The Impact of Access to Public Transportation on Residential Property Value: A Comparative Analysis of American Cities

Moses Snow Wayne

Dr. Patrick Bayer, Faculty Advisor

Dr. Kent Kimbrough, Seminar Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke

University

Durham, North Carolina

2018

Acknowledgements

I am incredibly thankful for the advice, insight and encouragement of Dr. Patrick Bayer, without whom this project would have been unfeasible. I am also immensely appreciative of the constant feedback and guidance of Dr. Kent Kimbrough, my honors thesis seminar advisor. I would also like to thank Jennifer Becker from the Economics Department for her aid in laying down the necessary groundwork for acquiring the data for this research project. I would also like to express my gratitude for the suggestions and unique perspectives of my classmates in my honors thesis seminar.

Abstract

This paper develops a consistent model for analyzing the impact of access to public transportation on property value applied to the four cities of Atlanta, Boston, New York, and San Francisco. This study finds a negative relationship between increasing distance to public transit and property value. Additionally, the elicited effects in each city generally align with geographic features and the degree to which a city is monocentric. This study also demonstrates the salience of using actual map-generated distances as proximity measures and characteristics of public transit systems in modeling the relationship between public transportation and residential property value.

JEL classification: C12; R14; R30; R41;

Keywords: housing; public transit; Atlanta; Boston; New York; San Francisco

I. Introduction

Public transportation plays a large role in the commutes of citizens in numerous metropolitan areas throughout the US. This fact holds true especially for the largest US metropolitan areas. The American Public Transportation Association reports that the ten largest cities by metropolitan population in the United States all rank within the top twenty most used public transportation systems (Neff & Dickens, 2015). Although these cities are similar in their characteristic of population size, each differ greatly in other aspects, such as in urban sprawl and degree of monocentricity. These varied characteristics can provide a framework for the analysis of the role that the public transportation system plays in each city.

The role of public transportation as a means of commuting to the central business district (CBD) drives the value of access to public transportation for individuals living in a city. The CBD is a given city's central location for employment and commerce activity. The access provided to this important locale by public transportation imparts a value to its access, and a significant volume of literature in the field of economics has attempted to quantify this value of public transportation access on nearby residential properties. However, methodology used in these studies has often been inconsistent. This narrow focus of individual cities in previous studies makes the findings between studies difficult to compare given these methodological differences.

The varied urban planning and development of cities lead to idiosyncratic features playing a role when analyzing the city effects on observed residential property values. Because of these varied processes of urban development in different cities across the United States, developing a consistent model across metropolitan areas of different urban characteristics can elucidate the questions in some of the gaps in the current economics literature. The staggered waves of urban development and decline in the second half of the twentieth century within US metropolitan areas have resulted in differences in the central city structures across regions of the United States; the timings of the rise of central city population and the later population losses occurred in different intervals for the cities in the Northern, Southern, and Western regions of the United States (McDonald 2015). Because of the varied paths toward urban development taken by each of these different regions, it is worthwhile to comparatively examine cities in each of these regions of the United States and the resulting differences in the impacts observed on residential property value from proximity to public transportation. The metropolitan areas of Atlanta, Boston, San Francisco, and New York City each represent cities that progressed along central city growth characteristic of their region. Additionally, each of the four cities have frequently used public transit systems, with each public transit system of each of these cities utilizing both bus and rail transit systems. This allows for parity in examining consistent modeling across each of the cities' public transportation systems. These factors establish the foundation for the intent behind using the four listed metropolitan areas and their public transportation systems in a comparative analysis of proximity of residential properties to public transit. While ideally a larger spectrum of cities would be examined to provide a more holistic view, the time intensity and cost of data collection provide barriers to this challenge. However, because of the varied histories and regions of the four examined cities, this research paper is able to examine a broad variety of types of urban development.

This paper furthers the analysis of access to public transportation's impact on property value by evaluating a rigorous hedonic pricing model that is attuned to controls on not only housing characteristics, but public transit characteristics. Applying this model to a spectrum of cities with differing aspects of urban planning can expand upon the discrepancy in effect observed from cities with idiosyncratic features. By examining features that impact the degree to which a city is monocentric, this paper analyzes the relationship between a city's monocentric characteristics and the magnitude of effect on property value derived from proximity to public transportation. Public transportation features are often uncontrolled in previous literature surrounding the impact of public transportation on property value. Through using public transportation features as independent variables in this research project, the effects derived from proximity to public transportation are better differentiated between cities. Public transportation characteristics, such as efficiency of transit commute compared to automobile commute, provides insight into city-specific value of public transit access. In this study, controlling for quality of public transit efficiency highlights the specific value of access to public transportation for an individual city irrespective of the public transit system. As a result, this research project provides a better comparative perspective of the impact of different forms of public transportation on changes in property value between different cities with differing public transit systems, emphasizing city characteristics.

This research project has numerous potential implications for both policy and the field of economics. While public transportation provides other documented benefits to cities, analyzing the proximity effects of public transportation stops on property value can highlight an even greater derived utility from public transportation stops. This is especially useful in determining policy for spatial planning in urban development. Moreover, in providing a comparative analysis on the impact of public transportation on changes in property value, this study furthers the economic literature on the subject and examines the role of city characteristics on the magnitude of this impact. The unique methodology of this research project also presents novel approaches to how distances are approximated and explores the salience of examining public transit system characteristics in modeling the impact of public transit access on property value.

The ramifications of the findings from this paper imparts a weight to the study of the public transportation-property value interaction. The following section, Section II, delves into the economic literature that exists on this topic to lay the foundation for the analysis in this paper. Section III provides a discussion on the monocentric city model and the bid rent curve, the theoretical models upon which the questions of this paper are constructed. In Section IV, the dataset that is used for the statistical analysis performed in this study is discussed, and the procedure for constructing this dataset is outlined. Section V will introduce the hedonic model explored in this paper along with the significant variables of the regression. Section VI will demonstrate the results of the research project. Section VII provides a discussion of the results from Section VI and provides analysis on the findings. Section VIII will conclude this paper.

II. Literature Review

A major motivation and source for inspiration for this research project comes from previous efforts in economic literature to quantify the impact of public transportation on property value. Despite the existence of literature on this subject, the findings of these studies vary. In their meta-analysis of railway station proximity and property value, Debrezion (2007) notes that results from studies on the subject of public transportation's impact on property value often differed due to inconsistent modeling and methodology. Moreover, in another review of empirical studies on property value and public transportation, Ryan (1999) explains inconsistencies found in literature as a possible result of the omission of commute time when analyzing the impact of proximity to public transportation; effects tended to be observed in studies accounting for this omission.

An analysis of rail transit proximity and residential property value by Bowes and Ihlanfeldt (2001) determined that the proximity of a station to a residential property directly affected that property's value. The model tested in the study was rigorous in the controls established, but methodological concerns exist for Bowes and Ihlanfeldt's study. The response variable measured in the study was residential property value rather than a change in property value over time. The use of value rather than the change in value as the response variable makes it difficult to determine causality. There could be an endogeneity concern over whether the placement of a public transportation stop resulted from property value or if the property value is derived from the proximity to a stop. Instead, a change in value evaluates a constant presence and effect over time, obviating this endogeneity concern. Furthermore, the study was conducted in the year 2000 and has the limited scope of only applying a pricing model on the metropolitan area of Atlanta. This research project provides an updated perspective of this effect while also removing dependency of the observed effect on the city in which it was studied. Furthermore, in removing an endogeneity concern from Bowes and Ihlanfeldt's study, this research project provides a better framework for determining causality.

Studies centering on the use of buses as public transportation have also been conducted, but with varied implementations of hedonic models and results. A study done by Wang, Potoglou, Orford, and Gong (2015) analyzed the impact of bus stops on property value in Cardiff, Wales as justification for the implementation of a land value tax. Instead of using proximity to the nearest bus stop as an explanatory variable, Wang, Potoglou, Orford, and Gong used number of bus stops within a preset distance deemed walkable. Their study found that the number of walkable bus stops surrounding a property had a positive effect on the property value of the residence in terms of sale price; this finding was especially true for residential properties with higher prices. This study especially emphasizes the importance of determining feasible walking distance over simply distance to the nearest public transportation stop. Furthermore, the study provides an important perspective on multiple points of access to the public transportation system.

Another study focusing on both buses as public transportation and walking distance from their stops is a study by Ramon Munoz-Raskin (2010) on property value in Bogota, Colombia. Munoz-Raskin's study examined a hedonic model using dummy variables on multiple walking times to different lines of the public bus system in Bogota. The study also controlled for socioeconomic characteristics of the housing, finding that the effect of walkability to a public transit stop was lower for lower-income areas. The study is significant in its breadth of analysis and unique in its methodology in assessing different levels of accessibility as dummy variables, but it does contain weaknesses. The lack of control for numerous variables, such as housing characteristics and distance to the CBD, hurt the explanatory power of the model, and this is reflected in the low R^2 values reported from the regressions on the model in Munoz-Raskin's study.

The monocentric city model focus of this research project draws motivation from the economic literature on the subject as well. While changes in urban planning over time have altered degrees of monocentricity for cities in the United States, a study conducted by Arribas-Bel and Sanz-Garcia (2014) demonstrated that the majority of metropolitan areas today still exhibit these characteristics. As a result of the continued importance of this model, efforts to develop models to quantify the monocentric characteristic exist in economics literature. The Urban Centrality Index, derived by Pereira, Nadalin, Monasterio, and Albuquerque (2012), is an index of how centered a city is around its CBD. While the index uses useful metrics like employment and population density, it does not include a measure of the bid rent curve. This research project examines the alignment of the bid rent curve with the general patterns of employment density in each of the four cities.

Research on this subject that I have previously conducted also drives the motivation behind this paper. Previously, I have examined the relationship between proximity to public transportation and changes in property value within the cities of Boston, Massachusetts and Durham, North Carolina, two cities with prolifically used public transportation systems on a per capita basis. The derived model controlled for size of residential homes, distance to the CBD, and census data of attributed block groups for each of the residential properties. There was a statistically and economically significant inverse relationship observed between distance to the nearest public transportation stop and yearly increase of property value in the city of Boston when the public transportation stop was less than a mile away. Namely, positive yearly property value changes diminished as a residential location was located further away from its nearest access point to public transportation. Conversely, no statistically significant relationship was observed for the city of Durham. The model in this previous analysis also expressed that Boston exhibited significantly stronger characteristics of a monocentric city than did the city of Durham. This sentiment is supported by the histories of how the cities were planned.

This paper builds upon and improves pervious literature on the relationship between access to public transportation and residential property value. This research project adopts a wider perspective on the effects of public transportation on property value by comparatively viewing the magnitudes of these effects across a span of varying histories of urban planning and development in the four cities examined. The additional analysis of city specific characteristics allows for an examination of city features that affect the value of a public transportation system to its surrounding residential properties. Features like walkability and transit efficiency can capture aspects of the central city's development and elucidate a relationship between a metropolitan area's planning and the value of access to public transportation. In providing a comparative analysis, this research project demonstrates a more robust model and examines overarching observations at the city level.

Furthermore, this research project adopts novel perspectives on the public transportation-property value relationship in exploring factors not considered in previous literature. Not observed in previous literature, efficiency of public transportation is determined by the difference in commute time to the central business district between the use of public transit and the use of an automobile. The inclusion of efficiency of a city's public transportation system as a variable provides a cleaner perspective on the consumption behavior of actors substituting transportation methods, and in turn, the value of access to a public transportation stop in determining the value of a residential property. Additionally, this research project utilizes actual distance to public transit instead of straight-line distances, an aspect that allows a better measure of distances and used public transit stops. Moreover, the use of change in value of residential properties rather than current value differs from some of the previous literature. It addresses the concern of determining added value to a residential property through being in a location with access to public transportation.

III. Theoretical Framework

A city that is very monocentric is highly concentrated around a central business district. The concept of the monocentric city was introduced in 1964 by economist William Alonso in *Location and Land Use*, but the basis for Alonso's theory was established in his prior analysis of rent and location. Alonso's model holds that actors choose their location for rent around the central business district as a result of their individual bid rent curve (1960). In maximizing their utility on this bid rent curve, individual actors must maximize their utility across proximity to the central business district and the cost of commuting to it; the equilibrium that actors reach in choosing their location around the CBD requires that the marginal cost of longer commute due to greater distance from the CBD must equal the marginal benefit of lower cost of rent from moving away from the CBD.

Figure 1. Log-log scatter plot demonstrating the relationship between changes in property value and changes in distance to the CBD for the city of Atlanta, regression table shown in Table A2.1

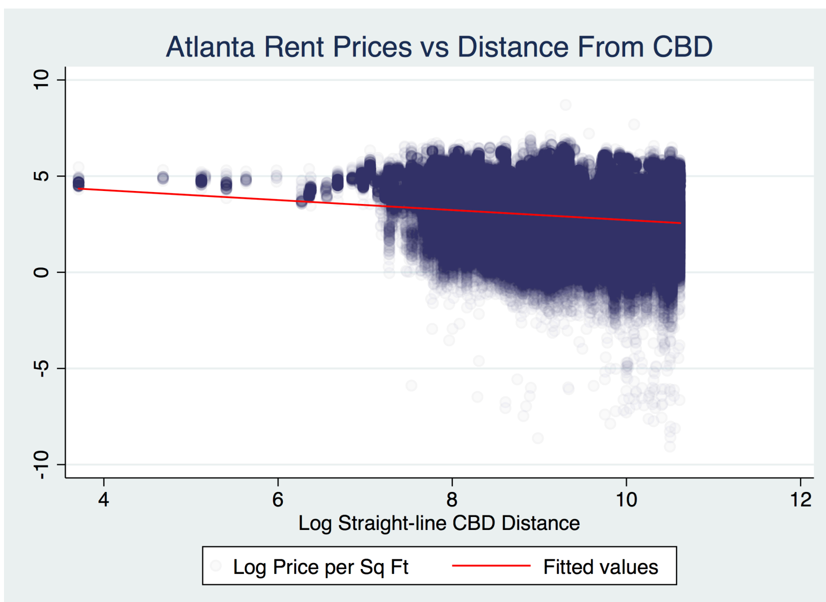


Figure 2. Log-log scatter plot demonstrating the relationship between changes in property value and changes in distance to the CBD for the city of Boston, regression table shown in Table A2.1

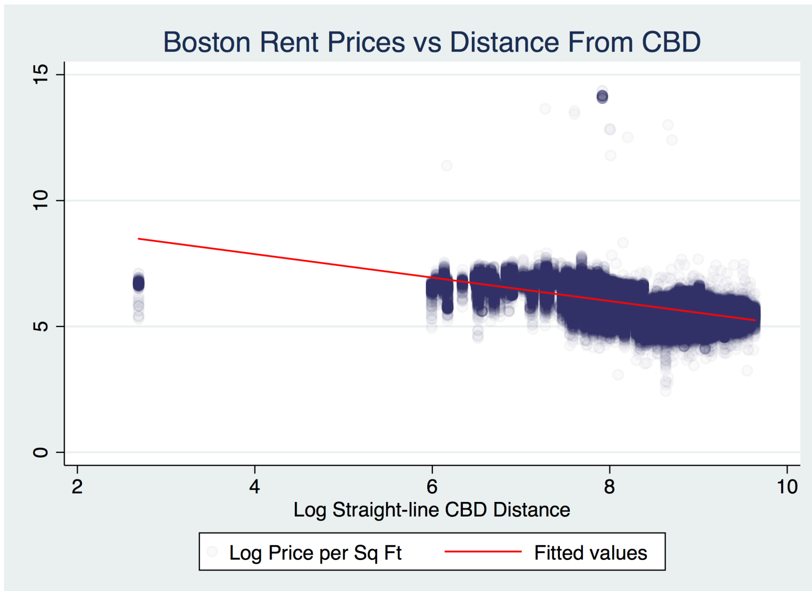


Figure 3. Log-log scatter plot demonstrating the relationship between changes in property value and changes in distance to the CBD for the city of New York, regression table shown in Table A2.1

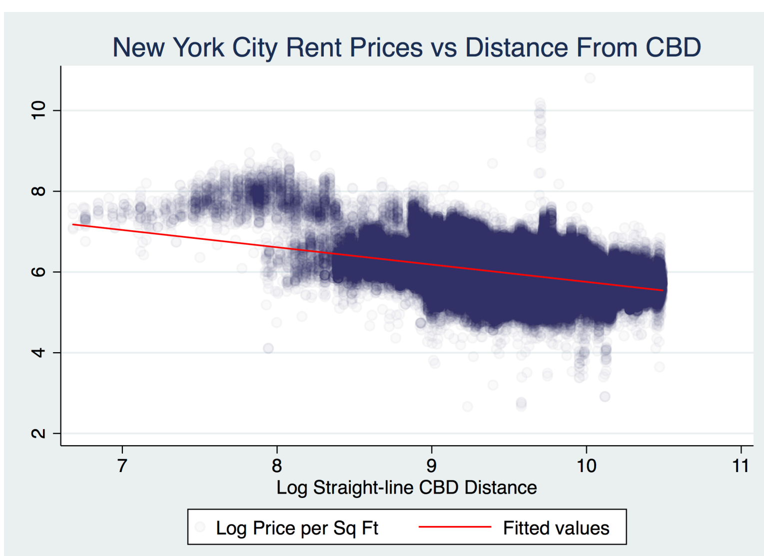
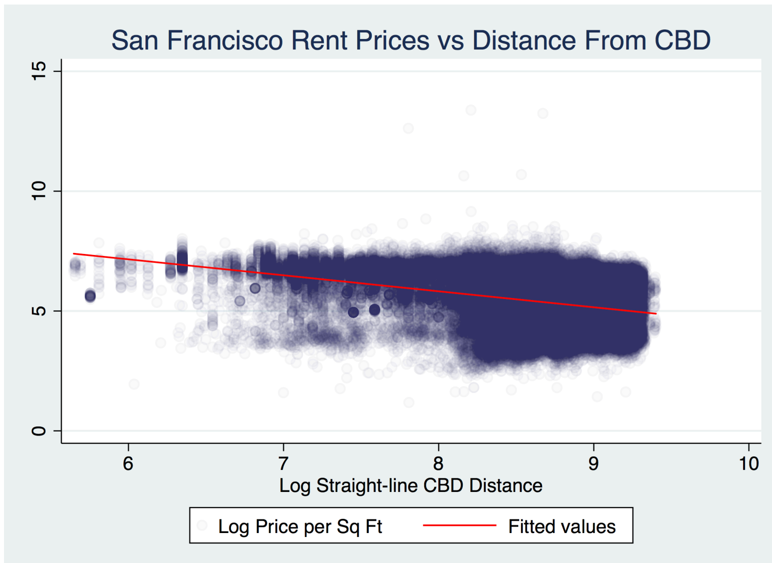


Figure 4. Log-log scatter plot demonstrating the relationship between levels of property value and distance to the CBD for the city of San Francisco, regression table shown in Table A2.1



Here, Figure 1, Figure 2, Figure 3 and Figure 4 operate as important pieces of the discussion on the continued relevance of the Monocentric City Model. The scatter plot graphs for each city descend in property value as the distance to the central business district increases. Absent controls for on the graphs for numerous variables, the above figures demonstrate generally decreasing price with larger distances away from the CBD, a characteristic of monocentric cities. The logarithmic axes aid in tempering heteroscedasticity for variables with non-uniform error terms, such as the assessed market value for residential properties outlined in the above figures. These figures provide evidence for the rent gradient of monocentric cities. As the distance from the central business district increases, the price per square foot of residential properties decrease in each of the above figures.

Introducing available public transportation into this system decreases the cost of commute to the CBD for a given location. The clearing of this market should increase the value of properties with easier access to public transportation. As the availability of public transportation shifts down the marginal cost curve of locating away from the CBD, the access to the service should be reflected as a paid premium in the property value. Moreover, as the value of the CBD lies in its offering of centralized employment and commerce, cities with more pronounced monocentricity should have a greater value in accessing the CBD. Because of this observation, the availability of public transportation would shift down an otherwise steeper marginal cost curve from the cost of locating away from the central business district. As a result, it should be expected that more pronounced monocentric cities have a higher premium on access to public transportation.

IV. Data

The dataset analyzed in this research project is based on the residential properties and public transportation systems of the four cities of Atlanta, GA, Boston, MA, New York, NY, and San Francisco, CA, as outlined in the introduction. Representing the Atlanta metropolitan area in the dataset is Fulton County, seated at Atlanta. The rest of the cities and their metropolitan areas are represented by city boundary rather than county. The years covered in the dataset for this research project are the most recently available tax assessed values and public transportation systems in present day. Across the dataset, the range in property value assessment years used is 2016 to 2017. For all datasets except for San Francisco, which spans 2014 to 2015, as the tax assessor of San Francisco has only released up to this year.

IV.1 Data Collection

The data used for this research project was derived from publically available tax assessor data. This data records the property value associated with the most recent assessed value used for taxation of a given property. Each of the examined cities provide prepared datasets for the property parcels and their tax assessor data. The data from the various tax assessors for property parcels in the cities being examined also provide a view of house characteristics and details about the property, such as the last sale price and date. Tax assessor data is acquired from open data portals published by each metropolitan area. The portals for each city are as follows: Atlanta - Fulton County GIS Data, Boston - Boston Open Data, New York City - NYC Open Data, San Francisco - DataSF. To clean up the large number of properties being examined, only residential properties are kept in the dataset. Residential properties here are defined as being buildings with any codes specifying usage as a non-short-term residence, ranging from detached single-family houses to condominiums. Buildings listed in tax assessor datasets contain building use codes that specify the specific usage of each building in the dataset. To isolate the dataset to residential properties, buildings with use codes not related to residency were excluded from the collected dataset for this research project. Furthermore, only residential properties lying inside the bounds of the city limits are kept in the dataset.

The use of tax assessed property value as a proxy for the actual property value of a residential property presents a few challenges. The method of assessment of property value by the city or county tax assessor likely varies across the four cities and does not necessarily serve as a strong predictor of future sale value of the residential property. In California, the existence of specific tax assessment laws for residential properties presents a challenge in the accurate assessment of property value. This modification of property assessment procedure through this California proposition will be discussed with the results of San Francisco. For the other three cities, given that tax assessment is done by one office across each city, property value should be consistent within the dataset. Additionally, the use of change in property value is useful in addressing this concern. While the individual present values for the assessments of different residential properties could be different from the explicit market values of these residences, the change in value of a residential property in a short period of time across a single city's tax assessor should be a consistent proxy for change in actual property value of that residence. As all residences in the city must have assessments from a single tax assessor done in a periodic cycle, the changes in the valuation should stay consistent in comparing across other residences of the city.

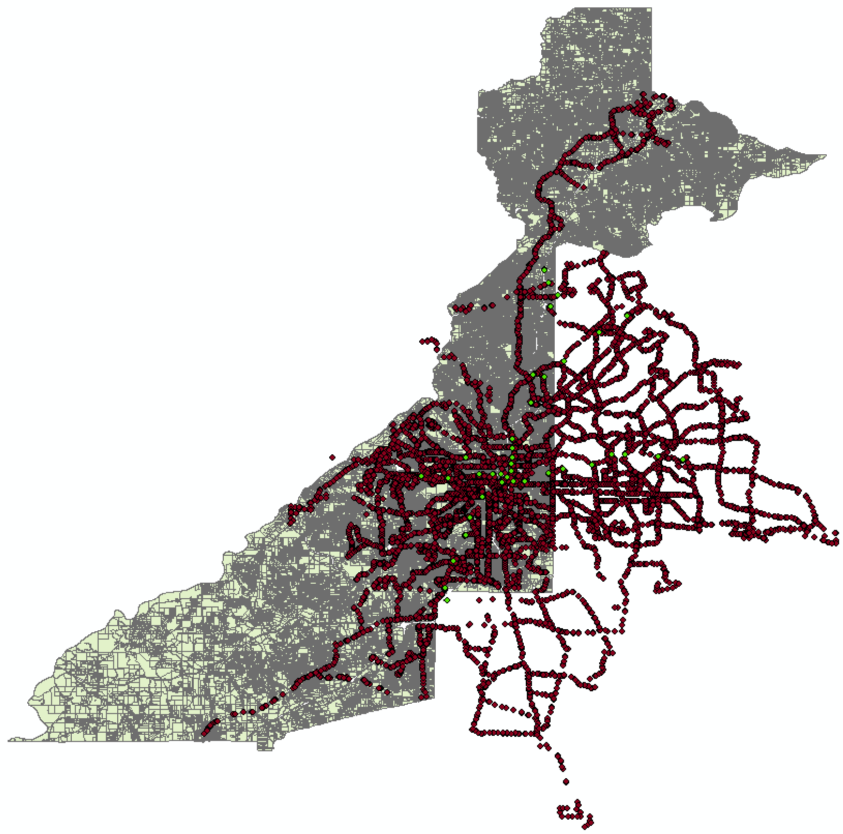
Change in property value and current property value serve as the dependent variables in this research project. In order to record change in property value rather than present value, a second dataset of a past valuation of residential properties is constructed for each of the analyzed cities. To merge this past dataset with the current value dataset, the datasets are joined by a common attribute between the two. Depending on the city in question, this could be parcel ID, or a combination of identification variables, such as the borough, block, and lot number. A derived variable of yearly change in property value is calculated from the difference in assessed value between the two most recent tax years reported by the office of that city's tax assessor, primarily 2016 and 2017.

The tax assessor data used in this research project presented some limitations. Formatting of the datasets was not consistent across the different cities, and the different variables present in each of the datasets posed a challenge in determining a consistent model across multiple cities. In the case of Atlanta, the city does not provide openly available information on the living area square footage of residential properties. Lot area of the property was used instead as a proxy for living area square footage in this research project.

Data is collected on the characteristics of public transportation. Public transportation services for metropolitan areas often provide geographic information system (GIS) maps containing layers of public transit locations. For each of the cities examined in this project, the dataset for the public transportation stops is either obtained as a GIS map layer from the transportation department or constructed as a GIS map layer from information on where these public transportation stops are located. The public transportation systems examined in each of the cities provide both rail and bus services. The Metropolitan Atlanta Rapid Transit Authority (MARTA) or Atlanta, the Massachusetts Bay Transit Authority (MBTA) of Boston, and the New York City Transit (NYCT) all run rail and bus public transit systems, while San Francisco divides its rail and bus services into the Bay Area Rapid Transit (BART) and the San Francisco Municipal Railway (Muni). The distinction between modes of public transportation here are significant, as the different forms of transportation are recorded.

In order to obtain usable data for statistical analysis, the data mappings of individual city public transportation systems need to be associated with the data on residential property value. To combine the data from public transportation GIS layers with the housing data from each city's property tax assessor, the locations of each individual residential property need to be spatially bound to the nearest public transportation stop. This is achieved with the spatial join operation within ArcGIS; this function uses the spatial location of elements from multiple layers to assign an element from one layer to an element from another layer based on a user specification. In this case, the specification for the join is the closest geographical distance, which binds each residential property data entry to all the data of the public transportation stop nearest to it.

Figure 5. Layers of residential parcels and public transportation stops for Fulton County and the city of Atlanta, GA



The above Figure 5 details one of the ArcMap projections of layers of residential parcels overlaid by the public transportation network of MARTA. The dark red points in the above mapping represent stops in the MARTA public transit system, and the bright green points represent the MARTA rail stations. The centroid of each residential parcel is spatially bound to both its nearest bus stop and its nearest rail station, in order to differentiate added value by method of public transportation. An observation of note with the above GIS mapping is the large separation between some of the fringe residential parcels of Fulton County and their nearest public transportation stops, which is significant when determining usable data from the dataset for analysis.

Demographic information is important in controlling for various external effects on property value in a given area. To capture this aspect of residential property value, US Census data is used to capture these demographic differences. The use of controls for census characteristics attributed to a given residential property's census block group requires that US Census data is incorporated into the dataset. To merge this data with the existing dataset, a spatial join with the specification of one layer's elements residing inside the other was used. This attributes census block data to a given residential property based on the block in which a residential property is located.

Efficiency of the transportation system can factor heavily in the willingness to pay to use the public transportation system. In order to model this characteristic of the transportation system for a given city, difference in travel time between use of public transportation and use of an automobile was recorded. To maintain consistency in this difference in time of commute, every residential property has two computed values, the time to commute by automobile to the central business district as well as the time to commute to the central business district by fastest method of public transportation available. The computed time to the central business district is computed consistently across all data points using the popular maps software Google Maps. Each data point has a departure time of 8 AM local time on a Wednesday, with the travel time being at what estimated time Google Maps indicates the trip to the CBD will end. The mode is set to driving for automobile transportation and transit for public transportation times. Both time to the CBD and distance to the CBD across the different commute option is recorded, and the differences in these values are used in this research project. This time to the central business district by automobile serves as distance to the CBD across all data points, and the difference in time between use of public transportation and use of automobile to commute to the CBD serves as a measure of the efficiency of the public transportation system.

To accurately record the commute times for both automobiles and public transportation, batched processes have to be run through a script for each data point in the dataset. Google Maps offers an Application Programming Interface (API) for their Directions service using an interface through HTTP. This allows a program on the client side to send requests for information to the Web API about specific direction information between two locations with a specified mode of transport. The returned data provides information such as the time taken to arrive at the first transit stop used on the transit path to the CBD and general commute distances and durations. This presents information that is not utilized in previous literature surrounding proximity to public transportation, as previous studies used straight-line distance metrics as measures for proximity. The returned data from the Google Maps API for Directions not only provides actual perceived distance and duration to arrive at the nearest public transit stop, it also selects the public transit stop that is relevant to the commute to the central business district. The closest public transit stop to a property parcel is not necessarily useful to the commuter.

While these map-generated distances are undoubtedly important, the procedure cannot be applied to all data points. The Web API provided for Google Maps has quotas placed on queries, presenting both limits to query load and cost for high volume. To address these quotas, a random sample of 50,000 residential properties was taken for each of the examined cities to get Google Maps queried data for the datasets.

IV.2 Data Description

The tables in Appendix 1 detail the summary statistics for assessed residential property values for each city. These tables in the appendix also depict the calculated distances from each residential parcel to its nearest rail and bus transit stations through the use of ArcMap's spatial join. These tables demonstrate that city rail and subway stations are typically further away from a given residential parcel than a bus stop. The study by Wang, Potoglou, Orford, and Gong demonstrated the significance of walkable transit stops, so a maximum walkability distance could provide similarly significant results in the regression. In terms of usable and consistent data, it is likely more meaningful to regress on walkable distances, as there is significant noise introduced from the long distances from public transit locations.

Summary statistics for map-determined distances are seen in the second set of tables in Appendix 1, depicting transit stop walk times, map-determined automobile commute, and the derived differences in commute time to the central business district between the two for random samples of 50,000 data points taken from the dataset. The values expressed are in seconds, the format returned by the Google Maps API. According to the summary statistics for the efficiency ratio of public transit, Boston has the most efficient public transit system of the four cities examined while Atlanta has the least efficient system. The value for this variable represents the percent change over the commute time of driving to the CBD; lesser public transit efficiency ratio values indicate more efficient public transportation. It is important here to observe the decrease in number of observations reported for Atlanta as compared to the other three cities. The sprawl of Atlanta made detecting public transit routes difficult for properties that lied too far in suburban areas, so the Google Maps API was unable to return meaningful data. Given the large number of observations without detected public transportation access, the efficiency of the Atlanta public transit system could be worse than indicated by the tables.

V. Empirical Specification

This study intends to analyze the impact of access to public transportation on the property value of residential locations in a city. To examine this effect, a hedonic regression of distance to the nearest public transportation stop on changes in residential housing prices was applied to the aforementioned four metropolitan centers of Atlanta, Boston, New York, and San Francisco. The derived hedonic model on changes in housing price controls for residential housing size, distance to the central business district, public transportation system characteristics, and demographic census characteristics attributed to a residential location's census block group.

As mentioned in the literature review, this research project uses change in property value over a time period in addition to current property value in order to address the possible endogeneity concern between property value and public transportation stop placement. In order to control for differences in housing size, the property value is adjusted by dividing the property value by the size in square feet of the living space of the residential property. The response variables for this hedonic model are both the current residential property value per square foot and the yearly change in property value per square foot over the time period being analyzed for a given city. The primary independent variable is the distance to the nearest public transportation stop, delineated in different ways. Straight line distance to the nearest public transportation stop is calculated for each property in the dataset and is divided across unique forms of public transit, such as rail and bus, and is measured in meters. Real distance to the nearest public transportation stop, in terms of time and distance to the public transit stop that provides access to the CBD, is calculated for a random sample of each city. Moreover, this model uses the measured efficiency of accessing public transportation to reach the CBD as an independent variable as well. The efficiency metric is measured as the percent change in transit duration from driving to the CBD and utilizing public transit to reach the CBD. Here, efficiency is measured as a ratio to control for initial distance from the CBD.

Demographic data is also used as independent variables in the model, as they can provide controls for unobserved characteristics. Consumption behavior influences property value by outlining preferences for housing consumption but is difficult to measure as an unobserved characteristic. Instead, census data on poverty rate and income level in the census block group can be used to control for this unobserved characteristic as independent variables. In the model, poverty rate, median income, and ratio of commuters utilizing public transit are used as independent variables. Poverty rate is measured as the ratio of people in a block group classified under the poverty status, median income is measured as the median income of the census block group of the property, and ratio of public transit commuters is measured as the ratio of workers above the age of 16 that utilize public transit as their main form of commute.

As mentioned in the Data section, the Google Maps API sets quotas on the volume of data that can be queried, preventing the real distance measure from being applied to the entire population of each city. As a result, the model used in this research paper was utilized in two different ways: measuring the impact of straight-line distance to public transportation across the entire population and measuring the impact of real distance to practical public transportation across a random sample of size 50,000 from each of the examined cities. As outlined, the straight-line distance model introduced by this research project is:

$$\frac{\Delta \text{Price}}{\text{ft}^2}, \frac{\text{Price}}{\text{ft}^2} = \beta_0 + \beta_1 * \text{BusDistance} + \beta_2 * \text{RailDistance} + \beta_3 * \log(\text{CBDDistance}) + \beta_4 * \text{Income} + \beta_5 * \text{PovertyRate} + \beta_6 * \text{TransitRatio} + \varepsilon$$

Whereas the map-determined distance model outlined for the random sample of actual distances is:

$$\frac{\Delta \text{Price}}{\text{ft}^2}, \frac{\text{Price}}{\text{ft}^2} = \beta_0 + \beta_1 * \text{TimeToPublicTransitStop} + \beta_2 * \log(\text{DriveTimeToCBD}) + \beta_3 * \text{Income} + \beta_4 * \text{PovertyRate} + \beta_5 * \text{TransitRatio} + \beta_5 * \frac{\text{TransitTimeToCBD} - \text{DriveTimeToCBD}}{\text{DriveTimeToCBD}} + \varepsilon$$

In the above models, "BusDistance" represents the straight-line distance to the nearest bus stop from the current residential parcel. "RailDistance" is similar to "BusDistance" but represents straight-line distance to rail stations. "CBDDistance" is the computed straight-line distance to the CBD from the residential parcel. "TimeToPublicTrasitStop" is the Google Maps API determined time in seconds to reach the first public transit stop on a route to the central business district. "DriveTimeToCBD" is the Google Maps API determined time in seconds to drive to the CBD. "TransitTimeToCBD" is the Google Maps API determined time in seconds to arrive at the CBD from the first public transportation stop. Time taken to reach the public transit stop was removed from the transit time to address collinearity between the two variables. "Income", "PovertyRate", and "TransitRatio" are all census block group variables that are used as controls for demographic factors on consumption. Here, "Income" refers to median household income for the block group, "PovertyRate" refers to the poverty rate of the block group as specified earlier, and "TransitRatio" refers to the ratio of workers commuting by public transit in the block group as specified earlier.

The empirical results of the model should demonstrate an inverse relationship between distance from the nearest public transportation stop and magnitude of increased property value over time. As the distance between a residential property and the nearest public transportation stop increases, the value of the residential property should decrease. Moreover, by examining city specific characteristics and the results of the regression, this research project expects to demonstrate a positive relationship between the magnitude of observed effect of proximity to public transportation and characteristics of monocentric cities. Cities with a greater coefficient on the logged distance to the central business district variable should see greater magnitude in the relationship between proximity to public transportation and property value.

In the control variables in the regression, the median household income of a block group should have a positive relationship with residential property value. The poverty rate of the census block group should also have a negative relationship with property value. While a causal relationship between these variables and the dependent variable may be difficult to discern due to possible simultaneous causality, the use of these variables in the regression is important to control for consumption behavior around a residential property.

The model examining real distance to the nearest practical public transit stop should outperform the model using straight-line distances. When optimizing for amenity access in finding a residence, it would be irrational to optimize toward an inaccessible public transportation stop even if it is geographically closer, so the Google Maps returned data should provide greater insight on preferences.

VI. Results

As specified before in the empirical model outlined in the above section, the dependent variable of the model was measured as both a variable indicating growth and a variable indicating gross value. In both instances, the natural logarithm of the dependent variable was used as a measure to counteract heteroscedasticity in prices. The difference in the natural logarithms of prices across two times is the percent growth from the initial time to the final time.

Because the time difference across the two periods is a year for each regression, the dependent variable indicating growth is the percent annual growth in property value.

VI.1 New York City

New York City's transit system differs from the other explored systems of public transportation in its inclusion of a subway system. The Long Island rail system was excluded from the regression for New York City because of the high distance of the rail stations from the majority of the data points in the population. There was little correlation between distance to the central business district, distance to the nearest bus stop, and distance to the nearest subway station, and these variables were used in the regression.

Straight-line Distance Model

The results of the OLS regression on residential properties in New York, NY using straight-line distances can be seen below in Table 1.

Table 1: Regression outcomes for the city of New York using straight-line distance measure for public transportation and the central business district

	Regression Models	
	log(price/ft^2) coefficient (std. error)	**Δlog(price/ft^2)** coefficient (std. error)
Distance to Bus Stop	.0000398*** (3.94e-06)	-.0000264*** (1.37e-06)
Distance to Subway Stop	9.42e-07*** (3.08e-07)	-7.29e-06*** (1.08e-07)
log(Distance to CBD)	-.555581*** (.0018572)	-.028503*** (.0006095)
Median Income	3.14e-06*** (2.64e-08)	1.77e-07*** (8.33e-09)
% of Transit Commuters	-.208747*** (.0037156)	.043928*** (.0012282)
Poverty rate	-.3474259*** (.0059329)	.0481876*** (.002006)
Constant	11.17913*** (.018686)	.3207879*** (.0060617)
Number Observations	553,575	553,573
R-squared	0.2498	0.0442

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

In the regression for the model based off of the dependent variable of gross property value per square feet, despite all being statistically significant, most coefficients on the regressors are not economically significant, with the exception of the logged distance to the central business district. The coefficient on the distance to the nearest bus stop for the gross value model indicates that a meter move away from a bus stop results in a .004% increase in property value, and the coefficient on distance to the nearest subway station demonstrates an even milder effect for moves away from subway stations. The coefficient on the CBD distance regressor indicates that a percent move away from the central business district results in a .56% decrease in property value. Other signs on coefficients are consistent with expectations, such as the positive relationship between median income and property value. With an r-squared value of .2498, the gross property value per square foot model for straight-line distances in New York account for around 25% of the variation in the data.

The growth model for residential property value does not have high explanatory power with an r-squared value of .0442 but has relatively consistent coefficient signs. Negative coefficients on the two distance from public transit variables indicate that moving away from public transportation stops decrease the annual growth rate of property value. However, the magnitudes of the coefficients are not high, demonstrating a lack of economic significance despite statistical significance. Additionally, the signs on the coefficients for the distance to public transportation regressors is the opposite in the growth model as compared to the gross property value model. This would suggest that the observed magnitude of the rese regressors I the gross property value model should diminish between years, but the economic significance is not high.

Map-determined Distance Model

With map-determined distances, specific times to reach the nearest useful public transit stop is recorded, presenting walkability as a variable to examine. Munoz-Raskin determined walkability to be ten minutes or around 800 meters in their study of public transportation in Bogota (2009). The same measure is applied to the mad-determined distance regression for New York and all subsequent regressions.

Table 2: Map distance regression outcomes for the city of New York, NY

	Regression Models			
	log(price/ft²)	**log(price/ft²) walkable**	**log(price/ft²) not walkable**	**Δlog(price/ft²)**
	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)
Time to Nearest Stop (sec)	-.0002204***	-.0001382***	-.0002928***	-.0000299***
	(5.83e-06)	(.0000127)	(.000014)	(2.08e-06)
log(Time to CBD)	-.7879977***	-.8245615***	-.6815757***	.0123606***
	(.0079168)	(.0098352)	(.012783)	(.0025016)
Median Income	2.96e-06***	2.85e-06***	3.29e-06***	9.45e-08***
	(8.72e-08)	(1.07e-07)	(1.44e-07)	(2.75e-08)
% Transit Commuters	-.3806696***	-.355304***	-.4623372***	.0367913***
	(.0119371)	(.0140269)	(.0224424)	(.0039901)
Poverty rate	-.4302227***	-.4837416***	-.2798699***	.0214001***
	(.0198141)	(.0231766)	(.0384528)	(.006583)
Public Transit Efficiency Ratio	-.497394***	-.4788904***	-.5645117***	-.072194***
	(.0061914)	(.0072224)	(.0120333)	(.0018797)
Constant	12.31053***	12.56016***	11.56438***	-.0087539
	(.0657148)	(.081801)	(.1054397)	(.0204705)
Number Observations	48,835	35,555	13,280	48,835
R-squared	0.3172	0.3245	0.3098	0.0626

*** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$

The above Table 2 illustrates the regression results of the map-determined distance model for the city of New York. Table A2.2 in the appendix demonstrates the hierarchical model of the gross property value model, measuring the significance of the change in r-squared with the addition of each explanatory variable. The signs on the coefficients for the logged gross value models are consistent with the expectations outlined in previous sections, deviating from the signs on the coefficients on public transit distance demonstrated in the straight-line distance model. Additionally, the r-squared measure of .3172 is significantly larger than the same measure for the straight-line distance model for New York. This observation is generally consistent in the rest of the examined cities, as the map-determined distance model demonstrates consistent or higher r-squared values in two other cities.

In the first regression of Table 2, a second move away from the nearest public transit stop decreased property value by .022%. For each minute move away from the nearest public transit stop, property value decreases by 1.32%. When separating the regressions by walkability, the magnitude of the regressor increased for the time to the nearest stop regressor when regressing on non-walkable distances. This could be an indication that the willingness to walk to public transportation is high in New York City, and that the measure used for walkability is not sufficiently large. Additionally, the statistically significant coefficient on the public transit efficiency ratio has relatively high economic significance, as property value decreases by 49% if the transit time to the central business district doubles the drive time.

Here, the growth model again has low explanatory power, but the statistical significance of the coefficients and the directions of the regressors provide some potential insight. The regressor for time to the nearest stop has a negative coefficient, indicating that growth rate slows with moves away from the nearest public transit stop. The positive coefficient on the time to the CBD regressor could be an indication that the magnitude of effect that city centrality has on property value is waning, but it could also be a signal of polycentricism in New York City. Many of the boroughs of New York have distinct central business districts, and this observation could be altering the CBD distance coefficient.

Comparison of Distance Models

As stated above, the map-determined distance model regressions outperform the straight-line distance model regressions in terms of sign consistency and significance. These models also generally have higher explanatory power than the straight-line distance models as well, with the exception of Boston. The distance measures returned from the Google Maps API are also a better approximation for CBD and public transportation access for individual residential parcels. These measures incorporate city traffic, transit stop significance and public transit scheduling and are not distorted by map projections performed in ArcGIS. These factors indicate that map-determined distances serve as more accurate measures for actual access to public transportation. Because the map-determined distances serve as better measures for public transportation access, only the map-determined distance model will be discussed in following sections. The regression tables for each of the straight-line distance models for the following sections are available in the appendix. The comparison between the straight-line distance model and the map-determined distance model for the other three cities is generally consistent with the comparisons made for the city of New York. Because of the difference in r-squared values for the city of Boston, a discussion of Boston's map-determined distance measure is included.

VI.2 Boston

As discussed before, Table A2.3 in the appendix demonstrates the results of the regression on Boston residential properties using the straight-line distance model. The statistical and economic significance of the regressors is generally similar to the results of the regression on the same model for New York, with the exception of the distance to the nearest rail stop, which has a negative coefficient in the Boston gross value regression, indicating a that meter move away from a rail stop yields a .005% increase in property value. While the coefficient on this regressor is statistically significant, it is not economically significant. Like New York in the same model, the distance to the CBD regressor has a statistically and economically significant coefficient, indicating that a one percent move away from the CBD results in a .36% decrease in property value.

Table 3: Map distance regression outcomes for the city of Boston, MA

	Regression Models			
	log(price/ft²)	**log(price/ft²) walkable**	**log(price/ft²) not walkable**	**Δlog(price/ft²)**
	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)
Time to Nearest Stop (sec)	-.0003096***	-.0001826***	-.0001772***	.0000214***
	(8.06e-06)	(.0000159)	(.0000193)	(1.59e-06)
log(Time to CBD)	-.6512646***	-.69526***	-.2924424***	-.0038965***
	(.0044282)	(.0047456)	(.0136335)	(.0008337)
Median Income	3.99e-06***	3.83e-06***	5.09e-06***	-1.57e-10
	(7.16e-08)	(7.65e-08)	(1.75e-07)	(1.33e-08)
% Transit Commuters	-.4914644***	-.5186758***	-.2913614***	.044204***
	(.0159733)	(.0178286)	(.0371347)	(.0029054)
Poverty rate	-.5110098***	-.4731369***	-.5402394***	.0127803***
	(.0247322)	(.0275024)	(.0417763)	(.003262)
Public Transit Efficiency Ratio	-.0661312***	-.0453842***	-.1164611***	.0192208***
	(.0053015)	(.005896)	(.0137387)	(.0008452)
Constant	10.39531***	10.68559***	7.534928***	.0753883***
	(.0339211)	(.0364293)	(.1072929)	(.0057482)
Number Observations	48,690	40,312	8,378	48,672
R-squared	0.4514	0.4451	0.3142	0.0146

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

The above Table 3 demonstrates the results of the regression for the city of Boston using the random sample of 50,000 parcels with Google Maps API derived distances. The hierarchical linear model for this gross value regression is demonstrated in Table A2.4 In the appendix. The change in the r-squared value between adding each explanatory variable is statistically significant.

Similar to the New York regression, there appears to be little impact in delineating walkable and non-walkable regressions, other than changing already high t values from standard errors. In general, a minute move away from the nearest practical public transit stop yields a 1.9% decrease in gross property value, calculated in the same fashion as New York. A departure from the New York model is the value of the coefficient on the efficiency ratio regressor. A 100%, or 1 unit, move in the regressor indicates that public transit is slower than driving by an additional drive time to the CBD. This move would yield a 7% decrease in property value, which is significantly smaller than the effect seen in New York. This could result from the vastly different city-wide efficiency ratios, as Boston has a city-wide mean efficiency ratio of .18 while New York has a city-wide efficiency ratio of .445. This indicates that the median length trip to the CBD in Boston increases the travel duration over driving time by 18% whereas the median length trip to the CBD in New York increases travel duration over driving time by about 45%. The descriptive statistics for these variables can be seen in Table A1.6 and Table A1.7 in the appendix.

The growth model for Boston does not have high explanatory power. Time to the CBD is statistically significant and has a negative sign on the coefficient. This indicates that moves away from the CBD negatively impact the growth in property value of a parcel. The coefficient on the time to the nearest public transit stop is positive, but this could be indicative of waning influence rather than negative influence of distance to public transportation on property value. Since the growth rate represents the slope of the property's value over time, measuring changes in the growth rates, as done by the growth model, measures acceleration of value.

Although the map-determined distance models in this research project generally have consistent or higher measures than the straight-line distance models for the same cities, Boston is the exception. The r-squared measure of the map-determined distance model for gross property value in Boston is .4514 and the r-squared measure for the straight-line distance model is .5091, greater in value than the map-determined distance model and the straight-line distance models of each of the other cities. The difference in r-squared measure likely arises from the higher explanatory power of the straight-line distance to the CBD compared to the map-determined time to the CBD. When substituting the map-determined distance measure for the straight-line distance, the r-squared measure for the map-determined distance model becomes .5112.

This suggests that location of a residential property is being optimized with respect to the straight-line distance to the CBD. A potential reason for this observation could be traffic. The correlation coefficient between driving time and driving duration to the CBD is .88 for the sample, significantly lower than the correlation coefficient of .986 between the two in New York. Because the Google Maps API considers traffic in its calculation of estimated time, it is possible that this relatively lower correlation coefficient is an indication of unpredictable traffic; the correlation coefficient of .88 for Boston indicates that the driving duration does not vary perfectly with the driving distance. The public transit efficiency ratio of Boston is the lowest in median compared to the other four cities, meaning that it has the most efficient public transportation service. The MBTA's transit system provides significantly better predictability in terms of commute time as the correlation coefficient between public transit distance and public transit duration is .965. Given better public transit predictability, commuters could be optimizing for straight-line distance to the CBD through optimizing for distance covered by the MBTA's rail transit, which operates fairly linearly toward the CBD. This is supported when altering the model to accommodate this observation. Substituting logged distance covered by transit for time to the CBD in the model further increases the r-squared value to .5171, demonstrating that logged transit distance is an even better measure than straight-line CBD distance for capturing variation in the data. Straight-line distances, then, do not outperform the map-determined distance model.

VI.3 San Francisco

As discussed earlier in the paper, the city of San Francisco poses a unique circumstance in the existence of California Proposition 13. Proposition 13 capped the rate at which property value would grow according to tax assessment to two percent, starting in 1978 when the proposition was passed (Cal. Const. Art. XIIIA Sec. 2). Additionally, the article indicates that the assessed value is only reset whenever the residential property changes ownership. Because tax-assessor data is being used in this research project, this legislation skews assessed property prices to values close to the previous sale price. Given the fluctuation of property value in San Francisco, assessed property value is not a good measure for properties in San Francisco with very old previous sale dates.

The scope of the regressions for the straight-line distance model and the map-determined distance model was limited to properties with previous sale dates dating after 2000, and the sell year was included in the model for San Francisco. Year sold is expected to be positive given the general trends of increasing property prices over time. Table A2.5 in the appendix illustrates the results of the straight-line distance model regression. Table 4 below holds the regression results for the map-determined distance model.

Table 4: Map-determined distance regression outcomes for the city of San Francisco, CA after 2000

	Regression Models			
	log(price/ft²)	**log(price/ft²) walkable**	**log(price/ft²) not walkable**	**Δlog(price/ft²)**
	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)
Time to Nearest Stop (sec)	-.0001744***	-.0001484***	-.0002009**	.0001182***
	(.0000183)	(.0000286)	(.0000502)	(.0000361)
log(Time to CBD)	-.3119633***	-.3159164***	-.0198465	.0774911***
	(.008747)	(.0087985)	(.0657611)	(.017037)
Median Income	2.79e-06***	2.77e-06***	2.98e-06***	-8.14e-07***
	(1.01e-07)	(1.04e-07)	(4.23e-07)	(1.84e-07)
% Transit Commuters	.2080581***	.2110075***	.0179254	-.0137853
	(.03599)	(.0372042)	(.1404475)	(.0674676)
Poverty rate	-.689767***	-.7204341***	-.3161716***	-.2604923***
	(.061432)	(.0635609)	(.2223983)	(.0990996)
Public Transit Efficiency	-.2284676***	-.2061237***	-.4211202***	-.0159505
Ratio	(.0104701)	(.0109932)	(.0324324)	(.020064)
Year Sold	.1027199***	.101998***	.1122741***	.1981237***
	(.0023357)	(.0024341)	(.0081517)	(.0047883)
Constant	-198.2086***	-196.7399***	-219.5133***	-398.7923***
	(4.705439)	(4.903065)	(16.45912)	(9.66676)
Number Observations	9,232	8,404	828	9,037
R-squared	0.4021	0.4072	0.3638	0.1969

*** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$

The two models using straight-line distance and map-determined distances have similar explanatory power but differ significantly in the coefficients on the regressors for distance to public transportation. While both of these coefficients are positive in the straight-line distance model, the time to the nearest useful transit stop variable in the map-determined distance model has a negative coefficient. The map-determined distance model indicates that a minute move away from the nearest useful public transit stop yields a 1.05% decrease in price per square foot of a residential property. This discrepancy is seen across New York and Boston models as well. The hierarchical linear model for this model is shown in Table A2.6 in the appendix. The table demonstrates significant F statistics for each of the changes in r-squared from the addition of independent variables.

Dividing the regression between walkable and not walkable distances to the nearest public transit stop presents very few observations for the model over non-walkable stops.

Because the dataset on San Francisco residential properties only covers the city limits of San Francisco, boundaries of the dataset limit the maximum distances that can be recorded. The regression over non-walkable stops has few statistically significant regressors, and the signs on the coefficients for these variables changes for distance to the central business district and the proportion of public transit commuters.

The growth models for property value has higher explanatory power than the previously explored models, but this likely stems from the year sold regressor. Because of the limit on growth imposed by Proposition 13, it is likely that the year sold variable is picking up the capped growth rate as its coefficient.

VI.4 Atlanta

Analyzing results from Atlanta, GA is difficult because of the aforementioned data availability inconsistency. As Fulton County does not disclose residential property living square footage, lot area square footage had to be utilized as a proxy. To provide a better comparison across cities with Atlanta, the gross property value model was recomputed for price per square foot in regard to lot area for each of the other cities.

Comparisons are encumbered further by the sprawl of Atlanta. The prevalence of suburban homes in the Atlanta dataset resulted in numerous suburban residential parcels in the random sample. As the MARTA transit system does not have good reach in suburban areas, the Google Maps API was unable to find public transportation to the CBD from a significant portion of the residential parcels. This results in only parcels with access to public transportation being included in the regression.

The results of the straight-line distance model regression on the entire population of residential parcels in Atlanta, GA can be seen in Table A2.7 in the appendix. The below Table 5 illustrates the results of the regression on the residential properties from Atlanta, GA. The hierarchical linear model is shown in Table A2.8 in the appendix. For each of the added independent variables, the F statistic on the change in r-squared value is significant. Unlike previous straight-line distance models, the regression for Atlanta demonstrates negative coefficients on the distance to bus and rail stops, consistent with the original expectations. For example, the model for straight-line distances holds that a meter move away from the nearest bus stop yields a .007% decrease in property value per square foot of lot area.

Table 5: Map-determined distance regression outcomes for the city of Atlanta, GA

	Regression Models			
	log(price/ft^2)	**log(price/ft^2) walkable**	**log(price/ft^2) not walkable**	**Δlog(price/ft^2)**
	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)	coefficient (std. error)
Time to Nearest Stop (sec)	-.0001862***	-.0003662***	-.0000828***	6.10e-07
	(.0000123)	(.0000613)	(.0000167)	(9.10e-07)
log(Time to CBD)	-.82578***	-1.107714***	-.3995468***	.0000191
	(.0190059)	(.0256861)	(.0276855)	(.0016701)
Median Income	8.46e-06***	.0000166***	4.31e-06***	6.32e-08***
	(1.94e-07)	(4.14e-07)	(2.02e-07)	(1.75e-08)
% Transit Commuters	-.8243911***	-.9574674***	-.294273**	-.0012267
	(.073115)	(.0908647)	(.1145994)	(.0095215)
Poverty rate	-3.785597***	-3.053274***	-3.858357***	-.0242306***
	(.0584539)	(.0800467)	(.0963548)	(.0072604)
Public Transit Efficiency Ratio	-.3583416***	-.3535749***	-.3445803***	-.0007184
	(.0122029)	(.0184304)	(.0163624)	(.0011927)
Constant	9.262188***	10.71136***	6.256045***	-.0046001
	(.1314172)	(.1768831)	(.1972295)	(.011933)
Number Observations	39,512	20,277	19,235	39,512
R-squared	0.3233	0.4214	0.2166	0.0024

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

Across the three gross price models, statistical significance is consistent for the coefficients on the regressors. However, when limiting the regression to walkable public transportation, the coefficient on the time to the nearest stop regressor doubles in magnitude. For walkable distance public transportation stops, a minute move away yields a 2.2% decrease in residential property value. The signs on other coefficients for the regressors are consistent with expectations as well, as median income is positively related to property value and poverty rate is negatively related.

The efficiency variable has a statistically and economically significant coefficient as well. Since more positive efficiency ratios indicate slower public transportation, the model indicates that adding the time of a drive commute to the CBD to the time of a transit commute to the CBD decreases property value by 35%.

The growth models of the map-determined distances and the straight-line distances have little explanatory power. This observation is consistent with the models examined in other cities, as the growth models rarely had high explanatory power. Given the small r-squared value of .0024, the model could be regressing on noise in the data.

Lot Area Model Comparison

To better facilitate cross-city comparisons with Atlanta, map-determined distance models regressing on gross property value were recomputed for each of the other three cities. Table 6 demonstrates the results of this regression.

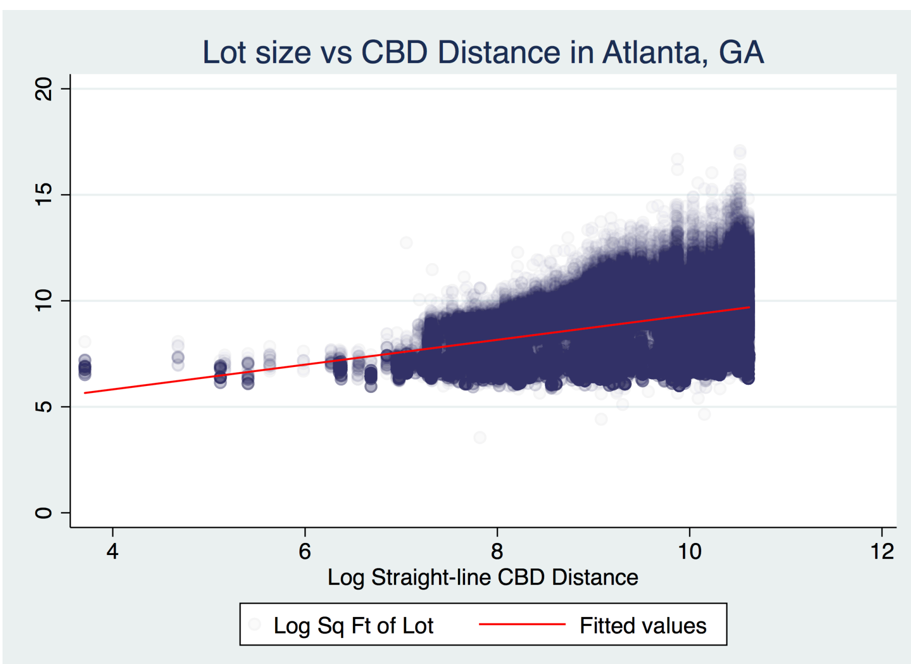
Table 6: Model comparison across the four examined cities using lot area as a proxy for living area square footage

	Regression Models			
	Atlanta **log(price/lot ft²)** coefficient (std. error)	**Boston** **log(price/lot ft²)** coefficient (std. error)	**New York City** **log(price/lot ft²)** coefficient (std. error)	**San Francisco** **log(price/lot ft²)** coefficient (std. error)
Time to Nearest Stop (sec)	-.0001862*** (.0000123)	-.0005286*** (.0000139)	-.0002204*** (5.83e-06)	-.0004172*** (.000037)
log(Time to CBD)	-.82578*** (.0190059)	-1.403127*** (.0071141)	-.7879977*** (.0079168)	-.7745336*** (.0357443)
Median Income	8.46e-06*** (1.94e-07)	4.36e-06*** (1.04e-07)	2.96e-06*** (8.72e-08)	5.02e-06*** (2.26e-07)
% Transit Commuters	-.8243911*** (.073115)	.3975185*** (.0230118)	-.3806696*** (.0119371)	.0639786 (.084734)
Poverty rate	-3.785597*** (.0584539)	-.3806989*** (.0299859)	-.4302227*** (.0198141)	-.7255296*** (.1132176)
Public Transit Efficiency Ratio	-.3583416*** (.0122029)	-.0536143*** (.0080118)	-.497394*** (.0061914)	-.3446908*** (.0273428)
Year sold				.090688*** (.0048699)
Constant	9.262188*** (.1314172)	15.11027*** (.0526281)	12.31053*** (.0657148)	-171.1927*** (9.789648)
Number Observations	39,512	48,720	48,865	5,018
R-squared	0.3233	0.5453	0.4550	0.3101

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

Among the four gross property value models in Table 6, Atlanta has the smallest magnitude coefficient on the time to the nearest stop regressor. Additionally, Atlanta also has the smallest magnitude coefficient on the time to the central business district regressor. These observations would indicate that the value of being closer to the nearest public transit stop and of proximity to the central business district is lowest for Atlanta among the four cities examined. The r-squared measure for each of the models is markedly larger than the models utilizing living area square footage. This could be a result of the bid rent curve. Land is more expensive in the city center and decreases in price as distance to the central business district increases. Since the measure is used in the denominator of the dependent variable, the increasing lot sizes away from the CBD depress the dependent variable further, and greater explanatory power is achieved through the relationship between land and distance to the central business district. Figure 6 below demonstrates this trend.

Figure 6: Atlanta lot size against distance to the central business district



VII. Discussion

On the whole, there are several observations to make about the results of the regressions from the results section. One primary observation is that the regressors used in the model are not salient regressors for the growth model. When using change in property value as the dependent variable, low explanatory power was coupled with economically insignificant coefficients in the results. This observation could indicate that proximity to public transportation augments property value but does not change the growth rate of the property value. This difficulty in generating a salient model could be driving the lack of literature on the impact of proximity to public transportation on the growth rate of property value.

Table 7: Table aggregating map-distance models of each city using price per living area square footage

	Regression Models			
	Atlanta	**Boston**	**New York City**	**San Francisco**
	log(price/lot ft^2)	**log(price/ft^2)**	**log(price/ft^2)**	**log(price/ft^2)**
	coefficient	coefficient	coefficient	coefficient
	(std. error)	(std. error)	(std. error)	(std. error)
Time to Nearest Stop (sec)	-.0001862***	-.0003096***	-.0003822***	-.0001744***
	(.0000123)	(8.06e-06)	(8.37e-06)	(.0000183)
log(Time to CBD)	-.82578***	-.6512646***	-1.125122***	-.3119633***
	(.0190059)	(.0044282)	(.0121843)	(.008747)
Median Income	8.46e-06***	3.99e-06***	2.57e-06***	2.79e-06***
	(1.94e-07)	(7.16e-08)	(1.37e-07)	(1.01e-07)
% Transit Commuters	-.8243911***	-.4914644***	.0997304***	.2080581***
	(.073115)	(.0159733)	(.0171171)	(.03599)
Poverty rate	-3.785597***	-.5110098***	-.1441025***	-.689767***
	(.0584539)	(.0247322)	(.0263292)	(.061432)
Public Transit Efficiency	-.3583416***	-.0661312***	-.9813536***	-.2284676***
Ratio	(.0122029)	(.0053015)	(.009696)	(.0104701)
Year sold				.1027199***
				(.0023357)
Constant	9.262188***	10.39531***	14.50212***	-198.2086***
	(.1314172)	(.0339211)	(.1011903)	(4.705439)
Number Observations	39,512	48,690	48,835	9,232
R-squared	0.3233	0.4514	0.3172	0.4021

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

The gross property value regression results for the map-determined distance model in each of the four cities is demonstrated above in the aggregated Table 7. The dependent variable utilizes living area square footage in each case except for Atlanta. For each of the four cities, the gross value model provides consistent and meaningful results. Each city demonstrates differing degrees of monocentricity and differing magnitudes of impact of proximity to public transportation, but the effects were consistent across each of the map-generated distance models. According to the Bureau of Labor Statistic's report on employment location in high population metropolitan areas, New York has the highest proportion of jobs located in the central city, followed by San Francisco and Boston, respectively (2012). While not very precise, this employment distribution serves as a measure for the monocentricity of a city, as it demonstrates the central city's importance as an employment center. The proxy for degree of monocentricity in the examined model, the regressor on logged time to the central business district, was consistent with expectation for Boston and New York, as the more monocentric New York had a higher coefficient on the regressor, but this was not consistent with San Francisco. This difference could stem from the geographic location of San Francisco as the city is surrounded by numerous city centers. Comparing across models using living area as the square footage measure, the coefficients on the time to the nearest transit stop regressor generally align with the usage rate of the public transportation system. The APTA reports that quarter 1 2017 weekday usage for these three cities are ordered by New York, Boston, San Francisco (2017). The ordering of magnitudes of the coefficients on the time to the nearest transit stop regressor is Boston, New York, San Francisco. In terms of public transit efficiency, the ordering of efficiency is Boston, San Francisco, New York, as seen in Table A1.6, A1.7, and A1.8.

While Atlanta is difficult to compare because of the use of lot area as a proxy for living area square footage, re-computing the regressions on the other three cities using lot area yielded interesting results. In this regression, the sprawled-out Atlanta had one of the smallest magnitude coefficients on the regressor for distance to the central business district and the regressor for distance to the nearest public transit stop. As mentioned before in the results section, the number of observations used in the regression was significantly smaller than the total number of observations. This results from the Google Maps API being unable to return routes using public transportation for many suburban parcels, as the distance to the nearest public transportation stop is very high. As a result, while the Atlanta model demonstrates some of the sprawl of the city, the exclusion of the over 10,000 parcels without access to public transportation likely diminished the significance of this observation. Because these parcels were on the fringes of the city of Atlanta, these data points would have likely decreased the value of access to the CBD even further, better highlighting the urban sprawl of Atlanta.

The straight-line distance models examined in this research project yielded conflicting results as often the signs on coefficients would differ between cities, whereas this did not occur in the map-determined distance regression results. While most control variables in the regressions remained consistent between cities, the distance measure for the nearest bus and subway or rail stop had coefficients of significantly different magnitudes and signs. This could potentially stem from how ArcGIS handles map projections and nearest data points. To determine distances from ArcGIS between two points, a map projection must be selected on which to ground distances on; these projections distort values differently depending on the map projection selected. Over large distances, this does not constitute a concern, but numerous small distances are used and compared for the bus and rail distance measures. This is supported by the findings from each of the straight-line distance regressions. The sign on the distance to the CBD regressor is consistent across all four cities and the ordinal magnitude of the coefficients is similar to the values determined for the map-determined distance models. The distances to the nearest bus and rail stops are significantly smaller than the distances to the CBD, making distortions in the distances more impactful. Additionally, correlating the map-determined times and distances with the computed straight-line distances highlights the problem further. The largest correlation coefficient between the two variables across all four cities is .2126. ArcGIS could be selecting transit stops that are inaccessible or non-useful to commuters based entirely on a distorted distance.

An observation made from the use of map-generated distances with straight-line distances is the importance of the use of actual distances over simulated distances. Proximity is altered not only by walkability, but by access to meaningful public transit. Many of the stops selected by straight line distance in the dataset were substantially different from the stop selected by the Google Maps API. As highlighted before in the paper, in using map-generated distances, the stop chosen for proximity better represents access to the rest of transit network. This is significant in that previous literature primarily utilizes straight-line distance as the measure for proximity. This research project provides demonstrable results supporting the use of map-generated distances.

VIII. Conclusion

The goal of this research project was to provide cross-city comparison on the impact of proximity to public transportation on residential property value by producing a model with consistent results across cities of varying urban development. In developing this goal, this research project offers insight into city characteristics that bolster the value of public transportation. The results from this research project are consistent and reflect on the narrative of the monocentricity of each city, and generally support the hypothesis that cities with greater monocentricity would reflect greater magnitude of impact from proximity to public transportation.

The current model could be strengthened by introducing more property-related variables and the use of market value of residential properties rather than assessed values. Inconsistencies in data availability limited the scope of variables that could be included in a consistent model, and irregularities in the assessment of property value, such as in the case of proposition 13 and California, limited usable data for some cities. Utilizing a real estate API, such as that of Zillow, could greatly improve the explanatory power of the models explored in this research project. Moreover, greater sensitivity to polycentric features of a city could elicit stronger results in cities like New York where city centers develop in different boroughs.

Use of the coefficient on the regressor for distance to the central business district is not a very precise measure of monocentricity and city centrality. Developing a gravity map for employment and commerce centers could aid in both demonstrating the centrality of different cities and highlighting amenity access for residential properties in the model.

To my knowledge, this research project is the first study to use map-determined distances as the distance metric for determining the impact of proximity to public transportation on residential property value. In demonstrating the importance of the use of actual experienced times and distances, this study introduces more accurate measures for distances and methodology to attain them. These measures have strong implications for how commuter optimization is perceived relative to public transportation.

References

Alonso, W. (1960). A theory of the urban land market. *Papers and Proceedings of the Regional Science Association,6*(1), 149-157.

American Community Survey. (2016). American Community Survey 5-year estimates. Retrieved from https://www.census.gov/geo/maps-data/data/tiger-data.html

American Public Transportation Association. (2017). Transit ridership report first quarter 2017. Retrieved from http://www.apta.com/resources/statistics/Documents/Ridership/2017-q1-ridership-APTA.pdf

Arribas-Bel, D., & Sanz-Gracia, F. (2014). The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. *Urban Geography,35*(7), 990-997. doi:10.1080/02723638.2014.940693

Bowes, D. R., & Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics,50*(1), 1-25. doi:10.1006/juec.2001.2214

Bureau of Labor Statistics. (2012). Employment statistics for selected cities and their metropolitan areas. Retrieved from https://socds.huduser.gov/Bls_laus/emplstat.pdf

C.A. Const. art. XIIIA, § 1

Department of Innovation and Technology. (2016). Property assessment FY2016. Retrieved from https://data.boston.gov/dataset/property-assessment

Department of Innovation and Technology. (2017). Property assessment FY2017. Retrieved from https://data.boston.gov/dataset/property-assessment

DataSF. (2016). Assessor historical secured property tax rolls. Retrieved from https://data.sfgov.org/Housing-and-Buildings/Assessor-Historical-Secured-Property-Tax-Rolls/wv5m-vpq2

Debrezion, G., Pels, E. A., & Rietveld, P. (2007). The impact of railway stations on residential and commercial property value: a meta analysis. *The Journal of Real Estate Finance and Economics,35*(2). doi:10.1007/s11146-007-9032-z

Fulton County Open Data. (2018). Parcels 2016. Retrieved from http://gisdata.fultoncountyga.gov/datasets/parcels-2016

Fulton County Open Data. (2018). Parcels 2017. Retrieved from http://gisdata.fultoncountyga.gov/datasets/parcels-2017

Google. (2018). Google Maps directions API.

Munoz-Raskin, R. (2010). Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy, 17*(2), 72-84. doi:10.1016/j.tranpol.2009.11.002

Neff, J., & Dickens, M. (2015). 2015 public transportation fact book. *APTA Transportation Factbook , 66*, 1-68.

NYC Department of Finance. (2016). FY16 database files by tax class. Retrieved from http://www1.nyc.gov/site/finance/taxes/property-assessment-roll-archives.page

NYC Department of Finance. (2017). FY17 database files by tax class. Retrieved from http://www1.nyc.gov/site/finance/taxes/property-assessment-roll-archives.page

Pereira, R. H., Nadalin, V., Monasterio, L., & Albuquerque, P. H. (2012). Urban centrality: a simple index. *Geographical Analysis*, 45(1), 77-89. doi:10.1111/gean.12002

Ryan, S. (1999). Property values and transportation facilities: finding the transportation-land use connection. *Journal of Planning Literature,13*(4), 412-427. doi:10.1177/08854129922092487

Wang, Y., Potoglou, D., Orford, S., & Gong, Y. (2015). Bus stop, property price and land value tax: A multilevel hedonic analysis with quantile calibration. *Land Use Policy, 42*, 381-391. doi:10.1016/j.landusepol.2014.07.017

Appendix:

Appendix 1

Table A1.1: Descriptive statistics for straight-line distance model variables in Atlanta

Variable	Description	Mean (SD)	Minimum	Maximum
val_2017	2017 assessed property value (USD)	254694 (306990.6)	0	1.62e+07
bus_distance	Distance to nearest bus stop (meters)	1451.145 (2060.277)	0	16098.95
rail_distance	Distance to nearest rail stop (meters)	7639.059 (6771.899)	32.72601	39378.82
cbd_dist	Distance to the central business district (meters)	18484.16 (12336.35)		
med_income	Median income of census block group (USD)	76974.37 (47537.82)	0	250001
transit_rate	Ratio of workers commuting with public transportation	. 1785692 (.1045642)	0	.8099662
pov_rate	Ratio of households ascribed poverty status	.116152 (.1447436)	0	.8717949

Table A1.2: Descriptive statistics for straight-line distance model variables in Boston

Variable	Description	Mean (SD)	Minimum	Maximum
val_2017	2017 assessed property value (USD)	594936.3 (584397)	10000	2.31e+07
bus_distance	Distance to nearest bus stop (meters)	152.1994 (117.9843)	1.651375	964.2753
rail_distance	Distance to nearest rail stop (meters)	1100.019 (1223.017)	7.179052	6105.083
cbd_dist	Distance to the central business district (meters)	6110.872 (3752.745)	14.73792	15301.67
med_income	Median income of census block group (USD)	73048.92 (36623.36)	0	250001
transit_rate	Ratio of workers commuting with public transportation	.3124059 (.1485409)	0	1
pov_rate	Ratio of households ascribed poverty status	.1092213 (.135557)	0	1

Table A1.3: Descriptive statistics for straight-line distance model variables in New York City

Variable	Description	Mean (SD)	Minimum	Maximum
val_2017	2017 assessed property value (USD)	755536.9 (1011576)	14990	8.04e+07
bus_distance	Distance to nearest bus stop (meters)	164.4544 (125.3675)	0	2051.418
subway_distance	Distance to nearest subway stop (meters)	1724.498 (1688.645)	7.387786	8158.844
cbd_dist	Distance to the central business district (meters)	16684.34 (5599.578)	797.379	35964.8
med_income	Median income of census block group (USD)	67098.46 (27324.79)	0	250001
transit_rate	Ratio of workers commuting with public transportation	.4407504 (.1803958)	0	1
pov_rate	Ratio of households ascribed poverty status	.1114835 (.1082304)	0	1

Table A1.4: Descriptive statistics for straight-line distance model variables in San Francisco

Variable	Description	Mean (SD)	Minimum	Maximum
val_2015	2015 assessed property value (USD)	942019.8 (6533391)	0	9.96e+08
bus_distance	Distance to nearest bus stop (meters)	440.582 (519.3443)	4.400222	3066.915
rail_distance	Distance to nearest rail stop (meters)	742.8917 (589.7947)	6.927428	3539.562
cbd_dist	Distance to the central business district (meters)	6107.588 (3009.331)	23.18891	12938.06
med_income	Median income of census block group (USD)	95030.79 (42949.19)	0	250001
transit_rate	Ratio of workers commuting with public transportation	.3066732 (.1103492)	.0296663	.7413249
pov_rate	Ratio of households ascribed poverty status	.0618233 (.0819204)	0	.7580645

Table A1.5: Descriptive statistics for map-determined distance model variables in Atlanta

Variable	Description	Mean (SD)	Minimum	Maximum
timetotransit	Time to the first transit stop on route to CBD (seconds)	760.2773 (643.2254)	0	3506
drivingduration	Time taken to drive to the CBD (seconds)	1475.895 (615.1925)	54	9817
pub_trans_eff_ratio	Percent change from drive time to CBD to transit time to CBD	.8714821 (.6179954)	-1.007874	3.270066

Table A1.6: Descriptive statistics for map-determined distance model variables in Boston

Variable	Description	Mean (SD)	Minimum	Maximum
timetotransit	Time to the first transit stop on route to CBD (seconds)	386.3698 (269.9838)	0	2662
drivingduration	Time taken to drive to the CBD (seconds)	1214.999 (2024.5)	171	167696
pub_trans_eff_ratio	Percent change from drive time to CBD to transit time to CBD	.1795244 (.4300663)	-1.00578	1.747212

Table A1.7: Descriptive statistics for map-determined distance model variables in New York

Variable	Description	Mean (SD)	Minimum	Maximum
timetotransit	Time to the first transit stop on route to CBD (seconds)	460.9777 (298.573)	0	1838
drivingduration	Time taken to drive to the CBD (seconds)	2330.25 (2465.624)	427	155427
pub_trans_eff_ratio	Percent change from drive time to CBD to transit time to CBD	.4452225 (.3906521)	-.6337778	1.882559

Table A1.8: Descriptive statistics for map-determined distance model variables in San Francisco

Variable	Description	Mean (SD)	Minimum	Maximum
timetotransit	Time to the first transit stop on route to CBD (seconds)	284.597 (221.2765)	0	3149
drivingduration	Time taken to drive to the CBD (seconds)	1255.773 (425.6908)	59	2175
pub_trans_eff_ratio	Percent change from drive time to CBD to transit time to CBD	.3753227 (.4005374)	-1.011494	2.661017

Appendix 2

Table A2.1: OLS Regression Statistics for rent curve visualizations

	Regression Models			
	Atlanta **log(price/ft²)** coefficient (std. error)	**Boston** **log(price/ft²)** coefficient (std. error)	**New York City** **log(price/ft²)** coefficient (std. error)	**San Francisco** **log(price/ft²)** coefficient (std. error)
log(Distance to CBD)	-.2591685***	-.4632976***	-.4288159***	-.6668513***
	(..0031651)	(.0022539)	(.0017824)	(.0055392)
Constant	5.312266***	9.711935***	10.04355***	11.15906***
	(.0302284)	(.0193964)	(.0173638)	(.047813)
Number Observations	287,681	121,724	555,537	92,847
R-squared	0.0228	0.4035	0.1562	0.1350

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

Table A2.2: NYC map distances hierarchical regression

	Regression Models			
	controls **log(price/ft²)** coefficient (std. error)	**transit efficiency** **log(price/ft²)** coefficient (std. error)	**time to CBD** **log(price/ft²)** coefficient (std. error)	**time to stop** **log(price/ft²)** coefficient (std. error)
Median Income	3.31e-06***	3.13e-06***	2.99e-06***	2.96e-06***
	(1.25e-07)	(1.14e-07)	(9.06e-08)	(8.72e-08)
Poverty rate	-.3018477***	-.4383987***	-.3886122***	-.4302227***
	(.0235225)	(.0222993)	(.0200252)	(.0198141)
% Transit Commuters	. 3888699***	.0795772***	-.287665***	-.3806696***
	(.0129303)	(.0128191)	(.0118762)	(.0119371)
Public Transit Efficiency		-.297786***	-.4057671***	-.497394***
Ratio		(.0061287)	(.0053705)	(.0061914)
log(Time to CBD)			-.7591856***	-.7879977***
			(.0079765)	(.0079168)
Time to Nearest Stop (sec)				-.0002204***
				(5.83e-06)
Constant	5.540381***	5.837231***	11.89865***	12.31053***
	(.0130892)	(.0126073)	(.0646792)	(.0657148)
Number Observations	48,835	48,835	48,835	48,835
R-squared	0.0655	0.1254	0.2965	0.3172
Δ R-squared	-	0.0599	0.1711	0.0206
F for Δ R-squared	589.56***	2360.86***	9058.71***	1429.35***

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

Table A2.3: Regression outcomes for the city of Boston using straight-line distance measure for public transportation and the central business district

	Regression Models	
	log(price/ft^2)	**Δlog(price/ft^2)**
	coefficient	coefficient
	(std. error)	(std. error)
Distance to Bus Stop	.0002207***	.0000123***
	(.0000103)	(1.78e-06)
Distance to Rail Stop	-.000055***	-4.46e-06***
	(1.38e-06)	(2.60e-07)
log(Distance to CBD)	-.3645776***	.0085418***
	(.002608)	(.0003034)
Median Income	3.80e-06***	-4.85e-09
	(4.97e-08)	(7.65e-09)
% of Transit Commuters	-.4916011***	.030313***
	(.0104453)	(.0021669)
Poverty rate	-.4428822***	.0125994***
	(.0148809)	(.0020331)
Constant	8.830275***	-.0048669*
	(.0233995)	(.0025296)
Number Observations	121,669	121,634
R-squared	0.5091	0.0139

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

Table A2.4: Boston map distances hierarchical regression

	Regression Models			
	controls	**transit efficiency**	**time to CBD**	**time to stop**
	log(price/ft^2)	**log(price/ft^2)**	**log(price/ft^2)**	**log(price/ft^2)**
	coefficient	coefficient	coefficient	coefficient
	(std. error)	(std. error)	(std. error)	(std. error)
Median Income	6.71e-06***	6.48e-06***	4.29e-06***	3.99e-06***
	(8.61e-08)	(8.64e-08)	(7.19e-08)	(7.16e-08)
Poverty rate	-.4644235***	-.4385314***	-.5112216***	-.5110098***
	(.0292823)	(.0287877)	(.025153)	(.0247322)
% Transit Commuters	-.1096861***	-.1382029***	-.4198968***	-.4914644***
	(.0174691)	(.0177049)	(.0161199)	(.0159733)
Public Transit Efficiency		-.1183455***	.0115936**	-.0661312***
Ratio		(.0059245)	(.0048685)	(.0053015)
log(Time to CBD)			-.6992582***	-.6512646***
			(.0043571)	(.0044282)
Time to Nearest Stop (sec)				-.0003096***
				(8.06e-06)
Constant	5.391314***	5.435371***	10.55377***	10.39531***
	(.0110319)	(.0114579)	(.0342981)	(.0339211)
Number Observations	48,690	48,690	48,690	48,690
R-squared	0.2183	0.2248	0.4376	0.4514
Δ R-squared	-	0.0065	0.2128	0.0138
F for Δ R-squared	3983.88***	399.03***	25756.27***	1475.44***

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

Table A2.5: Regression outcomes for the city of San Francisco using straight-line distance measure for public transportation and the central business district after 2000

	Regression Models	
	log(price/ft^2) coefficient (std. error)	**Δlog(price/ft^2)** coefficient (std. error)
Distance to Bus Stop	.0000646***	-.0000124
	(4.30e-06)	(9.09e-06)
Distance to Rail Stop	.0000751***	.0000148**
	(3.72e-06)	(6.96e-06)
log(Distance to CBD)	-.269119***	.0623494***
	(.0030618)	(.0064166)
Median Income	2.51e-06***	-6.47e-07***
	(4.91e-08)	(9.48e-08)
% of Transit Commuters	.3662352***	-.0781911**
	(.018787)	(.034628)
Poverty rate	-.6847012***	-.2130991***
	(.0301174)	(.0503445)
Year sold	.1022692***	.1986085***
	(.0011771)	(.0024596)
Constant	-197.4536***	-399.7226***
	(2.371323)	(4.963204)
Number Observations	34,498	33,717
R-squared	0.4159	0.1980

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

Table A2.6: San Francisco map distances hierarchical regression after 2000

	Regression Models			
	controls	**transit efficiency**	**time to CBD**	**time to stop**
	log(price/ft²)	**log(price/ft²)**	**log(price/ft²)**	**log(price/ft²)**
	coefficient	coefficient	coefficient	coefficient
	(std. error)	(std. error)	(std. error)	(std. error)
Median Income	3.42e-06***	3.21e-06***	2.76e-06***	2.79e-06***
	(1.05e-07)	(1.06e-07)	(1.00e-07)	(1.01e-07)
Poverty rate	-.7265817***	-.6610723***	-.6983612***	-.689767***
	(.0661655)	(.0656691)	(.0619091)	(.061432)
% Transit Commuters	.3191166***	.2053028***	.2513366***	.2080581***
	(.0391558)	(.0385862)	(.0360516)	(.03599)
Year sold	.1058027***	.1055955***	.1035257***	.1027199***
	(.0025078)	(.0024827)	(.0023484)	(.0023357)
Public Transit Efficiency		-.1600302***	-.1929614***	-.2284676***
Ratio		(.0102567)	(.0098681)	(.0104701)
log(Time to CBD)			-.3259788***	-.3119633***
			(.0086671)	(.008747)
Time to Nearest Stop (sec)				-.0001744***
				(.0000183)
Constant	-206.819***	-206.2905***	-199.8045***	-198.2086***
	(5.045709)	(4.994968)	(4.731031)	(4.705439)
Number Observations	9,232	9,232	9,232	9,232
R-squared	0.2839	0.3020	0.3968	0.4021
Δ R-squared	-	0.0181	0.0948	0.0053
F for Δ R-squared	858.89***	243.44***	1414.59***	91.28***

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

Table A2.7: Regression outcomes for the city of Atlanta using straight-line distance measure for public transportation and the central business district

	Regression Models	
	log(price/ft²)	**Δlog(price/ft²)**
	coefficient	coefficient
	(std. error)	(std. error)
Distance to Bus Stop	-.0000788***	6.48e-07***
	(1.74e-06)	(1.63e-07)
Distance to Rail Stop	-8.73e-08	-3.38e-07***
	(7.38e-07)	(5.90e-08)
log(Distance to CBD)	-.5003894***	.0010402**
	(.0045467)	(.0004113)
Median Income	6.38e-06***	6.67e-08***
	(6.84e-08)	(6.95e-09)
% of Transit Commuters	-.5698942***	-.0050486
	(.0287373)	(.0030914)
Poverty rate	-4.108907***	-.0144234***
	(.0239378)	(.0023717)
Constant	7.803973***	-.0134753***
	(.0414804)	(.0039144)
Number Observations	283,843	283,837
R-squared	0.2641	0.0018

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

Table A2.8: Atlanta map distances hierarchical regression

	Regression Models			
	controls **log(price/ft²)** coefficient (std. error)	**transit efficiency** **log(price/ft²)** coefficient (std. error)	**time to CBD** **log(price/ft²)** coefficient (std. error)	**time to stop** **log(price/ft²)** coefficient (std. error)
Median Income	5.92e-06***	7.45e-06***	7.73e-06***	8.46e-06***
	(1.82e-07)	(1.86e-07)	(1.92e-07)	(1.94e-07)
Poverty rate	-4.012272***	-3.656357***	-3.776681***	-3.785597***
	(.0618088)	(.0609519)	(.0584425)	(.0584539)
% Transit Commuters	1.198989***	.1565879***	-.8149345***	-.8243911***
	(.073408)	(.0743688)	(.0725643)	(.073115)
Public Transit Efficiency Ratio		-.5441928***	-.3266184***	-.3583416***
		(.0119994)	(.012097)	(.0122029)
log(Time to CBD)			-.9591513***	-.82578***
			(.0160731)	(.0190059)
Time to Nearest Stop (sec)				-.0001862***
				(.0000123)
Constant	2.788845***	3.303626***	10.08631***	9.262188***
	(.0243702)	(.0266584)	(.1157182)	(.1314172)
Number Observations	39,512	39,512	39,512	39,512
R-squared	0.2304	0.2674	0.3196	0.3233
Δ R-squared	-	0.0370	0.0522	0.0037
F for Δ R-squared	4580.32***	2056.76***	3561.03***	228.53***

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