OH, THE PLACES YOU'LL MOVE: URBAN MASS TRANSIT INVESTMENTS' EFFECTS ON NEARBY HOUSING MARKETS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Yue Ke

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2020

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF APPROVAL

Dr. Konstantina Gkritza, Chair

Lyles School of Civil Engineering

Dr. Brigitte Waldorf, Co-Chair

Department of Agricultural Economics

Dr. Jon Fricker

Lyles School of Civil Engineering

Dr. Michael Wilcox

Department of Agricultural Economics

Approved by:

Dr. Dulcy Abraham

Head of the School Graduate Program

To my friends and family

ACKNOWLEDGMENTS

I am deeply indebted to Dr. Gkritza for taking a chance on recruiting me to Purdue. Without her guidance and support throughout my studies and research, this dissertation would not have been possible. I'm extremely grateful for Dr. Waldorf, whose encouragement and patience was integral at pushing me to complete this work. In addition, I would like to extend my gratitude to my committee members, Dr. Fricker and Dr. Wilcox, whose advice was fundamental throughout this research.

To Davis Chacon-Hurtado, Christos Gkartzonikas, and Lisa Losada Rojas, your friendship and help throughout these years have been invaluable. Additionally, I appreciate the feedback from Matt Sloggy. Amy, Minh and Michael, thanks for keeping me sane during TRB Annual Meetings. Finally, I want to thank Joe Myers for helping me with data acquisition.

Last but certainly not least, I would like to thank my family for putting up with me during my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Ke, Yue Ph.D., Purdue University, August 2020. Oh, the Places You'll Move: Urban Mass Transit Investments' Effects on Nearby Housing Markets. Major Professor: Konstantina Gkritza.

The last couple of decades have seen a renewed interest among urban transportation planners in light rail transit (LRT) systems in large cities across the United States (US) as a possible means of addressing negative transportation externalities such as congestion and greenhouse gas emissions while encouraging the use of public transit [1]. LRT infrastructure investments have also gained traction as a means of revitalizing decayed urban centres because transportation infrastructure developments are highly correlated with economic growth in surrounding areas [2].

The primary objective of this dissertation is to examine the externalities associated with LRTs during its construction and operations phases. In particular, three areas of concern are addressed: (1) The effect that proximity to LRT stations have on nearby single family residences (SFRs) throughout the LRT life-cycle; (2) the effect that directional heterogeneity between LRT stations, the central business district (CBD), and the SFR has on SFR prices; and (3) the longer term effects on nearby populations due to LRT operations. To address the first two research objectives, quasi-experimental spatial econometric models are used, while a-spatial fixed effects panel models are developed to address the last objective. The analyses primarily uses SFR sales data from 2001-2019, publicly available geographical information systems data, as well as demographic data from eight 5-year American Communities Surveys. Charlotte, NC, a medium-sized US city, is chosen as the site of analysis, both due to the relative novelty factor of its LRT in the region and data availability.

The results show that SFR values are positively associated with proximity to LRT stations in the announcement and construction phases but negatively associated with

proximity to stations once the LRT is operational. Additionally, potential homeowners with prior experience with LRT do not behave any differently than potential homeowners with no prior experience with LRT in terms of willingness to pay to live a certain distance from LRT stations. Further, directional heterogeneity is shown to be a statistically significant source factor in deciding the extent to which house-buyers are willing to pay to be near LRT stations. Lastly, distance from LRT stations is found to have no statistically significant effect on changes in the racial composition of nearby areas but has significant positive effects on educational attainment and average median incomes of residents living in nearby areas over time.

The contributions of this research are twofold. First, in addition to highlighting the need to use spatial econometric methods when analyzing the effect that LRTs have on surrounding real estate markets, this research provides a framework by which directional heterogeneity can be incorporated into these analyses. Second, this research adds to the existing pool of knowledge on land use externalities of LRT through incorporating the life-cycle of LRT from announcement to operations. Furthermore, this research examines the effects that LRT have on surrounding populations in transit-adjacent areas to provide a look at the broader effects of LRT over time.

A major challenge in the analyses conducted in this dissertation is its reliance on SFR sales data. Urban areas near LRT may contain additional land uses. In order to fully determine LRT's effects on its surrounding area, one should examine the proximity effects on all land use types. Furthermore, LRT stations and rail lines are assumed exogenous, which may not be the case. Public hearings and town halls during the planning phase may influence stations' locations. Future research should seek to understand how the circumstances surrounding the planning process could indirectly affect the socio-demographic characteristics in transit adjacent areas over time. Finally, additional research is needed to better understand the extent to which LRT affects urban intra- and inter-migration. Knowing the population repulsion and attraction of LRT can help planners design facilities to better serve the public.

1. INTRODUCTION

Research in and the practice of transportation engineering have grown increasingly reliant on quantitative methods over the course of the last half century. In addition to increased data availability, innovations in statistical methods and econometric modelling, particularly in the areas of spatial statistics and urban economics, have enabled researchers to develop complex models that incorporate multiple aspects of land use, transportation, and socio-economic and environmental sustainability. Many of these models have been applied by practitioners seeking to mitigate the impacts of urban sprawl and transportation externalities, such as pollution and congestion, to develop comprehensive urban plans [3]. These plans typically forecast housing demand, travel demand, and estimate the impacts that policy programs may have on urban development across space. Additionally, as transportation agencies seek to allocate scarce resources, results from such models may be used as part of a benefit-cost analysis.

Moreover, the last couple of decades have also seen a renewed interest among urban transportation planners in light rail transit (LRT) systems in large cities across the United States (US) as a possible means of addressing negative transportation externalities such as congestion and greenhouse gas emissions while encouraging the use of public transit [1]. LRT infrastructure investments have also gained traction as a means of revitalizing decayed urban centres because transportation infrastructure developments are highly correlated with economic growth in surrounding areas [2].

1.1 Research Motivation

Prior to the construction of LRT systems, DOTs try to predict the effects an LRT may have on surrounding areas by submitting reports including environmental impact assessments as required by the National Environmental Policy Act and traffic impact analyses as recommended by the Institute of Transportation Engineers. Forecasting the effects on the environmental and local traffic conditions is an important first step but the presence of LRT systems can affect other dimensions of a city, including land use practices and local housing markets which in turn may lead to demographic changes in the urban environment. Additionally, after these systems become operational, there are no requirements for follow-up studies on the realized impacts of LRTs. Although this void has been partially filled by previous research, most studies only consider the effects of an LRT becoming operational without accounting for changes that may have occurred during the construction phase due to the announcement effect. Thus, the primary objective of this dissertation is to examine these additional effects that LRT may have on its surrounding areas during the construction and operations phases.

Other forms of urban mass transit, such as heavy rail transit (HRT; e.g., commuter rail lines) and bus rapid transit (BRT) have also regained popularity for these reasons. HRT lines are typically designed to move people from suburbs or edge cities into a central business district (CBD), rather than around a CBD as is typically the case of LRT. Because of this, HRT's effects on housing markets may be quite different since housing markets in suburbia may have different characteristics than housing markets in cities. Because this dissertation aims to better understand the effects of urban mass transit's effects on nearby housing markets, HRT lines are not considered in the analysis. Similarly, many BRT lines serve the same function as HRT. Further complicating the analysis, BRT systems can be implemented in a myriad of ways: BRTs may have dedicated busways, exclusive highway and street lanes, part-time exclusive lanes, heavy bypass lanes, use shoulders during congestion, share lanes with HOV traffic, as well as use a wide variety of payment systems. The heterogeneity in implementation may lead to varying degrees of effectiveness and efficiency between BRT systems causing homeowners to value proximity to BRT stations in different ways. Thus, for the purpose of this research, only LRT is considered.

1.2 Research Questions

The main objective of this dissertation is to better understand the effects that LRT may have on housing markets in surrounding areas during the construction and operations phases. To aid in the analysis of this topic, three research questions are posed:

1. **What are the (or are there any) differences in the capitalization of amenities access between residents living near a LRT in an area with no previous mass transit options (i.e., a new LRT line; not counting traditional bus services) compared to residents living near a LRT extension line (i.e., an area with previous LRT)?** (RQ1)

   The corresponding objective involves exploring the effect that homeowners' previous experience [or lack thereof] has on how they value LRT access as revealed in a hedonic price methods-approach (e.g., do past experiences with LRT affect homeowners' willingness to pay to be located near an LRT station?).

2. **What are the effects that a (new or existing) LRT line has on housing markets when one takes both distance from the line and relative directions to both the nearest LRT station and CBD into account?** (RQ2)

   Previous literature only looks at distance, and most of it is a-spatial. The objective herein is to expand on the spatial related literature by incorporating directional vectors in the analysis. For instance, assume there are two homeowners with homogenous characteristics equidistant from a LRT station but one is located north of the station and the other is south; if the CBD is further north of the LRT station, do these two homeowners value the LRT in the same way?

3. **How do LRT affect urban demographics?** (RQ3)

   As LRT is expected to affect housing markets in surrounding areas, it follows that changes in the affordability of housing will have a subsequent effect on the population make up in those areas as well. Alternatively, it could be that changes in demographics (e.g., LRT attracts certain types of people to move to be closer to its services) that affect housing market changes. This research question seeks to better understand LRT's proximity effects on the characteristics of homeowners across space and time.

1.3 Research Framework

Figure 1.1 shows how these three research questions, along with their respective hypotheses, fit within the context of the existing literature. Variables in diamonds are assumed to be exogenous factors. Variables in ovals are those that are of interest and are addressed in the research questions. Finally, variables in rectangles are affected by LRTs and can affect variables in ovals but are not the primary purpose of this research. The dotted lines indicate relationships that have been well explored in the literature, while the solid lines, color-coded by research question, indicate areas that are less-well explored.

1.3.1 LRT Distance, Prior Experience, and SFR Prices

To address the first research question, it is hypothesized that homeowners in new LRT markets will tend to overvalue its effects (i.e., buy into the "hype" or speculation surrounding a new LRT). This research question seeks to expand on prior research to understand whether homeowners' expectations due to [no] prior experience with living near LRTs have a significant impact on the prices they are willing to pay for a single family residence. A spatial hedonic price model using difference-in-differences and single-family residence sales data in two cities are used. In the context of this research, amenities refers specifically to the access of light rail transportation. An

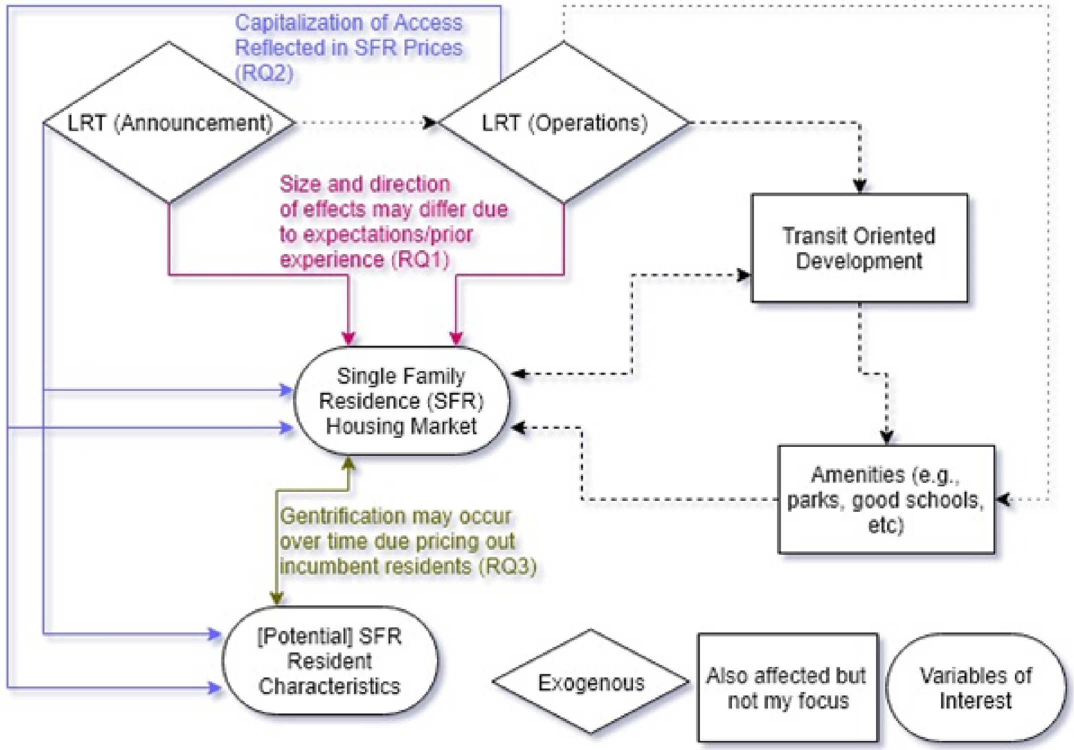

Figure 1.1. Research Framework

analysis of a new LRT line in Charlotte, NC, the first of its kind in that city, showed that the announcement led to an increase in adjacent property values [4]. However, prices of adjacent property values decreased once the LRT became operational—this result contradicts most results found in the literature. It could be that home-buyers overestimated the benefits of the LRT during the announcement phase and adjusted their valuation of living near the LRT once it became operational (i.e., speculation caused the increase in SFR prices).

To examine the effects LRT lines have on nearby housing markets, a-spatial difference-in-differences and spatial difference-in-differences are used. Model fit are calculated using Akaike's Information Criterion (AIC). The a-spatial model represents a base case of comparison (in terms of fit) for the spatially explicit models. It also allows for a direct comparison against the results of previous conducted a-spatial research. The spatial model accounts for the autocorrelation of errors due to spatial spillover effects and the heteroscedastic error commonly found in housing data.

In order to reveal the expectation effect of "inexperienced" home-buyers compared to that of "experienced" home-buyers on house prices near a new (or in the case of the latter, an extension of a) LRT system, a series of spatial difference-in-differences (DiD) models will be constructed. Charlotte, NC provides a natural case study, as it received its first LRT system in 2007, which was then extended nearly a decade later. Through the DiD framework, the effect of experience on home-buyers' willingness to pay to live close to LRT stations will be revealed. The spatial DiD procedure is outlined in Delgado and Florax (2015) [5].

1.3.2 LRT Directional Heterogeneity and SFR Prices

For the second research question, it is expected that direction as well as location/proximity will be significant factors in determining how homeowners value LRT. Building on the spatial models constructed in answering the first research questions, spatial directional (SD) models are used to estimate the effect that the relative direction LRT stations have vis a vis single family residences and city centers. While the spatial model developed for RQ 1 can account for autocorrelation due to spatial spillover effects, it cannot account for autoregressive errors due to directional effects. The SD model, in comparison, assigns directional neighbors to each observation and evaluates the global model. This method should reveal both the proximal and directional effects LRT has on nearby housing markets.

The SD procedure includes assigning directions to both the central business district (CBD) and closest LRT station relative to each housing parcel. To improve model results, several methods of assigning directions are considered, including using four angle and eight angle split methods. Additionally, robustness checks using alternative model specifications are implemented to isolate the SD component's sensitivity.

1.3.3 LRT and Demographic Change

The third research question analyzes demographic changes that may accompany LRT proximity induced changes in housing values over time. As data directly measuring intra- and inter-urban migration flows are not available, demographic change is proxied through the use of indicators measuring race, educational attainment, and median income. In this analysis, the unit of measurement is the census block group (CBG), the smallest geographic unit in publicly available data from the US Census Bureau.

Due to changing census block group boundaries over time and a lack of harmonized geographic information systems data, several a-spatial panel models are constructed to measure the effect proximity to LRT has on the three variables measuring demographic change. If distance to LRT station from CBG centroid is statistically significant, then it is an indication that the presence of LRT is associated with demographic change and thus affects migration.

1.4 Anticipated Contributions

This research is expected to make both methodological and empirical contributions, as well as have practical implications. The last decade has seen spatial statistical approaches slowly percolate from regional science to transportation engineering, allowing transportation engineers to explicitly account for proximal relationships in models. However, the distance between objects can only tell part of a story, particularly in terms of transportation engineering. Direction, in addition to distance, can affect transportation decisions. For instance, pedestrians may choose to brave mid-block traffic rather than going to a marked crosswalk at an intersection and backtrack to reach their intended destination. Directionally explicit spatial models, while common in fields such as genetics and ecology, have yet to be applied in transportation engineering: this study will be the first to do so. Finally, the directional spatial model will be extended to account for changes over time.

Empirically, this research will provide evidence of how "inexperienced" (i.e., no prior experience of living near an LRT system) home-buyers may exhibit different behaviors, as evidenced by the amount they purchase a home for, from "experienced" home-buyers, ceteris paribus. Previous research have only considered LRT systems individually, and therefore have not analyzed the role that consumer expectations may play in affecting housing prices. Additionally, past transportation engineering studies that have examined gentrification effects due to LRT primarily focused on changing land prices. However, gentrification also involves population turnovers, changes in neighborhood character and culture, and occupational changes by residents. This research will attempt to quantify LRTs effects on the additional aspects of gentrification[1].

A practical implication of this research is that it can help agencies explain to the public the effects of a proposed or operational LRT. Abetted by advancements in information archival and retrieval technologies and legislation like Freedom of Information Acts (FOIA), the public's desire to know what government agencies, such as DOTs, do with tax revenues has escalated, especially in the wake of macroeconomic downturns. Evidence of this includes a growth in the total number of federal FOIA requests over time and the increased adoption of asset management systems by state DOTs for simplified report preparations to comply with the General Accounting Standards Board Statement (GASB) 34 ("Basic Financial Statements-and Management's Discussion and Analysis-for State and Local Governments"). The results of this research can help the DOTs and the public understand the effects LRTs have on its surrounding environment. If, for instance, LRTs are found to cause gentrification after they become operational, then stakeholders can base gentrification mitigation measures such as rent controls or subsidized housing policies on the evidence presented in this research. Likewise, if LRT announcement effects are shown to significantly affect land parcel prices, then an agency interested in constructing a

[1]Herein, gentrification is exhibited by a statistically significant change in neighborhood occupant characteristics—including race/ethnicity, median income, and percentage of population with at least a bachelor's degree between time t and t-1.

LRT can perform a more informed cost-benefit analysis to decide if such a system is "worth it."

1.5 Organization

The following chapter includes a comprehensive literature review regarding each of the three research questions. Chapters 3, 4, and 5 seek to answer research questions 1, 2, and 3, respectively. Chapter 6 concludes.

2. REVIEW OF LITERATURE

Most existing studies concerned with transportation and residential land use are centred on cities with transit oriented developments (TODs), in which high rise apartments, department and grocery stores, and locations for recreational activities, such as city parks, are deliberately located near a light rail transit station—the presence of these can confound results (e.g., changes in mode choice may be due to residents of infill development using transit, or due to TOD zoning policies). These existing studies typically use a-spatial hedonic methods to determine the financial value-added to parcels of land, houses, apartments, and condos due to their proximity to transportation access points.

This chapter is organized as follows. First, it will provide a summary of the economic theory that the capitalization of accessibility benefits is based on. Then, an overview of previous research on the effect LRT has on housing markets is presented, including a meta-analysis of extant studies. An overview of studies focusing on the effects of directional heterogeneity follows. Finally, a selection of research studying changes in demographic composition as a result of LRT is summarized.

2.1 Urban Economics and the Built Environment

The capitalization of accessibility benefits' underlying theoretical pinning is derived from the location theory of the monocentric city model. These models predict that in equilibrium, income increases with distance from the central business district (CBD) because income elasticity of demand for land is expected to be greater than income elasticity of commuting costs and value of time [6, 7]. The monocentric city model was used to explain the migration of the wealthy from the city centre into the suburbs in the decades immediately following the Second World War [8]. This out-migration of the wealthy to the suburbs led to a relative decline of city centres as only lower income households remained.

The Alonso-Mills-Muth monocentric city model is powerful in that it is intuitively simple and easy to apply. However, it is limited due to its simplifying assumptions. These assumptions are as follows. The model assumes that all individuals have identical incomes and preferences. In reality, this is far from the case. The city is assumed to be monocentric with individuals only traveling in order to commute to work. However, many cities contain multiple business districts within municipal boundaries; in the real world, individuals also travel for reasons other than work. The model is a partial equilibrium model, so wages and capital markets are assumed to be exogenous. The model is also static, which means that it cannot account for changes in housing stock, changing demographics, or changes in the urban form.

The late 1970s and 1980s saw an in-migration of young upper income professionals move back into the CBD. Kern's model of residential choice argues that this demographic moved to the CBD to access a unique set of amenities, including social, cultural, recreational, and distinctive goods, as well as additional employment opportunities [9]. The urban amenities model was further enhanced with the inclusion of spatial patterns and density of amenities within a city, to further explain why certain demographics may cluster within a district of a city [10]. Ng added additional variables to the model of residential choice by including distances to amenities and heterogeneous amenities within the monocentric CBD framework [11].

In the context of these studies, households' access to transportation infrastructure, such as LRT stations, may be regarded as another type of amenity, which assists residents in reaching destinations of choice (i.e., the "traditional" amenities of the authors). Likewise, access to commercial or retail stores, as seen in the mixed-use development context, is another type of amenity. Because there tends to be a finite supply of properties (land parcels) that are close to these amenities, one may expect the price of acquisition of properties with this type of location advantage to rise in the face of high demand. In the economics and regional science literature, the value added due to a property's proximity to LRT stations is typically considered a premium, or the capitalization of accessibility benefits, as discussed next.

2.2 Capitalization of Accessibility Benefits: LRT and Housing

There have been many empirical studies on the value-added of having residences in proximity to transportation services (i.e., LRT). Previous research on the capitalization of accessibility benefits has primarily been a-spatial. However, as many home-buyers know, the price of a residence is partially dependent on neighbourhood characteristics, which are spatially heterogeneous. Additionally, the location of a residence relative to an LRT line or station is also spatial in nature. A-spatial methods estimating the effects of distances from LRT facilities are not able to fully control for these spatial characteristics, which leads to biased and inconsistent results. Spatial methods may be more computationally intense to utilize, but do not suffer from these drawbacks. This section will first summarize results from a-spatial studies, and then provide a summary of results from spatial studies.

2.2.1 A-spatial Studies

Using panel data and a-spatial, most studies agree that proximity to LRT stations add value to properties; the magnitudes of effects that rail transit have may differ significantly between studies due to the context, analysis period, and methodology used [12, 13]. Because of the rich literature on this topic, this section will only briefly highlight some selected studies. [14] and [15] provide more thorough literature reviews of this topic.

In one of the early studies on accessibility premiums, Baum-Snow and Kahn used a panel dataset of house values and rents in Boston, Atlanta, Chicago, Portland, and Washington, DC, from 1980 to 1990 to estimate the change in prices due to their distances from rail transit lines [16]. Using linear distance between housing tracts and the nearest rail transit lines, the authors used generalized least squares and found only small increases in rents and house values. A decrease in distance by two kilometres was associated with an increase of $19 in monthly rent and an increase of $4,972 in house value. The authors noted that residents who walk and ride transit would save on commuting costs at a disproportionately higher rate than the increase in housing and rent prices [16]. However, the authors did not control for differing house sizes (square feet).

In a study of Santa Clara County's TOD, [17] found that accessibility premiums for commercial properties were not as extensive as those for residential properties. Using data from 1998-1999, the authors estimated a hedonic price model for commercial, office, residential, and light industrial properties. Overall, properties farther than 0.25 miles from a rail station were not affected by accessibility premiums, and the premiums for being close to an LRT station were smaller than those for heavy rail and commuter rail. The authors surmised that this was due to the lower speeds and smaller coverage areas of LRT compared to other types of rail.

Focusing on a single city, Buffalo, New York, [14] examined the impact that LRT stations had on housing parcels within a 0.5-mile radius. Using the fourteen stations of the LRT, the authors found a positive effect on housing values for the 7,357 residential parcels in their dataset, which was based on a local assessor database. Unlike previous research, they included both Euclidean and network distance and controlled for location-specific amenities and household income. They found that high-income neighbourhoods experienced the highest increase in property values. Like [17], [14] found that premiums were lower for LRT than for heavy rail and commuter rail.

[18] also used parcel transactions data from assessor offices to examine the before (1995-1999) and after (2001-2007) effects of a new LRT line on nearby housing prices in Phoenix, Arizona . Separate hedonic price models for differing land use mixes depending on the type of housing and the density of nearby amenities were estimated. A dummy variable indicating whether a parcel was within walking distance of a transit station was used rather than a measure of distance. [18] found that condos were affected the most by LRT proximity; SFR-type parcels experienced an increase

in prices both before and after the opening of the LRT. Another study in Phoenix, Arizona found that premiums appeared during the planning and construction phases of LRT ([19]). These premiums affected residential, commercial, and vacant land parcels. Overall, these authors found that LRT had a positive impact, although SFRs close to transit tracks were negatively affected; the authors attributed this to a nuisance (noise) effect. These two studies indicate that all phases (planning, design, construction, and operation) of an LRT have positive impacts on housing prices.

International studies on rail transit's effect on property prices and land use have tended to focus on subway expansions. For instance, [20] examined the effect that distance from a single subway line in Seoul, Korea, had on property prices prior to and after the line's opening. The authors found a significant and positive effect prior to opening but not after opening. [21] examined the effect of a subway expansion in Beijing, China, on property values. Using a first-differenced hedonic model, they found that property values increased by 15% for every 1-kilometer increase in distance from a subway station. Using Naples, Italy, as a case study, Pagliara and Papa (2011) used descriptive statistics to show that while the presence of urban rail was associated with increases in residential and office property prices, the property prices of retail spaces were less affected [22].

In one of the few studies examining announcement effects, [23] showed that light rail plans have significant positive effects on prices for vacant land sales. However, urban areas typically have few vacant lots. [24] examined the pre-planning, announcement, and construction effects of the Charlotte LYNX LRT line but used SFR values aggregated at a census block level with a-spatial panel models. The authors reported that SFR values closer to LRT stations increased over time, indicating a greater desirability to be close to the LRT as it neared operational status. Due to the level of aggregation, [24] did not find using spatial methods to be necessary. However, the authors did not control for differing houses sizes (square feet) with respect to proximity to the LRT line. In another study examining the announcement effects of transit in Charlotte, North Carolina, [25] used a-spatial difference-in-difference estimators and found that neighbourhood property values within 1 mile of LRT stations increased by 4% and neighbourhood properties within 0.5 miles of LRT stations were unaffected by the announcement. However, because that study only focused on the announcement of the LRT, the analysis provided did not adequately capture capitalisation effects once the LRT line began operations.

Some more recent studies on the effects that LRT have on nearby residential units' prices indicate mixed results. For instance, [26] found that the LRT station in Bayonne, New Jersey had no significant impact on annual house price appreciation. However, this study focused on only a single station in a small city approximately nine miles southwest of New York City. Thus, it may be difficult to apply the results of this particular study to other locations. In another study, [27] use an a-spatial difference-in-differences estimator to assess home affordability in Portland, Oregon suburbs due to proximity of an LRT line. This study also did not find consistent significant changes in home affordability in transit-served neighborhoods. Because affordability was not significantly affected, one may infer that home prices did not significantly change. [27] is somewhat unique, however, in that the study area is suburban in nature; most research on this topic focuses on urban housing. [28]'s examination of LRT service on nearby property values in Seattle, Washington indicated that while some neighborhoods may experience a positive effect, others experience negative or null effects. Using an a-spatial difference-in-differences estimator, the estimated impact of LRT service was positive for one station, negative for two stations, and statistically insignificant for four stations. [28] however notes that the LRT line was constructed in an area that was already well served by local bus lines. Thus, home-buyers may not have valued the addition of LRT, especially if the LRT service interfered with pre-existing bus services. While [28] provides valuable insight into the capitalization of benefits of LRT in the face of good alternative public transit options, not all cities have the same bus transit ridership or level of service as Seattle. Additionally, spatial spillover effects in housing prices were not controlled for in the analysis.

2.2.2 Spatial Studies

As previously noted, most studies on this topic have used a-spatial hedonic price models. More recent studies have incorporated spatial econometric methods in examining the effects that proximity to transit has on housing prices. For instance, [29] used a spatial hedonic model to show that proximity to light rail stations led to an increase in single-family residences' prices in Phoenix, Arizona, for houses that were farther than 300 meters away from light rail track or within 900 meters of a station. However, the authors only examined the immediate effects of the LRT opening because the data were limited to housing values in 2009 (the LRT opened in 2008). Thus, it may be that the observed increases in property values were due to the initial excitement over the LRT and may not have been sustained over time.

[30] used a spatial difference-in-difference model to examine the planning impacts of transport improvements in Beijing, China; however, the difference in land density between Beijing, China, and most American cities may lead to significantly different results. [31] used a spatial panel model and found that proximity to light rail transit led to increases in housing prices in Portland, Oregon, more than a decade after the opening of the transit service. However, these results may be specific to Portland, Oregon, due to city- and state-specific zoning laws that have led to increased urban density, which may not be the case in other parts of the US.

In another recent study, [32] applied a spatial difference-in-difference estimator to analyse the effect that proximity to TOD has on housing prices in the Denver-Aurora-Lakewood Metropolitan Statistical Area, Colorado, and found that housing prices increased. However, the authors focused exclusively on TOD station locations rather than all transit stations.

2.3 Meta-analysis of previous LRT hedonic model studies

As seen from the previous discussion, the results from these studies vary greatly in direction, magnitude, and statistical significance. This section seeks to understand the variation of results among the ex-post studies. One method of reconciling the variation between previous researches is to perform a comprehensive review of the literature. However, literature reviews can be subjective since they are based on the researcher's interpretation of results from multiple case studies sans a systematic analysis. Further, even a systematic review can be prone to researchers' pre-existing biases. Alternatively, a meta-analysis approach can be used to systematically analyse the empirical results from multiple quantitative case studies. However, this is not to say that meta-analyses should replace literature reviews, but rather that meta-analyses can be used to complement them [33].

This section is organized as follows. The next subsection will include a brief review of pre-existing meta-analyses on this general topic area. After that, the methodology used in this study will be discussed, including the selection criteria of papers included in this meta-analysis. The fourth subsection presents and discusses the results of the meta-analysis while the fifth subsection concludes.

2.3.1 Previous meta-analyses

In addition to past research examining the capitalisation of accessibility effects of specific LRT lines, there have been two studies that conducted meta-analyses of rail transit's land use effects.

In the first, [34] examined 55 case studies on heavy rail (subway and commuter rail) and light rail stations' effects on nearby residential and commercial property markets in the US. Because all case studies examine systems in the US, the authors do not include a location variable. However, as transportation patterns may be regional (i.e., some areas may be more amenable to the use of private vehicles over public mass transit options), their analysis may have omitted an important variable. The authors control for several factors between studies, including whether or not the underlying studies included additional means of access, the presence of demographic variables, and whether or not the study took place after 1990. Because the authors examine both property types, much of their analysis attempts to explain the differing effects rail stations have between property types. Additionally, they note that heavy rail tends to have a larger effect on property values than light rail. In general, they suggest that studies that include additional accessibility modes show that rail stations have smaller magnitudes of effects than studies that do not. As this shows that there is a negative correlation between freeways and rail stations, claim that this is a source of omitted variable bias.

In the second meta-analysis, [35] examine the empirical results from 23 case studies that examined heavy and light rail systems' impacts on residential, commercial, and industrial property values from around the world. They include a regional variable to address differences across space; however, they chose to group studies by continents which may be problematic as it cannot account for differences in attitudes towards public transit ridership or private vehicle ownership rates between countries within a single continent. [35] build on [34]'s meta-analysis by expanding the scope of the study to include case studies outside of the US, as well as control for additional case study specific factors, including system maturity, geographic location, whether or not the unit of analysis was property or land, additional indicator variables for data type, analysis type (geographically weighted regression, hedonic price, or difference-in-difference), functional form of model (semi-log, log-log, or neither), and whether or not results were significant. Results of the meta-analysis show that geographical location, accessibility to roads, data specification, and methodological characteristics produce variation between studies. Additionally, characteristics of the rail system studied, such as type of rail service and rail system life cycle maturity, also effect case studies' results. Finally, the authors perform publication bias tests and find that while researchers tend report both positive and negative results, they are biased towards statistically significant estimates.

Although [34] and [35] conduct meta-analyses of rail transit's land use effects, they did not consider any level of service or ridership factors that could account for differences between studies' results (in the case of the former), nor do they focus exclusively on light rail transit's effects on US SFR parcel values (in the case of the latter). The level of service provided by a LRT may also contribute to differences between case studies. For instance, residents living near a station may have different perceptions of value added depending on the frequency of trains arriving and departing the station (i.e., headway). Additionally, LRT in an urban area with low public transit ridership may have a smaller effect on property prices than LRT in an urban area with high ridership rates. Thus, this meta-analysis, while focusing exclusively on LRT's effects on SFR prices in the US, will include these variables to test if local city-specific contextual factors also partially account for the heterogeneity in results between studies. The next subsection describes the method used in this meta-analysis and provides a brief description of the data.

2.3.2 Methodology and Data Used

Method used

While not commonly found in the transportation engineering literature, meta-analyses have seen extensive use in the experimental sciences, where it originated, and the social sciences. A meta-analysis is a statistical analysis of empirical research results on a topic and allows researchers to synthesize and evaluate empirical results across multiple studies. Typically, this involves a regression-type analysis, or meta-regression, in which the dependent variable is the estimated effect reported in a study and independent variables include variables measuring the features of the original study. Thus, a meta-analysis can distinguish the statistically significant reasons underlying the variations found in empirical results across multiple case studies. It is important to note, however, that a meta-analysis should complement rather than replace a literature review. Literature reviews typically consist of commentary on the findings of previous research and is qualitative in nature. Compared to a quantitative meta-regression, the qualitative nature of such commentary may lead to questionable conclusions because it is subject to the author's opinion. However, meta-analyses suffer from several limitations. In order to yield defensible results, they must only include studies that are relatively homogeneous in research design; literature reviews, being narrative in nature, can cover a much broader range of studies. Another problem with meta-analyses is that they may violate the independent and identically distributed random variables assumption (iid), especially when they include multiple effect estimates from a single underlying study leading to inefficient estimators. Finally, because the dependent variable includes effect sizes with different variances, meta-regressions tend to be heteroskedastic.

[36] note that the basic meta-analysis equation is:

$$Y = f(P, X, R, T, L) + \varepsilon \tag{2.1}$$

in which Y is the variable under study, P is the set of causes on the outcome of Y, X are the characteristics of the set of objects under examination affected by P, R includes the characteristics of the underlying research methods, T is the time periods covered by the studies, L is the location of each study conducted, and ε is the error term .

In the case of this meta-analysis, all underlying studies included similar features of properties (P) in their analyses; thus this is dropped to avoid multicollinearity. X includes city-specific contextual variables, such as level of service indicators and population density measures. The underlying studies used a variety of estimators and functional forms (R). To account for this, dummy variables including ordinary least squares (OLS), hedonic model, did, semi-log, and loglog were developed. Additionally, because studies occurred at both different time periods and system maturity periods, indicator variables including system maturity, prey2kstudy, and prey2kdata have been created. Finally, to account for additional location-specific effects (L), the factor variable study area denotes the city in which each underlying study occurred. Table 2.1 further explains the variables developed for the meta-analysis. To determine the impacts these variables have on the observed variation in effect sizes across studies, a hedonic model using a simple linear regression specification is used in this meta-analysis.

The Stata command, metareg, was used to conduct the random effects meta-regression. While the current default of metareg is to calculate the tau-squared statistic using a maximum likelihood estimator (MLE), this study elects to use the method of moments (MoM) estimator. This was primarily done in order to have results that could be directly compared and verified against [34] which also used the MoM estimator.

To test for differences between effect sizes within the included studies, a Cochran's Q-test was conducted. It is a non-parametric statistical test to verify whether treatments have identical effects [37]. The significant Cochran's Q-test outcome indicated that there was heterogeneity in effect sizes of LRT station proximity on property values. This indicates that a random effects model is required, as random effects assumes that the variance associated with each effect size is due to within- and between-studies variance. In other words, a random effects meta-analysis assumes that deviations of individual studies from the centre of a distribution of true effects represent true heterogeneity. Had Cochran's Q-test been insignificant, a fixed effects model would have been appropriate. In that case, the deviations of individual studies from a single "true" effect is assumed to be due to random variations due to sampling error.

Selection Criteria

For the purposes of this meta-analysis, only studies that quantify LRT's effects on SFR prices in the US using regression techniques, are considered. The majority of the underlying studies were found through searching for the terms "light rail transit", "single family residence", and "property price" in Google Scholar. Some additional studies were found by perusing the works cited sections of the previously located studies. Research that were omitted include those that did not include empirical results, did not use regression-based analyses, specify if rail stations were of LRTs or

Table 2.1. Meta-analysis Data Summary

Variable	**Description**	**Type**	**Average**
extrapresults	Effect size of underlying study observation extrapolated to 1 mile expressed as percentage change; dependent variable	Continuous	-0.0128
	Contextual variables		
studyarea	Factor variable consisting of city underlying study used	Factor	N/A
linelength	Length of LRT line in miles	Continuous	21.6037
peakhrheadway	Peak hour headway in minutes	Discrete	9.8125
nonpeakheadway	Non-peak hour headway in minutes	Discrete	17.4107
vehiclespeed	Average speed of LRT in miles per hour	Discrete	22.1054
rawlosindex	Index consisting of average station density, peak and non-peak hour headways, vehicle speeds; from 0 to 1	Continuous	0.6481
popdensity	Population density of study area at time of study in persons per square mile	Continuous	1991.4960
	Methodological Variables		
samplesize	Sample size of underlying study	Discrete	20024
systemmaturity	Factor indicating whether study occurred before or after LRT began operations	Factor	N/A
prey2kdata	If study used data from before year 2000: 1 = yes	Binary	0.3036
prey2kstudy	If study occurred before year 2000: 1 = yes	Binary	0.1429
pooledcrossdata	If study used [pooled] cross sectional data: 1 = yes	Binary	0.1964
paneldata	If study used panel data models: 1 = yes	Binary	0.2143
transactionsdata	If study used transactions data (as opposed to tax assessor data): 1 = yes	Binary	0.5714
euclideandist	If study used Euclidean distances (rather than network distance): 1 = yes	Binary	0.7857
ols	If study used OLS linear regression: 1 = yes	Binary	0.8929
hedonicmodel	If study used hedonic model: 1 = yes	Binary	0.9643
did	If study used difference-in-differences estimator: 1 = yes	Binary	0.1788
semilog	If study used semi-log functional form: 1 = yes	Binary	0.3750
loglog	If study used double-log functional form: 1 = yes	Binary	0.2679
freewaydummy	If study controlled for SFRs' distance to freeway access points: 1 = yes	Binary	0.5000

other types rail transit (in the case that an urban area had both), or if authors did not specify between SFRs, condos, and rental units. Additionally, studies that examined LRT effects on SFR prices in transit-oriented developments (TOD) were omitted as confounding factors attributed to TOD-specific contexts could not be easily controlled for. Finally, studies whose results did not include standard errors or z-scores were also omitted. In total, 16 studies were found to fit these criteria. Many studies reported multiple estimates of the dependent variables, such as including estimates at different distance bands or during different phases of the LRT system's life cycle. Each of these estimates are treated as separate observations. In total there are 54 observations from these 16 studies. Table 2.2 outlines the studies and observations used in this meta-analysis. Unlike a "true" meta-analysis, however, which includes unpublished or working papers, the underlying studies used in this analysis were all published in peer-reviewed journals. Thus, this research does not include a test for publication bias as may be found in other meta-analyses.

Data

As different authors reported their results in a variety of ways, including change in dollars per square foot, elasticities, and percentage change in house values, it can be difficult to conduct an analysis with differing units. For simplicity, this analysis converts all published results into percentage change in house values. Further, different studies used different distance bands. To create a single standard unit, the results from the underlying researches have been extrapolated to the 1-mile mark. Thus, all results are interpreted in terms of percent change in house values per mile away from LRT stations. Table 2.2 shows the number of observations generated from each included paper as well as the mean, minimum, and maximum effect sizes of those observations. Based on the range of effect sizes, one can easily see that there is wide variation between reported results in the underlying studies. It is important to note, however, that although mean, minimum, and maximum effect sizes are shown, not all observations are associated with statistically significant results.

Table 2.2. Underlying Studies' Effect Size Summary Statistics

Author (Year)	**City**	**Obs.**	**Effect Size Per Mile**		
			Mean	**Min**	**Max**
Hess and Almeida (2007)	Buffalo	8	-10.25%	-260.10%	52.74%
Billings (2011)	Charlotte	2	3.35%	-1.24%	7.94%
Wagner et al. (2017)	Hampton Roads, VA	6	-10.34%	-21.57%	-4.84%
Cervero and Duncan (2002)	L.A.	2	2.30%	-5.00%	9.60%
Zhong and Li (2016)	L.A.	2	-18.21%	-19.66%	-16.76%
Pilgram and West (2018)	Minneapolis	2	1.21%	1.17%	1.24%
Atkinson-Palombo (2009)	Phoenix	8	-1.55%	-23.60%	11.80%
Golub et al. (2012)	Phoenix	2	2.03%	0.37%	3.70%
Seo et al. (2014)	Phoenix	10	6.52%	0.89%	18.88%
Chen et al. (1997)	Portland, OR	1	-42.00%	-42.00%	-42.00%
Lewis-Workman and Brod (1997)	Portland, OR	2	1.11%	0.00%	2.22%
Cervero and Duncan (2002c)	San Diego	4	4.40%	-9.00%	24.60%
Duncan (2008)	San Diego	1	22.80%	22.80%	22.80%
Duncan (2011)	San Diego	1	-8.41%	-8.41%	-8.41%
Landis et al. (1994)	San Jose, Sacramento, San Diego	3	10.12%	1.44%	26.80%
Overall	**N/A**	**54**	**-1.28%**	**-260.10%**	**52.74%**

The dependent variable, the change in price of SFRs, is regressed against vectors of contextual factors and methodological characteristics from the underlying study areas and study analyses, respectively. Contextual factors are as follows: a location variable specifying the region of the country the study took place, LRT system maturity, an index representing the level of service provided by the LRT which includes peak hour and non-peak hour headways, and average vehicle speed, population density of surrounding area at the time of the study, length of rail line(s), and the number of stations per line. Methodological factors include whether the underlying studies used demographic variables, time dummies, or inclusion of other means of access, data type, analysis type, and model type. Table 2 summarizes these variables. The next section presents the final model estimation and discusses the results.

Results and Discussion

Because the dataset includes only 54 observations, it is difficult to estimate the effects of all contextual and methodological variables concurrently without encountering a severe loss in degrees of freedom due to over-fitting the model. To account for this, an iterative process based on variables' estimated statistical significance was applied in determining which variables belonged in the model. Table 2.3 presents the preliminary results of the meta-regression using method of moments estimation of between-study variance percentage due to heterogeneity. All variables are significant to at least the 90th percentile.

Based on these results, the length of the LRT line, whether studies used data from before year 2000, whether or not studies utilized ordinary least squares linear regression models, and whether or not they controlled for access to alternative transportation facilities such as freeways have significant effects on the published results. The sign of the estimated coefficient on line length is interesting. Typically, longer LRT lines are associated with increased accessibility. However, increased lengths of LRT lines may be an indication that a city is sprawled. In general, urban sprawl in American cities is a land use trend that tends to be more favourable to automotive transportation options. This could partially explain why the estimated coefficient on line length is negative: since residents favour the use of autos over rail due to the built environment, they are, in effect, valuing the proximity to LRT stations less, ceteris paribus. Alternatively, longer line lengths may be associated with longer travel

Table 2.3. Meta Regression Results

Variable	**Estimate (Std. Err.)**
linelength	-0.003***
	(-0.001)
prey2kdata	-0.060*
	(-0.033)
ols	-0.108**
	(-0.043)
freewaydummy	0.093**
	(-0.039)
(intercept)	0.116**
	(-0.044)
τ^2	0.00085
I^2	100%
Adj-R^2	99.91%
Model F (4, 49)	8.85
Prob > F	0.0000
*Note: p<0.001 '***', 0.05 '**', 0.1, '*'*	

times between origin and destination locations. Since travellers typically desire to minimize travel times, this may also help explain why there appears to be a negative relationship between line length and LRT stations' proximity effect on SFR prices.

The pre-2000 dummy variable is estimated to be weakly significant and negative. While this could simply be a reflection of larger macro-economic trends during those periods, there is a lack of data to better understand this finding. For comparison, both [34] and [35] include dummy variables to control for the time period of case study data collection; however, neither of those two studies found their respective dummy variable to be statistically significant. Studies using an OLS estimator tended to underestimate effects of light rail transit compared to studies using difference-in-differences, repeat sales, and any other estimator types. However, the functional form variables were never estimated to be statistically significant, regardless of the inclusion of additional or alternate variables. This runs contrary to that of [35], who found that semi-log and log-log models produced lower estimates than linear models. Studies that do not include controls for alternative transportation access also underestimate effect sizes. This is consistent with both prior meta-analyses.

An alternate method of interpreting the reported effects of multiple studies is using forest plots. Figure 2.1 shows a forest plot of the results of studies analysing LRT's effects on nearby SFR prices before the lines are operational, while Figure 2.2 displays the results of studies analysing LRT's effects on nearby SFR prices after lines become operational. As was the case in the meta-regression, all studies' estimates have been transformed into percentage change in SFR sales value per mile away from LRT. The left column indicates which study corresponds to the box and line. The boxes represent the estimated effect, while the lines indicate the [95%] confidence interval of the estimates. Additionally, the size of the boxes represents the relative weight of each study. Larger boxes indicate studies have larger weights. Weights are calculated as the proportion of the study's sample size divided by the total sample size of all studies belonging in either the before or after group. To aid visual inspection, a vertical line is placed at 0, which represents the null. In other words, if the confidence interval (i.e., horizontal line) intersects the 0-line, then the statistic is not significantly different from 0 at the 0.05 level.

Conclusions and Further Work

Because the current literature contains wide variations in estimated impacts, the results of this meta-analysis can help explain such variation. Using contextual variables, stakeholders and practitioners in affected areas may be better able to view the results of the underlying studies with a more critical eye.

However, due to a relatively small sample size of 54 observations, it is difficult to evaluate all contextual and methodological factors jointly. With so few observations, it is difficult to include enough control and explanatory variables in the meta-regression. This may be alleviated by expanding the scope of the meta-analysis

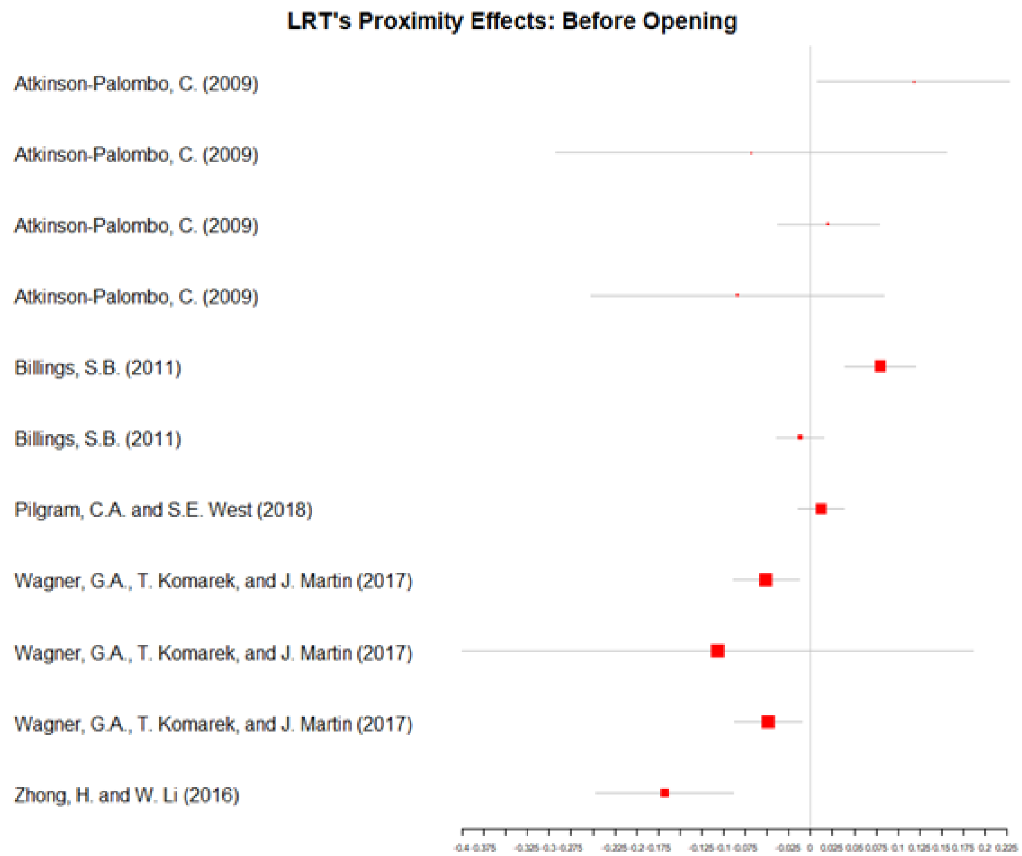

Figure 2.1. LRT's Proximity Effects Before Operations

to include additional types of parcels, including multifamily residences, rental properties, and commercial and industrial land uses. This would allow one to consider the benefits in terms of total development, both commercial and residential, and could help to justify the costs of having a LRT in a particular urban area. For instance, commercial development may be the main type of real estate development nearest to the station with decreasingly dense residential development further away. However, the extant literature on LRTs' effects on these property types is also scant and may partially be due to the difficulty in acquiring the associated data. As this remains a lively research area, more recent work can also be added to the meta-analysis.

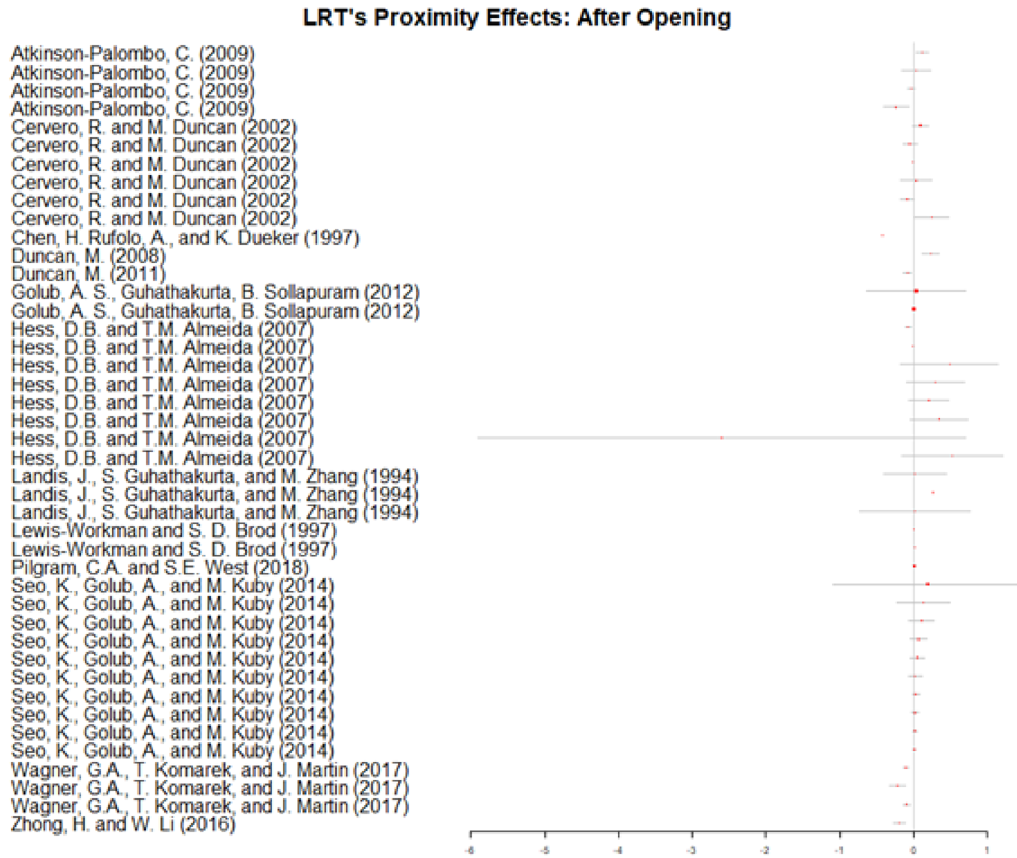

Figure 2.2. LRT's Proximity Effects After Operations

Alternatively, the geographic scope may be expanded to include additional countries or continents. This approach is not without its own set of complexities that may be difficult to control for; for instance, public perceptions and attitudes towards public rail transit may vary significantly. Another possibility is to include studies that examined LRTs' effects on housing in TODs. As of this writing, however, there are no existing databases of locations of TODs in the US. Short of using satellite imagery or apps like Google Maps' Streetview, it is difficult to determine whether urban mixed-use developments are indeed "TODs." Finally, the scope of the meta-analysis could be expanded to include additional types of rail transit, including heavy intercity rail, commuter rail, subway, and traditional LRTs.

2.4 Directional Heterogeneity

Another, hitherto unexamined, reason for variation between LRT study results might be attributed to the effects of directional heterogeneity. LRT is typically conceived as an alternative transportation mode that allows residents to travel from their homes to a central business district (CBD). However, despite a house being located near a LRT station, if the LRT station is in the opposite direction of the CBD relative to the house, residents may choose to use a more direct mode of transportation (e.g., private vehicle) to travel to the CBD. If this were the case, then these homeowners may not value the LRT service as much as homeowners for which the LRT station is "on the way" to the CBD. In other words, potential home-buyers could value LRT services if the location of the LRT station is downstream, rather than upstream, of the SFR location. By not including these directional effects, research may suffer from omitted variable bias leading to biased estimates on the coefficient on the distance variable if distance and direction are correlated. Using housing data from Charlotte, NC, this chapter examines the significance and magnitude of directional heterogeneity's effects on SFR prices near LRT stations.

While studies utilizing directional heterogeneity are rather limited in transportation engineering, researchers in other fields, particularly in the biological sciences, have long recognized that in addition to spatial (neighbour) relationships affecting outcomes within a study population, directional effects may also affect outcomes. In an early theoretical paper, [38] extend spatial autocorrelation analyses intended to characterize geographic variation in biological populations to include directional spatial autocorrelation using directional correlograms, which describe the similarity between pairs of localities as a function of distance between the localities. Their motivation was to be able to characterize geographic trends in biological variables using compass directions when biological populations span large areas . While [38] note that their procedure is not new, citing a much earlier paper studying wheat yields ([39]), the application of two-dimensional correlograms as a means to indicate directional autocorrelation was novel. [40] present an alternative method based on canonical ordination to parse out the directional component of ecological variation within a study species population. A third method, commonly taught in numerical ecology courses, involves using Mantel tests to detect directional autocorrelation [41].

As computing power has improved, so too have methods to account for directional heterogeneity. [42] apply asymmetric eigenvector maps, loosely based on Moran's eigenvector maps, to model directional spatial processes in the variation of brook trout diet across 42 lakes in Canada. [43] develop a spatial-directional model to examine the role of direction and non-directional forces on yellow perch larval dispersal in Canada. The effect of direction in addition to space on variables of interest are of concern to researchers in other fields as well. For instance, [44] develop a spatial panel model with a directional variable to measure wave heights and wave directions in the Adriatic Sea.

Although directional heterogeneity is not often considered in the transportation literature, it has received attention in environmental and urban economics. [45] shows that a failure to account for directional heterogeneity can obscure otherwise statistically significant distance effects in hedonic property value models through an empirical analysis of housing prices in an area that had been associated with chemical manufacturing and noxious odours. Employing an alternative methodology, [46] show that SFRs downwind of a crematory are affected more than residences upwind of a crematory. Directional heterogeneity has also been shown to affect urban rental markets. Using a rent shift model, [47] finds that a failure to account for the relative location to the CBD hinders one's ability to accurately estimate the impact of most economic predictors on rent changes.

Additionally, there is a rich literature on the estimation of the land value gradient in urban economics that show directional gradients affect housing prices. In the simple monocentric city model, land rent falls off rapidly near the city centre and less rapidly as the distance to the city centre increases. The standard model's assumption of uniform access to the CBD implies a constant land value gradient. This is an unrealistic assumption, however, as transportation facilities and services are not uniformly distributed in cities, leading to nonuniform access to the CBD. Concerned with a radial transportation network, [48] use piecewise linear functions to estimate directional land price gradients in Chicago while generating parametric estimates of the gradients to show that allowing for directional variation when measuring the effect of distance in land price gradients is important. [49] find that directional spatial irregularities, measured as angle direction from CBD, can help explain local differences in population densities in cities. If one assumes that housing is a limited resource in urban areas, then their finding suggests that directional heterogeneity can affect SFR prices.

2.5 LRT and Demographic Changes

First used in 1964 by [50], gentrification originally referred to middle- and upper-class households purchasing and renovating older homes for personal use in poor urban neighbourhoods. Neighbourhoods were typically close to the CBD with low home prices due to deteriorating infrastructure. As more middle-class households moved in, the incumbent working class residents were displaced. Since then, the meaning of the term gentrification has evolved to fit the observed patterns of behaviour at points of time.

[51] analyzes data from 110 US urban cores between 2000-2015 to determine the degree of gentrification experienced through sixteen variables that had been previously used in the gentrification literature. The proportion of white residents, young adults, and professional workers, as well as educational attainment, income, rent, and home values increased, while the proportion of working class residents, children, and average household sizes decreased. Further, the cost of housing and rent increased faster than income, which suggests an increase in housing stress within urban cores. These results indicate that gentrification has occurred in all urban cores.

Gentrification can be categorized as having had four waves [52]. [50] describes the first wave, where an influx of wealthy households displaced poorer households. The second wave occurred when developers in the 1970s-80s increased investments in urban core areas. In this context, [9]'s young professionals and [53]'s "creative class" who moved into urban areas are considered gentrifiers. A third wave of gentrification, occurring during the 1990s, is characterized by increases in private and public capital being used in urban development projects [54]. [55] note that the state-backing of corporate redevelopment projects led to an intensification of gentrification trends in many cities. [56] describe an ongoing fourth wave of gentrification characterized by the diffusion of gentrification trends further away from urban downtowns. Unlike the first wave, the fourth wave occurs as new-build developments, including retail, housing, and transportation infrastructure, lead to the displacement of the incumbent lower socioeconomic-tier residents.

Economic Perspectives of Demographic Change

Despite these trends of displacement, there is no single definition of gentrification. [57] note that many studies recognize that neighbourhood turnover is one aspect of gentrification despite not always explicitly noting displacement. Some studies suggest that it is due to wealthy residents moving into an area previously occupied by poorer residents [58, 59]. Others include changes in a neighbourhood's racial and ethnic makeup in their definition (e.g., [60]). [61] refers to gentrification without displacement as neighbourhood upgrading. While the actual residents of the neighbourhood may not change, the existing residents are able to improve the quality of their neighbourhoods or change their occupational status from blue- to white-collar workers [62]. Despite agreement that gentrification is generally categorized by an influx of the wealthy moving into a poorer neighbourhood, [63] argues that displacement is not necessarily a "given" when gentrification does occur. [64], for instance, find that low-income families with young children in gentrifying areas are typically not displaced.

Alternatively, [65] argue that displacement is an inherent effect of gentrification. [66] suggests that even if there is a lack of empirical evidence pointing to displacement, mitigation measures should still be implemented to ensure lower income households are not priced out of their neighbourhoods.

In addition to various definitions in the literature, there is no agreement over the mechanisms of gentrification. The literature includes both supply-side and demand-side partial equilibrium explanations. Supply-side activities are usually explained by rent gap theory, an extension of bid rent theory. [67] describes a rent gap as the difference between the actual value of a land parcel and the parcel's potential value given its opportunity cost. According to bid rent theory, gentrification is the result of land and housing markets seeking equilibrium. The movement of capital to suburbs and the depreciation of inner-city capital lead to a rent gap. By investing in urban infrastructure, such as a LRT system, inner city land can reach its potential value [68]. Alternatively, cities can update property tax codes to close the rent gap. This can also lead to gentrification if incumbent homeowners are unable to pay. However, [69] find that cities can also implement tax relief policies for long-term homeowners to mitigate the risk of tax deliquencies and displacement.

Demand-side explanations of gentrification typically focus on who causes gentrification. Some indicate that white, middle-class, urban professionals cause gentrification as they move into gentrifying neighbourhoods [70,71]. Others examine gentrifiers' racial diversity [72]. The lack of consensus suggests that gentrification can occur in various ways within many contexts [57]. In addition to people causing gentrification, public investments may also lead to gentrification. Brownfield redevelopment practices [73], charter schools [74] , and housing programs [75] may lead to gentrification.

Neighborhood Change and LRT

Transit improvements may also lead to gentrification. In terms of rent gap theory, the opening of a LRT station, for instance, can cause undervalued areas to gain value and close the rent gap. Land prices increase, due to the increased accessibility that transit provides and the increased land use intensity (e.g., TODs), to meet their potential value. This causes increases in housing values, and can potentially lead to low-income residents being priced out from areas near a LRT station [76].

Neighborhood change includes not only changes in demographic characteristics such as median income, age of residents, and race, but also a sense of place. [77] analyze survey responses of residents within a half-mile of a LRT station in Denver and find that respondents on average held positive views on the impacts and changes to their neighborhoods due to the presence of the LRT station. Respondents with negative views mentioned increases in density, traffic, and higher rents. Additionally, homeowners expressed frustrations with the changing character of their neighborhoods due to an increasing percentage of rental units. These concerns were highly associated with perceptions of social class, similar to what other studies on gentrification have found (e.g., [62]).

Using three decades of data from nine metropolitan areas with light rail, [78] find that impoverished neighborhoods are most likely to gentrify after LRT operations begin while affluent neighborhoods are less likely to change. Affluent neighborhoods that do change tend to densify. They further find that the socioeconomic ascent following station openings tends to not be associated with changes in racial composition of neighborhoods. Interestingly, while LRT may have played a role in neighborhood upscaling, new residents of those neighborhoods may not actually use the available transit options. For instance, [79] finds that commuting via transit has declined in urban cores while the proportion of workers working from home and cycling to work have grown faster than other commuting modes.

3. THE ROLE OF EXPERIENCE IN HOME-BUYERS' WILLINGNESS TO PAY TO LIVE NEAR LRT STATIONS

As documented in Chapter 2, there has been a myriad of research efforts characterizing the effects LRT has on nearby land use. However, the role of home-buyers' prior experience with, or exposure to, LRT systems in conjunction with the construction and operations of light rail remains unexplored. This chapter is divided into three major sections[1]. The first section analyzes the capitalization of amenities access with respect to single family residence purchase prices in Charlotte, NC, a city in which a brand-new LRT line was opened. Amenities in this context refers to the accessibility of light rail. Construction of the line, the LYNX Blue Line (hereafter referred to as the Blue Line), began in 2005 and operation commenced in 2007. The second section conducts a similar analysis for single family residences on the LYNX Blue Extension Line (hereafter referred to as the Extension Line) in Charlotte. Construction on the extension started in 2011 and operations began in 2019. The final section compares the results from the preceding two sections to determine whether residents' experience with LRT influenced their willingness to pay to live near LRT access points.

3.1 Background and Motivation

Based on the literature review (Chapter 2), much work has been done and is ongoing in quantifying the potential capitalization benefits of residential proximity to LRT stations. However, there has been little work done in examining whether or not previous experience with LRT affect home-buyers' willingness to pay to live near a

[1]Portions of the content of this chapter are published in the author's scholarship: Ke. Y, & Gkritza, K. (2019). "Light rail transit and housing markets in Charlotte-Mecklenburg County, North Carolina: Announcement and operations effects using quasi-experimental methods", *Journal of Transport Geography, 76,* 212-220.

LRT station. This analysis establishes a baseline for such a comparison to take place by using the original segment of the LYNX Blue Line in Charlotte, NC.

Settled in the mid-18th century,Charlotte is a medium-sized US city with a single public school district and uses a lottery scheme to determine which schools children attend as a means to promote racial integration within its schools. Thus, school quality is not considered as an amenity. Additionally, the city is automobile-centric, with lower public transit ridership than other similarly sized cities.

From the meta-analysis in the preceding chapter, SFR premiums may differ significantly based on the location analysed, variables controlled for, and the methodology used. Thus, rather than use previous research on this topic, a new analysis, described below, was undertaken. This section is structured as follows: The next subsections describe the methodology and data, respectively. Next, the estimation results of the ordinary least squares (OLS) and spatial econometric models are presented in sections 3.1.3 and 3.1.4, respectively. Section 3.1.5 concludes and includes suggestions for further study.

3.1.1 Methodology

Based on the review of previous studies on the capitalization of accessibility benefits, the price per square foot for SFRs is dependent on house-specific characteristics, including the year built; the number of half baths, full baths, and bedrooms; the presence of other house-specific amenities; the year the house was sold; and the distance to an LRT station. The price per square foot measure only includes the area of a house that is walled in and roofed. The number of days a house is on the market is also thought to affect price per square foot, in that houses generally drop in price the longer they stay on the market. In addition to these characteristics, year dummy variables were included in order to capture some of the year-specific effects and determine whether the potential price gradient is due to macroeconomic conditions such as the 2008-2009 recession.

It may seem counter-intuitive to use price per square foot as the dependent variable rather than the selling price of SFRs as the dependent variable and the total square footage of houses as an independent variable. However, this decision was driven by data availability. All observations in the data included values for the price per square foot variable. However, only roughly one third of observations included data for the total square foot variable (6,906 out of 20,586 and 8,480 out of 25,669 in the announcement and operations phases, respectively). Attempts to use spatial econometric models with only these observations with closing price as the dependent variable and total square footage of houses as an independent variable led to singular matrix issues. Similarly, imputing the missing values of square foot also led to singular matrix errors. Thus, the price per square foot variable was chosen as the dependent variable in the models that follow.

Furthermore, to specifically identify the effect the LRT had on house prices, a difference-in-differences approach is used. Like the difference-in-differences identification in [25], the treatment group is defined as SFRs within a 1-mile radius of Blue Line LRT stations. The control group is defined as SFRs within a 1-mile radius of the proposed but not built Red Line LRT stations. Because the Red Line was not built due to the city government's inability to reach an agreement with the owners of the rail infrastructure, rather than characteristics associated with nearby neighborhoods, the SFRs within a 1-mile radius of the proposed stations are chosen as the control group. In order to assess the announcement and operations effects of the LRT, separate analyses were run using respective subsets of the data.

A-spatial Models

Using OLS, quasi-experimental hedonic pricing models were estimated for the price per square foot of single-family residences for the periods before and after operations began on the LRT system. In addition to regressing the price per square foot of sold single-family residences against the house and lot characteristics, the price per square foot was also regressed against the interaction term consisting of the treatment indicator variable and distance and distance squared between the house and the nearest LRT station as well as distance between the house and the nearest freeway access point. The distance between the house and the nearest freeway access point is included to help isolate the effect that LRT stations have on nearby residential units.

Year dummy variables were included to capture year-specific fixed effects, including otherwise unobservable changes that may have occurred during that time, such as macroeconomic shocks, as well as temporal distance to and from the opening of the LRT system for operations. An a-spatial Chow test was conducted to test whether separate models needed to be developed to capture announcement and operations effects; the test results (test statistic of 230.17, p-value $<$ 2.2e-16) indicated that this was indeed the case because there existed a structural break in the housing price data once the LRT station began operations.

Spatial Models

In order to create a spatial weights matrix using the rook strategy, the point data were converted into Thiessen polygons. The more widely used queen-based weight matrix was also considered. However, because some of the SFR locations were so densely clustered, using a queen matrix was not optimal because every polygon had too many neighbors. Additionally, using a queen matrix led to nearly 5,000 data points being removed because SFRs were so closely located that they were given duplicate IDs.

As potential home-buyers can attest, the price of a house depends on the price of nearby houses. This is likely due to realtors' behavior of pricing property based on similar properties within a neighborhood which creates positive spatial autocorrelation. Thus, there is a conceptually based rationale for using spatial models. To statistically test whether there was a case for using the spatial analysis approach, Moran's I was performed on the price per square foot of the polygons in the pooled data using a rook weights matrix (I=0.998, z=530.382, p-value=0.000). The p-value was calculated by performing 99,999 Monte-Carlo permutations. Based on the significance of the p-value in Moran's I test, the null hypothesis was rejected. The positive z-value indicated the presence of clustering. Had the z-value been negative, it would have indicated the presence of dispersion .

Just as in the a-spatial case, a spatial Chow test was conducted to test whether separate models needed to be developed to capture the announcement and operations effects of the LRT on housing prices. Following Anselin (1988), the spatial Chow test was implemented as a likelihood ratio test because the simple form of the Chow test was not appropriate as the regression error terms are spatially autocorrelated. Test results (test statistic of 222.21, p-value $< 2.2e\text{-}16$) showed that the null hypothesis should be rejected, suggesting that even after correcting for spatial lag and spatial error, there are significantly different coefficients in each of the regimes (announcement and operations phases).

Using a specific-to-general approach, robust Lagrange multiplier tests for spatial error and spatial lag were performed to determine which spatial model would be appropriate. Test results indicated that spatial autocorrelation in both the lag and error terms is present. Typically, the more significant result is used to determine the type of model to use. However, in this case the statistical significance of the models was identical. Because of this, a spatial autoregressive model with autoregressive disturbances (SARAR) is selected. Additional details related to running a spatial differences-in-differences model may be found in [80].

Because of the inclusion of spatial weights matrices in the model formulation, the estimated coefficients of spatial models are not as straightforward to interpret as in the a-spatial case. The direct effect of an independent variable on one observation's dependent variable also influences the dependent variables of neighboring observations, which in turn affects the original observation's dependent variable via indirect effects. Therefore, properly accounting for these feedback effects is important in the interpretation of spatial model results. Following [81], average direct marginal effects are calculated by finding the average value of the diagonal of the partial derivatives matrix. Similarly, average indirect marginal effects are calculated by finding the average value of all off-diagonal entries in the partial derivatives matrix. Total marginal effects are the sum of the direct and indirect marginal effects.

3.1.2 Data

The data for this study primarily came from two sources, the Charlotte Department of Transportation (CDOT) and the Multiple Listing Service (MLS) Listings database. CDOT data include shapefiles of all LRT lines in the Charlotte/Mecklenburg County area, including those that are planned, under construction, and operational. The Blue Line has been operational since November 2007 and includes 15 stations. The stations on this line are used in the analysis that follows. Additional lines are the Red Line (planned but ultimately not constructed) and the Blue Line Extension (opened in 2019). Two additional lines, the Gold Line and Gold Line Extension, also exist as streetcar lines. Because streetcars are considered a different mode of transportation from LRT, they are not included in the analysis. Figure 3.1 shows a map of the Blue Line service and the planned Red Line. Both lines use pre-existing track owned by a single freight company. The MLS database is maintained by a third-party company for use by real estate agents. All realtors are required to list the properties they show in this database or risk forfeiture of their realtor licenses. The MLS data include all sales of SFRs from January 2001 to December 2016. Additional data used included a shapefiles from NC Department of Transportation (NCDOT), which contained the locations of highways and access points in the study area.

Characteristics of residences listed in the MLS data included the number of bedrooms, bathrooms, and half baths; square footage; and street address. Additionally, there were two fields containing the realtor's private remarks and public comments about each property. These remarks and comments included directions to access the property, whether the previous owners had completed significant renovations or re-

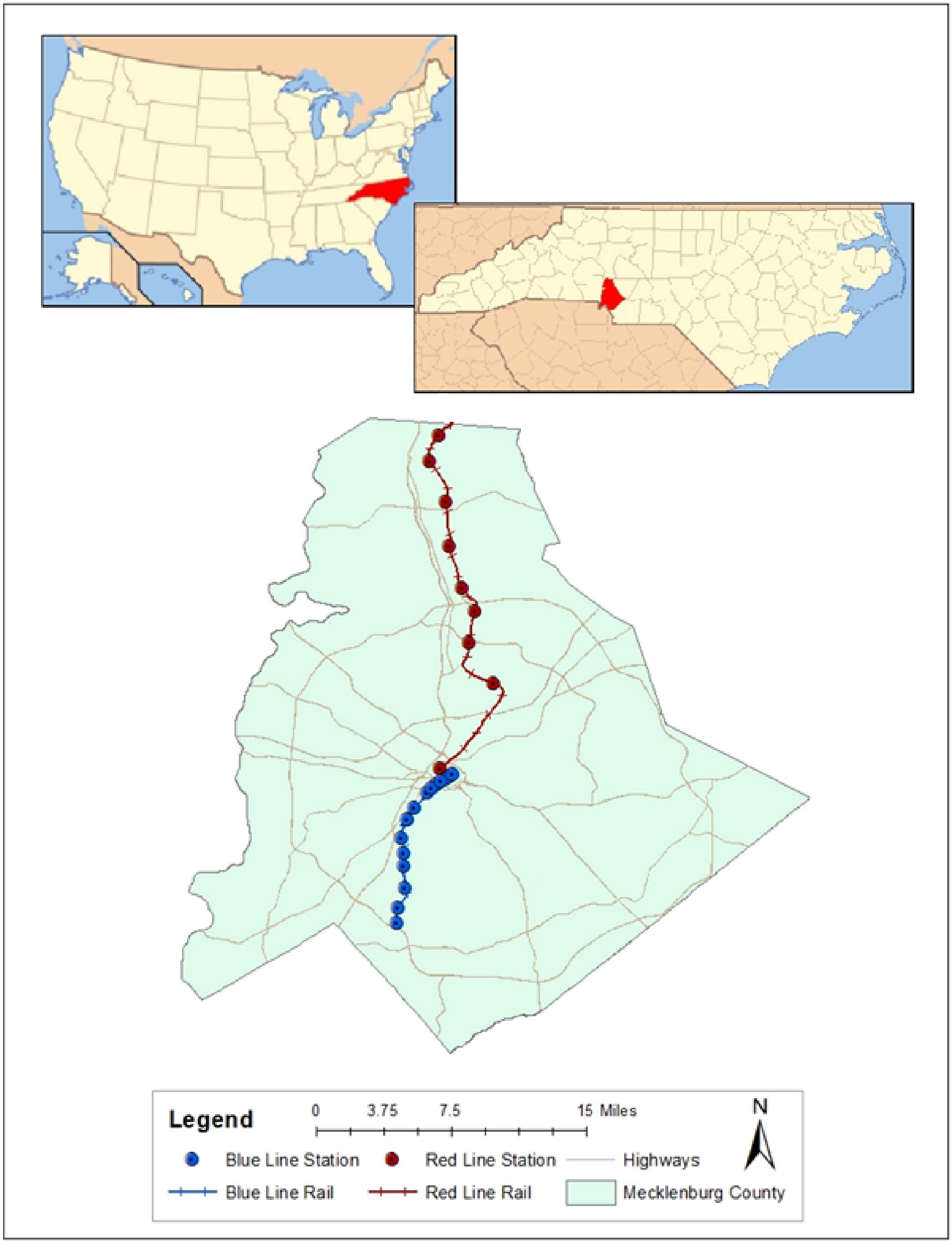

Figure 3.1. Blue Line Service and Planned Red Line in the Charlotte/Mecklenburg County Areas

modelling, the quality of nearby schools, the proximity to amenities such as parks or waterfronts, and details concerning whether garages had automatic doors. In its original form, the MLS database contained only the characteristics of each SFR and its street address. In order to import these data into GIS to create a shapefile, a geolocation database was created using the US Census's 2011 Address Feature TIGER/Line shapefiles, and addresses were matched with a 97% success rate. Street addresses may have gone unmatched for a variety of reasons. Some of the streets were not included in the TIGER/Line shapefile or did not exist according to Google Maps. Because the addresses were manually entered into the MLS database by realtors, it is also possible that street numbers were mistyped or street names misspelled.

Because the MLS data were for individual properties, the source shapefile contained a collection of points. These were filtered according to proximity to the LRT stations. As this study is concerned primarily with the effect that the distance to an LRT station has on SFR prices (price per square foot), only properties within a one-mile radius were considered. For distance measurements, the Euclidean distance between an SFR and an LRT station was used. Similarly, the Euclidean distance between an SFR and highway on-ramp was used.

The average characteristics of residences sold by year for the treatment and control groups are presented in Table 3.1. The treated group includes 14,383 observations. The control group includes 28,936 observations. As Table 3.1 shows, most properties sold within three to four months of being listed. SFRs tended to have two and a half baths and three bedrooms. Although price per square foot values vary from year to year, some of this variation is due to inflation, which was not accounted for. While old houses typically sell for less than newly built houses, the average SFR was built in the 1950s or 1960s. This indicates that the area in which the LRT was constructed was already developed, which may have implications for land use patterns and residents' attitudes toward LRT stations. The values for distance to an LRT station, measured in feet, indicate that most houses sold were more than half a mile away from an LRT station. Although the number of houses sold varies from year to year, some of the differences may be due to the economic recession and subsequent recovery period between 2008 and 2012.

3.1.3 OLS Estimation Results

The OLS specification results for the effects of the announcement are shown in Table 3.2. With an adjusted R^2 of 0.412 and F-stat of 601.4, the model fits the data relatively well. Examination of variance inflation factors indicated that there are not any collinearity problems. Consistent with real estate literature, an increase in the number of bedrooms is significant and negatively associated with the price per square foot, while an increase in the number of full bathrooms and the presence of a detached garage is positively associated with the price per square foot. As the planning literature suggests, an increase in the distance to an LRT station leads to a slight drop in price per square foot, indicating that consumers are willing to pay more for the expected increase in accessibility due to the planned LRT. As expected, properties that have a view of the water were found to be positively associated with price per square foot. Properties not within a subdivision are also associated with a lower price per square foot; it may be that subdivisions offer unobserved amenities that are otherwise inaccessible. The year fixed effects are relative to the omitted year, 2007. Similar to [24]'s findings, there is an increase in the price per square foot of residences as the opening of the LRT approaches. However, given the available data, it is difficult to determine whether this was due to macroeconomic conditions at the time or due to the imminent opening of the LRT.

Table 3.3 presents the results of the OLS specification for the effects of LRT operations. The adjusted R^2 is 0.317; the F-stat is 426.3 and highly significant. According to the OLS specification results, the number of bedrooms, the year in which the property was built, the distance from highway access and LRT stations, the presence of an attached garage, and not being located within a subdivision had negative effects on price per square foot. Houses with many bedrooms or a large lot size due to the

Table 3.1. Average Characteristics of Residences Sold by Year, Treatment (T) and Control (C) Groups

Year	**$/sqft**		**Beds**		**Half Baths**		**Full Baths**		**Year Built**		**Days on Market**		**Distance to Station (ft)**
	T	**C**	**T**	**C**	**T**	**C**	**T**	**C**	**T**	**C**	**T**	**C**	**T**
2001	$110.69	$89.80	2.81	3.45	1.02	0.64	1.71	2.06	1965	1988	85	61	2,294
2002	$116.31	$88.01	2.84	3.44	1.02	0.62	1.77	2.1	1968	1991	119	78	2,327
2003	$123.81	$87.16	2.51	3.53	1.03	0.63	1.77	2.12	1974	1991	141	75	2,128
2004	$126.45	$88.43	2.85	3.45	1.02	0.63	1.8	2.11	1968	1992	127	71	2,288
2005	$143.40	$90.43	2.76	3.48	1.02	0.67	1.77	2.08	1973	1993	110	67	2,223
2006	$166.96	$97.74	2.68	3.41	1.02	0.65	1.76	2.07	1972	1993	101	47	2,197
2007	$172.72	$103.93	2.69	3.41	1.03	0.68	1.81	2.07	1975	1994	120	44	2,216
2008	$163.42	$97.91	3.18	3.45	0.41	0.67	1.81	2.11	1960	1995	120	50	2,612
2009	$139.61	$89.12	3.35	3.48	0.51	0.67	1.84	2.14	1964	1996	128	59	2,778
2010	$137.02	$85.77	3.34	3.5	0.41	0.64	2.03	2.15	1958	1995	134	102	2,734
2011	$130.60	$84.12	3.15	3.5	0.38	0.68	1.86	2.19	1954	1995	137	111	2,493
2012	$142.39	$85.94	3.3	3.52	0.38	0.7	2.01	2.19	1957	1997	132	101	2,574
2013	$152.48	$94.92	3.19	3.48	0.42	0.69	1.86	2.18	1957	1995	116	78	2,476
2014	$178.32	$102.96	3.2	3.47	0.4	0.68	1.99	2.17	1956	1994	104	64	2,568
2015	$189.36	$113.33	3.28	3.53	0.42	0.66	1.96	2.19	1957	2009	95	46	2,611
2016	$208.92	$118.21	3.25	3.49	0.41	0.67	1.95	2.15	1961	1994	84	38	2,558
Average	$150.34	$96.57	2.52	3.47	0.85	0.66	2.19	2.13	1967	1995	114	65	2,333

Table 3.2. OLS and SARAR Results of Announcement Effects

	OLS		SARAR	
	Estimate	*Std. Error*	*Estimate*	*Std. Error*
(Intercept)	101.200***	2.071	204.95***	6.367
	SFR Characteristics			
Distance to highway access	-0.001***	0.000	0.008***	0.002
Distance to LRT station	-0.013***	0.002	-0.002***	0.000
Distance to LRT station^2	4.403E-06***	5.95E-07	6.365E-07*	2.86E-07
Number of bedrooms	-7.392***	0.447	-4.5932***	0.312
Number of full baths	15.74***	0.488	2.1282***	0.286
Number of half baths	-0.1611	0.628	-2.2948***	0.338
Days on market	.007**	0.003	0.007***	0.002
Has attached garage	-0.237	0.652	1.9035***	0.448
Has fireplace	2.469***	0.646	0.586*	0.27
Has granite countertop	26.68***	1.157	11.761***	0.735
Has hardwood floor	9.212***	0.619	1.9064***	0.371
Not in a subdivision	-9.102***	1.102	-3.800***	0.841
On a waterfront	35.17***	3.145	18.414***	2.809
Age of SFR	0.010**	0.003	-0.007***	0.001
Age of SFR2	2.07E-06***	6.07E-07	-1.385E-06***	1.76E-07
	Year Fixed Effects (2007 omitted)			
Y2000	-4.096	7.953	-38.135***	4.872
Y2001	-19.42***	1.161	-24.267***	0.712
Y2002	-20.09***	0.934	-23.268***	0.577
Y2003	-20.39***	1.312	-22.266***	0.805
Y2004	-19.67***	0.858	-20.822***	0.527
Y2005	-13.93***	0.807	-14.072***	0.488
Y2006	-4.674***	0.802	-4.6325***	0.45
Rho	—	—	-0.94318***	0.018
Lambda	—	—	0.96838***	0.002
AIC	206130		192240	
Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1				

presence of an attached garage may suffer from a dilution effect (i.e., bigger houses tend to have lower prices per square foot than smaller houses, ceteris paribus). The negative signs on the estimated coefficients for distance to highway access and distance from LRT stations suggest that homeowners may value transportation access. The latter result is consistent with the majority of the literature reviewed. Houses that are not located within a subdivision may lack access to certain amenities, such as swimming pools or country clubs, causing the price per square foot of those houses to be lower. As in the announcement effects model, the operations effect model includes year fixed effects to account for additional effects not captured by the data, such as the 2008-2009 recession and subsequent recovery period. The resulting coefficient estimates are relative to the omitted base year, 2016.

Moran's I tests for autocorrelation in the residuals from these estimated OLS specifications were significant, which indicates that the F-test and t-test results are inflated. This calls into question the goodness of fit of the a-spatial model. Robust Lagrange Multiplier tests revealed the presence of spatial lag and spatial error; the results of these tests were of the same statistical significance level. Jarque-Bera (JB) tests were also performed on the dependent variables and yielded statistically significant results; the null hypothesis that the data followed a normal distribution was rejected. JB test results indicate that OLS estimators may lead to misleading conclusions. This led to a formulation of a SARAR model.

3.1.4 Spatial Econometric Model Estimation Results

The results of the models for the announcement effects and operations effects of the LRT are shown in Tables 3.2 and 3.3, respectively. The Akaike information criterion (AIC) values indicate that the spatial models are superior to the a-spatial models. The statistically significant (95th percentile) direct, indirect, and total marginal effects are shown in Tables 3.4 and 3.5 for the announcement and operations effects, respectively. The following discussion will focus on the total effects columns. The discussion will

Table 3.3. OLS and SARAR Results of Operations Effects

	OLS		**SARAR**	
	Estimate	*Std. Error*	*Estimate*	*Std. Error*
(Intercept)	129.000***	2.555	238.54***	6.712
SFR Characteristics				
Distance to highway	-0.002***	0	0.001	0.001
Distance from LRT	-0.048***	0.002	0.003***	0.001
Distance from LRT^2	1.435E-05***	7.18E-07	-8.789E-07.	4.83E-07
Number of bedrooms	-11.050***	0.55	-6.5383***	0.316
Number of full baths	27.040***	0.609	2.8541***	0.381
Has attached garage	-2.498***	0.799	-0.63223.	0.368
Has carport	7.802***	1.513	-2.5322*	1.139
Has detached garage	21.700***	1.094	3.342***	0.855
Has driveway	11.040***	1.073	5.416***	0.596
Has fireplace	7.558***	0.791	0.997.	0.529
Has granite countertops	14.370***	0.796	6.715***	0.423
Has hardwood floor	8.810***	0.752	1.840*	0.823
Not in a subdivision	-7.090***	1.524	-2.659*	1.16
Has a patio	5.176***	1.28	2.308***	0.561
Has aesthetic views	11.500***	3.396	3.234	2.678
On a waterfront	42.640***	4.695	23.412***	2.916
Has view of water	8.475***	3.747	-0.206	0.328
Age of SFR	0.013***	0.002	0.001**	0.000
Year Fixed Effects (2016 omitted)				
Y2008	-18.570***	1.34	-14.054***	0.792
Y2009	-29.760***	1.431	-25.410***	0.794
Y2010	-35.270***	1.4	-34.380***	0.81
Y2011	-40.090***	1.383	-38.141***	0.807
Y2012	-35.480***	1.303	-36.762***	0.784
Y2013	-26.370***	1.236	-27.901***	0.753
Y2014	-18.860***	1.188	-20.043***	0.652
Y2015	-8.460***	1.158	-11.293***	0.697
Rho	—	—	-0.809***	0.016
Lambda	—	—	0.973***	0.002
AIC	274070		246970	

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1* first take note of the effects of the controls in each model before focusing on the variable of interest, distance to LRT station.

Announcement Effects

In the case of the announcement effects (Table 3.2), residence-specific characteristics that have a negative effect on the price per square foot include the total number of bedrooms, number of half-bathrooms, and the age of the property unit. The number of bedrooms may contribute to a dilution effect in that it substantially increases the square feet a house occupies while the dependent variable is price per square foot. The negative effect of number of half baths may reflect potential homeowners' preference for full baths. However, the data lacks additional information to ascertain whether or not this is the case. As discussed for the OLS case, houses that are not located within a subdivision may lack access to certain amenities that are exclusive to subdivisions; this may explain the negative effect of that variable.

The number of full bathrooms, the presence of an attached garage, proximity to highway access, the number of days on the market, and the presence of a fireplace, granite countertop, hardwood floor, and waterfront have positive effects on the price per square foot. Potential home-buyers may view the presence of multiple full bathrooms in a house to be worth a premium. Additionally, they may be willing to pay a premium for the convenience of having an attached garage. The effects of the year dummy variables are also consistent with those in the OLS-based model and indicate an increase in price per square foot over time; however, it may be that some of the indirect effects are due to macroeconomic conditions or other unobserved variables not captured in the MLS data. It is important to note that these yearly fixed effects are relative to 2007, the year before the Great Recession began. Finally, the control variable for distance to highway access point has a positive effect on SFR price per square foot. This indicates that the further away from highway access a property is, the more expensive it is in terms of price per square foot. This could be due to

Table 3.4. Marginal effects of SARAR Model of Announcement Effects

	Direct	**Indirect**	**Total**
	SFR Characteristics		
Distance to highway access	0.009	-0.004	0.004
Distance to LRT station	-0.002	0.001	-0.001
Distance to LRT station^2	6.83E-07	-3.56E-07	3.28E-07
Has attached garage	2.043	-1.064	0.98
Number of full baths	2.284	-1.189	1.095
Number of half baths	-2.463	1.282	-1.181
Number of bedrooms	-4.930	2.566	-2.364
Days on market	0.008	-0.004	0.004
Has fireplace	0.629	-0.327	0.302
Has granite countertop	12.624	-6.571	6.053
Has hardwood floor	2.046	-1.065	0.981
Not in a subdivision	-4.079	2.123	-1.956
Age of SFR	-0.008	0.004	-0.004
Age of SFR^2	-1.49E-06	7.74E-07	-7.13E-07
On a waterfront	19.764	-10.288	9.476
	Year Fixed Effects (2007 omitted)		
Y2000	-40.933	21.307	-19.625
Y2001	-26.047	13.559	-12.488
Y2002	-24.974	13.000	-11.974
Y2003	-23.899	12.441	-11.458
Y2004	-22.349	11.634	-10.715
Y2005	-15.105	7.863	-7.242
Y2006	-4.972	2.588	-2.384

potential home-buyers wishing to avoid nuisances associated with highways access points, such as increased vehicle traffic and noise.

As in the OLS case, the negative effect distance from LRT stations, or alternatively stated, the positive effect of proximity to an LRT station, indicates that home-buyers are willing to pay more to live close to the planned LRT route. Although the \$0.001 per square foot total marginal effect may seem small, it is important to note that distance was measured in feet. Furthermore, the statistically significant distance to LRT station squared indicates that the price gradient is not linear; however, the magnitude of the total effect for this variable suggests that the curvature is quite small. These results show that a one mile increase in distance from an LRT station leads to a -$4.68 per square foot premium in the study area.

Operations Effects

According to the total marginal effects of the spatial model examining the operations effects on property values (Table 3.4), characteristics that have a negative effect on the price per square foot include the total number of bedrooms, the presence of a carport, whether the SFR has a view of the water, and whether the SFR is located in a subdivision. The number of bedrooms may contribute to a dilution effect because they increase the square feet a house occupies while the dependent variable is price per square foot. The negative effect associated with the presence of a carport may reflect home-buyers' views regarding carports compared to alternative methods of storing one's car, such as in an attached garage. Properties with a view of the water are typically considered to be more expensive than properties without such amenities. That they have a negative effect on price per square foot indicates that further study may be required to determine the cause. It could be that these properties are located near waterfront properties, and as such are viewed as inferior relative to properties in their immediate vicinity. Indeed, that the indirect effect is positive for this variable suggests that nearby property values are high and have a positive effect on properties with a view of the water but not direct access. Finally, as discussed for the OLS case, houses that are not located within a subdivision may lack access to certain amenities that are exclusive to subdivisions; this may explain the negative effect of that variable (-$1.47 per square foot).

The number of full bathrooms, the presence of a detached garage or driveway, location within a cul-de-sac or on a private lot, whether the house included a fireplace, granite countertop, hardwood floor, or had a patio, and location on a waterfront had positive effects on the price per square foot. Potential home-buyers may view the presence of multiple full bathrooms in a house to be worth a premium. Although

Table 3.5. Marginal Effects of SARAR Model of Operations Effects

	Direct	**Indirect**	**Total**
SFR Characteristics			
Distance from LRT station	0.004	-0.002	0.002
Distance from LRT station^2	-9.24E-07	4.38E-07	-4.86E-07
Has attached garage	-0.665	0.315	-0.350
Number of full baths	3.002	-1.424	1.578
Number of bedrooms	-6.876	3.261	-3.615
Has carport	-2.663	1.263	-1.400
In a cul-de-sac	1.445	-0.685	0.760
Has detached garage	3.514	-1.667	1.848
Has driveway	5.696	-2.702	2.995
Has fireplace	1.049	-0.497	0.551
Has granite countertops	7.062	-3.349	3.713
Has hardwood floors	1.935	-0.918	1.017
Not in a subdivision	-2.797	1.326	-1.470
Has a patio	2.427	-1.151	1.276
Age of SFR	0.001	-0.001	0.001
On a waterfront	24.622	-11.677	12.944
Has view of water	-0.216	0.103	-0.114
Year Fixed Effects (2016 omitted)			
Y2008	-14.78	7.01	-7.77
Y2009	-26.723	12.674	-14.049
Y2010	-36.156	17.148	-19.008
Y2011	-40.112	19.024	-21.088
Y2012	-38.661	18.336	-20.325
Y2013	-29.342	13.916	-15.426
Y2014	-21.079	9.997	-11.082
Y2015	-11.876	5.633	-6.244

houses may have both attached and detached garages, the square footage of a detached garage is not included in the calculation of the dependent variable; this may explain the differing effects between these two garage types. An alternate explanation may be a consequence of the age of the SFRs that were purchased. Note that the average SFR purchased was built in the mid-1960s. Attached garages are a relatively new development compared to detached garages. It could simply be that attached garages when added to older homes result in a less visually appealing SFR, whereas properties with detached garages would preserve their original aesthetics. Similarly, home-buyers may view having a driveway to be worth paying a premium for, relative to not having a driveway. Relative to houses on a street, houses in a cul-de-sac are in shorter supply, and potentially offer home-buyers more privacy that they are willing to pay a premium for. Finally, additional amenities such as a fireplace, granite countertops, hardwood floors, or a patio may contribute to higher square footage prices as these are generally considered desirable features. Likewise, waterfront properties tend to command a higher price ceteris paribus than properties not located on a waterfront since water access is typically viewed as a luxury. Interestingly, the distance to highway access variable is insignificant in the spatial model. This result highlights the problem of having inconsistent and biased estimates due to not properly accounting for spatial lag and spatial error. Year fixed effects helped to control for unobserved variables that affected price per square foot of SFRs. Most notably, this includes the 2008-2009 economic recession. The magnitude and direction of these fixed effects are with respect to the reference year, 2016. The total marginal effects associated with these year fixed effects are consistent with the recession that began mid-2008 and continued through 2009. The subsequent negative marginal effects suggest that the housing market experienced a lag as well as a gradual recovery at a rate generally consistent with that which is depicted in the Case-Shiller House Price Index for Charlotte during the same time period. However, the dataset used for this analysis is insufficient to further analyze the recession's effect on the local housing market.

Unlike the a-spatial OLS results, the distance from LRT station's total marginal effect is positive. This indicates that an inverse relationship exists between proximity to an LRT station and price per square foot of SFRs. In other words, home-buyers were willing to pay a premium to be located further away from stations. An alternative model specification using categorical distances consisting of 200-foot bands between SFRs and LRT stations was also evaluated. Despite having much faster computational times, no significant differences in the results were found. These results run contrary to the findings in the previous literature, where houses closer to LRT stations have a premium associated with the greater access to transportation infrastructure. However, unlike other study locations, residents in Charlotte on average use public transit less often. Additionally, the average age of properties near LRT stations may be greater in Charlotte than in Phoenix or Denver. On average the houses in the treatment group were built in the 1950s and 1960s. This higher average age indicates that land use patterns have been set for a long period of time, which may affect potential residents' attitudes towards the LRT and lead to a preference for houses situated away from stations. However, more research is required to fully explain the causes of this result. The squared distance from LRT station variable was weakly significant and negative, indicating that the price gradient may not be linear. However, as in the case of the announcement model, the magnitude of the marginal effect associated with this term is small. Based on these results, a one mile increase in distance from the LRT station leads to a \$9.78 per square foot premium in the study area.

3.1.5 Conclusions

This section examined the effect that proximity to a planned LRT station has on the price per square foot of transit-adjacent single -family residences in Charlotte, North Carolina, using a quasi-experimental spatial econometric model. After controlling for spatial lag and error effects, our findings in the announcement model are consistent with the planning literature and indicate that houses farther away from planned LRT stations had lower prices per square foot than houses closer to these stations. These results demonstrate that the announcement of an LRT line can have significant effects on land values even before construction and operations begin. However, our results regarding the capitalisation benefits of LRT operations are contrary to the results found in the literature, in that we found houses farther away from an LRT station commanded higher prices per square foot than houses closer to an LRT station. The differences in findings may be attributed to residents' attitudes towards transit. If residents near LRT stations do not value the presence of those stations, then they may be willing to pay a premium to be located farther away from such stations. For instance, potential residents near a LRT station may view increased traffic due to non-residents using the facility as a nuisance that outweighs the potential benefits of living in close proximity to the station. This may especially be the case in park-and-ride type stations. These effects affected only the treatment group, the 14,383 SFRs within one mile of the proposed, and eventually constructed, LRT stations along the Blue Line corridor.

As LRT becomes an increasingly popular transportation option, these results can help urban planners better understand its effects for residents near proposed lines. Most notably, in this study the announcement of an LRT led to an increase in the sale values of SFRs close to the proposed LRT stations. This increase in the price per square foot of houses was not sustained, however; model results from the operations stage indicate that properties farther away from LRT stations increased in value when SFR-specific characteristics were controlled for. These results suggest that there may have been price inflation due to speculation by home sellers and developers leading up to the opening of the LRT. Further, the decrease in residential property values with respect to increased proximity to the stations after the LRT began operations indicate that residents may have gone from viewing the system as a beneficial service (i.e., one that increases accessibility) to a nuisance (e.g., a source of noise pollution or a cause of increased traffic through residential neighborhoods).

These results also highlight the importance of developing spatial models to examine the effect of distance. In the a-spatial operations effects model, for instance, the distance to LRT station variable was statistically significant and negative; a naive interpretation would be that home-buyers were willing to pay a premium on SFR price per square foot in order to live closer to LRT stations. Upon controlling for spatial autocorrelation and spatial lag however reveals that opposite to be true.

Due to the data set used in this study, only residential values were considered. Additionally, rather than evaluate housing prices directly, the price per square foot had to be used due to dataset limitations. However, commercial and industrial land values may also be affected by distance to proposed LRT stations. Methodological improvements include using propensity score matching to improve the identification of the control group. Additionally, the effect of distance from LRT stations on SFR values may be anisotropic; future research evaluates the extent to which relative directions between SFRs, the central business district, and LRT stations affect home prices. Finally, location theory suggests that household demographics (e.g., income) play an important role in where people live. The research can also be enhanced with the inclusion of census or survey data that better explain the demand for housing relative to LRT station locations.

3.2 Extension LRT Line - Charlotte, NC

In order to facilitate ease of comparison, the same methodology is applied to real estate data for single family residences along the Lynx Blue Line Extension. The extension increases track mileage by 9.7 miles and includes another 11 stations. The preliminary planning for the extension began in 2011, and in 2013 construction of the line began. In March 2018, the Blue Line Extension began operations. Figure 3.2 shows the location of the extension track and LRT stations with respect to the original Lynx Blue Line track.

3.2.1 Methodology

As in the previous case study concerning the Blue Line, this analysis uses a quasi-experimental spatial model to account for the design, construction, and operations periods for the extension. However, unlike the prior case study, all major event (preliminary planning, construction, and operations) periods have been combined into a single model. This was primarily due to a lack of data during the operations phase because the Extension Line has only one and a half year's worth of data.

Figure 3.2. Blue Line and Extension Services in the Charlotte/Mecklenburg County Area

The quasi-experimental difference-in-differences structure is as follows: The treated group includes houses within a 1-mile radius of LRT stations while the treated group includes houses between a 2- and 3-mile radius of LRT stations. For distance measurements, the Euclidean distance between an SFR and an LRT station was used. Similarly, the Euclidean distance between an SFR and highway on ramp was used. Because there were no planned additional LRT lines during this time period, the control group selection differs from that of the previous section. Thus, it is assumed that all SFRs within a 3-mile radius of LRT stations are similar. Indeed, many of the homes in the control and treated groups share the same [home developers'] neighborhoods. Additional robustness checks in the form of alternative model specifications, including the use of difference-in-difference-in-differences, were conducted as well, particularly to determine whether or not grouping the three phases of the Extension Line's life cycle affected model results.

3.2.2 Data

The data used in this analysis includes all SFR sales from 2011, when the Extension Line was announced and preliminary planning began, to 2019, the most recent calendar year of data available from MLS.com of all SFRs within a 3-mile radius of LRT stations. Thus, the data includes the full planning and construction periods, as well as 19 months of operations. Table 3.6 shows the average characteristics, by year sold, of the treated and control groups. As with the data in section 1, characteristics of residences included the number of bedrooms, bathrooms, and half baths; square footage; and street address; two fields containing the realtor's private remarks and public comments about each property also exist, from which several dummy variables such as the presence of marble countertops, hardwood floors, and others., could be created. Street addresses were geolocated using the US Census's 2018 TIGER/Line shapefile with a 96% success rate. Reasons for unmatched SFRs include misspelled addresses and addresses that appeared on newer streets not captured by the 2016 TIGER/Line shapefile. The data cleaning process also removed SFRs with missing characteristics, such as those lacking square feet values, number of bathrooms, and sales close dates. This led to a total of 17,836 households used in the analysis, with 2,916 SFRs in the treated group and 14,920 in the control group.

Examination of Table 3.6 reveals that, on average, houses in the control group tended to be nearly a decade older and slightly larger than houses in the treatment group. At the same time, the average ages of the houses, 51 and 42, suggest that on average, houses in both the treatment and control groups, respectively, were built in relatively established neighborhoods that have existed long before the construction of the Extension Line. Furthermore, despite adjusting for inflation using the Case-Shiller housing index, the average price per square feet of houses sold are on average similar. Additionally, houses in both groups spent an average of 8 months on the market before selling. The average distance from the closest LRT station in the treatment group is approximately 6/10ths of a mile, while that value is a little over to two miles in the control group.

3.2.3 OLS Estimation Results

Table 3.7 presents the results of the a-spatial model, which controls for SFR characteristics and uses year fixed effects. The fitted model has an adjusted R^2 of 0.47 and F-statistic of 448.3. Variance inflation factors did not reveal collinearity problems.

Results indicate that the presence of an attached garage, an increase in the number of bedrooms, location within a cul-de-sac, and not being in a subdivision are significant factors affecting price per square feet of an SFR; these variables are negatively associated with the price per square foot. An increase in the number of bathrooms, presence of a detached garage or driveway, and indoor amenities such as a fireplace, granite countertops, and hardwood floors are associated with adding value to the SFR. These results are similar to the findings in section 1.3. The OLS results also indicate that older SFRs are worth more in terms of dollars per square foot, although at a declining rate. Interestingly, SFRs with aesthetic views are more expensive but those on a waterfront or located on a wooded lot are worth less. Increasing the distance to highway access is associated with a lower price per square feet value. This suggests that residents account for vehicle network accessibility when choosing where to live. The OLS model also indicates that for the treatment group, increasing distance from the LRT station is associated with a lower price per square feet. The year fixed effects are relative to the omitted year, 2011. As recovery from the 2008-2009 recession was still ongoing in 2011, it is no surprise that relative to 2011, the year fixed effects are all positive and significant. However, without additional data it is difficult to determine to what extent the magnitude of the coefficient estimates may

Table 3.6. Average Characteristics of Residences Sold By Year, Treatment and Control Groups

Year	**$/sqft**		**Beds**		**Half Baths**		**Full Baths**		**Built**		**Days on Mkt**		**LRT Dist.**	
	T	**C**	**T**	**C**	**T**	**C**	**T**	**C**	**T**	**C**	**T**	**C**	**T**	**C**
2011	$61.80	$91.27	4.01	4.31	0.35	0.52	1.74	1.99	1969	1975	228	241	3,480	11,055
2012	$70.79	$100.34	3.94	4.33	0.42	0.52	1.68	2.02	1974	1974	238	256	3,393	11,043
2013	$93.97	$116.90	3.98	4.3	0.29	0.55	1.69	2.03	1961	1975	253	253	3,206	10,986
2014	$112.93	$120.71	3.95	4.32	0.31	0.54	1.72	2.02	1965	1977	251	247	3,191	11,014
2015	$133.48	$129.37	3.91	4.36	0.33	0.53	1.73	2.04	1965	1978	251	237	3,402	11,070
2016	$150.95	$137.45	4.1	4.37	0.35	0.53	1.85	2.04	1964	1978	214	246	3,358	11,096
2017	$157.55	$143.05	4.09	4.34	0.35	0.51	1.75	2.03	1971	1976	218	234	3,403	11,186
2018	$176.94	$166.73	4.09	4.39	0.32	0.51	1.83	2.05	1969	1978	244	231	3,344	10,973
2019	$200.34	$179.21	4.13	4.39	0.34	0.53	1.86	2.08	1973	1979	236	234	3,363	11,084
Average	$136.29	$135.42	4.03	4.35	0.34	0.53	1.77	2.04	1968	1977	238	242	3,341	11,049

be attributed to macroeconomic conditions in the region at the time or other local activities.

3.2.4 Spatial Econometric Model Estimation Results

Upon running the a-spatial model, Moran's I (I = 0.684, p < 2.2e-16) and Lagrange multiplier tests were conducted. As the robust Lagrange multiplier test for error and lag were equally significant, a spatial Durbin model (SARAR) was deemed most appropriate. Further, a Jarque-Bera test (χ^2 = 4326.8, p < 2.2e16) indicated the presence of heteroskedasticity. Thus, a general moments estimation of a Cliff-Ord type model with heteroskedastic innovations based on [82] was used.

The results of the spatial model are also shown in Table 3.7. Upon controlling for spatial lag, spatial error, and heteroskedasticity, some variables have lost statistical significance while others have gained significance. These results indicate the importance of using the appropriate model specification, as solely relying on the a-spatial model would have led to incorrect conclusions. In particular, the variable indicating whether an SFR has a patio is positive and significant. Variables that are no longer significant in the spatial model include the SFR located within a cul-de-sac, the age and age squared of the SFR, whether or not the SFR is on a waterfront, the year 2012 fixed effects dummy, and, interestingly, the distance from the LRT station for houses in the treatment group.

While null results can be difficult to explain, there are several possible reasons behind why the distance from the closest LRT station to an SFR has no statistically significant effect. Through the lenses of utility maximisation theory and revealed preferences via purchases, one explanation for the null result is that on average people who purchased homes in the treatment area were not interested in the LRT. If they do not view proximity to LRT stations as an amenity, then it makes sense that they would not pay a premium to live near it. Less polarizing, it could be that buyers did value the increased mobility that living near a LRT station would provide but were also

Table 3.7. OLS and SARAR Results of Extension Line Model

	OLS		SARAR	
	Estimate	Std. Error	Estimate	Std. Error
(Intercept)	29.000***	4.048	-16.141***	2.277
	SFR Characteristics			
Has attached garage	-15.640***	1.323	-2.200***	0.620
Number of full baths	48.510***	0.855	9.001***	0.693
Number of half baths	15.760***	0.996	1.237*	0.624
Number of bedrooms	-17.470***	0.849	-8.400***	0.486
Has detached garage	46.870***	1.791	10.080***	1.180
Has driveway	10.160***	1.183	6.6096***	0.711
Has fireplace	9.270***	1.181	3.683***	0.663
Has granite countertops	16.000***	1.309	8.114***	0.687
Has hardwood floor	13.830***	1.181	2.894***	0.694
Not in a subdivision	-47.370***	2.592	-12.812***	1.443
Has a patio	3.350	2.182	2.558*	1.026
Private lot	4.639	3.516	3.851.	2.129
Age of SFR	1.177***	0.020	-0.001	0.015
Age of SFR^2	-0.001***	0.000	0.000	0.000
Has trees	-2.261	2.172	2.503.	1.299
Has aesthetic views	23.570***	4.864	8.670*	3.268
On a waterfront	-38.290**	14.240	-0.836	6.961
Wooded lot	-19.740***	3.342	-7.422***	1.569
	Year Fixed Effects (2011 omitted)			
Y2012	5.611*	2.279	1.802	1.288
Y2013	18.430***	2.208	13.397***	1.228
Y2014	23.530***	2.191	20.241***	1.196
Y2015	33.150***	2.189	29.796***	1.169
Y2016	43.100***	2.192	43.442***	1.225
Y2017	50.680***	2.736	55.763***	1.579
Y2018	70.100***	2.209	72.365***	1.258
Y2019	82.700***	2.191	83.350***	1.288
	Distance Variables			
Treated	5.491	4.060	1.073	1.946
Distance to highway access[1]	-0.078***	0.000	-0.113*	0.000
Distance from LRT station[1]	-0.076***	0.000	0.009	0.000
Treated:LRT Distance[1]	-0.366***	0.001	-0.561	0.000
lambda	—	—	0.934***	0.007
rho	—	—	-0.399***	0.023
AIC	111,921.80		93,995.40	

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1*

Distance measured in thousands of feet

concerned with the possibility of dealing with nuisances associated with being near a transit hub, such as increased neighborhood traffic and sound and noise pollution. Perhaps these opposing preferences cancelled each other out, and the null result is reflective of both factors. Without further survey data however, it is impossible to know for certain.

A third reason unrelated to buyer preferences could be attributed to the physical urban geography in the local area. The average age of SFRs in the treatment area is around 60. This indicates that the neighborhoods in the region have been established for some time. Additionally, Charlotte is a vehicle-oriented city. These two factors could mean that proximity to an LRT station would be viewed as a redundant transportation option by potential buyers, as the road network would suffice—indeed, the spatial model results concerning the significant and positive variables associated with vehicle ownership (e.g., presence of a driveway, detached garage, etc.) suggest that home-buyers already own personal vehicles. Finally, the location of the Extension Line and the destinations offered may affect how home-buyers view the worth of living near a station. The Extension Line's southernmost station is in the downtown central business district, while its northernmost station is located near UNC-Charlotte (UNCC). Indeed, one of the primary goals stated in the planning phase was to connect these two areas of the city. However, if residents along the line do not attend UNCC or work in the CBD, then they may appreciate the worth of having an LRT line but be unwilling to pay a premium to live near one of its stations as it provides little practical utility.

As stated in Section 1.4, one cannot directly use estimated coefficients from spatial econometric models as marginal effects. Thus, Table 3.8 shows the calculated direct, indirect, and total effects of the significant variables used in the model. Direct effects are characteristics of an SFR's price per square foot that affects it directly, whereas indirect effects are the characteristics of nearby SFRs that affect the SFR's price per square foot. The total effect is the sum of direct and indirect effects. It should be noted that values have been rounded to the nearest hundredth while the dependent variable, price per square feet is in dollars. Some of the total effects may seem rather large. For instance, the inclusion of an additional full bathroom in an SFR contributes to an average effect of an increase of $135.50 in price per square feet in SFRs in the study area; this value can be decomposed into a direct effect ($12.64; the addition of a bathroom in the SFR itself) and an indirect effect ($122.85; the presence of an additional bathroom in houses neighboring the SFR). All values have been adjusted to be in 2016 dollars to ease comparisons in later sections. As with the case of the discussion in subsections 1.4.1 and 1.4.2, it is more difficult to interpret the average effect values of the year fixed effect variables. These variables capture much variation, including changes in macroeconomic conditions within each year (e.g., recovering from the Great Recession), as well as any changes in the local subdivision a SFR is in that occurred in a particular year (e.g., the opening of a local pool). Without further data, it is difficult to ascertain what exactly is being measured within the fixed effects. In any case, the pattern of growth in total effects between the years 2013-2019 is consistent with the Case-Shiller Housing Index[2] for the Charlotte, NC area. It is also useful to examine the average total effects to determine the extent of impact a significant result may actually have on price per square feet. A prime example of this is with the distance from highway access variable. While it is statistically significant, as seen in Table 3.8, the direct effect is a small number. However, it is important to keep in mind that distance is measured in feet. These results indicate that the for every foot a SFR is closer to a highway onramp, it's value in price per square feet to potential home-buyers drops by 0.002. Assuming a constant or linear relationship, this becomes a decrease of $10.36 in price per square feet for every mile closer to a highway onramp.

[2]The S&P/Case-Schiller Housing Index tracks monthly changes in the value of residential housing via the purchase price and resale value of SFRs from 1980 onward.

Table 3.8. Spatial Model Average Direct, Indirect, and Total Effects

Variable	Direct	Indirect	Total
SFR Characteristics			
Has attached garage	-3.089	-30.028	-33.117
Number of full baths	12.636	122.854	135.49
Number of half baths	1.736	16.880	18.616
Number of bedrooms	-11.792	-114.644	-126.435
Has detached garage	14.151	137.579	151.730
Has driveway	9.279	90.213	99.492
Has fireplace	5.170	50.261	55.431
Has granite countertops	11.390	110.741	122.131
Has hardwood floor	4.062	39.496	43.558
Not in a subdivision	-17.987	-174.875	-192.862
Has a patio	3.590	34.907	38.497
Has aesthetic views	12.199	118.603	130.801
Wooded lot	-10.419	-101.302	-111.721
Year Fixed Effects (relative to Y2011)			
Y2013	18.807	182.855	201.662
Y2014	28.415	276.259	304.674
Y2015	41.829	406.680	448.509
Y2016	60.985	592.925	653.910
Y2017	78.283	761.102	839.385
Y2018	101.589	987.692	1089.281
Y2019	117.010	1137.625	1254.635
Distance Variables			
Distance to highway access	-0.159	-1.549	-1.708

Distance is measured in thousands of feet.

3.2.5 Robustness Checks

The Treated:LRT Distance variable in Table 3.7 is the main variable of interest to this case study. It could be that the result shown is sensitive to the model specification. Thus, two additional spatial models were run using different specifications (Table 3.9). The model shown in Table 3.7 is labelled "Original Model," while the additional models are marked "Model 1" and "Model 2." For ease of comparison, only the spatial versions of these robustness check models are shown. Distances are measured in thousands of feet.

Table 3.9.: Robustness Checks

	Original	Model 1	Model 2
	Estimate	Estimate	Estimate
	Std. Err.	*Std. Err.*	*Std. Err.*
(Intercept)	-16.141***	-15.963***	-16.334***
	2.277	*2.256*	*2.309*
SFR Characteristics			
Has attached garage	-2.2001***	-1.400*	-2.240***
	0.620	*0.633*	*0.620*
Number of full baths	9.001***	8.434***	9.088***
	0.693	*0.722*	*0.694*
Number of half baths	1.237*	1.177.	1.272*
	0.624	*0.658*	*0.625*
Number of bedrooms	-8.400***	-8.758***	-8.468***
	0.486	*0.498*	*0.486*
Has carport	-0.430	-1.455	-0.433
	1.104	*1.14*	*1.108*
On a corner lot	0.850	0.387	0.869
	0.949	*0.967*	*0.949*
In a cul-de-sac	-0.570	-0.853	-0.616
	0.712	*0.713*	*0.712*
Days on market	0.000	0.001	0.000
	0.002	*0.002*	*0.002*
Has detached garage	10.080***	10.135***	10.120***
	1.18	*1.193*	*1.185*
Has driveway	6.670***	5.223***	6.532***
	0.711	*0.699*	*0.713*

continued on next page

Table 3.9.: *continued*

	Original Estimate *Std. Err.*	Model 1 Estimate *Std. Err.*	Model 2 Estimate *Std. Err.*
Has fireplace	3.683***	3.525***	3.73***
	0.663	*0.679*	*0.663*
Has granite countertops	8.114***	8.287***	8.269***
	0.687	*0.690*	*0.686*
Has hardwood floor	2.894***	2.533***	2.850***
	0.694	*0.703*	*0.694*
Not in a subdivision	-12.812***	-11.220***	-12.984***
	1.443	*1.456*	*1.440*
Has a patio	2.558*	2.500*	2.626*
	1.026	*1.05*	*1.026*
Private lot	3.851.	4.148*	4.065.
	2.129	*2.066*	*2.144*
SFR Age	-0.001	0.001	0.003
	0.015	*0.016*	*0.015*
SFR Age^2	0.000	0.000	0.000
	0.000	*0.000*	*0.000*
Has trees	2.503.	2.951*	2.397.
	1.299	*1.251*	*1.298*
Has aesthetic views	8.670**	7.544*	7.67*
	3.268	*3.197*	*3.265*
On a waterfront	-0.837	-0.604	-0.634
	6.961	*7.268*	*7.077*
Has view of water	9.266	4.648	9.128

continued on next page

Table 3.9.: *continued*

	Original	Model 1	Model 2
	Estimate	Estimate	Estimate
	Std. Err.	*Std. Err.*	*Std. Err.*
	5.900	*4.835*	*5.842*
Wooded lot	-7.422***	-6.237***	-7.479***
	1.569	*1.499*	*1.570*
Year Fixed Effects (2011 omitted)			
Y2012	1.802	—	1.834
	1.288	—	*1.286*
Y2013	13.397***	11.288***	13.421***
	1.228	*1.213*	*1.225*
Y2014	20.241***	18.31***	20.350***
	1.196	*1.169*	*1.194*
Y2015	29.796***	27.863***	29.835***
	1.169	*1.145*	*1.168*
Y2016	43.442***	41.330***	43.566***
	1.225	*1.199*	*1.226*
Y2017	55.763***	53.155***	55.883***
	1.579	*1.537*	*1.581*
Y2018	72.365***	69.024***	72.583***
	1.258	*1.227*	*1.259*
Y2019	83.350***	—	—
	1.288	—	—
Distance Variables			
Treated	1.073	-0.370	-2.249
	1.946	*1.903*	*2.151*

continued on next page

Table 3.9.: *continued*

	Original	Model 1	Model 2
	Estimate	Estimate	Estimate
	Std. Err.	*Std. Err.*	*Std. Err.*
Distance to highway access	-0.113*	-0.085.	-0.112*
	0.000	*0.000*	*0.000*
Distance from LRT station	0.009	0.044	0.072
	0.000	*0.000*	*0.000*
Built	—	77.399***	85.462***
	—	*1.305*	*3.504*
Treated:LRT Distance	-0.561	-0.823.	-0.114
	0.000	*0.000*	*0.001*
Treated:Built	—	12.715***	21.160**
	—	*2.270*	*7.232*
LRT Distance1:Built	—	—	0.402
	—	—	*0.000*
Treated:LRT Distance:Built	—	—	-2.918
	—	—	*0.002*
lambda	0.934***	0.946***	0.932***
	0.007	*0.007*	*0.007*
rho	-0.399***	-0.486***	-0.391***
	0.023	*0.024*	*0.023*
AIC	93,995.40	93,935.90	93,930.60

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1*

The key difference between the original model and the two additional models shown in Table 3.9 is that in the latter models the timing aspect of the Extension Line is more fully taken into account. This is done through the use of the interaction terms Treated:Built, LRT Distance:Built, and Treated:LRT Distance:Built. Because the Extension Line began operations in 2019, the year fixed effects dummy for 2019 was omitted due to collinearity. Additionally, in Model 2, the year dummy for 2012 was omitted as its inclusion led to a singular matrix error when attempting to run the spatial model. In general, the SFR characteristics and year fixed effects indicators were consistent in direction and statistical significance between the three models, although some SFR characteristics did become more significant in the robustness check models. Additionally, the magnitude of the year fixed effects' estimated coefficients are similar across the models.

In the robustness check models, the Treated:LRT Distance variable measures the effect that distance from the LRT has on SFR sales prices in the treated group. The variable Treated:Built controls for the effects of the announcement and construction phases of the LRT on SFR sales. The LRT Distance:Built interaction is included to see if SFR distance from LRT stations during the pre-operations phases affected sales prices. Finally, the Treated:LRT Distance:Built interaction tests whether or not sales of SFRs in the treatment group specifically were affected during the pre-operational phases of the LRT. While the interaction variables that include SFR distance from LRT stations remain insignificant, indicating that the results of the original model as shown in Table 3.7 are fairly robust, it is interesting to note that in both Models 1 and 2, the Treated:Built interaction is positive and significant. This suggests that during the planning and construction phases, the price per square foot of houses within one mile of the station construction locations were positively associated with the ongoing activities. However, it is difficult to determine whether there lies a causal relationship between the construction of the stations and housing sales prices. In other words, there could have been ongoing local efforts to improve neighborhoods' amenities that occurred concurrently with the construction, leading to the positive effect shown in the results. Without additional data that captures the "on-the-ground" changes in that part of Charlotte between 2011 and 2018, one cannot claim that the construction of LRT stations definitively led to an increase in price per square foot of the houses that were on the market at that time. Using currently available secondary data sources, one can only observe the positive associations.

Despite the insignificant results for the spatial model(s), it is interesting to observe the differences in model outcomes between Sections 3.1 and 3.2. The next section will examine these differences in more detail.

3.3 Homebuyer's willingness to pay due to LRT experience and policy implications

As mentioned in the introductory section, a goal of this chapter is to examine the effect that potential home-buyers' prior experience with LRTs in a particular region may have on their purchasing decision of a SFR. Section 3.1 examined a case in which the Blue Line was established in Charlotte, NC. Because this was the first LRT line to have been built and operated in the metro area, it was assumed that residents of the city had little direct experience with light rail.

To aid the comparison, Table 3.10 shows the average total effect of the statistically significant accessibility variables in each of the three spatial models. Distance variables are measured in feet. The key finding in Section 3.1 is that home-buyers paid a \$0.001 per square foot premium for every foot a property was closer to the closest planned LRT station during the construction phase on each square foot of the SFR. In the operations phase, however, home-buyers paid a premium of \$0.002 per square foot for every foot a property was further from the closest LRT station. One explanation of the shift is that home purchasers in the construction phase were excited by the prospect of living near one of these new transit centers while buyers in the operations phase were more concerned with the nuisances associated with being near a transit center.

Section 3.2 examined the effect that the Extension Line had on its nearby housing market. The Extension Line's construction began nearly four years after the Blue Line had been in operations, and itself began operations eleven years after the Blue Line had been running. With both the Blue Line and Extension Line operating in the same city (and physically connected), it is assumed that potential home-buyers interested in properties near the Extension Line would have some degree of familiarity with LRT. Unlike the first case study, data limitations did not allow for separate models capturing construction and operations phase effects on SFR sales. The key finding from the second case study is that the distance between an SFR and its closest LRT station had no effect on the price per square feet of the SFR.

Table 3.10. Average Total Effects of Statistically Significant Accessibility Variables

	Blue Line		Extension Line
	Construction	Operations	
Station distance	-0.001	0.002	Not significant
Station distance^2	3.28E-07	-4.86E-07	—
Onramp distance	0.004	Not significant	Not significant

These results have some important policy implications for both cities with existing LRT lines and cities considering the extension or addition of LRT services. LRT systems are typically constructed using local, state, and federal funds. Locally, urban areas may impose additional taxes on residents to help pay for the construction and operations of LRT services. Tax revenue may also be used to pay back state and federal loans. If, however, local residents do not value proximity to LRT stations in their housing location decision, then they may also not value the LRT service itself. This raises two issues. For cities with LRT systems, it could be that local taxes are assessed on non-users to pay for others to use, bringing into question whether such taxes are just. For cities without LRT systems but considering their implementation, it begs the question whether or not having LRT is truly the best use of transportation funding. Additionally, if residents become ambivalent towards LRT, as seems to be the case shown in Table 3.10, it may not bode well for cities with LRT systems who need to pay back city and county bonds. These are typically paid back via local value capture initiatives; however, value capture only works if there are location premiums associated with proximity to LRT stations [83].

3.4 Conclusions, Limitations, and Further Work

This chapter uses Charlotte, NC SFR sales data to determine the role proximity to LRT stations has on nearby SFR prices. Using three quasi-experimental models and twenty years of sales data, the construction and operations phases of the Lynx Blue Line and its extension are examined. Unlike many studies that focus on a brief period of time following the opening of a LRT system, this chapter demonstrates that LRT's effects on neighboring residential parcels are fluid over time. During the construction of the Blue Line, SFR buyers were willing to pay a premium for homes near station locations. Once the Blue Line became operational, home-buyers paid a premium to live farther away. Finally, as shown with the Extension Line model, proximity to LRT stations had no associated premium or discount. In addition to contributing to the body of knowledge surrounding the land use impacts LRT has on surrounding areas via two case studies, this chapter also highlights the need for using spatially explicit methods when examining real estate data, particularly when also considering proximity to transportation accessibility nodes.

In total, the findings in this chapter suggest that the average home-buyer in Charlotte, NC as captured through the MLS data had preconceived notions of the advantages of living near a LRT station (Blue Line construction) which changed over time to be overshadowed by disadvantages (Blue Line operations). According to the models, by the time the Extension Line was announced and built a decade later, the average home-buyer in the area was able to apply personal experience with light rail to their decision-making process for purchasing an SFR.

However, reality may not be as straightforward. Without further analysis beyond the scope of this work, one cannot determine to what extent home-buyers in the Blue Line treated group are similar or dissimilar from home-buyers in the Extension

Line treated group. Likewise, with just housing sales data, one cannot determine the extent to which those who purchased homes during the Blue Line construction period were like those who bought homes in the Blue Line operations period. Without demographics and survey data, for instance, one cannot determine whether buyers in the Extension Line model had lived in Charlotte before and were already familiar with the Blue Line as a result or were recent migrants to the area; in the case of the latter, additional heterogeneities could arise depending on the level of familiarity they had with LRT systems in their previous place(s) of residence. In addition to possible disparities in the underlying SFR purchaser characteristics, the analyses span two decades (2000-2019) in a rapidly evolving urban area. In population alone, for instance, the city has almost doubled in size. During this time, social, economic, and environmental attitudes have changed. Likewise, city planners' priorities may have shifted. Without being able to account for these longitudinal structural changes, it is difficult to draw definitive conclusions based on the three models presented in this chapter. Another challenging aspect of this research is that it is limited to only single-family residences. In urban areas, however, there are also commercial and business properties, and multi-family residences such as condominiums and apartments.

These limitations highlight opportunities for further work. The shift from a positive premium to a negative premium in the construction and operations of the Blue Line, respectively, to finally no premium in SFRs near the Extension Line has direct urban planning and policy implications and bears further examination. These results seem to imply that potential homeowners have lost interest in LRT over time. Additional research is needed to understand what has caused this attitudinal shift. It could be that ridesharing, ride hailing, and other mobility options have replaced trips that were originally completed via light rail. Indeed, survey data gathered by [84] indicate that ride-hailing services in California replaced 12.5% of trips that would have otherwise been made by light rail transit. By extension, would autonomous vehicles eventually render LRT obsolete? Perhaps autonomous vehicles could serve as first-mile last-mile solutions for potential LRT users who do not live near LRT stations. In that scenario, LRT stations could have reduced effects on nearby housing values. Furthermore, what are the implications for cities who have invested heavily in light rail and other public transit options? It may be that the light rail is used by suburban commuters rather than local urban residents. If that were the case, perhaps certain city taxes used for funding LRT ought to be changed to better reflect actual usage. Alternatively, perhaps the destinations along the route are unappealing to locals.

Furthermore, a more comprehensive approach is needed to gauge the extent to which prior experience with light rail affects how much more or less future homeowners are willing to pay in order to live closer to LRT stations. In addition to the SFR sales data, one would need to incorporate secondary data sources such as US Census and American Community Survey products as well as conduct surveys of current and potential homeowners to obtain a more holistic data set with which one can utilize both revealed and stated preferences methods to better get at the role of prior experience in decision making with regards of SFR purchases. Although evidence exists in consumer goods and some durables markets that previous experiences play a significant role in purchase decisions, the sheer magnitude of buying a home compared to buying particular brand of coffee grounds makes the former a much more complex purchase—especially given that one typically does the former less often than the latter. For urban planners and transportation engineers to meet their goal of facilitating the safe, effective, and efficient movements of people, goods, and services, it is important to therefore to understand the factors that drive household location decisions in order to design transportation systems that can better serve the populace.

4. HOW DOES DIRECTIONAL HETEROGENEITY AFFECT PRICES OF SINGLE FAMILY RESIDENCES NEAR LRT STATIONS?

Along with the renewed interest in the last couple of decades amongst urban transportation planners in light rail transit (LRT) systems in large cities across the US, there have been many research activities attempting to forecast the effects LRT systems may have on local traffic conditions, environmental systems, land use practices, and local housing markets. In the case of the last, previous research have mostly used hedonic price models and quasi-experimental methods to estimate the changes in single-family residence (SFR) parcel values due to proximity to LRT lines. Indeed, the previous chapter applies a spatial difference-in-differences model to uncover the effect that distance from LRT stations have on the price per square foot of SFRs in Charlotte, NC during the construction, operations, and extension phases of the line.

4.1 Research Motivation

The results from the various studies differ greatly in direction, magnitude, and statistical significance. While most studies reviewed in Chapter 2.2.2 indicate that increased proximity to LRT increases home values, for instance, the results in Chapter 3 show the opposite effect can also be true. A meta-analysis on LRT's effects on SFR prices indicate that point estimates from the multitude of existing studies depend on methodological factors such as whether the underlying data came from before the year 2000, whether studies controlled for highway access, and the length of the LRT line in the city (see Chapter 2.2.3). Other meta-analyses concerned with heavy and commuter rail transits' effects on SFRs point to additional factors such as the analysis type (e.g., studies using OLS, GWR, or other spatial econometric methods), functional form of models (semi-log, log-log, or neither), and type of rail transit being considered as sources of variation between studies' results [34, 35].

Another, hitherto unexamined, reason for variation between LRT study results might be attributed to the effects of directional heterogeneity. LRT is typically conceived as an alternative transportation mode that allows residents to travel from their homes to a central business district (CBD). However, despite a house being located near a LRT station, if the LRT station is in the opposite direction of the CBD relative to the house, residents may choose to use a more direct mode of transportation (e.g., private vehicle) to travel to the CBD. If this were the case, then these homeowners may not value the LRT service as much as homeowners for which the LRT station is "on the way" to the CBD. In other words, potential home-buyers could value LRT services if the location of the LRT station is downstream, rather than upstream, of the SFR location. By not including these directional effects, research may suffer from omitted variable bias leading to biased estimates on the coefficient on the distance variable if distance and direction are correlated. Using housing data from Charlotte, NC, this chapter examines the significance and magnitude of directional heterogeneity's effects on SFR prices near LRT stations.

This chapter is organized as follows. The next section will describe the methodology and outline the data used. Section 4.3 will present and analyze the results of the spatial hedonic model. Section 4.4 outlines some of the limitations and challenges of the research and section 4.5 concludes and suggests avenues of further research.

4.2 Methodology and Data

This research uses a spatial hedonic model to measure the significance and magnitude of directional heterogeneity in determining the price per square foot of SFRs located near LRT stations in Charlotte, NC. Recent work using this dataset indicate that a spatial ARAR model is the most appropriate [4].

To account for directional heterogeneity, three measuring methods are used (Figure 4.1). The first approach builds on [46] by considering direction explicitly by subdividing the adjacent SFRs into one of four possible 90-degree regions (North, South, East, West; Fig. 1A) relative to the closest LRT station. The LRT is between (indicated by the dummy variable isBetween) the SFR and CBD if the relative direction variable is identical and the distance to the CBD from the SFR is longer than the distance to the LRT station from the SFR. Additionally, if the SFR is within 660 feet (1/8 mile), or approximately a 3-minute walk [85], of a LRT station it is also included in the between group. This model uses a dummy variable that indicates whether the closest LRT station to a nearby SFR is upstream (Fig. 4.1C) or downstream (Figure 4.1D) relative to the CBD. "Nearby" is defined as being within 1.5 miles of the LRT station.

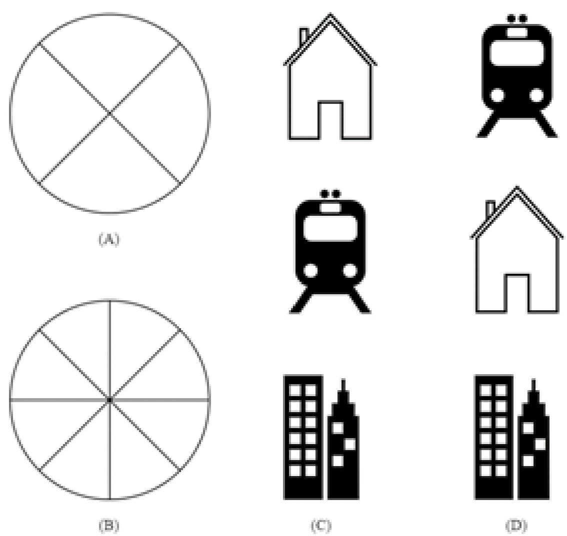

Figure 4.1. Model Assumptions

The second model utilizes the difference-in-differences (DiD) method in which the treatment group consists of SFRs for which the LRT is between the SFR and CBD and the control group consists of all other SFRs. As with the first model, direction is measured using four 90-degree regions corresponding to the cardinal directions. In this specification, a dummy variable is used to account for SFR sales during the recession and recovery periods instead of year fixed effects. The recession and recovery periods are defined as being from 2008-2012. This was done to preserve additional degrees of freedom.

A third model, in which SFRs are subdivided into one of eight possible 45-degree regions (Figure 4.1B) relative to the closest LRT station and CBD, using difference-in-differences is also constructed. Additionally, if the SFR is within a quarter mile of any LRT station it is included in the between treatment group. It is assumed that for homeowners in SFRs within such a short distance of LRT stations the directional effect is drowned out by the proximity effect.

The data for this study primarily came from two sources, the Charlotte Department of Transportation (CDOT) and the Multiple Listing Service (MLS) Listings database. Additional information on the data used can be found in Chapter 3.1.2. Because the MLS data were for individual properties, the source shapefile contained a collection of points. These were filtered according to proximity to the LRT stations. Further, because this study is concerned primarily with the effect that the distance to an LRT station has on SFR prices (price per square foot), only properties within a one and half mile radius were considered. For distance measurements, the Euclidean distance between an SFR and an LRT station was used. Table 4.1 shows a summary of average SFR characteristics for SFRs that are in the between (indicated by columns labelled (T)) and not between groups (indicated by columns labelled (C)).

All sales prices are in 2016 dollars, using interest rates derived from the Case-Shiller Composite Home Price Index for Charlotte, NC. Year fixed effects are included to control for housing market fluctuations over time, as the data includes the 2008-2009 recession and subsequent recovery period. Additional controls include SFR characteristics, such as number of bedrooms and baths; the presence of a fireplace, granite countertops, hardwood floors, and patio; the age and age squared of the SFR; the number of days on market before being sold; and whether or not the SFR is in a subdivision. The distance from the CBD to the SFR is also controlled for, as urban land use models suggest land rents increase the closer one is to city centers.

Table 4.1. Average Characteristics of SFRs in Treatment and Control Groups

Year Sold	Close Price		Bedrooms		Full Baths		Days on Mkt		SFR Age		Distance to CBD	
	(T)	(C)	(T)	(C)	(T)	(C)	(T)	(C)	(T)	(C)	(T)	(C)
2007	$ 278,640	$ 261,078	4.1	4.3	1.8	1.8	250	259	42	40	4,806	7,954
2008	$ 283,114	$ 254,378	4.2	4.3	1.8	1.9	246	249	49	38	5,449	8,261
2009	$ 270,186	$ 229,294	4.3	4.3	1.9	1.8	220	241	42	40	5,831	8,123
2010	$ 251,807	$ 231,484	4.1	4.3	1.9	1.9	222	238	51	43	4,825	7,702
2011	$ 248,450	$ 207,854	4.1	4.3	1.9	1.9	215	230	55	43	5,259	7,756
2012	$ 292,391	$ 241,831	4.2	4.3	2.0	2.0	225	231	56	45	4,816	7,817
2013	$ 309,813	$ 259,437	4.1	4.3	1.9	2.0	223	234	56	45	4,535	8,005
2014	$ 336,000	$ 284,535	4.1	4.3	1.9	2.0	224	222	53	48	4,612	7,829
2015	$ 395,542	$ 297,132	4.3	4.3	2.0	2.0	207	209	53	48	4,467	7,920
2016	$ 430,388	$ 339,590	4.3	4.3	2.0	2.0	209	214	52	48	4,564	7,832
2017	$ 410,954	$ 271,884	4.2	4.2	2.0	1.9	204	212	58	50	4,865	7,796
Average	$ 328,918	$ 269,731	4.2	4.3	1.9	1.9	220	227	52	45	4,820	7,913

Figure 4.2 shows the study area and SFRs used in the data analysis. The green area shows the central business district while the blue area shows the rest of Charlotte. Areas in Mecklenburg County not within city limits are shown in purple. The spectrum of red to green dots represent SFRs, with red SFRs having low price per square foot sales values and progressively lighter colors representing higher price per square feet sales values. Even visually, one can see that high- and low-value SFRs are clustered, further justifying the need to use a spatial econometric method of analysis. Consistent with bid rent theory, SFRs closer to the CBD tend to have higher price per square foot values than SFRs further away. Additionally, SFRs north of LRT stations tend to have lower price per square foot values than SFRs south of the stations.

4.3 Results

The following three tables (Tables 4.2, 4.3, and 4.4) show the results of the three approaches outlined in the methodology section, respectively. In each table, the first column shows the results of the a-spatial ordinary least squares model while the second column displays the results of the spatial model. Using just a dummy variable (Table 4.2), whether or not the nearest LRT station is in between a SFR and the CBD is shown to be weakly significant (90th percentile). When using the four-angle

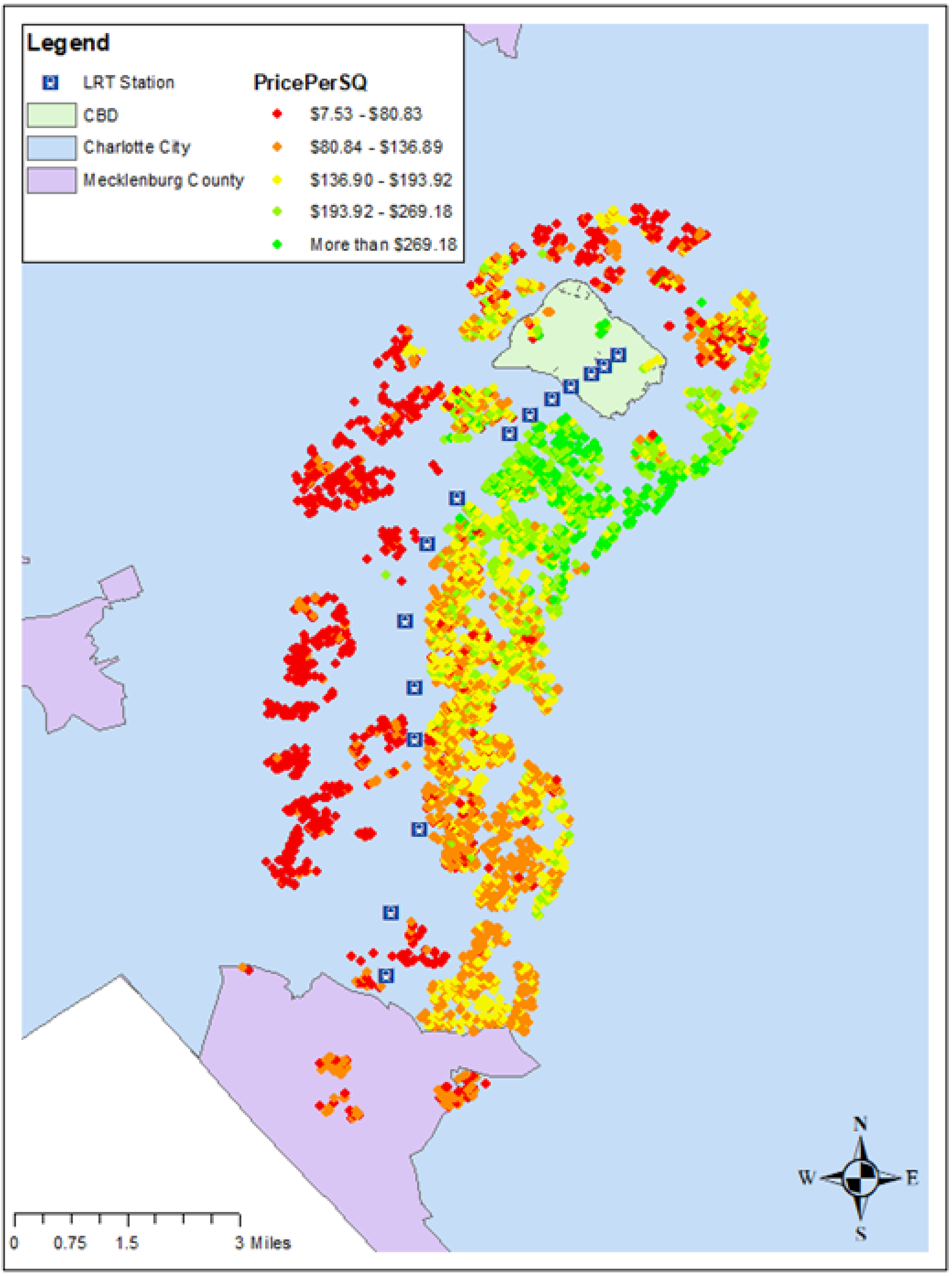

Figure 4.2. Map of Study Area

split method, the spatial DiD estimator is also only weakly significant (Table 4.3). Instances in which variables appear to be significant in the a-spatial model but are not significant in the spatial model and vice-versa are illustrative of the problems that can arise when using a-spatial regressions in the presence of spatial spillover effects as not properly accounting for spatial lag and spatial error may lead to biased and inconsistent estimates.

Based on the AIC, the spatial difference-in-differences model using the eight-angle split method is the best fit. Thus, subsequent discussions will focus solely on those results (Table 4.4). To better interpret the results for the spatial model, the direct, indirect, and total marginal effects are calculated for the significant variables (Table 4.5).

As seen in Table 4.4, the control variables for number of bedrooms, the presence of a fireplace, granite countertops, and patio, whether the SFR belonged in a subdivision, the age and age-squared of the SFR, and the recession dummy variable were significant in the spatial DiD model. That the number of bedrooms have a negative effect on price per square foot may seem counterintuitive. It could be that a high number of bedrooms may contribute to a dilution effect because they increase the square feet a house occupies when the dependent variable is price per square foot.

As expected, the presence of additional SFR amenities, including patio, fireplace, and granite countertops added value to the property and had a positive impact on the price per square foot. Similarly, by not being in a subdivision SFR owners would not have access to amenities that are typically found in neighbourhoods—thus it is expected that the estimated coefficient on the no subdivision dummy variable has a negative sign.

The negative sign on SFR age indicates that older houses tend to sell for a lower price per square foot. The positive sign on SFR age squared indicates that this effect is lessened for very old houses. The estimated coefficient for the dummy variable indicating that the SFR was sold during the recession-affected years is significant, and as expected, negative.

Finally, the distance to CBD is also controlled for. As one would expect, and consistent with monocentric city models, the price per square foot of SFRs decrease

Table 4.2. Dummy Variable with Four Angle Split Method

	A-spatial OLS model		Spatial model	
	Estimate	**Std. Error**	**Estimate**	**Std. Error**
(Intercept)	142.20***	7.52	443.85***	17.04
CBD Distance	-6.23***	0.00	-13.41***	0.00
isBetween	-2.23	1.49	4.51·	2.71
Attached Garage	14.50***	2.38	-3.30*	1.41
Full Baths	32.37***	1.18	1.24·	0.73
Total Beds	-7.20***	1.17	-10.46***	0.71
Days on Mkt	0.01·	0.00	0.00	0.00
Driveway	0.41	1.95	2.73*	1.07
Fireplace	14.81***	1.66	2.51*	0.93
Granite	23.29***	1.60	11.21***	0.91
Hardwood Floor	8.64***	1.44	2.15**	0.81
No Subdivision	-53.23***	3.91	-9.75***	2.44
Patio	15.79***	2.86	6.21***	1.57
SFR Age	0.39***	0.09	-0.60***	0.07
SFR Age^2	0.00	0.00	0.01***	0.00
Y2007	-10.28	7.53	-24.33***	4.07
Y2008	-14.48*	5.96	-28.42***	3.25
Y2009	-31.70***	6.06	-44.46***	3.32
Y2010	-48.53***	6.03	-57.50***	3.29
Y2011	-52.31***	5.99	-60.69***	3.3
Y2012	-45.29***	5.9	-58.36***	3.23
Y2013	-31.73***	5.78	-43.26***	3.16
Y2014	-19.69***	5.72	-32.61***	3.11
Y2015	-7.00	5.67	-19.82***	3.07
Y2016	8.19	5.65	-2.89	3.06
Rho			0.84***	0.03
Lambda			0.96***	0.00
AIC	88766		82355	

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table 4.3. DiD Model, Four Angle Split Method

	A-spatial OLS model		Spatial model	
	Estimate	**Std. Error**	**Estimate**	**Std. Error**
(Intercept)	130.50***	5.94	425.16***	19.78
isBetween4	15.56***	3.44	12.78*	6.04
CBD Distance[1]	-6.27***	0.00	-14.25***	0.00
LRT Distance[1]	1.76	0.00	7.40	0.01
isBetween4:LRT_Dist[1]	-11.19***	0.00	-6.55·	0.00
Attached Garage	16.77***	2.43	-0.19	1.53
Full Baths	31.85***	1.21	1.37	0.80
Total Bedrooms	-6.65***	1.20	-9.60***	0.77
Days on Market	0.01	0.00	0.00	0.00
Fireplace	15.52***	1.7	3.25**	1.01
Granite Countertop	22.21***	1.64	10.64***	0.99
Hardwood Floor	7.41***	1.48	1.58.	0.88
No Subdivision	-53.18***	4.00	-10.86***	2.65
Patio	14.70***	2.94	5.23**	1.71
SFR Age	0.28**	0.09	-0.69***	0.08
SFR Age^2	0.00*	0.00	0.01***	0.00
Recession	-27.18***	1.38	-26.57***	0.82
Rho			0.93***	0.03
Lambda			0.95***	0.00
AIC	89177		83807	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '·' 0.1 ' ' 1

[1] Measured in thousands of feet

as distance from CBD increases. The variables of interest to this research are the space variable and the interaction between distance and space variables. There is a weakly significant and negative association between price per square foot of the SFR and whether a station is between a SFR and the CBD. Furthermore, when the LRT station is between a SFR and the CBD, distance from the LRT station has a weakly significant and positive effect on SFR prices.

Table 4.5 shows the average direct, indirect, and total marginal effects of the significant proximal and locational variables from the spatial eight-angle split method.

Table 4.4. DiD Model, Eight Angle Split Method

	A-spatial OLS model		Spatial model	
	Estimate	**Std. Error**	**Estimate**	**Std. Error**
Constant	150.10***	7.28	454.47***	20.86
IsBetween8	-18.78***	5.02	-25.68*	10.48
CBD Distance[1]	-6.01***	0.00	-14.11***	0.00
LRT Distance[1]	-12.30***	0.00	-7.23	0.01
isBetween8:LRT Distance[1]	15.55***	0.00	12.54*	0.01
Attached Garage	17.21***	2.43	-0.27	1.53
Full Baths	31.43***	1.21	1.40.	0.80
Bedrooms	-6.54***	1.20	-9.66***	0.77
Days on Market	0.01	0.00	0.00	0.00
Fireplace	15.61***	1.70	3.25**	1.01
Granite Countertop	22.20***	1.64	10.60***	0.99
Hardwood Floor	7.41***	1.48	1.55·	0.88
No Subdivision	-53.01***	4.00	-10.66***	2.65
Patio	14.13***	2.93	5.16**	1.71
SFR Age	0.40***	0.09	-0.69***	0.08
(SFR Age)2	0.00	0.00	0.01***	0.00
Recession	-27.25***	1.37	-26.59***	0.82
Rho			0.93***	0.03
Lambda			0.95***	0.00
AIC	89146		83806	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
[1] Measured in thousands of feet

SFRs that are within a quarter mile of an LRT station or have the station between them and the CBD are \$13.32 less expensive per square foot than SFRs further away or do not have a station between them and the CBD. Interestingly, the distance from an LRT station for all SFRs (treatment group and control group) variable was not found to be significant. This result suggests that not accounting for directional effects, as was the case for the models presented in Chapter 3, could lead to potentially erroneous inferences.

However, the results also show that for SFRs in the treated group, the price per square foot increases by \$6.50 (in 2016 dollars) per one thousand feet, or \$8.58 per quarter mile, away from the closest LRT station. While these results may seem counterintuitive, especially as one would assume that potential home-buyers would value having LRT as an additional mobility option to reach the CBD, they are consistent with the results of the previous chapter. These results could reflect the degree to which the urban transportation system is automobile oriented. Alternatively, it may be that the SFR buyers do not regularly travel to the CBD, or when they do travel to the CBD they prefer a more direct route that does not include frequent stops, as would be the case of light rail.

Table 4.5. Marginal Effects: DiD Model, Eight Angles Method

	Direct	**Indirect**	**Total**
IsBetween8	-28.14	14.82	-13.32
CBD Distance[1]	-15.46	8.14	-7.32
isBetween8:LRT Distance[1]	13.74	-7.24	6.50

[1] Measured in thousands of feet

4.4 Limitations and Further Work

Due to the data used in this study, only SFR values were considered although commercial and industrial land values may also be affected by directional heterogeneity with respect to LRT stations and the CBD. Population density was not directly controlled for but may have an impact on the price per square foot of houses in urban areas. Another challenge attributed to the data is that the dataset is a pooled cross section consisting of eleven years of data. Despite applying interest rates to unify the dollar values of sales, the temporal relationships of sales are not adequately accounted for. It may be that "historical" sales (e.g., in period t-1) may have an effect on sales in the "current" (i.e., period t) year for SFRs that are proximally close. Similarly, "future" sales (i.e., period t+1) cannot have an effect on sales in the "current" year for SFRs in close proximity. By pooling the eleven annual cross sections together and not controlling for the temporal effect appropriately, measurement error is introduced in the model as both paths (past effects present and future effects present) of temporal dependence are allowed. To account for this, several modelling alternatives were attempted. First, a pure panel model was attempted; however, as SFRs are not frequently bought and resold even within a decade, the panel dataset was too unbalanced to for a model to run. The second alternative was to pursue a repeated sales model. This method was also limited by the data. The underlying MLS data includes a parcel identification variable. This variable's format, however, was so inconsistent across time that it was impossible to track parcel sales across time. A final alternative was to create standalone models for each year of data to completely disallow temporal effects. This approach was also difficult to implement as many parcels' nearest neighbors were not proximally close; further, the model fits suffered significantly when restricting the model parameters to be consistent across each model, leading to an increased difficulty in interpreting the results as a whole.

Additionally, there may be other reasons, in addition to access to a transportation network that takes them to a CBD, that draw potential residents to live near transit stations. Further, this research assumes that all residents travel to the CBD, which may not be the case—especially if their place of work is not in the CBD. The research can be enhanced by taking into consideration commuting pattern data from the Census Transportation Planning Products (CTPP). Finally, location theory suggests that household demographics (e.g., income) play an important role in where people live. The research can be enhanced with the inclusion of census or survey data that better explain the demand for housing relative to LRT station locations and the CBD. Ongoing work includes considering these additional factors as well as methodological improvements to the operationalization of directional heterogeneity.

Despite these shortcomings, robustness testing using thirty-two additional model specifications (sixteen for the 4-angle split method and sixteen for the 8-angle split method) yielded consistent results. Year fixed effects rather than a recession dummy variable, not including SFRs within a quarter-mile radius of stations regardless of directional component as being in the treatment group, and only allowing for SFRs within an eighth-mile radius of stations were amongst the differences in model specifications. The latter two models were run to help distinguish between the negative nuisance effect due to proximity to LRT stations revealed in the previous chapter and any treatment effects. However, despite separating these two effects, the treatment effect remained positive and significant. Additionally, some robustness check models also controlled for distance to highway onramps. In those models, while the distance to highway variable was significant, the variable of interest, the DiD estimator, remained significant and positive.

4.5 Conclusions

This chapter examined the effect that directional heterogeneity has on the price per square foot of LRT-adjacent single-family residences in Charlotte, North Carolina, using a robust spatial econometric model. In this case, directional heterogeneity is operationalized as whether the nearest LRT station to an SFR is between the SFR and the CBD or not. Two methods for determining direction based on the angles from which the LRT station and CBD are relative to SFRs are used. Moreover, two estimation methods, including the use of a dummy variable and difference-in-differences, are applied. The quasi-experimental nature of difference-in-differences in conjunction with the eight-angle split approach led to the best fit model. In the latter model, the spatial DiD estimator is shown to be significant, indicating that directional heterogeneity influences SFR price per square foot if the LRT is between the SFR and the CBD. Thus, a failure to properly account for this factor could lead to biased inferences.

As LRT becomes an increasingly popular transportation option, these results may be useful for urban planners interested in understanding its effects on surrounding housing markets. Additionally, the results may be of use to transportation agencies planning to implement light rail in the design and spacing of transit facilities. These results also highlight the importance of developing spatial models that also examine directional effects in addition to distance-based effects.

5. PROXIMITY TO LRT AND DEMOGRAPHIC CHANGE IN NEIGHBORHOODS

Previous chapters have highlighted the changes to real estate markets by LRT stations as a function of proximity (Chapter 3) and directional heterogeneity (Chapter 4). For the case of Charlotte, NC, the price per square foot of SFRs is shown to be positively associated with distance; as distance away from LRT stations increases, the price per square foot of SFRs also increases. Interestingly, the empirical results of the Charlotte case study show that prior experience with LRTs does not have a significant effect on home-buyers' valuation of property prices. The research on directional heterogeneity indicates that the price per square foot of SFRs in Charlotte, NC indicates that the proximity effect is nuanced. When a LRT station is located between an SFR and the central business district, the price per square foot of that SFR decreases at a slower rate per unit distance further from the LRT station, indicating that the relative position of the LRT station offsets to some degree the change in housing prices.

5.1 Organization

This chapter is organized as follows. The next sections cover the concepts and hypotheses to be tested and the methods and models used. After that, the dataset used in the analysis is described. Section five discusses the results of the models used. Finally, section six notes the limitations of the study and suggests avenues for further research.

5.2 Concepts and Hypotheses

Changes in nearby housing markets due to LRT stations affect housing affordability. This, in turn, affects the composition of such areas as potential home-buyers may be priced out of purchasing certain houses over time (see Chapter 2.5). However, there is a gap in the empirical literature on the long-term demographic changes in neighborhoods and census block groups (CBGs) adjacent to LRT stations that are not part of transit-oriented developments (TODs). Changes over time in socio-demographic characteristics including race, education, and income in these areas indicate that the presence of LRT stations can have lasting effects in neighborhood composition due to in- and out-migration. While migration flows are difficult to directly measure, inferences regarding migration may be drawn from changes in characteristics collected by the American Communities Survey (ACS). Thus, this chapter analyses the effects that proximity to LRT has on nearby residents in CBGs over time based on population composition traits between 2011 and 2018.

In particular, the traits examined include racial composition, educational attainment levels, and average median income of people living within CBGs. Based on the literature summarized in Chapter 2.5, it is hypothesized that there will be a decline in the percentage of population that identify as African American due to proximity to LRT stations. While the majority of LRT and gentrification effects literature is focused on transit-oriented developments, similar effects may also occur in transit adjacent areas. Based on the results in Chapters 3 and 4 regarding housing values and distance from LRT stations, it is hypothesized that educational attainment levels and average median incomes will increase with distance from LRT stations.

5.3 Methods and Models

Although data is available to perform spatial analyses, it should be noted that the geographical boundaries for CBGs changed in 2015. Harmonized CBG data was not publicly available. Since spatial panel models require consistency in the spatial component, this would have led to constructing separate models for 2011-2014 and 2015-2018. However, because the research is more interested in the longitudinal effect of proximity to LRT stations than the spatial effect between CBGs, a-spatial panel is an appropriate alternative that would allow for the analysis of all CBGs between 2011-2018 within a single model.

Because the research is primarily concerned with changes in characteristics of each CBG over time, fixed effects are preferred over random effects. This would allow for CBG-specific, time invariant, idiosyncrasies to drop out. To confirm this statistically, Hausman tests were also conducted for each of the three models. In all cases, the Hausman test resulted in a rejection of the null hypothesis that random effects were more appropriate. Furthermore, Fisher unit root tests based on augmented Dickey-Fuller tests indicated that at least one panel was stationary for each of the three models [86]. In addition to specifying fixed effects, autoregressive correlation was tested for using a Wooldridge test [87].

In all three models, the Wooldridge test was significant, indicating the presence of AR(1) disturbances. Both the inclusion of a lagged dependent variable as an independent variable and the use of a distributed lag model failed to adequately account for these disturbances. Thus, the method outlined by [88] is employed. All model estimations are conducted in Stata using the xtregar command.

The three demographic factors include racial composition, income, and education. These characteristics were chosen because they are often used in the literature (summarized in Chapter 2.5) as indicators for gentrification and displacement due to the presence of TODs. Because this research is concerned with transit adjacent CBGs, a related topic, using similar indicators may help with conducting comparisons across studies. Specifically, the dependent variables chosen to measure racial composition, income, and education are percentage African American, median income, and percentage of bachelor's degree holders, respectively.

Linear regressions are used for both the racial composition and education models despite the dependent variables being percentages. Although a Tobit model may be better in some respects for handling bounded values, the likelihood estimator for fixed effects is biased and inconsistent in nonlinear functions. As both the theoretical basis and statistical tests indicate, a fixed effects model is preferred over random effects.

Alternative regression methods including Poisson and negative binomial regression, in which these two dependent variables were converted to population count values, were also examined. However, these two methods were unable to properly deal with the autoregressive structure of the error terms, leading to biased estimates.

In addition to the theoretical basis for variable selection, these three variables also have some of the highest degrees of within and between variability within the panel dataset for each of their demographic characteristic categories (see Table 5.1). The following three equations show the specification of each model. Note that these are independent models and not a system of equations.

$$perAA_{it} = \beta_0 + \beta_1 Dist_i + \beta_2 Unemp_{it} + \beta_3 HU_{it} + \beta_4 lag.perAA_{it} + \varepsilon_{it} \tag{5.1}$$

$$\begin{gathered} Bacdeg_{it} = \beta_0 + \beta_1 Dist_i + \beta_2 Unemp_{it} + \beta_3 perWhite_{it} + \\ \beta_4 perLess30k + \beta_5 HU_{it} + \varepsilon_{it} \end{gathered} \tag{5.2}$$

$$medInc_{it} = \beta_0 + \beta_1 Dist_i + \beta_2 Bacdeg_{it} + \beta_3 Unemp_{it} + \beta_4 HSDip_{it} + \varepsilon_{it} \tag{5.3}$$

in which,

i = CBG identifier;

t = time;

$perAA$ = Percent of population that identify as African American;

$Dist1k$ = Distance from CBG centroid to closest LRT station in thousands of feet;

$Unemp$ = Percentage of civilians aged 16 and older that are unemployed;

HU = Total housing units available;

$lag.perAA$ = $perAA$ lagged by 1 year;

$medInc$ = Average median income;

$Bacdeg$ = Percentage of people whose highest educational attainment level is bachelor's degree in the population aged 25 and over;

$HSDip$ = Percentage of people whose highest educational attainment level is high school diploma in the population aged 25 and over;

$perWhite$ = Percent of population that identify as White;

$perLess30k$ = Percent of population aged 16 and over with household income of less than \$30,000; and

ε = error term.

Because household location decisions include many factors unaccounted for by the estimated models, the overall model fit (e.g., R^2) is expected to be low. However, the sparse models may still reveal meaningful relationships between the demographic variables and distance from LRT stations. In all three models, the total number of housing units available is included as a supply-side control.

As a historically disadvantaged population, the percentage of population that is African American is expected to be positively associated with the unemployment rate. Additionally, the proportion of African Americans living in a CBG each year will likely be positively influenced by the proportion of African Americans living in that CBG in previous years; thus, the lagged term is also included.

The second model relates CBG average educational attainment with distance from LRT stations, the unemployment rate, percent of population that is White, the percent of population with annual household incomes of less than \$30,000, and also controls for the total number of housing units available.

Educational attainment is measured as the percent of the population aged 25 or older whose highest level of education is a bachelor's degree. Although the ACS also collected data on higher education levels, including number of persons with master's, professional, and doctoral degrees, these variables tended to be time-invariant. Because both the theoretical basis and statistical tests indicate that panel fixed effects regression is the appropriate estimation technique, having a time-invariant dependent variable would have been problematic. High educational attainment is commonly associated with an increased likelihood of being employed. A low unemployment rate may be indicative of a higher percentage of bachelor's degree holders in a CBG. Similarly, high educational attainment is typically associated with well-paying jobs; one would expect a negative association between the percentage of bachelor's degree holders and the percentage of population in a CBG earning an annual income of less than

$30,000. To aid the analysis, all dollar values have been inflated to 2018 dollars. The percentage of population in a CBG that is White is included as that racial group has historically been the advantaged majority, particularly in the American South, and may therefore have had increased access to educational opportunities that led to obtaining bachelor's degrees.

The final model relating median incomes in a CBG with distance from LRT stations also controls for total housing units. That income is related to education has been proven in other literature; therefore, the percentage of high school diploma holders and bachelor's degree holders are included as independent variables. Another factor that can influence median income is the unemployment rate, as CBGs with high unemployment may have a lower median income.

5.4 Data

The data used in the analysis comes from eight 5-year ACS's (2007-2011, 2008-2012, 2009-2013, 2010-2014, 2011-2015, 2012-2016, 2013-2017, and 2014-2018; hereafter referred to as the ACS2011, ACS2012, ACS2013, ..., ACS2018). The 5-year ACS is chosen as it is publicly available and contains a wealth of demographic information. While the 1- and 3-year ACS can provide similar data, the sample sizes for these data are much smaller than the 5-year version of the ACS. Further, not all census block groups in the study area were consistently surveyed in the other two versions of the ACS. The 2006-2010 5-year ACS was omitted from the analysis as it did not ask the same questions as the other eight included in the study.

CBGs were chosen as the unit of analysis as it is the smallest geographic unit available in the ACS. Rather than include all CBGs in the Charlotte/Mecklenburg County area, the study area is restricted to the CBGs within five miles of any LRT station along the Blue Line (Fig. 5.1). The five-mile threshold was chosen primarily to ensure an adequate sample size without being too large as to dilute the proximity effect of LRTs. As mentioned in the previous section, the geographic boundaries of

CBGs were altered in 2015. In data collected before 2015 there are 293 CBGs while there are only 292 CBGs in data collected after 2015.

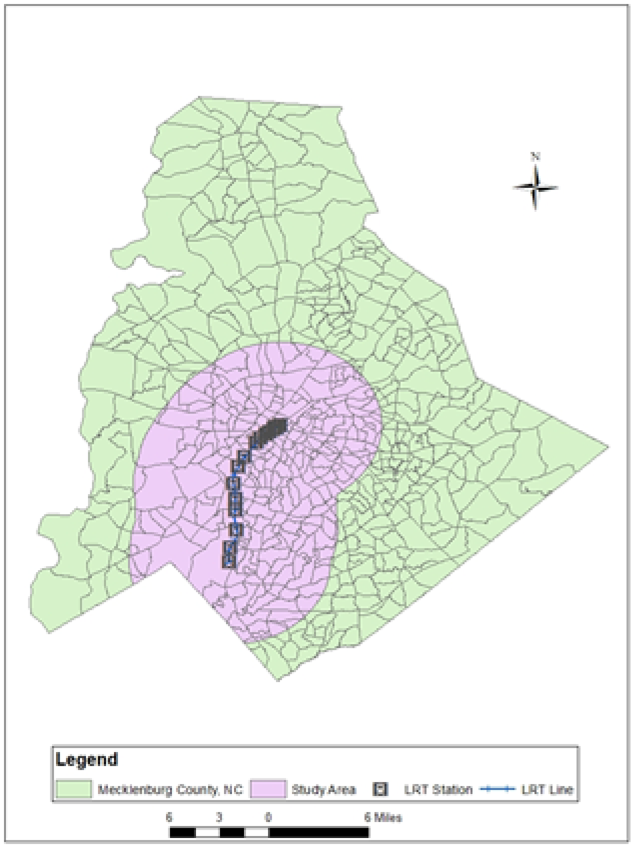

Figure 5.1. Map of Study Area

Table 5.1 shows the summary statistics of the variables included in the analysis. As the dataset is comprised of eight cross sections, the within and between variations are also shown. In this context, within refers to the variability of an indicator over time within CBGs, while between refers to the variability of an indicator between CBGs in the same period. It should be noted that the ACS masks data to preserve confidentiality of respondents at times, which explains why some variables have only 290 or 291 observations (n) despite there being more CBGs than that. As seen in Table 5.1, despite all CBGs being in a relatively small area, there is a great deal of heterogeneity between them in racial composition–some CBGs are 100% White while others are 100% African American. Additionally, as noted by the summary statistics, there is a large degree of variability in incomes, measured as median income and percentage of population aged 16 and over with household income less than $30,000, both between CBGs and within a CBG over time. Educational attainment levels also vary greatly across time and space. Some CBGs do not have anyone with a bachelor's degree, while other CBGs have 74% of the population aged 25 and over with a bachelor's degree. Finally, the data includes both the Great Recession (captured in ACS2011, ACS2012, ACS2013) as well as the subsequent recovery period (i.e., ACS2014, ACS2015, ACS2016). This contributes to the large range in values of the unemployment rate.

5.5 Results

5.5.1 Exploratory Data Analysis

As part of the exploratory data analysis, the percent changes for perAA, medInc, and Bacdeg were examined. While a longitudinal spatial analysis is beyond the scope of this research, it nonetheless can be helpful to see where the heterogeneity occurs within the study area. Because the underlying geography changed in 2015, the analysis includes two figures. Figures 5.2 and 5.3 show these changes between the ACS2011-ACS2014 and ACS2015- ACS2018, respectively, in quintiles. The values shown on these figures are percentages (represented in decimals). In general, red colored CBGs indicate a decrease while green values indicate an increase.

There is a great degree of heterogeneity in percentage change in median income between 2011-2014. Some CBGs experienced a decrease of as much as 63%, while other CBGs saw an increase of 123%. Visual inspection indicates that the changes do not seem to be spatially clustered. Educational attainment at the CBG level seems to have declined over the 5-year period captured between ACS2011-ACS2014. This indicates that the residents in the CBGs in the study area have changed over time, especially as educational attainment levels appear to have, in general, declined.

Table 5.1. Summary Statistics

Variable		Mean	Std. Dev.	Min	Max	Obs
perAA	overall	30%	31%	0%	100%	N = 2,317
	between		26%	0%	85%	n = 291
	within		16%	55%	69%	
medInc	overall	61,904	41,284	6,190	250,001	N = 2,297
	between		39,564	11,390	245,450	n = 290
	within		11,591	39,558	158,683	
Bacdeg	overall	26%	17%	0%	74%	N = 2,317
	between		16%	0%	58%	n = 291
	within		5%	1%	51%	
Dist1k	overall	12.11	8	0	26.36	N = 2,340
	between		8.03	0	26.36	n = 293
	within		0	12.05	12.16	
Unemp	overall	10%	9%	0%	74%	N = 2,340
	between		8%	0%	36%	n = 293
	within		5%	0%	48%	
HSDip	overall	17%	11%	0%	63%	N = 2,317
	between		10%	0%	42%	n = 291
	within		5%	7%	52%	
HU	overall	128	238	0	1,650	N = 2,327
	between		60	0	462	n = 293
	within		230	271	1383	
perWhite	overall	52%	32%	0%	100%	N = 2,317
	between		24%	12%	100%	n = 291
	within		22%	36%	100%	
perLess30k	overall	32%	20%	0%	100%	N = 2,317
	between		19%	0%	83%	n = 291
	within		7%	15%	81%	

Another indication that the population has changed can be seen in the map for racial composition. Nearly all CBGs experienced a 100% decrease in percent of African Americans in the five years. However, without regression analysis and additional data it is difficult to ascertain why these demographic changes occur.

Figure 5.3 shows the changes in the three dependent variables captured in the ACS2015-ACS2018 data. During this time, most CBGs experience a rise in median annual income. This could be indicative of improved macroeconomic conditions after

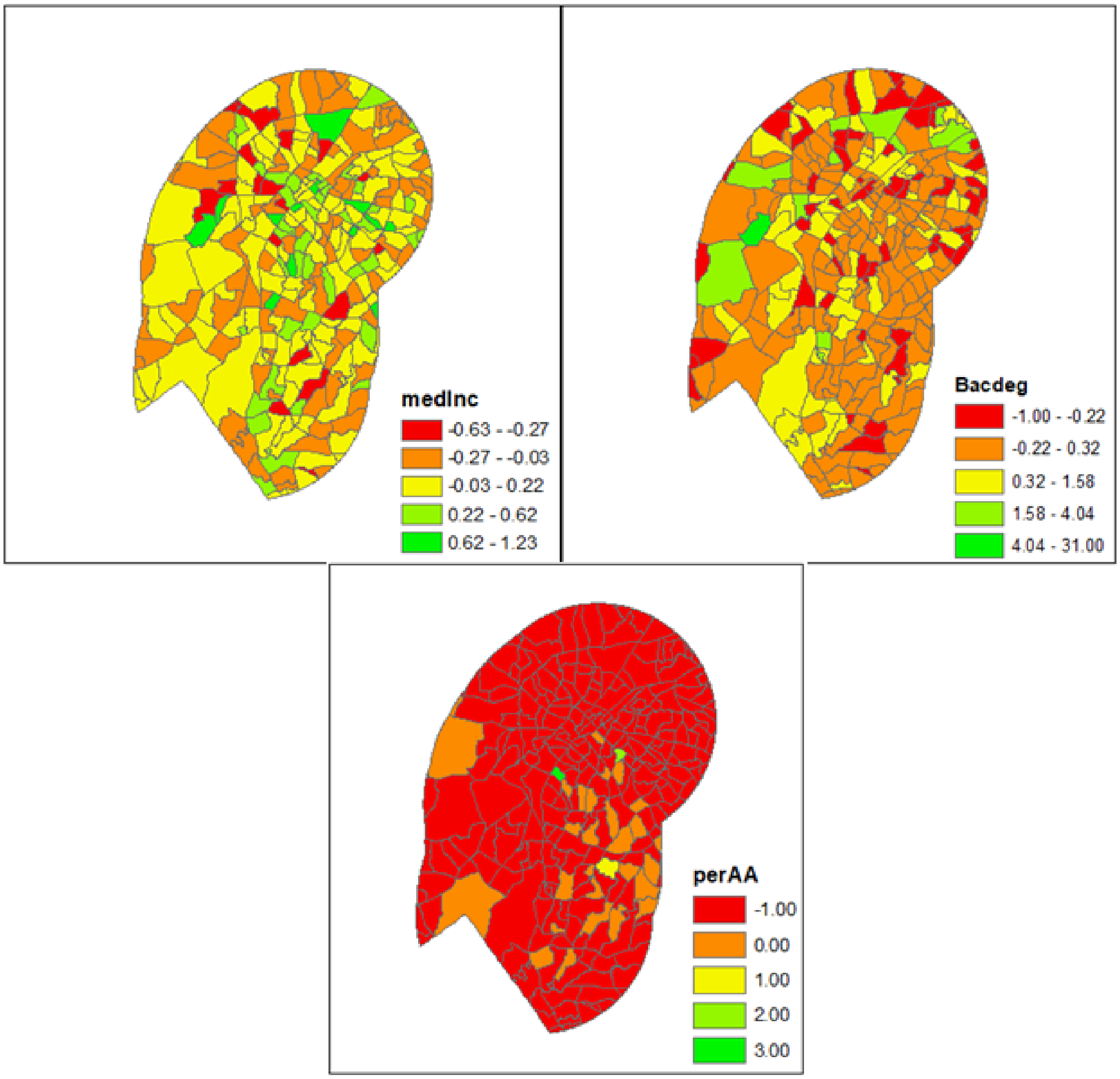

Figure 5.2. Change in Dependent Variables, 2011-2014

the recessionary and recovery periods, or it could be due to population changes. The percentage of bachelor's degree holders also increased over time. This could be due to population changes as well with higher educated people moving in while lower educated people moving out. Alternatively, it may be that more of the incumbent residents of those CBGs finish college. However, the changes in percentage of CBG that are African American indicate that changes in the CBG demographic variables may be due to changes in the underlying populations.

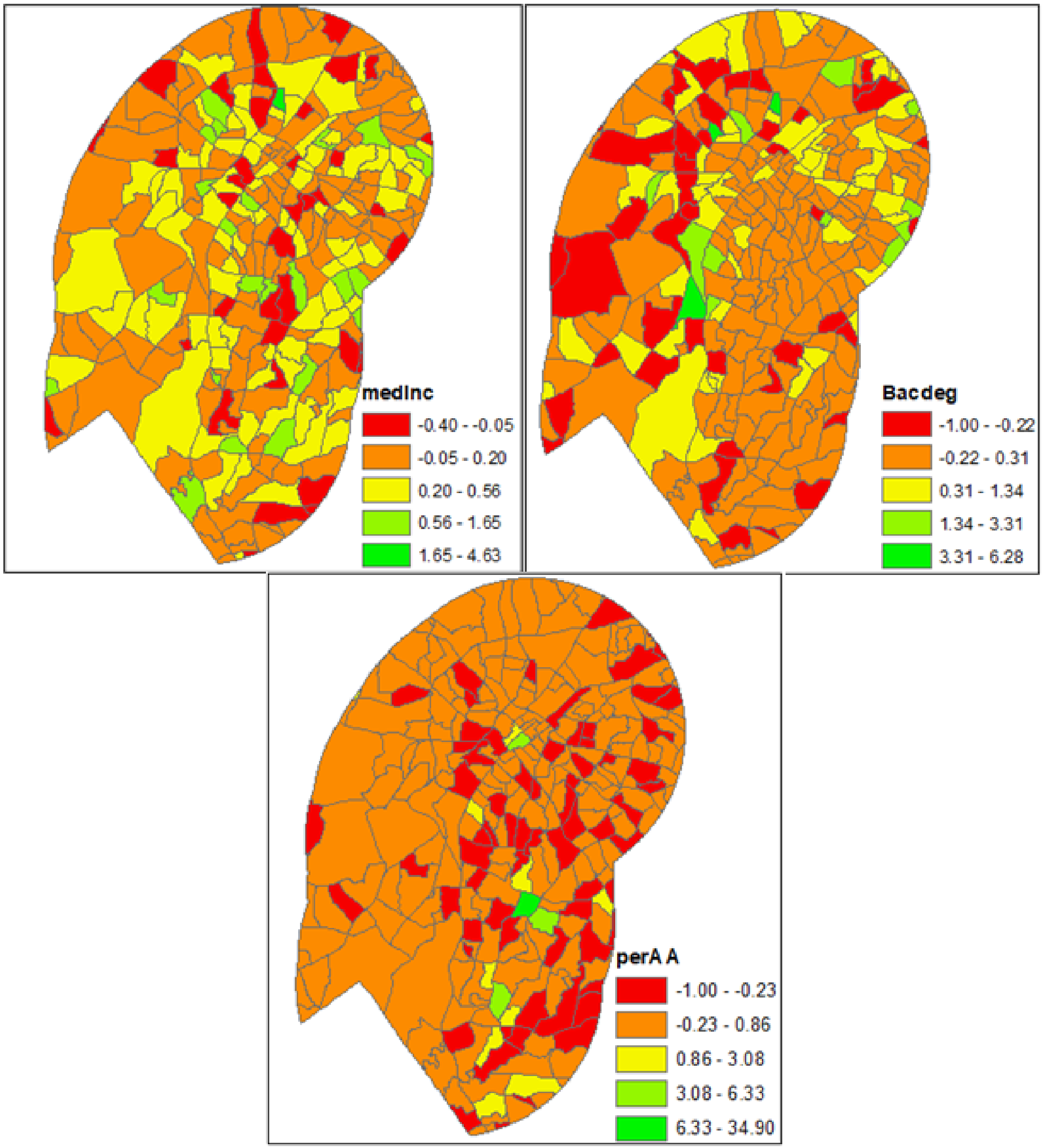

Figure 5.3. Change in Dependent Variables, 2015-2018

5.5.2 Regression Results

The following table (5.2) presents the results for the three estimated models. As noted in the methodology section, the location decision of households can include many other factors in addition to the variables included in the models. This leads to the models having low overall goodness of fit statistics. However, the F-stats are significant, indicating that the models are significantly different from zero and have some explanatory power.

Table 5.2. Model Estimation Results

		(1)		*(2)*		*(3)*	
Dependent variable		perAA		Bacdeg		medInc^{+}	
		Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
(Intercept)		-0.6692	1.4588	-0.6540*	0.0855	3.6766	7.0564
Dist1k		0.0836	0.1093	0.0792*	0.0182	4.6674*	1.4115
HU		0.0014*	0.0007	1.27E-5*	0.0000	0.0063*	0.0015
lag.perAA		-0.0981*	0.0256	–	–	–	–
Unemp		-0.4383*	0.0796	-0.0484'	0.0255	-26.6690	6.2070
perWhite		–	–	0.0034	0.0032	–	–
perless30k		–	–	-0.1260	0.0179	–	–
HSDip		–	–	–	–	-14.2894*	6.3443
Bacdeg		–	–	–	–	36.1909*	5.9078
ρ_{ar}		-0.0820		0.6134		0.5913	
σ_{μ}		0.7153		0.6785		49.127	
σ_{ε}		0.1767		0.0410		9.9897	
ρ_{fov}		0.9425		0.9963		0.9603	
R^2	Within	0.0385		0.0540		0.6490	
	Between	0.0004		0.0153		0.3770	
	Overall	0.0002		0.0139		0.0370	
		F(3,1721)	22.98	F(5,1719)	19.62	F(5, 1704) =	23.66
		Prob >F	0.000	Prob >F	0.000	Prob >F	0.000

*Significance codes: * 0.05, ‘ 0.10*

$^{+}$ medInc in thousands of dollars

Based on the model results, the first hypothesis, that the percentage of CBG population that identify as African American would decrease with proximity to LRT stations is rejected. The second and third hypotheses, that educational attainment levels and median income, respectively, are supported by the results.

The first model estimated examines the relationship between racial composition of the CBG and the distance from CBGs' centroid to LRT stations. These results indicate that based on the data used, proximity to LRT had no significant longitudinal effect on racial makeup of CBGs in the study area. Additionally, this result suggests that changes in race are not necessarily a sign of gentrification or displacement. In the context of this research, the implication is that changes in racial mix may not be the best indicator of migration. On the other hand, the lagged dependent variable is negative and significant, indicating that the racial composition of CBGs is changing over time regardless of the distance a CBG may be from the closest LRT station. Additional significant variables include the number of housing units available and the unemployment rate. Increases in the supply of housing is positively associated with the percentage of African Americans in the CBG, and likely speaks to the improved ability to find a place to live due to a housing supply change. The unemployment rate has a large and negative effect on percentage of African Americans. This could be primarily driven by the 2008-2009 recession and subsequent recovery period. Alternatively, it could demonstrate that an increase in unemployment within a CBG may provide the stimulus to leave the CBG and find work elsewhere.

The middle columns present the estimation results for the educational attainment model. When measuring educational attainment as the percentage of population whose highest level is bachelor's degree, the distance to LRT station and number of housing units are statistically significant, while unemployment is weakly significant. The results show that for every 1,000 feet further away from an LRT station, there are nearly eight percentage points more college graduates; this implies that residents in CBGs close to stations are not as well educated on average. Although total housing units has a positive and significant effect on the percentage of population who have finished college in a CBG, the value itself is so close to zero that the effect may as well be negligible. Although being weakly significant, unemployment is negatively associated with high educational attainment levels. The race variable is not significant; this may be a result of using fixed effects.

The last pair of columns in Table 5.2 present estimation results for the income model. Distance from LRT station, total housing units in the CBG, percent of population who finished high school, and percent of population who finished college are found to be significant. Residents in CBGs that are further away from LRT stations have higher median annual incomes. For every thousand feet further from LRT stations a CBG centroid is, the average median annual income increases by $4,667. When the percentage of population whose highest educational attainment level is a high school diploma increases, the median income decreases. Alternatively, when the percentage of population whose highest educational attainment level is a college degree increases, the median income increases. These latter two results are consistent with the education-income literature and are as expected.

5.6 Conclusions

Based on the results of these three models, some preliminary conclusions can be drawn regarding the CBGs in the study area and their distance to LRT stations. In general, it seems that distance has no effect on racial composition of CBGs. The racial composition of CBGs in the present are affected more by the racial composition in the past than by other factors. The model suggests that LRT stations did not attract non-Black populations to move near them over the period studied. Both educational attainment and median income increases with distance from LRT. This implies that while well-educated and high-income households may have moved to be closer to LRT stations over time, there was not a statistically significant number of these households. Finally, due to the use of fixed effects, this research removes CBG-specific time-invariant characteristics that may also have affected racial composition, educational attainment, and median income. In other words, it assumes that all CBGs are similar. This may not be the case. It could be that CBGs that are closer to LRT stations are innately different from CBGs further from stations, which may affect households' decisions to relocate to CBGs that are even closer to the stations.

Additional data-related limitations impede the quality of the models estimated. The use of ACS data meant that the most granular unit of data is at the census block group level. A more thorough analysis would require even smaller geographic units, such as a city block. Further, the ACS questionnaire is prone to changing over time. In addition to phrasing questions in different ways, which could skew survey results by introducing measurement error, the ACS also asks additional questions over time. Later versions of the ACS include questions regarding house-level amenities such as number of bedrooms, heating fuels, and monthly rent as well as household broadband access. Since these questions were not asked in earlier versions, it is difficult to incorporate the responses in the panel models. Finally, as mentioned in the data section, CBG boundaries can change over time leading to potential spatial mismatch without harmonized data. Spatial panel models may reveal more insights than the a-spatial panel models that were estimated as they capture indirect spatial interactions. Similarly, spatial panel structural equation models may be even better, as the use of latent variables can also help inform researchers on how factors that cannot be directly measured affect population change in an area.

The models can be improved in additional ways. One method is to include additional data sources. For instance, EMSI employment and jobs data could be added to better understand who live in the CBGs in the study area. If people follow jobs, then these data would likely improve overall model fit. Further work should also incorporate data on rental properties. As these CBGs are all located in heavily urban areas, changes in the rental market due to distance from LRT can have a significant effect on who lives where.

6. CONCLUSIONS, RESEARCH CONTRIBUTION AND RECOMMENDATIONS FOR FUTURE RESEARCH

As with other transportation technologies with the potential to change how people travel and, as a result, where they might live, the discussion surrounding light rail transit's effects on nearby populations has remained lively. Proponents of urban investments in LRT service point to direct benefits such as a reduction in congestion and emissions, while opponents of LRT argue that potential negative externalities, including increased noise pollution and decreased housing affordability near stations, outweigh such benefits. The research presented in this dissertation examines the potential externalities associated with LRT service on nearby areas in a medium sized American city, Charlotte, NC. Specifically, the research addresses the following three questions.

1. What are the effects of LRT stations during the announcement, construction, and operations phases of its life-cycle on nearby SFR prices, and do home-buyers' experience with LRT mitigate these effects?

2. How does directional heterogeneity in LRT station and CBD locations affect SFR prices?

3. How does LRT affect urban demographics over time?

This chapter is organized as follows. The next section will briefly summarize the key findings for each research question. After that the contributions of the research will be discussed. The chapter concludes with suggestions for future work that may overcome the limitations encountered in this research.

6.1 Key Findings

Gaps identified in the literature review and meta-analysis (Chapter 2) indicated that previous research on this topic failed to explicitly consider LRT life-cycle effects as well as tended to be a-spatial analyses. Through quasi-experimental spatial modeling, Chapter 3 shows that during the announcement and construction phases, proximity to proposed LRT stations is positively associated with SFR price per square foot. During the operations phase however, proximity to LRT stations is negatively associated with SFR prices. Furthermore, as seen in the analysis on LRT extension stations, home-buyers' prior experiences with LRT operations have a significant effect on SFR prices. These results show that to fully assess the impacts LRT have on nearby real estate, researchers need to consider the full life-cycle as well as whether or not the system being studied is an extension or a new system.

Another gap in the literature is a failure to account for the effects of directional heterogeneity with respect to LRT stations, SFRs, and CBD locations. Directional heterogeneity is operationalized as whether or not LRT stations are in the same direction as the CBD, and whether the station is between the SFR and the CBD, from an SFR's perspective. Chapter 4 presented that when this is the case, distance from the LRT station is only weakly associated with increased SFR prices. The implication of this result is that homeowners' willingness to pay to live near LRT stations is affected by the convenience by which they have access to LRT; thus, to assess LRT impacts one should also account for directional heterogeneity.

The majority of prior research surrounding the externalities associated with LRT have concentrated on its effects on real estate. Chapter 5 examines its effects on intra-urban migration. Such migration is proxied by changes in three demographic characteristics, percentage of African Americans, percentage of bachelor's degree holders, and the average median income of CBG residents in CBGs within a five-mile radius of LRT stations. While distance between CBG centroids and LRT stations have no effect on race, this variable has a negative and significant effects on educational attainment and median income. These results suggest that LRT stations have affected urban migration over time.

6.2 Research Contributions

The contributions of this research are in both the methodological and knowledge domains. Spatial econometric methods are under-utilized when conducting analyses on the effect that LRTs have on surrounding real estate markets, despite spatial clustering effects due to realtors' house showing behaviors. The failure to adequately account for such effects may lead to biased and inconsistent estimates, thereby limiting the value of inferences made from such studies.

Additionally, directional heterogeneity remains an under-examined component that can also affect housing prices in conjunction with distance. As much of transportation engineering research inherently involves both distance and direction via the movement of people and goods from point to point, the failure to account for directional effects can result in misleading inferences. While adequately operationalizing directional heterogeneity can be difficult, this research provides a framework by which it can be done.

This research adds to the extant pool of knowledge regarding the land use externalities of LRT. At the time of the conception of the dissertation research questions, most studies used only one or a few years' worth of housing data collected during the operations phase of newly opened light rail systems and their extensions. By using data from announcement, construction, and operations phases of an LRT, the dissertation research presents a more holistic picture of how LRT stations can impact nearby SFR markets. Furthermore, this research takes advantage of a natural experiment in the form of an extension line to evaluate the way in which prior exposure to LRT may have on potential home-buyers' willingness to pay for a single-family residence.

Finally, this research delves into the effects that LRT have on surrounding populations in addition to housing markets. While much has been published in recent years on displacement and gentrification due to transit-oriented developments centered on LRT stations, transit adjacent areas have received scant attention. By examining the demographic characteristics of census block groups, this research provides a look at the broader effects of LRT over time.

6.3 Limitations and Recommendations for Future Research

Due to data limitations, the analyses conducted in this dissertation have primarily focused on LRT's effects on nearby single-family residence sales. However, urban areas near LRT may include additional land uses, such as rental housing, condominiums, commercial and retail properties, and industrial parks. Indeed, according to [89], nearly 41% of Charlotte residents reside in rental units. In order to fully ascertain the life-cycle effects of LRT on nearby areas, one should analyze the proximity effects on all types of land uses. There may be heterogeneous effects due to differently purposed land. For instance, renters in downtown apartments with limited access to private vehicle parking may be willing to pay more per month to live closer to an LRT station. Similarly, commercial and retail businesses with downtown locations that lack parking facilities may seek to remove first-mile last-mile barriers to their potential customer base by preferring to locate near LRT stations. As land near stations are limited in supply, the preferences of these users may increase land rents which would be reflected in appraised land values.

The siting and location decisions of LRT stations and rail lines are assumed to be exogenous in this research, yet cities interested in constructing LRT typically conduct a series of town hall meetings with stakeholders as part of the planning process to determine where the physical LRT infrastructure ought to be located. Station locations are not randomly assigned, but rather, a product of a series of negotiations and compromises with incumbent landowners, transportation planners, residents, local businesses, and others. City officials, for instance, may be interested in acquiring affordable tracts of land to build on while transportation planners may seek to optimize station locations to maximize potential ridership. Business owners may want LRT to be located near their shops to attract more customers. Homeowners may prefer stations to be located further from their residences due to a fear of noise pollution and increased traffic.

Thus, participation at such town hall meetings would inform transportation agencies where to best locate LRT stations. However, there may be opportunity costs associated with attending these meetings, leading to an under-representation of certain viewpoints. For instance, low-income shift workers may be less likely to attend due to conflicting work schedules compared to medium- and high-income salaried employees who can attend evening planning meetings. Since housing opportunities and income are intrinsically linked, further work is needed to better understand how the circumstances surrounding the planning process of LRT systems may have indirect longer-term effects on the socio-demographic characteristics and real estate markets in transit adjacent areas.

Moreover, this research examines some of the longer run effects that LRT have on nearby areas. Short run effects are not accounted for. These may include noise pollution, other environmental consequences, and concurrent LRT-motivated commercial land developments experienced by nearby populations during the construction phases. Additional research is needed to know how LRT construction affects surrounding areas outside of just the real estate market, particularly as the construction of LRT may take many years. Further, a better understanding of such effects can help planners and policy makers educate locals of the potential advantages and disadvantages of LRT station siting decisions.

Finally, additional research is needed to better understand the extent to which LRT affects urban migration, both in terms of incumbent residents moving within a city as well as newcomers moving to the city. Although the data requirements for such an analysis may be difficult to meet, the results could have important implications for transportation planners. Such an analysis may reveal that some population segments are more likely to relocate to other parts of the city, or away from the city altogether, than others. Conversely, the presence of LRT may attract outsiders to move to the city from the suburbs or elsewhere. Knowing the population repulsion and attraction potential of LRT can help planners design transportation facilities to better serve the public.

REFERENCES

[1] Kenneth A Small and Camilla Kazimi. On the costs of air pollution from motor vehicles. *Journal of Transport Economics and policy*, pages 7–32, 1995.

[2] Richard D Knowles and Fiona Ferbrache. Evaluation of wider economic impacts of light rail investment on cities. *Journal of Transport Geography*, 54:430–439, 2016.

[3] Brian Gregor. Use of the greenstep model for scenario planning in oregon. 2012 National Scenario Planning Peer Exchange, 2012.

[4] Yue Ke and Konstantina Gkritza. Light rail transit and housing markets in charlotte-mecklenburg county, north carolina: Announcement and operations effects using quasi-experimental methods. *Journal of Transport Geography*, 76:212–220, 2019.

[5] Michael S Delgado and Raymond JGM Florax. Difference-in-differences techniques for spatial data: Local autocorrelation and spatial interaction. *Economics Letters*, 137:123–126, 2015.

[6] William Alonso et al. Location and land use. toward a general theory of land rent. *Location and land use. Toward a general theory of land rent.*, 1964.

[7] Edwin S Mills. An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 57(2):197–210, 1967.

[8] Richard F Muth. Cities and housing; the spatial pattern of urban residential land use. 1969.

[9] Clifford R Kern. Racial prejudice and residential segregation: The yinger model revisited. *Journal of Urban Economics*, 10(2):164–172, 1981.

[10] Jan K Brueckner, Jacques-Francois Thisse, and Yves Zenou. Why is central paris rich and downtown detroit poor?: An amenity-based theory. *European economic review*, 43(1):91–107, 1999.

[11] Chen Feng Ng. Commuting distances in a household location choice model with amenities. *Journal of Urban Economics*, 63(1):116–129, 2008.

[12] Thomas A Garrett. Light rail transit in america: Policy issues and prospects for economic development. *Unpublished Manuscript, Federal Reserve Bank of St. Louis, Research Department*, 2004.

[13] Michael Duncan. Comparing rail transit capitalization benefits for single-family and condominium units in san diego, california. *Transportation Research Record*, 2067(1):120–130, 2008.

[14] Daniel Baldwin Hess and Tangerine Maria Almeida. Impact of proximity to light rail rapid transit on station-area property values in buffalo, new york. *Urban studies*, 44(5-6):1041–1068, 2007.

[15] Miriam Zuk, Ariel H Bierbaum, Karen Chapple, Karolina Gorska, Anastasia Loukaitou-Sideris, Paul Ong, and Trevor Thomas. Gentrification, displacement and the role of public investment: a literature review. In *Federal Reserve Bank of San Francisco*, volume 79, 2015.

[16] Nathaniel Baum-Snow and Matthew E Kahn. The effects of new public projects to expand urban rail transit. *Journal of Public Economics*, 77(2):241–263, 2000.

[17] Robert Cervero and Michael Duncan. Benefits of proximity to rail on housing markets: Experiences in santa clara county. *Journal of Public Transportation*, 5(1):3, 2002.

[18] Carol Atkinson-Palombo. Comparing the capitalisation benefits of light-rail transit and overlay zoning for single-family houses and condos by neighbourhood type in metropolitan phoenix, arizona. *Urban studies*, 47(11):2409–2426, 2010.

[19] Aaron Golub, Subhrajit Guhathakurta, and Bharath Sollapuram. Spatial and temporal capitalization effects of light rail in phoenix: From conception, planning, and construction to operation. *Journal of Planning Education and Research*, 32(4):415–429, 2012.

[20] Chang-Hee Christine Bae, Myung-Jin Jun, and Hyeon Park. The impact of seoul's subway line 5 on residential property values. *Transport policy*, 10(2):85–94, 2003.

[21] Shanjun Li, Jun Yang, Ping Qin, and Shun Chonabayashi. Wheels of fortune: Subway expansion and property values in beijing. *Journal of Regional Science*, 56(5):792–813, 2016.

[22] Francesca Pagliara and Enrica Papa. Urban rail systems investments: an analysis of the impacts on property values and residents' location. *Journal of Transport Geography*, 19(2):200–211, 2011.

[23] Gerrit J Knaap, Chengr Ding, and Lewis D Hopkins. Do plans matter? the effects of light rail plans on land values in station areas. *Journal of Planning Education and Research*, 21(1):32–39, 2001.

[24] Sisi Yan, Eric Delmelle, and Michael Duncan. The impact of a new light rail system on single-family property values in charlotte, north carolina. *Journal of Transport and Land Use*, 5(2):60–67, 2012.

[25] Stephen B Billings. Estimating the value of a new transit option. *Regional Science and Urban Economics*, 41(6):525–536, 2011.

[26] Jacob Camins-Esakov and Donald Vandegrift. Impact of a light rail extension on residential property values. *Research in Transportation Economics*, 67:11–18, 2018.

[27] Hongwei Dong. Rail-transit-induced gentrification and the affordability paradox of tod. *Journal of Transport Geography*, 63:1–10, 2017.

[28] Michael R Ransom. The effect of light rail transit service on nearby property values. *Journal of Transport and Land Use*, 11(1):387–404, 2018.

[29] Kihwan Seo, Aaron Golub, and Michael Kuby. Combined impacts of highways and light rail transit on residential property values: a spatial hedonic price model for phoenix, arizona. *Journal of Transport Geography*, 41:53–62, 2014.

[30] Wenjie Wu, Guanpeng Dong, and Bing Wang. Does planning matter? effects on land markets. *The Journal of Real Estate Finance and Economics*, 50(2):242–269, 2015.

[31] Timothy F Welch, Steven R Gehrke, and Fangru Wang. Long-term impact of network access to bike facilities and public transit stations on housing sales prices in portland, oregon. *Journal of Transport Geography*, 54:264–272, 2016.

[32] Eleni Bardaka, V Dimitra Pyrialakou, and Konstantina Gkritza. Assessing the socioeconomic effects of transit rail development: An exploration of displacement and gentrification. Technical report, 2016.

[33] Brigitte Waldorf and Pillsung Byun. Meta-analysis of the impact of age structure on fertility. *Journal of Population Economics*, 18(1):15–40, 2005.

[34] Ghebreegziabiher Debrezion, Eric Pels, and Piet Rietveld. The impact of railway stations on residential and commercial property value: a meta-analysis. *The Journal of Real Estate Finance and Economics*, 35(2):161–180, 2007.

[35] Sara I Mohammad, Daniel J Graham, Patricia C Melo, and Richard J Anderson. A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice*, 50:158–170, 2013.

[36] Raymond JGM Florax, Henri LF de Groot, and Ruud A De Mooij. Meta-analysis: A tool for upgrading inputs of macroeconomic policy models. Technical report, Tinbergen Institute Discussion Paper, 2002.

[37] William G Cochran. Estimation of bacterial densities by means of the" most probable number". *Biometrics*, 6(2):105–116, 1950.

[38] Neal L Oden and Robert R Sokal. Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Zoology*, 35(4):608–617, 1986.

[39] WB Mercer and AD Hall. The experimental error of field trials. *The Journal of Agricultural Science*, 4(2):107–132, 1911.

[40] Daniel Borcard, Pierre Legendre, and Pierre Drapeau. Partialling out the spatial component of ecological variation. *Ecology*, 73(3):1045–1055, 1992.

[41] Pierre Legendre, Marie-Josée Fortin, and Daniel Borcard. Should the mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, 6(11):1239–1247, 2015.

[42] F Guillaume Blanchet, Pierre Legendre, and Daniel Borcard. Modelling directional spatial processes in ecological data. *Ecological Modelling*, 215(4):325–336, 2008.

[43] Andrea Bertolo, F Guillaume Blanchet, Pierre Magnan, Philippe Brodeur, Marc Mingelbier, and Pierre Legendre. Inferring processes from spatial patterns: the role of directional and non–directional forces in shaping fish larvae distribution in a freshwater lake system. *PLoS One*, 7(11):e50239, 2012.

[44] Fangpo Wang, Alan E Gelfand, and Giovanna Jona-Lasinio. Joint spatio-temporal analysis of a linear and a directional variable: space-time modeling of wave heights and wave directions in the adriatic sea. *Statistica Sinica*, pages 25–39, 2015.

[45] TA Cameron. Directional heterogeneity in distance profiles in hedonic property value models. *Journal of Environmental Economics and Management*, 51(1):26–45, 2006.

[46] Mark D. Agee and Thomas D. Crocker. Directional heterogeneity of environmental disamenities: the impact of crematory operations on adjacent residential values. *Applied Economics*, 42(14):1735–1745, 2010.

[47] S. Waldorf, Brigitte. A spatial analysis of rent shitfs in the chicago rental housing market, 1970-1980. *Urban geography*, 12(5):450–468, 1991.

[48] Peter F Colwell and Henry J Munneke. Directional land value gradients. *The Journal of Real Estate Finance and Economics*, 39(1):1–23, 2009.

[49] Gershon Alperovich and Joseph Deutsch. An application of a switching regimes regression to the study of urban structure. *Papers in Regional Science*, 81(1):83–98, 2002.

[50] Ruth Glass. *London: aspects of change*. Number 3. MacGibbon & Kee, 1964.

[51] Bradley Bereitschaft. Gentrification central: A change-based typology of the american urban core, 2000–2015. *Applied Geography*, 118:102206, 2020.

[52] Nathan Eric Rochester. On both sides of the tracks: Light rail and gentrification in portland, oregon. 2016.

[53] Richard Florida. The flight of the creative class: The new global competition for talent. *Liberal Education*, 92(3):22–29, 2006.

[54] Loretta Lees. Gentrification and social mixing: towards an inclusive urban renaissance? *Urban studies*, 45(12):2449–2470, 2008.

[55] Jason Hackworth and Neil Smith. The changing state of gentrification. *Tijdschrift voor economische en sociale geografie*, 92(4):464–477, 2001.

[56] Mark Davidson and Loretta Lees. New-build 'gentrification'and london's riverside renaissance. *Environment and planning A*, 37(7):1165–1190, 2005.

[57] Dwayne Marshall Baker and Bumsoo Lee. How does light rail transit (lrt) impact gentrification? evidence from fourteen us urbanized areas. *Journal of Planning Education and Research*, 39(1):35–49, 2019.

[58] Daniel Monroe Sullivan. Reassessing gentrification: Measuring residents' opinions using survey data. *Urban Affairs Review*, 42(4):583–592, 2007.

[59] Tom Slater. Missing marcuse: On gentrification and displacement. *City*, 13(2-3):292–311, 2009.

[60] Peter Marcuse. Gentrification, abandonment, and displacement: Connections, causes, and policy responses in new york city. *Wash. UJ Urb. & Contemp. L.*, 28:195, 1985.

[61] Phillip L Clay. *Neighborhood renewal: middle-class resettlement and incumbent upgrading in American neighborhoods*. Free Press, 1979.

[62] Rowland Atkinson. Measuring gentrification and displacement in greater london. *Urban studies*, 37(1):149–165, 2000.

[63] Margaret Kohn. What is wrong with gentrification? *Urban Research & Practice*, 6(3):297–310, 2013.

[64] Kacie Dragan, Ingrid Gould Ellen, and Sherry Glied. Does gentrification displace poor children and their families? new evidence from medicaid data in new york city. *Regional Science and Urban Economics*, 83:103481, 2020.

[65] Annelise Grube-Cavers and Zachary Patterson. In order to keep mass transit accessible, we must understand the relationship between gentrification and public transportation. *LSE American Politics and Policy*, 2014.

[66] Lance Freeman. Displacement or succession? residential mobility in gentrifying neighborhoods. *Urban Affairs Review*, 40(4):463–491, 2005.

[67] Neil Smith. Gentrification and the rent gap. 1987.

[68] Neil Smith and Peter Williams. *Gentrification of the City*. Routledge, 2013.

[69] Lei Ding, Jackelyn Hwang, and Mohammad Tahsildoost. Effects of gentrification on homeowners: Evidence from a natural experiment. *Regional Science and Urban Economics*, page 103536, 2020.

[70] Frank JB Stilwell and Frank JB Stilwell. *Economic inequality: Who gets what in Australia*. Pluto press Sydney, 1993.

[71] Japonica Brown-Saracino. Social preservationists and the quest for authentic community. *City & Community*, 3(2):135–156, 2004.

[72] Damaris Rose. Rethinking gentrification: beyond the uneven development of marxist urban theory. *Environment and planning D: Society and Space*, 2(1):47–74, 1984.

[73] Lindsey Dillon. Race, waste, and space: Brownfield redevelopment and environmental justice at the hunters point shipyard. *Antipode*, 46(5):1205–1221, 2014.

[74] Tomeka Davis and Deirdre Oakley. Linking charter school emergence to urban revitalization and gentrification: A socio-spatial analysis of three cities. *Journal of Urban Affairs*, 35(1):81–102, 2013.

[75] Gary Bridge and Tim Butler. *Mixed communities: Gentrification by stealth?* Policy Press, 2011.

[76] Casey Dawkins and Rolf Moeckel. Transit-induced gentrification: Who will stay, and who will go? *Housing Policy Debate*, 26(4-5):801–818, 2016.

[77] Sara L Jackson and Joshua Buckman. Light rail development with or without gentrification?: Neighborhood perspectives on changing sense of place in denver, colorado. *Journal of Transport Geography*, 84:102678, 2020.

[78] Isabelle Nilsson and Elizabeth Delmelle. Transit investments and neighborhood change: On the likelihood of change. *Journal of Transport Geography*, 66:167–179, 2018.

[79] Bradley Bereitschaft. Gentrification and the evolution of commuting behavior within america's urban cores, 2000–2015. *Journal of Transport Geography*, 82:102559, 2020.

[80] Eleni Bardaka, Michael S Delgado, and Raymond JGM Florax. Causal identification of transit-induced gentrification and spatial spillover effects: The case of the denver light rail. *Journal of Transport Geography*, 71:15–31, 2018.

[81] J LeSage and R Pace. Introduction to spatial econometrics. 2009.

[82] Harry H Kelejian and Ingmar R Prucha. Hac estimation in a spatial framework. *Journal of Econometrics*, 140(1):131–154, 2007.

[83] Francesca Medda. Land value capture finance for transport accessibility: a review. *Journal of Transport Geography*, 25:154–161, 2012.

[84] Giovanni Circella, Grant Matson, Farzad Alemi, and Susan Handy. Panel study of emerging transportation technologies and trends in california: Phase 2 data collection. 2019.

[85] Karen Aspelin and Nick Carey. Establishing pedestrian walking speeds. *Portland State University*, pages 5–25, 2005.

[86] Graham Elliott. Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution. *International Economic Review*, 40(3):767–784, 1999.

[87] Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

[88] Badi H Baltagi and Ping X Wu. Unequally spaced panel data regressions with ar (1) disturbances. *Econometric Theory*, pages 814–823, 1999.

[89] U.S. Census Bureau. 2018 5-year american communities survey, 2018.

VITA

Yue Ke is a PhD Candidate in the Lyles School of Civil Engineering in the area of transportation and infrastructure systems at Purdue University in West Lafayette, Indiana. He grew up in Davidson, North Carolina and attended the University of North Carolina - Chapel Hill where he received his bachelor's of arts in economics and international studies with a minor in computer science in 2008. After completing his Peace Corps service in Tanzania in 2011 and tenure at the UN World Food Program in 2013, Yue resumed his studies at Oregon State University (OSU) under the supervision of Dr. B. Starr McMullen and Dr. Mallory Rahe. In 2015 he received his M.S. in Applied Economics with a concentration in trade and development; his thesis was titled "Economic Analysis of a Road Usage Charge." Afterwards, he began his PhD in Civil Engineering at OSU. In 2016, he transferred to Purdue University to continue his PhD studies.

During his stay at Purdue University, Yue was involved with various research projects including the evaluation of wider economic impacts of road and bridge preservation projects, incorporating economic resilience into transportation agency decision-making, and the impacts transportation has on regional economic resilience. His research interests include transportation and economic development, transportation economics, pedestrian safety, and public health. Specifically, he applies spatial econometric models to assess the externalities due to transportation investments and policies. Yue has also served on the council of the Purdue Student Chapter of the Institute of Transportation Envineers (ITE) as Treasurer (2017-2018) and Secretary (2018-2019).

PRESENTATIONS & PUBLICATIONS

PUBLICATIONS

Ke, Y. & Waldorf, B., (2020). Rural Elderly Mobility: A Case Study of Oregon, USA. *Travel Behaviour and Society (under review).*
Ke, Y. & Gkritza, K., (2019). Light rail transit and housing markets in Charlotte-Mecklenburg County, North Carolina: Announcement and operations effects using quasi-experimental methods. *Journal of Transport Geography, 76*, 212-220.
Ke, Y., & Gkritza, K. (2018). Safety ramifications of a change in pedestrian crosswalk law: A case study of Oregon, USA. *International Journal of Transportation Science and Technology.*
Ke, Y., & Gkritza, K. (2018). Income and spatial distributional effects of a congestion tax: A hypothetical case of Oregon. *Transport Policy*, *71*, 28-35.
Lavrenz, S. M., Vlahogianni, E. I., Gkritza, K., & **Ke, Y**. (2018). Time series modeling in traffic safety research. *Accident Analysis* & *Prevention*, *117*, 368-380.
Ke, Y., & McMullen, B. S. (2017). Regional differences in the determinants of Oregon VMT. *Research in Transportation Economics*, *62*, 2-10.

PRESENTATIONS

2019 North American Regional Science Conference. Yue Ke and Konstantina Gkritza. H*ow does directional heterogeneity affect prices of single-family residences near LRT stations?*
2018 North American Regional Science Conference. Yue Ke and Konstantina Gkritza. *A meta-analysis of light rail transit's effects on nearby single-family residence land values in the United States.*
2018 International Transportation Economic Development Conference. Yue Ke and Konstantina Gkritza. *Transportation Infrastructure and Poverty: A Case Study of Ethiopia.*
2018 Transportation Research Forum. Yue Ke, Konstantina Gkritza, and Brigitte Waldorf. *Rural Elderly Mobility in Oregon: A Case Study.*
2018 TRB 97th Annual Meeting. Yue Ke and Konstantina Gkritza. *Geospatial Trends in Light Rail Transit's Effects on Nearby House Prices.* 2017 Road Safety Simulation Conference. Yue Ke and Konstantina Gkritza. *Safety Ramifications of a Change in Pedestrian Crosswalk Law: A Case Study of Oregon, USA.*
2017 North American Regional Science Conference. Yue Ke and Konstantina Gkritza. *The Announcement of Light Rail Transit's Signaling Effect on Nearby Property Values in Charlotte, NC.*

2017 Applying Census Data for Transportation Workshop. Yue Ke and Konstantina Gkritza. *Conspicuous Consumption: Geospatial trends in vehicle choice and travel behavior.*
2017 TRB 96th Annual Meeting. Yue Ke. *Congestion: A Time Series Analysis.*
2017 TRB 96th Annual Meeting. Yue Ke and B. Starr McMullen. *Determinants of VMT in Oregon.*
2017 Transportation Research Forum. Yue Ke and Konstantina Gkritza. *Income and Spatial Distributional Effects of a Congestion Tax.*