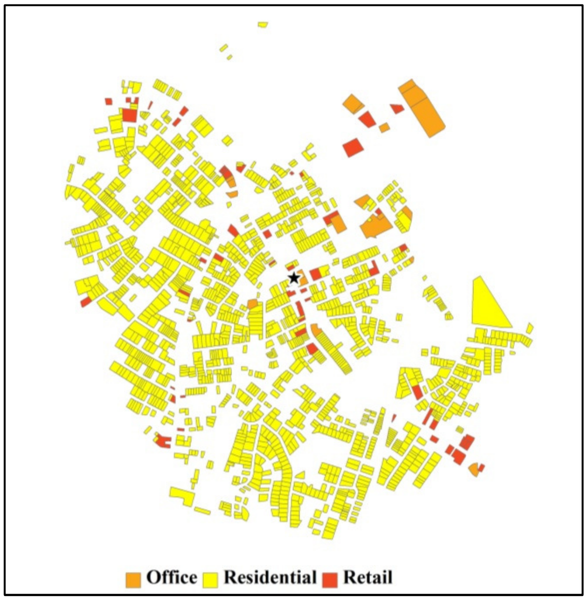

Figure 18: Selected 1km line-based network buffer parcels classified by 3 land use categories.

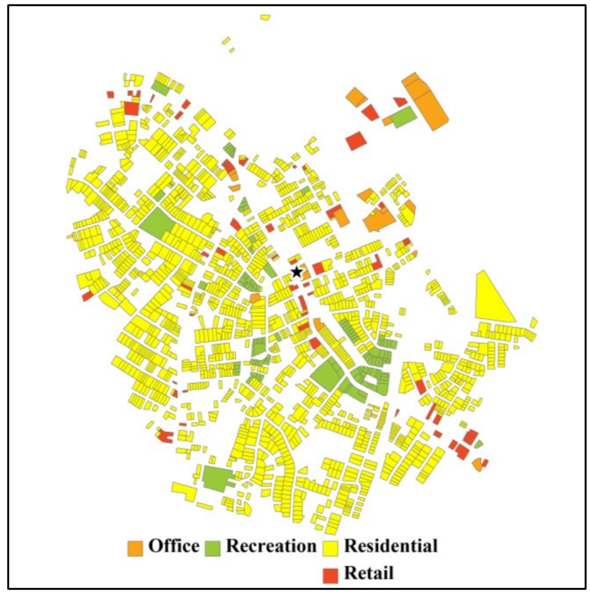

Figure 19: Selected 1km line-based network buffer parcels classified by 4 land use categories

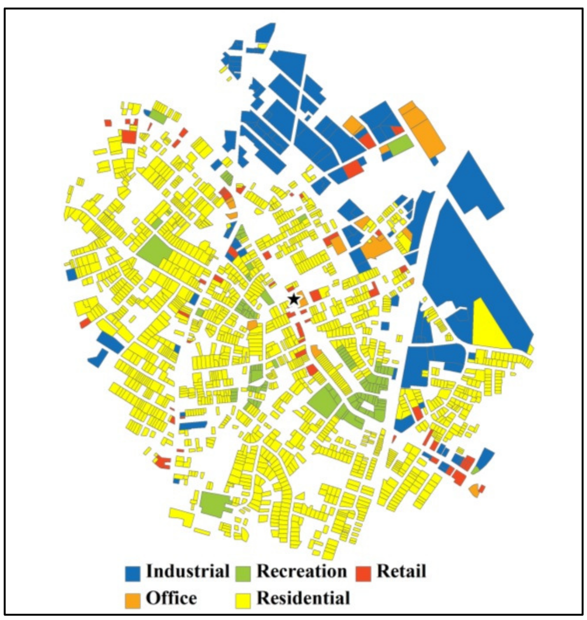

Figure 20: Selected 1km line-based network buffer parcels classified by 5 land use categories

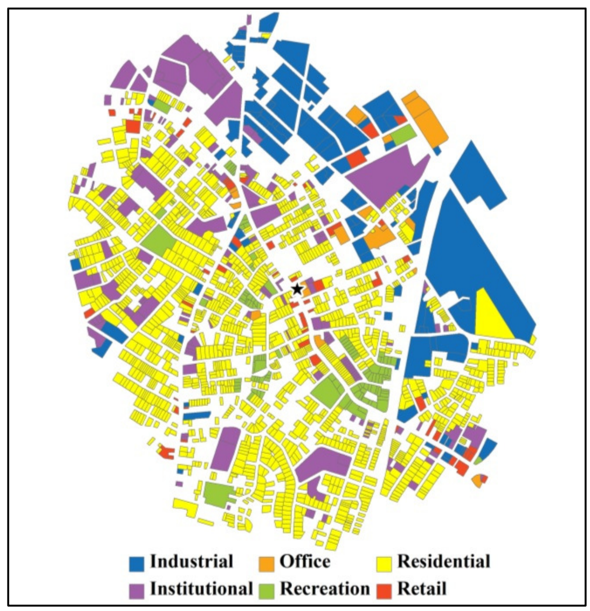

Figure 21: Selected 1km line-based network buffer parcels classified by 7 land use categories

These graphics all considered together, it is clear that the area of capture technique and the land use categorization scheme both change the actual data which is used for analysis. Accurate land use grain analysis requires on the part of the researcher a knowledge of which pieces of the study area are actually being included in the final measurements, in order to better inform analysis and potential decision-making.

Measurement Results & Analysis

The entropy-based equation and the HHI can both be calculated in a spreadsheet application using the information in the above maps. As was stated in Chapter III, the HHI is calculated using the proportion of land taken up by each individual land use, as opposed to the percentage, to ensure a result on a 0–1 scale. The jobs-to-housing ratio requires different data from the entropy-based equation and the HHI. For this analysis, jobs data is gleaned from the *Doing Business As* database of licensed businesses in Boston, from the Boston Office of the City Clerk.[92] Jobs numbers have been checked against the 2009 County Business Patterns data at the zip code level,[93] with total number of employees estimated using the proportional representation of each zip code in the appropriate buffered area.[94] The number of housed workers is estimated using the methods in the E.P.A.'s Smart Growth Index indicators manual entry on the jobs-to-housing ratio[95], by estimating two housed workers per residential parcel. This estimate differs from the E.P.A. standard of 1.4 housed workers per parcel, and is based on local knowledge on the part of the author.

DEMONSTRATION ANALYSIS

Below is a series of tables summarizing results from the demonstration. These tables list the results for each calculation and each relevant buffer style and land use categorization scheme, along with a mean measurement for each buffer style.

---

[92] (City of Boston Office of the City Clerk, 2012)
[93] (US Census Bureau, 2009)
[94] For a tutorial on proportional representation using Census boundaries, see (Yale University Library).
[95] (United States Environmental Protection Agency, 2003)

Table 3: Entropy-based equation results

Simple Buffer	
3 Land Use Categories	0.52
4 Land Use Categories	0.49
5 Land Use Categories	0.62
7 Land Use Categories	0.70
Mean Given Simple Buffer	**0.59**
Network Buffer	
3 Land Use Categories	0.46
4 Land Use Categories	0.46
5 Land Use Categories	0.53
7 Land Use Categories	0.61
Mean Given Network Buffer	**0.52**
Line-based Network Buffer	
3 Land Use Categories	0.49
4 Land Use Categories	0.48
5 Land Use Categories	0.61
7 Land Use Categories	0.70
Mean Given Line-based Buffer	**0.57**
Mean Score for all Calculations	**0.56**

Table 4: Herfindahl-Hirschman Index Results

Simple Buffer	
3 Land Use Categories	0.13
4 Land Use Categories	0.13
5 Land Use Categories	0.17
7 Land Use Categories	0.19
Mean Given Simple Buffer	**0.16**
Network Buffer	
3 Land Use Categories	0.12
4 Land Use Categories	0.12
5 Land Use Categories	0.13
7 Land Use Categories	0.14
Mean Given Network Buffer	**0.13**
Line-based Network Buffer	
3 Land Use Categories	0.14
4 Land Use Categories	0.14
5 Land Use Categories	0.17
7 Land Use Categories	0.19
Mean Given Line-based Buffer	**0.16**
Mean Score for all Calculations	**0.15**

The entropy-based equation provides a middle-of-the-road value of land use grain in this particular instance. A mean score of **0.56** indicates that this particular neighborhood is not entirely composed of an even mixture of land uses at the parcel level, as a score closer to 1 would indicate. However, this result is closer to 1 than to 0, which would indicate that this neighborhood is more mixed than homogenous.

As for the measurement contingencies of area of capture and land use categorization scheme, the entropy-based equation appears to be change more across land use categorization. This shows that the land use categorization scheme used is an important consideration when using the entropy-based equation. The network-based buffer does return a lower value of land use grain than the other two buffering types, but the result is not large.

The HHI runs on an opposite scale to the entropy-based equation, where a score of 0 indicates an entirely mixed use neighborhood and a score of 1 indicates an entirely homogenous land use. This is because the HHI measures concentration, as opposed to measuring mixture. However, the HHI results are not linear, as was described in Chapter III. A neighborhood with a moderate concentration of land uses will score between 0.1 and 0.18. As seen in Table 4, the HHI provides a mean land use grain score of **0.15**. This would indicate that this neighborhood is moderately concentrated. This result runs somewhat counter to intuition, which would tell us that a score of 0.15 would indicate a very high level of mixture among land uses. This intuitive response is important to keep in mind when reporting the HHI—particularly in a professional planning context—because it can be easily misinterpreted.

The variation given different measurement contingencies is somewhat minor. Contrary to the entropy-based equation, the network-based buffer here shows greater mixture rather than less mixture. Because both the entropy-based equation and the HHI are at their core an addition of shares with different weighting standards, the network-based buffer must be returning lower individual shares for a variety of land uses. Across land use categorization schemes, the HHI increases for each land use category added as expected. This property of the HHI is important to keep in mind when deciding on a land use categorization scheme. The HHI will always increase for every land use category added, in proportion to the share of total land taken up by that particular land use category.

Table 5: Jobs-to-Housing Ratio Results

Simple Buffer	2.96
Network Buffer	0.91
Line-based Network Buffer	1.09
Mean Score for all Calculations	**1.65**

The jobs-to-housing ratio results are reported in Table 5 to the left. There are no values for differing land use categorization schemes, as the jobs-to-housing ratio does not rely upon land use categories. The jobs-to-housing ratio reports a mean score of **1.65**, indicating approximately 1.65 jobs for every member of the workforce in the community. However, the results for differing buffer techniques range widely, from slightly more workers than jobs to nearly three times as many jobs as workers. This wide range of scores is likely due to the proximity of industrial neighborhoods in the easternmost portion of the study area, where street connectivity is low (see Figures 8, 14, and 20 in the previous section). This result is indicative of flaws in the simple buffer method of capturing a study area for this particular location. In general, the less well-connected street networks are in the area of capture, the less accurate the simple buffer will be as a proxy for walkable area.[96]

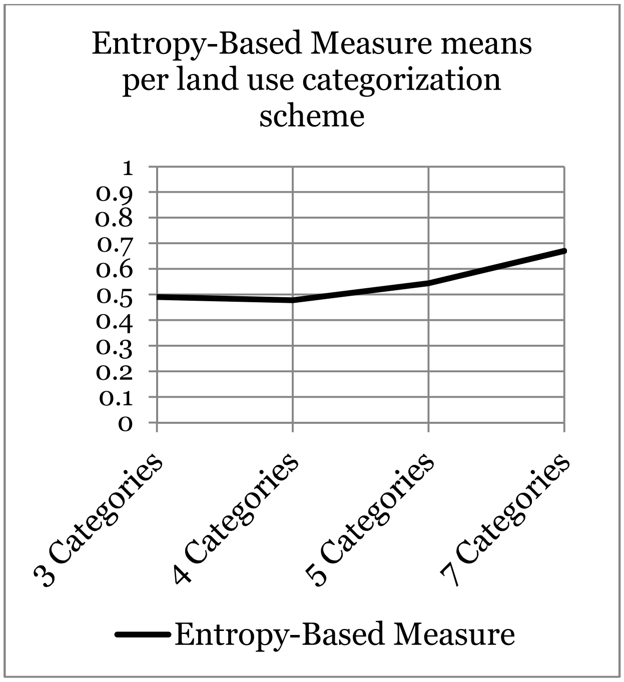

Figure 22: Entropy-based equation per categorization scheme

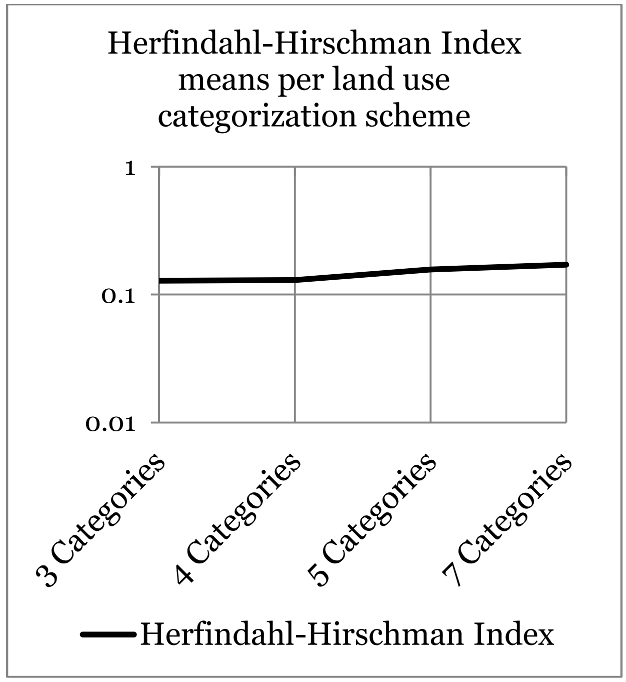

Figure 23: HHI per categorization scheme

From these results, the entropy-based equation would appear to be the most intuitive measure of land use grain, while providing a somewhat consistent result regardless of land use categorization scheme or buffering technique. The HHI may be more useful to indicate extreme levels of mixture, such as a situation where only one or two land uses make up the entire area of capture. The HHI is also susceptible to greater manipulation through land use categorization. For instance, if residential and commercial properties were further summarized into distinct categories, the HHI would decrease significantly. Likewise, if all commercial properties in this study were categorized as “commercial” instead of being broken into “retail” and “office,” HHI would increase. The entropy-based equation does not suffer from this potential problem as strongly.

96 (Oliver, Schuurman, & Hall, 2007)

The jobs-to-housing ratio does not represent land use grain itself as an attribute of the urban fabric, but is rather a proxy for one of the impacts of land use grain: access to work within a certain distance of the home. As a proxy measure, the jobs-to-housing ratio appears to be stable, given known employment levels, with the caveat that particularly large employers or dense residential developments may have an undue effect on the final result. While the jobs-to-housing ratio is useful in transportation studies as a measure of potential work accessibility, it does not give a picture of the grain of land uses itself.

Chapter V: Research Questions Revisited/Conclusions

Having looked at the meaning and potential implications of land use grain, various land use grain measurement standards, and samples of various land use grain measurements in a single context, the gathered information must be assembled into a cohesive whole. Using the information collected and evaluated in this thesis, the research questions from Chapter I will be individually addressed as best as possible. Limitations of this work and options for future research will be outlined after addressing the research questions, to allow the formulation of next steps.

Research Question 1

> *What is meant by the term "land use grain" in academic literature, and how does this meaning inform measurement of land use grain?*

RQ1 is addressed in Chapter II, where the concept of land use grain as a component of urban grain as a whole was discussed, and the importance of land use grain to aspects of the urban fabric was demonstrated using contemporary literature. Grain itself—the distribution of some aspect of the urban fabric across the city—is a fundamental attribute of all aspects of the urban fabric, of which the grain of land uses and activities is an important piece.

Land use grain refers to the distribution of multiple land uses in the urban fabric. This idea is intuitive. In layman's terms, an area that appears to be walkable and detailed is likely to have a fine mixture of multiple land uses over small areas, whereas an area that appears to be auto-oriented or homogenous is likely to have a coarse mixture of land uses.

The different possible mixtures of land uses have impacts on how people interact with the urban fabric, altering transportation, recreation, and health-related outcomes. Research has shown that the distribution of land uses within walking distance has an impact on transportation mode choice, daily activity levels, and a broad variety of other factors composing the experience of the urban fabric as a whole.

Because land uses are generally not very evenly distributed in neighborhoods in the U.S., and because research has shown that coarse land use grains can result in negative transportation, recreation, and health-related outcomes, it is of importance to better understand land use grain quantitatively, in order to measure changes in land use grain over time.

Due to the potential effects on particular behaviors, land use grain is most often studied as a predictor for some other behavior or attitude rather than as an independent phenomenon. As seen in Chapter III, this encourages researchers to measure land use grain in very particular manners as the concept pertains to their specific interests. For example, public health-related research may look at the distribution of pollution point sources or active recreation spaces in an environment, whereas transportation-related research is more likely to look at the distribution of workplaces relative to housing and retail. These specific measurements have shown themselves to be useful, but often these different academic disciplines obscure land use grain by focusing heavily on particular land use categories of interest to the exclusion of others. A more unified concept of land use grain may help connect research across subfields, while providing a more coherent view of land use grain as an independent phenomenon.

RESEARCH QUESTION 2

How do standard measurement techniques and their results compare with one another when performed in context?

RQ2 is addressed in Chapter III, which looked at several land use grain measurement techniques and discussed considerations which may alter their results.

Measuring land use grain is a recent development in planning and in public health, resulting in several different metrics in use in contemporary literature. This thesis looks closely at three of the most commonly used metrics: the entropy-based equation, the jobs-to-housing ratio, and the Herfindahl-Hirschman Index (HHI). The entropy-based equation and the HHI both come from fields outside of planning and public health, but have been applied to measure the distribution or concentration of land uses in particular environments. These two metrics both aim to uncover the distribution of land uses in a single environment by comparing the proportion of each individual land use to the proportions of the other land uses present in the area of concern. The jobs-to-housing ratio is used by transportation planners and smart growth advocates as a stand-in for land use grain in particular situations, and may indicate a stronger relationship with non-motorized transportation than other measures. The jobs-to-housing ratio does not attempt to uncover a measure of land use grain itself.

The entropy-based equation uses the mean of the natural logs of the shares of land taken up by each individual land use category studied. This results in a scale of 0 to 1, where 1 indicates a high level of mixture and 0 represents a single land use. The HHI uses the square of the proportion of space dedicated to each land use category, and results in a similar scale of 0 to 1. For the HHI, a 0 indicates a high level of mixture and 1 represents a single land use. The HHI scores increase exponentially, resulting in a moderate land use mixture being represented by a relatively low score range of 0.1 to 0.18. The jobs-to-housing covers the range of real numbers from 0 to infinity, with a score of 1 indicating exactly the same number of jobs as house workers.

These different metrics are subject to certain contingencies in data collection which may have a large effect on their outcomes. The two most important contingencies are the area of capture and the land use categorization scheme. Variations in these may be hidden in indices which include a land use grain metric, despite their potential to affect measurement results.

The area of capture is largely a question of which geographic areas produce an effect in the individual(s) being studied. Research has tended to focus on capturing the population of parcels to which the individual(s) can theoretically walk, bicycle, or otherwise physically access. Capturing this parcel population requires knowledge of how far individuals in the sample are willing to walk or bicycle. Deeper questions of how far individuals are willing to walk or bicycle for different land use categories, or whether out-of-reach parcels may impact land use grain-related decisions, were not addressed, but may require closer attention.

The land use categorization scheme used often differs by the purpose of the study. While using particular land use categorization schemes based on differing purposes may be more likely to provide statistically significant results than relying on a holistic land use categorization scheme, this practice may hinder transferability of studies across disciplines and may clouds true relationships between land use grain as a whole and other expected outcomes.

Research Question 3

How do standard measurement techniques and their results compare with one another when performed in context?

RQ3 is addressed in Chapter IV, which demonstrated the three focus metrics (the entropy-based equation, the jobs-to-housing ratio, and the HHI) in a single context, using a variety of buffering techniques and land use categorization schemes.

Chapter IV showed that the entropy-based equation holds somewhat consistent across buffer types and land use categorization schemes for this demonstration. The entropy-based equation also provides an intuitive and descriptive value of land use grain compared to other measurements. The HHI provides a consistent result as well, but the score is unintuitive because of the metric's scale. The HHI is dependent upon the number of land use categories measured. The HHI may work best at showing extreme levels of concentration among land uses, but does not intuitively distinguish between highly mixed and somewhat mixed environments. The jobs-to-housing ratio can be influenced by individual businesses with large numbers of employees, and is not particularly helpful if the businesses of interest require a different workforce than that which lives nearby. The jobs-to-housing ratio does not rely at all on land use data or land use categorization on the part of the researcher. This removes potential error by removing some human effort in the execution of the metric, and this removes the burden of finding parcel-level land use data. However, this also obscures potential effects of the distribution of employment among different land uses or employment categories.

The results indicate that the entropy-based equation is the most intuitive measure of land use grain across buffer types and land use categorization schemes. Results also imply that the entropy-based equation tells a more complete story about the mixture of land uses in a particular urban environment than does the HHI, because the HHI only distinguishes between extreme land use concentration and not-extreme land use concentration. The HHI may be useful for demonstrating an outlier in the data by sorting extremely concentrated environments from mixed and somewhat mixed environments. The jobs-to-housing ratio is not a measure of land use grain, but is often a proxy for the assumed relationship between land use grain and transportation mode choice. The demonstration shows that the jobs-to-housing ratio is generally stable, when outlier employers with a large number of employees are removed. But the influence of outlier businesses and the lack of attention to other aspects of the grain of land uses in the urban fabric together prevent the jobs-to-housing ratio from being a convincing proxy for land use grain itself.

The conclusion, then, is that the entropy-based equation of land use grain is the most useful metric of land use grain in analyzing land use grain itself, given a broad view of land use categorization. The HHI and the jobs-to-housing ratio may be useful in particular instances, but the entropy-based equation is more likely to be accurate regardless of the buffering techniques available, the land use categorizations employed, or the type of urban environment being studied. The entropy-based equation is also more likely to distinguish between a wide variety of contexts in a numerically-useful manner.

Limitations

As was stated in Chapter I, this work measures land use grain on its own merits, instead of measuring land use grain as an indicator in a larger index. While land use grain is an interesting phenomenon on its own, it is most useful when paired with other data for research. Primary examples of this include the various transportation and public health related studies discussed in Chapter III. However, this thesis did not discuss how land use grain should be included in larger indices, or how land use grain should be treated in relation to other variables. For example, some research has chosen to classify land use grain scores into a categorical variable, and other research has used land use grain scores as a response variable. When included in an index, different research has used different weights for the land use grain variable. The decision of how to use land use grain scores in a particular study is to be made at the researcher's discretion, given the topic and architecture of the study.

As a corollary to the above, this thesis did not set out to define which measurement practice is *best*. Changing measurement practices changes the statistical outcome of the research being conducted. Therefore, measurement practices must be made based on the best information possible. This thesis provides information, rather than making a judgment as to which method should be used in every following work. However, this thesis does recommend that the entropy-based equation be employed in large, cross-context studies or in general work. This is because the entropy-based equation is intuitive and sensitive to a variety of contexts.

This thesis focuses on the holistic measure of land use grain, rather than on the impacts of particular mixtures on particular desired outcomes. While targeted studies focusing only on retail or recreational land may have higher associations with particular public health- or transportation-related outcomes, they do not capture land use grain itself as a whole. More research on individual land use categories, other categorization schemes, and their relationship with particular outcomes is required before definite conclusions can be drawn.

Future Research

Future research analyzing the variance across different measurement standards could be performed to decide more precisely whether using different land use categorization practices or measurement standards provides consistently distinct results, given standards in other areas of measurement. This analysis was outside the scope of this study, but would be useful to researchers aiming to justify the use of different standards more quantitatively.

Future research should also look more closely at sub-categorizations of land use, including multi-family vs. single-family and particular forms of industry or retail trade. For example, in this particular study area there exists a range of residence types, from multi-family apartment buildings to single-family detached homes. There is also a range of retail, from used car lots to upscale boutiques. Current measures of land use grain would have a difficult time providing useful results at this level of detail (the HHI, for instance, would entirely break down), but these specific differences may create more lasting impressions on individuals and on decision-making within the urban fabric.

BIBLIOGRAPHY

Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (1993). *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall.

Badger, E. (2012, March 30). The Simple Math That Can Save Cities From Bankruptcy. *The Atlantic Cities*.

Berrigan, D., Pickle, L., & Dill, J. (2010). Associations between street connectivity and active transportation. *International Journal of Health Geographics, 9*(20).

Bhadeshia, H. (n.d.). Thermodynamics and Phase Diagrams, Lecture 3: Models of Solutions. *Masters of Philosophy, Materials Modeling, Course MP4*. United Kingdom. Retrieved from http://www.msm.cam.ac.uk/phase-trans/mphil/MP4-3.pdf

Boarnet, M. G., & Sarmiento, S. (1998). Can Land-use Policy Really Affect Travel Behavior? A Study of the Link between Non-work Travel and Land-use Characteristics. *Urban Studies Journal Limited*, 1155-1169.

Børrestad, L., Anderson, L., & Bere, E. (2011, February 1). Seasonal and socio-demographic determinants of school commuting. *Preventative Medicine, 52*(2), 133-135.

Brownson, R., Hoehner, C., Day, K., Forsyth, A., & Sallis, J. (2009). Measuring the Built Environment for Physical Activity: State of the Science. *American Journal of Preventative Medicine, 36*(4S), S99-S123.

Buehler, R., & Pucher, J. (2012). Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. *Transportation*, 409-432.

Caliper. (2012, June). *Caliper - Mapping software, GIS, and transportation planning software*. Retrieved from Caliper: http://www.caliper.com/

Cervero, R. (1988). *America's Suburban Centers: A Study of the Land Use-Transportation Link*. Washington: US Department of Transportation.

Cervero, R. (1989). Land-use Mixing and Suburban Mobility. *Transportation Quarterly*, 429-446.

Cervero, R., & Kockelman, K. (1997). Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research*, 199-219.

Christian, H. E., Bull, F. C., Middleton, N. J., Knuiman, M. W., Divitini, M. L., Hooper, P., . . . Giles-Corti, B. (2011). How important is the land use mix measure in understanding walking behavior? Results from the RESIDE study. *International Journal of Behavioral Nutrition and Physical Activity, 8*(55).

City of Boston Office of the City Clerk. (2012). *City Clerk | City of Boston*. Retrieved from City of Boston.gov: http://www.cityofboston.gov/cityclerk/

Design for Health. (2012, January). NEAT-GIS Protocols. (5.1). (A. Forsyth, Ed.) Retrieved May 2012, from http://www.designforhealth.net/GISprotocols.html

DiNardo, J. (2008). Natural Experiments and Quasi-Natural Experiments. *The New Palgrave Dictionary of Economics, Second*. (S. N. Durlauf, & L. E. Blume, Eds.) Palgrave Macmillan. doi:10.1057/9780230226203.1162

Duany, A. (2002). Introduction to the Special Issue: The Transect. *Journal of Urban Design, 7*(3), 251-260.

Duany, A., & Plater-Zyberk, E. (1995). Neighborhoods and Suburbs. *Design Quarterly*.

Duany, A., & Talen, E. (2001-2002). Making the Good Easy: The Smart Code Alternative. *Fordham Urban Law Journal, XXIX*, 1445-1468.

Duncan, M. J., Winkler, E., Sugiyama, T., Cerin, E., duToit, L., Leslie, E., & Owen, N. (2010). Relationships of Land Use Mix with Walking for Transport: Do Land Uses and Geographical Scale Matter? *Journal of Urban Health, 87*(5), 782-796.

Durlauf, S., & Blume, L. (Eds.). (2008). *The New Palgrave Dictionary of Economics* (Second ed.). New Palgrave.

Enos, R. (n.d.). What tearing down public housing projects teaches us about the effect of racial threat on political participation. *Unpublished*. Retrieved from http://ryandenos.com/papers/chicago_threat.pdf

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, 265-294.

Ewing, R., Pendall, R., & Chen, D. (2003). Measuring Sprawl and Its Transportation Impacts. *Transportation Research Record*, 175-183.

Fainstein, S. (2011). *The Just City*. Ithaca, NY: Cornell University Press.

Frank, L. D., & Pivo, G. (1994). An analysis of relationships between urban form (density, mix, and jobs: housing balance) and travel behavior (mode choice, trip generation, trip length, and travel time). 301 pages. University of Washington.

Frank, L. D., Saelens, B. E., Powell, K. E., & Chapman, J. E. (2007). Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? *Social Science & Medicine*, 1898-1914.

Frank, L. D., Schmid, T. L., Sallis, J. F., Chapman, J., & Saelens, B. E. (2005). Linking Objectively Measured Physical Activity with Objectively Measured Urban Form. *American Journal of Preventive Medicine*, 117-125.

Guo, Z., Wilson, N. H., & Rahbee, A. (2007). Impact of Weather on Transit Ridership in Chicago, Illinois. *Transportation Research Board of the National Academies, 2034*, 3-10.

Harvard Geospatial Library. (2011). Parcels. Cambridge, MA. Retrieved March 2012

Hettinga, F., De Koning, J., Vrijer, A., Wust, R., Daanen, H., & Foster, C. (2007). The effect of ambient temperature on gross-efficiency in cycling. *European Journal of Applied Physiology*, 465-471.

Ingram, G. K., Carbonell, A., Hong, Y.-H., & Flint, A. (2009). *Smart Growth Policies: An Evaluation of Programs and Outcomes*. Cambridge: Lincoln Institute of Land Policy.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Random House, Inc.

Key, V. (1949). *Southern politics in state and nation*. New York: Knopf.

Lawlor, M. J. (2011). Gaining ground in the final frontier: surveying legal issues raised by New England's form-based codes. *The Urban Lawyer*, 839.

Lu, H. Q., & Nimbole, P. (2008). *Intro to TransCAD GIS*. Model Research and Development Unit. Raleigh, NC: NCDOT. Retrieved from http://www.ncdot.gov/doh/preconstruct/tpb/mrd/training/PDF/IntroAll.pdf

Lynch, K. (1960). *The Image of the City*. Cambridge: The MIT Press.

Lynch, K. (1981). *A Theory of Good City Form*. Cambridge: MIT Press.

Mannering, F. L., Washburn, S. S., & Kilareski, W. P. (2009). *Principles of Highway Engineering and Traffic Analysis* (4th ed.). Wiley.

Massachusetts Department of Revenue. (2012, March). Property Type Classification Codes. Boston, MA. Retrieved from http://www.mass.gov/dor/docs/dls/bla/classificationcodebook.pdf

Matthews, J. W., & Turnbull, G. K. (2007). Neighborhood Street Layout and Property Value: The Interaction of Accessibility and Land Use Mix. *Journal of Real Estate and Financial Economics*, 111-141.

New Orleans Regional Planning Commission. (2010). *Analysis of Regional Vision Software for Metropolitan New Orleans*. New Orleans, LA.

Oakes, J. M., Forsyth, A., & Schmitz, K. (2007). The effects of neighborhood density and street connectivity on walking behavior: the Twin Cities walking study. *Epidemiologic Perspectives & Innovations, 4*(16).

Oliver, L. N., Schuurman, N., & Hall, A. W. (2007). Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. *International Journal of Health Geographics*.

Oliver, L., Schuurman, N., Hall, A., & Hayes, M. (2011). Assessing the influence of the built environment on physical activity for utility and recreation in suburban metro Vancouver. *BioMed Central Public Health*.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. Norwich, England: Geo Books.

Pucher, J., Garrard, J., & Greaves, S. (2011). Cycling down under: a comparative analysis of bicycling trends and policies in Sydney and Melbourne. *Journal of Transport Geography*, 332-345.

Pushkarev, B., & Zupan, J. M. (1977). *Public transportation and land use policy*. Bloomington, Indiana: Indiana University Press.

Pushkarev, B., Zupan, J. M., & Cumella, R. (1982). *Urban Rail in America: An Exploration of Criteria for Fixed-Guideway Transit*. Bloomington, Indiana: Indiana University Press.

Schiller, P., Bruun, E., & Kenworthy, J. (2010). *An Introduction to Sustainable Transportation*. Washington: Earthscan LLC.

Steiner, R. L., Bejleri, I., Wheelock, J. L., Cahill, M., & Perez, B. O. (2008). Understanding and Mapping Institutional Impediments to Walking and Bicycling to School. *Transportation Research Record, 2074*, 3-11.

Sundquist, K., Eriksson, U., Kawakami, N., Skog, L., Ohlsson, H., & Arvidsson, D. (2011). Neighborhood walkability, physical activity, and walking behavior: The Swedish Neighborhood and Physical Activity (SNAP) Study. *Social Science & Medicine*, 1266-1273.

United States Department of Justice. (n.d.). *USDOJ Antitrust Division: Herfindahl-Hirschman Index*. Retrieved May 5, 2012, from Herfindahl-Hirschman Index: http://www.justice.gov/atr/public/guidelines/hhi.html

United States Environmental Protection Agency. (2003). *EPA's Smarth Growth INDEX In 20 Pilot Communities*. Development, Community, and Environment Division. Washington: US EPA. Retrieved from http://www.epa.gov/dced/pdf/Final_screen.pdf

United States Environmental Protection Agency. (2012, May 03). *Smart Growth | US EPA*. Retrieved from US EPA: http://www.epa.gov/smartgrowth/

US Census Bureau. (2009). Table CB0900A1: 2009 County Business Patterns: Geography Area Series: County Business Patterns. Washington, DC. Retrieved Spring 2012, from http://factfinder2.census.gov/

US Census Bureau. (2010). TIGERLine Shapefile. Boston, MA, USA.

USDOT Federal Highway Administration. (2011, April 6). *Financing Federal-Aid Programs - Highway Trust Fund*. Retrieved April 8, 2012, from Federal Highway Administration Website: http://www.fhwa.dot.gov/reports/fifahiwy/fifahi05.htm

Walk Score. (n.d.). *Apartments and Rentals - Find a Walkable Place to Live*. Retrieved 2012, from Walk Score: http://www.walkscore.com/apartments/

Wehrmeyer, S. (n.d.). *Mapnificent - Dynamic Public Transport Travel Time Maps*. Retrieved 2012, from Mapnificent: http://www.mapnificent.net

Winters, M., Friesen, M., Koehoorn, M., & Teschke, K. (2007, January). Utilitarian Bicycling: A Multilevel Analysis of Climate and Personal Influences. *American Journal of Preventative Medicine, 32*(1), 52-58.

Wood, L., Frank, L., & Giles-Corti, B. (2010). Sense of community and its relationship with walking and neighborhood design. *Social Science & Medicine*, 1381-1390.

Yale University Library. (n.d.). *Overlay of Demographic Datasets with Unlike Boundaries*. Retrieved from http://www.library.yale.edu/MapColl/files/docs/Overlay%20of%20Demographic%20Datasets%20with%20Unlike%20Boundaries.pdf

Yang, Y., Roux, A. V., & Bingham, C. R. (2011). Variability and seasonality of active transportation in USA: evidence from the 2001 NHTS. *The International Journal of Behavioral Nutrition and Physical Activity, 8*(96).

APPENDIX I: MASSACHUSETTS PROPERTY TYPE CLASSIFICATION CODES

See the following pages for a copy of the full Massachusetts Property Type Classification Codes definitions. These classification codes are used in this thesis to create smaller land use categories according to the following scheme:

New Land Use Categorization	**MA Property Type Classification Codes**
Residential	101, 102, 104, 105, 106, 109, 111, 112, 113, 121, 125, 304, 309, 970
Retail	320, 321, 325, 326, 328, 329
Office	340, 341, 342, 343, 344, 353
Industrial	310, 311, 316, 317, 318, 319, 332, 334, 400, 402, 404, 405, 406, 407, 414, 422
Recreation	387, 995
Institutional	900, 902, 904, 905, 908, 985, 986
Vacant	130, 131, 132, 336, 337, 390, 391, 392, 396, 441, 442, 974, 976, 979

Note that in this reclassification, major property type classifications (indicated by the first digit in the code) were broken into the various land use categories defined in this study. For example, commercial properties (major classification 3) had to be classified into retail and office, while some commercial lands more accurately could be described as industrial in effect on the urban environment, and some commercial lands were vacant. These distinctions were judgment calls on the part of the author; categorization of land uses will always be a contentious subject, and this work does not attempt to make a statement on classification other than to say different categories are important for different fields. See below for a full description of the classification code system.[97]

[97] (Massachusetts Department of Revenue, 2012)

PROPERTY TYPE CLASSIFICATION CODES,
Non-arm's Length Codes and
Sales Report Spreadsheet Specifications

Prepared by the Bureau of Local Assessment
Revised March 2012

CHANGES

- New Non-arm's Length Code, "X"
- Expanded Non-arm's Length Code Descriptions
- Required *Property Sales Report Spreadsheet Specifications*

Changes are Highlighted in Yellow

INTRODUCTION

These Guidelines are intended to assist the Board of Assessors in determining the proper classification of property according to its use.

The coding structure has three digit level of detail. The first digit indicates a major classification. The second digit is a major division and the third digit is a subdivision, both within the major classification of property.

If the guidelines do not include a three digit code for a specific property use, the assessor should use the code that most appropriately identifies the property's use. For example, the assessors would use codes 321-326 to classify a retail condominium, based on the use of the property.

TABLE OF CONTENTS

CHANGES

New Non-Arm's Length Code "X",

Expanded Non-Arm's Length Code Descriptions

Required *Property Sales Report Spreadsheet Specifications*

MULTIPLE-USE PROPERTY

CODE 0

Real property used or held for use for more than one purpose, including parcels with multiple detached or attached buildings, are considered multiple-use property for classification purposes. Any necessary related land on a multiple-use property must be allocated among the classes of property within the building.

The first digit of multiple-use property is always a zero (0). The second and third digits are the major classification of the property represented. The digits following zero (0) are listed in the order of major importance.

Examples

Since the guidelines for coding multiple-use property are unique, several specific examples of how to identify such property with these codes are listed here. These are only examples and do not represent all possible multiple use codes.

013 Multiple-Use, primarily Residential

A building with a retail store on the first floor, apartments on the upper floors, and a major portion of the related land is reserved for tenant parking.

031 Multiple-Use, primarily Commercial

A building with retail use on the first floor, office space on the second and third floors, apartments on the fourth floor and a major portion of the related land is allocated for commercial use.

037 Multiple-Use, primarily Commercial with part of land designated under Chapter 61A use

A farm property with land and buildings predominantly used for commercial farming with part of land (at least 5 acres) designated horticulture/agricultural under Chapter 61A.

021 Multiple-Use, primarily Open Space

A single-family house with substantial acreage designated open space by the assessors.

RESIDENTIAL

CODE 1

M.G.L. Chapter 59 §2A: All real property used or held for human habitation containing one or more dwelling units including rooming houses with facilities assigned and used for living, sleeping, cooking and eating on a non-transient basis, and including a bed and breakfast home with no more than three rooms for rent. Such property includes accessory land, buildings or improvements incidental to such habitation and used exclusively by the residents of the property or their guests. Such property shall include: (i) land that is situated in a residential zone and has been subdivided into residential lots, and (ii) land used for the purpose of a manufactured housing community, as defined in Chapter 140, §32F. Such property shall not include a hotel or motel.

Incidental accessory land, buildings or improvements would include garages, sheds, in-ground swimming pools, tennis courts, etc. Non-incidental accessory land, classified and coded differently, would include mixed use properties, such as a variety store, machine shop, etc. on a residential parcel.

10 Residences

101 ......Single Family
102 ......Condominium
103 ......Mobile Home (includes land used for purpose of a mobile home park)
104 ......Two-Family
105 ......Three-Family
106 ......Accessory Land with Improvement - garage, etc.
107 ......(Intentionally left blank)
108 ......(Intentionally left blank)
109 ......Multiple Houses on one parcel (for example, a single and a two-family on one parcel)

11 Apartments

111 ......Four to Eight Units
112 ......More than Eight Units

12 Non-Transient Group Quarters

121...... Rooming and Boarding Houses
122...... Fraternity and Sorority Houses
123...... Residence Halls or Dormitories
124...... Rectories, Convents, Monasteries
125...... Other Congregate Housing which includes non-transient shared living arrangements

13 Vacant Land in a Residential Zone or Accessory to Residential Parcel

130...... Developable Land
131...... Potentially Developable Land
132...... Undevelopable Land

14 Other

140...... Child Care Facility (M.G.L. Chapters 59 §3F; 40A §9C) (see also Code 352)

OPEN SPACE

CODE 2

M.G.L. Chapter 59 §2A: Land which is not otherwise classified and which is not taxable under the provisions of Chapter 61, 61A or 61B, or taxable under a permanent conservation restriction, and which land is not held for the production of income but is maintained in an open or natural condition and which contributes significantly to the benefit and enjoyment of the public.

For land designated as Forest, Agricultural/Horticultural and Recreational under Chapters 61, 61A, 61B, see Codes 6, 7, 8.
Land placed under conservation restriction according to Chapter 184, §31 is to be classified according to its use as residential, commercial or industrial property.

20 Open Land in a Residential Area

201 ......Residential Open Land
202 ......Underwater Land or Marshes not under public ownership located in residential area (typically, privately owned ponds, lakes, salt marshes or other wetlands of non-commercial use)

21 Open Land in Rural Area

210 ......Non-Productive Agricultural Land (that part of an operating farm not classified as Chapter 61A Agricultural/Horticultural or Chapter 61 Forest Land)
211 ......Non-Productive Vacant Land

22 Open Land in a Commercial Area

220 ......Commercial Vacant Land (acreage without site improvements and not in commercial use)
221 ......Underwater Land or Marshes not under public ownership located in commercially zoned area

23 Open Land in an Industrial Area

230...... Industrial Vacant Land (acreage without site improvements and not in commercial or industrial use)
231...... Underwater Land or Marshes not under public ownership located in industrial area

Chapter 61, 61A, 61B Property Being Classified as Open Space

Forest, Agricultural/Horticultural and Recreational lands valued according to M.G.L. Chapters 61, 61A 61B and is being classified as open space. (Without an Open Space Classification they must be placed in Codes 6, 7 or, see page 8.)

26 Forest Land

261...... All land designated under Chapter 61
262...... Christmas Trees

27 Agricultural/Horticultural

All land that designated under Chapter 61A. (Land devoted to this use must be in excess of 5 acres and meet other requirements of the law and is being classified as open space.) Note Non-Productive land is being coded as 29.

Productive Land

270...... Cranberry Bog
271...... Tobacco, Sod
272...... Truck Crops - vegetables
273...... Field Crops - hay, wheat, tillable forage cropland etc.
274...... Orchards - pears, apples, grape vineyards etc.
275...... Christmas Trees
276...... Necessary related land-farm roads, ponds, land under farm buildings
277...... Productive Woodland - woodlots
278...... Pasture
279...... Nurseries

Non-Productive Land

290...... Wet land, scrub land, rock land

28 Recreational Land

All property designated under Chapter 61B. (If an area has more than one use according to the codes below, use the code which represents the primary use of the land and is being classified as open space.).

280 ......Productive woodland -woodlots
281 ......Hiking - trails or paths, Camping - areas with sites for overnight camping, Nature Study - areas specifically for nature study or observation
282 ......Boating - areas for recreational boating and supporting land facilities
283 ......Golfing - areas of land arranged as a golf course
284 ......Horseback Riding - trails or areas
285 ......Hunting - areas for the hunting of wildlife and Fishing Areas
286 ......Alpine Skiing - areas for “downhill” skiing and Nordic Skiing - areas for “cross-country” skiing
287 ......Swimming Areas and Picnicking Areas
288 ......Public Non-Commercial Flying - areas for gliding or hand-gliding
289 ......Target Shooting - areas for target shooting such as archery, skeet or approved fire-arms

COMMERCIAL

CODE 3

M.G.L. Chapter 59 §2A: All real property used or held for use for business purposes and not specifically included in another class, including but not limited to any commercial, business, retail, trade, service, recreational, agricultural, artistic, sporting, fraternal, governmental, educational, medical or religious enterprise for non-profit purposes.

30 Transient Group Quarters

300...... Hotels
301...... Motels
302...... Inns, Resorts or Tourist Homes
303...... (Intentionally left blank)
304...... Nursing Homes - includes property designed for minimal care with or without medical facilities
305...... Private Hospitals
306...... Care and Treatment Facilities - designed and used on a transient basis, including half-way houses or other types of facilities that service the needs of people

31 Storage Warehouses and Distribution Facilities

310...... Tanks Holding Fuel and Oil Products for Retail Distribution, either Above Ground or Underground (Underground tanks of service stations would be real estate; however, above ground tanks that rest on concrete saddles or steel frames that can be separated without damage are personal property.)
311...... Bottled Gas and Propane Gas Tanks
312...... Grain and Feed Elevators
313...... Lumber Yards
314...... Trucking Terminals
315...... Piers, Wharves, Docks and related facilities that are used for storage and transit of goods
316...... Other Storage, Warehouse and Distribution facilities (see also Industrial Code 401)
317...... Farm Buildings - barns, silo, utility shed, etc.
318...... Commercial Greenhouses

32 Retail Trade

321 ......Facilities providing building materials, hardware and farm equipment, heating, hardware, plumbing, lumber supplies and equipment
322 ......Discount Stores, Junior Department Stores, Department Stores
323 ......Shopping Centers/Malls
324 ......Supermarkets (in excess of 10,000 sq. ft.)
325 ......Small Retail and Services stores (under 10,000 sq. ft.)
326 ......Eating and Drinking Establishments - restaurants, diners, fast food establishments, bars, nightclubs

33 Retail Trade - Automotive, Marine Craft and Other Engine Propelled Vehicles, Sales and Service

330 ......Automotive Vehicles Sales and Service
331 ......Automotive Supplies Sales and Service
332 ......Auto Repair Facilities
333 ......Fuel Service Areas - providing only fuel products
334 ......Gasoline Service Stations - providing engine repair or maintenance services, and fuel products
335 ......Car Wash Facilities
336 ......Parking Garages
337 ......Parking Lots - a commercial open parking lot for motor vehicles
338 ......Other Motor Vehicles Sales and Services

34 Office Building

340 ......General Office Buildings
341 ......Bank Buildings
342 ......Medical Office Buildings

35 Public Service Properties (see Code 9 for Exempt Public Service Properties)

350...... Property Used for Postal Services
351...... Educational Properties
352...... Day Care Centers, Adult (see also Code 140)
353...... Fraternal Organizations
354...... Bus Transportation Facilities and Related Properties
355...... Funeral Homes
356...... Miscellaneous Public Services - professional membership organizations, business associations, etc.

36 Cultural and Entertainment Properties

360...... Museums
361...... Art Galleries
362...... Motion Picture Theaters
363...... Drive-In Movies
364...... Legitimate Theaters
365...... Stadiums
366...... Arenas and Field Houses
367...... Race Tracks
368...... Fairgrounds and Amusement Parks
369...... Other Cultural and Entertainment Properties

37 Indoor Recreational Facilities

370...... Bowling
371...... Ice Skating
372...... Roller Skating
373...... Swimming Pools
374...... Health Spas
375...... Tennis and/or Racquetball Clubs
376...... Gymnasiums and Athletic Clubs
377...... Archery, Billiards, other indoor facilities

38 Outdoor Recreational Properties (excluding those classified under General Laws 61B)

380 ......Golf Courses
381 ......Tennis Courts
382 ......Riding Stables
383 ......Beaches or Swimming Pools
384 ......Marinas - including marine terminals & associated areas primarily for recreational marine craft
385 ......Fish and Game Clubs
386 ......Camping Facilities - accommodations for tents, campers or travel trailers
387 ......Summer Camps - children's camps
388 ......Other Outdoor facilities - e.g., driving ranges, miniature golf, baseball batting ranges, etc.
389 ......Structures on land classified under Chapter 61B Recreational Land

39 Vacant Land - Accessory to Commercial parcel or not specifically included in another class

390 ......Developable Land
391 ......Potentially developable Land
392 ......Undevelopable Land
393 ......Agricultural/Horticultural Land not included in Chapter 61A

INDUSTRIAL

CODE 4

M.G.L. Chapter 59 §2A: All real property used or held for use for manufacturing, milling, converting, producing, processing, extracting or fabricating materials unserviceable in their natural state to create commercial products or materials; the mechanical, chemical or electronic transformation of property into new products and any use that is identical to or an integral part of such use, whether for profit or non-profit purposes; property used or held for uses for the storage, transmitting and generating of utilities.

40 Manufacturing and Processing

400...... Buildings for manufacturing operations
401...... Warehouses for storage of manufactured products
402...... Office Building - part of manufacturing operation
403...... Land - integral part of manufacturing operation
404...... Research and Development facilities

41 Mining and Quarrying

410...... Sand and Gravel
411...... Gypsum
412...... Rock
413...... Other

42 Utility Properties

420...... Tanks
421...... Liquid Natural Gas Tanks
423...... Electric Transmission Right-of-Way
424...... Electricity Regulating Substations
425...... Gas Production Plants
426...... Gas Pipeline Right-of Way
427...... Natural or Manufactured Gas Storage
428...... Gas Pressure Control Stations

43 Utility Properties - Communication

430 ......Telephone Exchange Stations
431 ......Telephone Relay Towers
432 ......Cable TV Transmitting Facilities
433 ......Radio, Television Transmission Facilities

44 Vacant Land - Accessory to Industrial Property

440 ......Developable Land
441 ......Potentially Developable Land
442 ......Undevelopable Land

45 Electric Generation Plants

450 ......Electric Generation Plants
451 ......Electric Generation Plants, Transition Value
452 ......Electric Generation Plants, Agreement Value

PERSONAL PROPERTY

CODE 5

M.G.L. Chapter 59 §2: All personal property...wherever situated, unless expressly exempt, shall be subject to taxation...

501...... Individuals, Partnerships, Associations and Trusts

All personal property is taxable and includes: stock in trade, machinery used in the conduct of the business, personal property used in connection with any cleaning or laundry processes, machinery used in the refrigeration of goods or in the air conditioning of premises, all furnishings and fixtures and owner non-domicile furnishings.

502...... Domestic Business Corporations or a Foreign Corporations, as defined in Chapter 63 §30

Taxable personal property includes only: underground conduits, wires and pipes wherever located; poles and wires on private ways and machinery used in the conduct of the business, except stock in trade or machinery directly used in connection with dry cleaning or laundering processes, refrigeration of goods, air conditioning of premises or in any purchasing, selling, accounting or administrative function.

503...... Domestic and Foreign Corporations Classified Manufacturing, as defined in Ch. 63, §38C & §42B

Taxable personal property includes only: underground conduits, wires and pipes wherever located, poles and wires on private ways.

504...... Public Utilities -- Transmission and Distribution

Taxable personal property includes underground conduits; wires and pipes wherever located; poles and wires on private ways and machinery used in manufacture.

505 ......Machinery, Poles, Wires and Underground Conduits, Wires and Pipes of all Telephone and Telegraph Companies, as determined by the Commissioner of Revenue.

508 ...... Cellular/Mobile Wireless Telecommunications Companies

506 ......Pipelines Of 25 Miles Or More In Length For Transmitting Natural Gas Or Petroleum**,** as determined by the Commissioner of Revenue.

550 ......Electric Generation Plants Personal Property

551 ......Electric Generation Plant P.P., Transition Value

552 ......Electric Generation P. P., Agreement Value

CHAPTER 61, 61A, 61B PROPERTY

Forest, Agricultural/Horticultural and Recreational lands valued according to M.G.L. Chapters 61, 61A 61B are not specifically included in any of the four major classifications. The commercial property tax rate, however, is the applicable rate for land under these chapters.

CODE 6
Forest Land

601...... All land designated under Chapter 61
602...... Christmas Trees

CODE 7
Agricultural/Horticultural

All land that has been designated under Chapter 61A. (Land devoted to this use must be in excess of 5 acres and meet other requirements of the law.)

71 Productive Land (Including Necessary and Related Land)

710...... Cranberry Bog
711...... Tobacco, Sod
712...... Truck Crops - vegetables
713...... Field Crops - hay, wheat, tillable forage cropland etc.
714...... Orchards - pears, apples, grape vineyards etc.
715...... Christmas Trees
716...... Necessary Related Land-farm roads, ponds, Land under farm buildings
717...... Productive Woodland - woodlots
718...... Pasture
719...... Nurseries

72 Non-Productive Land

720...... Wet land, scrub land, rock land

CODE 8
Recreational Land

All property that has been designated under Chapter 61B. (If an area has more than one use according to the codes below, use the code which represents the primary use of the land).

801 ......Hiking - trails or paths
802 ......Camping - areas with sites for overnight camping
803 ......Nature Study - areas specifically for nature study or observation
804 ......Boating - areas for recreational boating and supporting land facilities
805 ......Golfing - areas of land arranged as a golf course
806 ......Horseback Riding - trails or areas
807 ......Hunting - areas for the hunting of wildlife
808 ......Fishing Areas
809 ......Alpine Skiing - areas for “downhill” skiing
810 ......Nordic Skiing - areas for “cross-country” skiing
811 ......Swimming Areas
812 ......Picnicking Areas
813 ......Public Non-Commercial Flying - areas for gliding or hand-gliding
814 ......Target Shooting - areas for target shooting such as archery, skeet or approved fire-arms
815 ......Productive Woodland - woodlots

EXEMPT PROPERTY

CODE 9

All property which is totally exempt from taxation under various provisions of the law and owned by:

90 Public Service Properties

900...... United States Government
901...... (Intentionally left blank)

91 Commonwealth of Massachusetts – Reimbursable Land

910...... Department of Conservation and Recreation, Division of State Parks and Recreation
911...... Division of Fisheries and Wildlife, Environmental Law Enforcement
912...... Department of Corrections, Division of Youth Services
913...... Department of Public Health, Soldiers' Homes
914...... Department of Mental Health, Department of Mental Retardation
915...... Department of Conservation and Recreation, Division of Water Supply Protection
916...... Military Division – Campgrounds
917...... Education – Univ. of Mass, State Colleges, Community Colleges
918...... Department of Environmental Protection, Low-level Radioactive Waste Management Board
919...... Other

92 Commonwealth of Massachusetts – Non Reimbursable

920 ......Department of Conservation and Recreation, Division of Urban Parks and Recreation
921 ......Division of Fisheries and Wildlife, DFW Environmental Law Enforcement, Department of Environmental Protection
922 ......Department of Corrections, Division of Youth Services, Mass Military, State Police, Sheriffs' Departments
923 ......Department of Public Health, Soldiers' Homes, Department of Mental Health, Department of Mental Retardation
924 ......Mass Highway Dept
925 ......Department of Conservation and Recreation Division of Water Supply Protection (conservation restrictions and sewer easements), Urban Parks
926 ......Judiciary
927 ......Education – Univ. of Mass, State Colleges, Community Colleges
928 ......Division of Capital Asset Management, Bureau of State Office Buildings
929 ......Other

GASB 34 Codes

93 Municipal or County Codes

930 ......Vacant, Selectmen or City Council
931 ......Improved, Selectmen or City Council
932 ......Vacant, Conservation
933 ......Vacant, Education
934 ......Improved, Education
935 ......Improved, Municipal Public Safety
936 ......Vacant, Tax Title/ Treasurer
937 ......Improved, Tax Title/ Treasurer
938 ......Vacant, District
939 ......Improved, District

94 Educational Private

940...... Elementary Level
941...... Secondary Level
942...... College or University
943...... Other Educational
944...... Auxiliary Athletic
945...... Affiliated Housing
946...... Vacant
947...... Other

95 Charitable

950...... Vacant, Conservation Organizations
951...... Other
952...... Auxiliary Use (Storage, Barns, etc.)
953...... Cemeteries
954...... Function Halls, Community Centers, Fraternal Organizations
955...... Hospitals
956...... Libraries, Museums
957...... Charitable Services
958...... Recreation, Active Use
959...... Housing, Other

96 Religious Groups

960...... Church, Mosque, Synagogue, Temple, etc.
961...... Rectory or Parsonage, etc.
962...... Other

97 Authorities

970...... Housing Authority
971...... Utility Authority, Electric, Light, Sewer, Water
972...... Transportation Authority
973...... Vacant, Housing Authority
974...... Vacant, Utility Authority
975...... Vacant, Transportation Authority

98 Land Held by other Towns, Cities or Districts

980 ......Vacant, Selectmen or City Council, Other City or Town
981 ......Improved, Selectmen or City Council, Other City or Town
982 ......Vacant, Conservation, Other City or Town
985 ......Improved Municipal or Public Safety, Other City or Town
988 ......Vacant, Other District
989 ......Improved, Other District

99 Other

990 ......121A Corporations
991 ......Vacant, County or Regional
992 ......Improved, County or Regional, Deeds or Administration
993 ......Improved Count or Regional Correctional
994 ......Improved County or Regional Association Commission
995 ......Other, Open Space
996 ......Other, Non-Taxable Condominium Common Land
997 ......Other

PROPERTY SALES REPORT INSTRUCTIONS

The Property Sales Reports (LA-3) are used in conducting assessment/sales ratio studies. In order to conduct an accurate study, the following information needs to be completed on all sales over $1,000. The Board of Assessors must sign, date and submit the LA-3 via DLS Gateway. See *Property Sales Report Spreadsheet Specifications* on pages 13 and 14 for submission requirement standards. (Codes can be used for all programs, only Code X is restricted to Interim Years.)

NON-ARM'S LENGTH CODES

An "arm's length" sale is a sale between a willing buyer and a willing seller with no unusual circumstances involved in the sale. Listed below are the codes for sales that are considered non-arm's length.

A. Sale between members of the same family

B. An intra-corporation sale, e.g. between a corporation and its stockholder, subsidiary, affiliate or another corporation whose stock is in the same ownership

C. Sale of any real property which includes personal property, machinery, equipment, inventories or "good will".

D. Sale of property substantially changed before the sale occurred but after the assessment date, *i.e. sale price includes change, whereas assessed value does not.*

E. Sale to / from a federal, state, or local government

F. Transfer of convenience, e.g., correcting defects in a title, a transfer by a husband either through a third party or himself and his wife to create a tenancy by the entirety, etc.

G. Sale of only a portion of the assessed unit, e.g., a parcel sold from a larger tract and the assessment is for the larger tract, or a portion is in another municipality

H. Sale resulting from a court order, e.g., a divorce settlement, estate sale

I. Sale in proceedings of bankruptcy

J. Sale of an undivided interest

K. Sale to / from an educational, charitable, or religious organization

L. Repossession of a foreclosed property by a financial institution or other lender.

M. Sale of property, the value of which has been materially influenced by zoning changes not reflected in current assessments

N. Other, when a non-arm's length sale does not fall into any other category, this code is used, accompanied by a written explanation and/or comparable sales analysis.

O. Sale of property substantially changed after the sale occurred but before the assessment date. *i.e. sale price does not include change, whereas the assessed value does*

P. Sale of property with a change in use when compared to its use on the assessment date.

Q. Sale of property which includes both a trade of property and cash for the property conveyed

R. Sale of property which has been sold more than once in the same analysis period. Only the most recent valid sale closest to the assessment date is used for analysis purposes.

S. Sale of a foreclosed property by a financial institution or other lender. (If considered arm's length, must be supported by detailed documentation.)

T. Property sold to an abutter

U. Private sale not put on the market

V. Sale of multiple parcels

W. Sale affected by deed restriction, e.g., 40B housing

X. Sale of parcel where no value exists for prior assessment because the parcel ID is new. (Used for coding in interim years only.)

PROPERTY SALES REPORT - LA3

Spreadsheet Specifications

Data Layout Example

Columns

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
jur_ code	**sale_date**	**parcel_id**	**seller**	**buyer**	**st_ num**	**st_ alpha**	**st_name**	**prop_ type_ id**	**nal_ code**	**sale_ price**	**assessment _value**	**proposed _value**	**as_ ratio**	**outlier**	**time_ trend**	**comments**
001	05/03/2011	8-0-28	Smith John	Jones Paul	121		Woodland St	101		470,000	390,000	447,500	0.95		485,000	
001	12/22/2011	12-0-160A	Harrison W.	Raycroft B.	83	A	Forest St	102		320,000	270,000	332,000	1.04		320,000	
001	07/12/2011	6-0-156	Johns P	Bradley A	13		Ralph Ave	104	N	125,000	185,000	170,000	1.36		125,000	Short Sale
001	06/18/2011	3-0-66	Bartlett Co.	Miller William	175		Maple St	101	P	225,000	220,000	475,000	2.11		230,900	

Row Headings should be on one line (wrapped if necessary) labeled exactly as above ***see note below***

	Column Heading	Description	Format
Column A	jur_code	DOR community ID number	Text column – Three digits
Column B	sale_date	Date of sale	Date column - mm/dd/yyyy
Column C	parcel_id	Community identification	No special format – up to 30 Characters*
Column D	seller	Grantor of the property	No special format – up to 40 Characters*
Column E	buyer	Grantee of the property	No special format – up to 40 Characters*
Column F	st_num	Street number of the property	Numeric – up to 10 digits
Column G	st_alpha	For any text character part of st_num	Text Column up to 5 Characters
Column H	st_name	Name of the street, road etc.	Maximum Length – 40 Characters
Column I	prop_type_id	State use code of property	Text column – 3 Characters **
Column J	nal_code	Non-arms Length Code	Text column – up to 3 Characters ***
Column K	sale_price	Sale Price of the property	Numeric *
Column L	assessment_value	Prior Fiscal Year Assessment	Numeric *
Column M	proposed_value	Proposed current Fiscal Year Assessment.	Numeric *
Column N	as_ratio	Assessment Sales Ratio	Numeric with 2 place decimal
Column O	outlier	DOR use only, should be blank for all entries	
Column P	time_trend	(If applicable) Time-Adjusted Sales Price.	Numeric ****
Column Q	Comments	Explanation of "N" codes or other as needed	Text

* No entry can be blank.
** This should reflect the property's class code as of the proposed assessment date, not what it was at the time of the sale.
*** Must be left blank for all valid sales.
**** If using a time adjustment for any or all classes, entire column may be filled. (Use actual selling price for those sales not time adjusted.) If a community is not using a time-adjustment, column can be left blank.

Note: In the example above, the original sale of $225,000 is arms length since a vacant piece of land (class 130) sold and the prior FY assessed value reflects this ($220,000). However, the same sale, when compared to the current FY assessed value of a single family home ($475,000), becomes a non-arms length sale with the NAL code of "P". The usage class changes from a 130 to a 101.

(Over)

PROPERTY SALES REPORT - LA3

DATA UPLOAD DIRECTIONS

BULK UPLOAD

Certification

While you are in Gateway, click on the LA3 tab, and in the LA3 Upload Program, click on the Bulk Upload screen. Select your jurisdiction. Correctly identify the Fiscal Year and Process, which will be Certification – not Interim Year Adjustment. Click on "Go".
Note: If data has already been entered or uploaded, the system will prompt you that data already exists and do you want to proceed. If you proceed the system will modify existing data or add new data records.

In your Excel file, copy the LA3 data that meets the format prescribed in Spreadsheet Specifications – without the header - and paste the data into the template spreadsheet with built-in macro programs. This template spreadsheet with built-in macros for cleansing the data of problematic characters and confirming field formats can be accessed by clicking on the link above the Bulk Upload box. See: Download Correctly Formatted Excel Template with Macros for Pre-Submission File Cleanup (See help)

After cleaning, copy the data from this spreadsheet - without the headers – and go back to the LA3 Bulk Upload screen, paste that data into the box for the bulk data upload. Click on the Process Bulk Data button. The system will show the number of correctly formatted records and any errors. If the file has errors, correct and select Re-process Incorrect Data. Once the data is correct, click on the Save button at the bottom of the screen. You must then go to the Sign and Submit LA3 Data screen, on the menu, to complete submission of the LA3. When you are ready to formally submit the file and lock the file from further local changes, click the Assessor signature box at the bottom of the screen, make any appropriate comments, and click Submit.

Interim Year Adjustment

Uploads can be made in the manner described above, except the Process selected will be Interim Year Adjustment. The resulting LA3 will contain two extra columns automatically inserted by the Gateway system - Prior Year Use Code and Prior Year NAL Code - after the LA3 is saved. It is necessary to review these added columns to insure the Class Code and Non Arms Length (NAL) codes are applicable to the prior year assessment. It is also possible to upload the data from an Excel file that contains the two additional columns inserted between the st-name (H) and the prop-type-id (I). The line above the bulk upload box labeled: "Upload includes Prior Year Use Code and Prior NAL Code columns" has a check box for this purpose. Once checked the cleaning template will also change to accommodate the new data.

SINGLE RECORD UPLOAD

While in Gateway at the LA3 tab, go to the Single Record Upload screen and select your jurisdiction. Correctly identify the Fiscal year and Process (Certification or Interim Year Adjustment). Enter the data in the correct format as listed in Spreadsheet Specifications. If a field format is incorrect, the system will prompt Data formats are not valid in the highlighted field(s). Please correct. Click the Save button to add the record for that community, process, and fiscal year. Click Add New to add an additional record, as opposed to overwriting the information on the screen and clicking Save. The latter action will simply overwrite one record's information with different information.

UPDATES ON-LINE

THIS INFORMATION IS UPDATED ONLY ON AN AS NEEDED BASIS

.

FOR FURTHER INFORMATION, CONTACT YOUR BUREAU OF LOCAL ASSESSMENT ADVISOR OR

EMAIL

bladata@dor.state.ma.us

VITA

Benjamin Williams is a graduate student of Urban and Regional Planning at the University of New Orleans (UNO). His concentration is Land Use Planning, with a focus on the intersection of land use, urban design, and sustainable transportation. He is a data management intern at the Dudley Street Neighborhood Initiative in Boston, MA, formerly a graduate planning assistant at the New Orleans Regional Planning Commission. Benjamin holds a B.S. in Urban and Regional Studies from Cornell University. He hopes to use his education and expertise to encourage context-driven and sustainable land use, design, and transportation planning decisions.