4-8-2016

Assessing Property Value Impacts of Access to Bus Rapid Transit (BRT): Case Study of the Cleveland HealthLine

Victoria A. Perk
University of South Florida, vperkw@gmail.com

Assessing Property Value Impacts of Access to Bus Rapid Transit (BRT):

Case Study of the Cleveland HealthLine

by

Victoria A. Perk

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Economics
College of Arts and Sciences
University of South Florida

Major Professor: Joseph S. DeSalvo, Ph.D.
Xuehao Chu, Ph.D.
Murat Munkin, Ph.D.
Christopher R. Thomas, Ph.D.

Date of Approval:
April 7, 2016

Keywords: bus rapid transit, hedonic regression, housing prices, property values

DEDICATION

This dissertation is dedicated to my family; my wonderful and supportive husband, David J. Wells, and our two amazing children, Ryan John Franklin Wells and Allison Ann Elizabeth Wells. The completion of this work took many hours of my time away from them, and I will always appreciate their patience, support, and encouragement.

ACKNOWLEDGMENTS

I would like to acknowledge the support, guidance, and feedback of my dissertation committee: Joseph DeSalvo, Xuehao Chu, Murat Munkin, and Christopher Thomas. In addition, I would like to thank Dennis Hinebaugh and Joel Volinski of the National Center for Transit Research (NCTR) at the Center for Urban Transportation Research (CUTR) at the University of South Florida, for providing support for the completion of this work. Finally, I would like to recognize Martin Catalá, a colleague at CUTR, for assisting with construction of the data set and related geographic information systems (GIS) analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The nation's economy depends heavily on mobility of goods and people. As communities look to improve mobility, many options can be considered, including roadway improvements, congestion-pricing options such as dynamic tolling and toll lanes, and public transit. Investment in public transit services can come in the form of increased and enhanced bus services, including bus rapid transit (BRT), as well as rail transit investments. As BRT continues to grow in popularity in the United States, a better understanding of the mode's impacts on land uses and economic development is needed. One method of assessing the mode's impacts is by examining the market value of properties with access to BRT stations. Based on land-rent theory, it is hypothesized that people will be willing to pay a premium for convenient and reliable access via BRT to the central business district (CBD) or other locations with employment, educational, recreational, and shopping opportunities.

Very little research has been conducted on BRT as it operates in the present day in the United States. For this work, the hypothesis is that the BRT stations have a positive impact on the market value of residential properties. To test this hypothesis, hedonic price regression models are used to estimate the impact of access to BRT

stations on the sale prices of surrounding single-family homes using a case study of the HealthLine BRT system in Cleveland, Ohio that began operating in 2008. Three time periods were examined: 2004, the year construction began; 2008–2009, after the HealthLine BRT service began operation; and 2010–2011, the latest year for which sales data are available. Despite a documented decline in median sale prices of single-family homes in the city of Cleveland from 2005 to 2011, overall results of the analysis were mixed. Although it was prior to the opening of the BRT system, the 2004 data did not show any impacts of the stations on surrounding home sale prices. For the 2008–2009 data, positive and statistically significant impacts were found; however, the positive impacts did not persist in the 2010–2011 data. It would likely be necessary to seek out additional years of data to fully answer the question posed by this research.

It is important for decision-makers to have the most accurate and most recent information on the benefits and costs of all transportation alternatives, including BRT. The research presented herein makes a significant contribution to filling the current gap in quantitative research on the subject and provides planners, policymakers, and the transit industry with the best information possible to make sound transit investment decisions in their communities.

CHAPTER ONE:

INTRODUCTION

The nation's economy depends heavily on mobility of goods and people. According to the US General Accounting Office (US GAO), increased levels of investment are needed to improve and maintain the mobility currently provided by the nation's highways and transit systems [1]. As communities look to improve mobility, many options can be considered, including roadway improvements, congestion-pricing options such as dynamic tolling and toll lanes, and public transit. Investment in public transit services can come in the form of increased and enhanced bus services, including bus rapid transit (BRT), as well as rail transit investments. Communities seeking to invest in public transit infrastructure are expecting benefits such as decreased travel times, decreased greenhouse gas emissions, and economic development [1] [2].

In addition to expected costs and benefits, the political climate in a given area can heavily influence the alternatives considered and ultimately selected. Some areas may not approve of paying for improved transit services[1]. Other areas may wish to invest in

---

[1] As examples local to the Tampa Bay area, the Greenlight Pinellas initiative failed in Pinellas County, Florida in November 2014. Hillsborough County, Florida experienced a failed sales tax referendum for transit in 2010, but the County is hoping to get another plan on the ballot for 2016.

transit but have a modal preference, often toward the more expensive rail transit because it is generally perceived by some as more attractive, cleaner, and faster than buses, and because it is perceived to have greater economic impacts [3] [4] [5]. Clearly, local decision-makers should be going beyond perceptions and have available the best and most reliable information regarding expected costs and benefits of various transit alternatives. While the importance of perceptions, particularly by the public cannot be downplayed, a solid body of research in this area can help determine whether perceptions are true or need to be adjusted. The research presented in this document aims to contribute to the literature on this topic by analyzing the property-value impacts of access to BRT stations for the HealthLine BRT system operating in Cleveland, Ohio.

Background

Because benefits such as increased property values are an important factor in the selection and funding of a transit alternative, a brief description of the major federal funding mechanism is included here, along with definitions of the relevant transit modes. Under the MAP-21 legislation of 2012 (Moving Ahead for Progress in the 21st Century Act), the Federal Transit Administration (FTA) has continued its "New Starts" Fixed Guideway Capital and Investment Grants program. A "fixed guideway" is a separate right-of-way used exclusively for public transportation (or that includes a rail or catenary system). Eligible projects include new fixed guideways or extensions, BRT projects operating in mixed traffic that also represent a significant investment in the selected corridor, and other projects that improve capacity on an existing fixed guideway [6]. Thus, both rail and BRT projects are eligible for funding under this program. An application to the New Starts program consists of many required elements, one of which is a review of alternative analyses conducted at the local or regional level. Project costs and expected benefits are also required to be reported. The US GAO has addressed common pitfalls in alternative analyses and benefit/cost analyses and provides guidance on how to improve the results of such exercises [1]. To be eligible for federal grants for capital costs of such projects (maximum federal share is 80 percent), local match funding is necessary, which may require a vote by citizens to tax themselves to pay for such improvements. Though typically much more expensive to construct and implement, in some cases heavier political favor has been given to new rail projects, typically light rail transit (LRT) or streetcar projects, along with the hope of revitalizing older areas and spurring new economic development. However, BRT is also becoming more popular, with more than 20 cities in the U.S. having implemented some form of BRT and many more planning such services [7]. In some cases, as with the recent Greenlight Pinellas initiative in Florida, proposed transit improvements contain both LRT and BRT projects.

A short description of the major public transit modes can be helpful in better understanding the context of the research presented herein. In some cases, particularly among the public, there can be some confusion about the different types of transit and what exactly a proposed transit improvement will look like upon implementation. Briefly discussed on the following pages are commuter rail, heavy rail, light rail transit/streetcar, and bus rapid transit. The source for this information is the FTA [8].

Commuter Rail

Commuter rail service typically consists of local short haul travel between a central city and its suburbs. The service uses locomotive hauled or self-propelled passenger cars operating on mostly current or former freight railroad track. Stations are spaced out widely, and the service is also characterized by station-to-station or zone fares. Examples include Virginia Railway Express and Caltrain in California. In Florida, SunRail, newly opened in 2014, connects Volusia County and Orange County through Downtown Orlando and Tri-Rail in Florida connects Miami-Dade, Broward, and Palm Beach Counties. Figure 1.1 shows these systems.

Heavy Rail

Heavy rail is most often synonymous with transit services called the "metro," "subway," or "rapid transit." This type of service operates on an electric railway with high capacity passenger cars and is always on a separate right-of-way. It can operate underground, at grade, or on an elevated track. Examples are the New York subway system, the Washington D.C. Metro rail system, and Chicago's "L" system (short for "elevated"). The only example in Florida is Miami-Dade Transit's Metrorail service. Examples of these heavy rail systems are shown in Figure 1.2.

For photo credits see Appendix C)

Figure 1.1. Examples of Commuter Rail Systems

Light Rail Transit/Streetcar

Light rail transit (LRT) is a rail mode that operates on fixed tracks yet is not necessarily grade-separated. It can operate in its own right-of-way or can sometimes operate in mixed traffic. Single cars or trains of multiple cars are used depending on capacity requirements. LRT and streetcar vehicles are usually electrically powered using an overhead electric line or catenary system. Examples include the San Diego Trolley, the Portland MAX, and the Lynx in Charlotte, North Carolina. The TECO Line Streetcar in Tampa is a local example of this service (and the only service of this kind currently operating in Florida). These systems are shown in Figure 1.3.

For photo credits see Appendix C)

Figure 1.2. Examples of Heavy Rail Systems

For photo credits see Appendix C)

Figure 1.3. Examples of Light Rail/Streetcar Systems

Bus Rapid Transit

Additional discussion is devoted to BRT because it is the focus of this research. According to the FTA, BRT is an "*enhanced bus system that operates on bus lanes or other transitways in order to combine the flexibility of buses with the efficiency of rail. By doing so, BRT operates at faster speeds, provides greater service reliability and increased customer convenience. It also utilizes a combination of advanced technologies, infrastructure and operational investments that provide significantly better service than traditional bus service*" [9]. This definition, however, applies to a wide variety of rapid bus services currently operating and in the planning stages in the U.S. A BRT system comprises an integration of various service characteristics including the type of running way, stations, vehicles, fare collection method, intelligent transportation systems (ITS, which can include technology such as real-time information), service plans, and unique branding [10]. It is important to note that, to be regarded as a form of BRT, the service must incorporate some form of each of these seven elements. Typical express bus services or limited-stop services are not considered BRT.

Currently, more than 20 cities in the U.S. are operating some form of BRT. These systems range from what is referred to as BRT "Lite" services such as the Metro Rapid in Los Angeles and the MAX in Kansas City, to the full-featured rail-like operations of Cleveland's HealthLine and the EmX in Eugene, Oregon.

BRT can be constructed and implemented relatively quickly, has operational flexibility, and can have elements tailored specifically to the needs and characteristics of the community. In general, BRT systems are less expensive to construct and operate than LRT systems; however, the more rail-like the BRT system, the higher the costs. Interestingly, it is BRT's flexibility that can result in the assumption that it is not as "permanent" an investment as a rail mode and, therefore, some believe it cannot attract economic development to the extent that rail transit, with its fixed tracks, can [5]. Even so, if decision-makers consider the marginal return per dollar of investment, even if LRT was to generate more development in absolute terms, BRT could still look more favorable given its lower costs. Further, the extent to which public transit in general, and rail specifically (particularly LRT), can spur economic development is itself a subject of debate [5] [2].

Figure 1.4 shows some of the operating BRT systems in the U.S. Two BRT systems operating in Los Angeles are shown in Figure 1.4, including the BRT "Lite" Metro Rapid which operates in mixed traffic and the full-featured, more rail-like Metro Orange Line, which is branded with the name of a color like Metro's other rail system lines, and operates in an exclusive guideway. Also shown in Figure 1.4 is the Las Vegas MAX, which uses stylized vehicles that appear to be rail cars. The Kansas City MAX is considered to be another BRT "Lite" system, operating in mixed traffic with typical 40-foot transit buses that are branded differently from the rest of the transit system. In

Eugene, Oregon, the Lane Transit District operates the Emerald Express, or EmX, which is a more full-featured BRT system operating for most of its alignment in an exclusive median guideway. The EmX features stylized transit vehicles and other rail-like characteristics including real-time customer information at stations, signal priority at intersections, and off-board fare collection. Distinct branding of the service is coupled with attractive public art in and around the stations. The other three BRT systems exhibited in Figure 1.4 are also the subjects of the only research (completed and ongoing) to date in the U.S. on the mode's impact on residential property values [11] [12]. As described in a later chapter of this dissertation, it is the Cleveland HealthLine system that is the selected case study site for this research.

The Cleveland HealthLine is considered one of the country's most successful BRT systems to date. It was also one of the most expensive to construct, as the work was coupled with a complete renovation of the city's famed Euclid Corridor, including the relocation of utilities. However, it is also one of the more rail-like BRT systems, with stylized vehicles, an exclusive median guideway for most of the alignment, off-board fare collection, signal priority at intersections, real-time information at stations, and level boarding on elevated platforms. The station areas include public art as well as lush landscaping. The Boston Silver Line (Washington Street Corridor) does operate in its own marked lane, but not on a separated guideway. It also does not have stylized vehicles, instead operating with more traditional 40- and 60-foot transit buses. Still, the

For photo credits see Appendix C)

Figure 1.4. Select BRT Systems Operating in the U.S.

Silver Line service, which is branded as part of the Massachusetts Bay Transportation Authority's (MBTA) rail system, does serve several neighborhoods of multi-unit housing and provides direct access to the Boston Central Business District (CBD). Lastly, the Pittsburgh East Busway is an exclusive guideway on which several of the transit system's bus routes travel. The busway was implemented in 1983, but it was several years later that the routes operating on it began to include more typical BRT characteristics. As such, it is still considered one of the oldest operating BRT systems and was the subject of the first study of property values impacts around stations in the U.S. [11].

While a mode such as LRT has a straightforward definition, the information above provides evidence that the BRT mode is very broadly defined and applied. BRT has such a wide variety of applications that it can be more difficult to draw conclusions about its impacts, since no two systems are alike. Additional research needs to be conducted on the various types of BRT operations to provide some sense of the mode's overall impacts. To date, research for U.S. operations is scant.

Focus of Research

As BRT continues to grow in popularity in the United States, a better understanding of the mode's impacts on land uses and economic development is needed. One method of assessing the mode's impacts is by examining the market value of properties with access to BRT stations. Based on land-rent theory (discussed in

Chapter Two), it is hypothesized that people will be willing to pay a premium for convenient and reliable access via BRT to the CBD or other locations with employment, educational, recreational, and shopping opportunities.

Most of the previous research on this topic has been focused on rail transit modes and is both qualitative and quantitative in nature. It is often the anecdotal, qualitative work that attracts the most attention in the media and is used by proponents of rail transit to advance their cause. As will be further discussed in this dissertation, when rigorous quantitative studies are conducted it is often found that closer access to rail transit does increase property values in a statistically significant way, but the increases are relatively small in magnitude.

Very little research has been conducted on BRT as it operates in the present day in the United States. Studies have been conducted on the topic of property value impacts of BRT operating in other countries, including Colombia, South Korea, and Australia. However, because of various social, cultural, political, and institutional differences, it is unlikely that the experiences in those countries will correlate to the U.S. experience, as discussed in Chapter Three of this document. To date, only two studies have been published on property value impacts of BRT systems in the U.S., both authored by Perk, et al., and they alone are not sufficient to provide enough information on the topic [11] [12].

Further, to address those who advocate for rail investments over BRT based on the available rail research and relative lack of BRT research, it should be noted that a lack of research does not equate to an absence of impacts. Simply because some groups may believe that BRT would not have impacts comparable to those of rail modes does not make it true. Little research exists because the BRT mode is relatively new, with systems that have been operating in the U.S., for the most part, only within the last 10 to 12 years. Thus, there exists a need for more research on this subject. An interesting anecdote related to this involves the failed Greenlight Pinellas sales tax referendum in Pinellas County, Florida. The LRT component to the plan was apparently chosen because it was believed it could attract significant benefits [13]. However, it seems that one of the reasons it failed was due to strong opposition to the expensive LRT. A representative of the opposing group No Tax for Tracks indicated that the group did support a stronger bus system [14]. One can wonder if the outcome of the referendum might have been different if the plan had focused on allocating the LRT funds to increased bus services, including BRT (although some BRT was part of the plan). What if the planners behind Greenlight Pinellas, as well as its supporters and detractors, had more research to reference regarding the impacts of BRT? Of course, it is not clear that the outcome would have been different, but this example illustrates the need for a larger body of research on this topic.

For this research, the hypothesis is that the BRT stations have a positive impact on the market value of residential properties. To test this hypothesis, hedonic price regression models are used to estimate the impact of access to BRT stations on the sale prices of surrounding residential properties.

It is important for decision-makers to have the most accurate and most recent information on the benefits and costs of all transportation alternatives, including BRT. The research presented herein makes a significant contribution to filling the current gap in quantitative research on the subject and provides planners, policymakers, and the transit industry with the best information possible to make sound transit investment decisions in their communities.

CHAPTER TWO:

THEORETICAL BACKGROUND

The theoretical foundation for the expectation that property values will increase with access to public transit is based in urban economics. Over 50 years ago, economists became more interested in studying and explaining urban form and spatial structure, and how advances in communications and transportation helped shape modern cities. Much research has been focused on agglomeration economies, which refer to decreasing average costs as production increases in a specific geographical location, and can result in a premium on land in certain areas [15].

Some early work focused on firms converting rural land to urban land for production. A 1961 paper by Richard Muth addressed the issue of converting land from rural to urban uses by deriving rent-distance functions which show the rent firms would offer for land at any given distance from the market [16]. He used an example of firms in two industries and how their locations would change with changes in demand and supply conditions for their two commodities. Muth considered four kinds of equilibrium conditions:

1. Firm equilibrium requires profit maximization.
2. Locational equilibrium requires profits of identical firms be equal no matter the location.
3. Market equilibrium for land requires land being allocated to the use that yields the highest rent.
4. Industry equilibrium requires the quantity supplied of the good to be equal to the quantity demanded at the market price.

Muth showed that the rent-distance functions are derived from maximizing profits subject to the production function. He defined a rent-distance function for each industry, showing the rent that firms would offer at various distances from the market.

From this initial work, subsequent important contributions were made by Alonso, Muth, Mills, and Wheaton, leading to a synthesis and additional contributions by Brueckner. Below is a brief description of Alonso's model, as discussed in Anas, et al.

William Alonso's monocentric city model, which began as an evolution of von Thünen's theory of agricultural land use (1826), incorporates production, transportation, and housing [17]. As described in Anas, et al., Alonso's closed city case envisions a city as a circular residential area that surrounds a CBD. All jobs are located in the CBD. There are N identical households that receive utility, $u(z, A)$, from some good z and residential lot size A. A household x miles from the CBD incurs annual transportation costs, $T(x)$, which represent commuting costs. The household's exogenous income, y, must pay for z, transportation, and land rent $r(x)$. As shown in

Anas, et al., residential bid-rent, *b*, at a location x can be defined as the maximum rent (per unit of land area) that a household can pay and still keep utility constant at $\bar{u}$ [15]. This is shown by:

$$b(x, \bar{u}) = \max_{z,A} \frac{y - T(x) - z}{A} \qquad s.t.\ u(z, A) \geq \bar{u}$$

It is also shown that, by the envelope theorem, the slope of the bid-rent function is represented by:

$$\frac{db(x, \bar{u})}{dx} = \frac{-T'(x)}{A[y - T(x), \bar{u}]}$$

The above is considered a basic result of the monocentric model and demonstrates that a household at a small additional distance (*dx*) from the CBD incurs additional transportation costs, $T'(x)dx$ [15]. To keep the household indifferent between two locations (i.e., to keep it at $\bar{u}$), land rent must be lower at the farther location by the same amount as the increase in transportation costs, $Adb = -T'(x)dx$. For each household, there exists a family of residential bid-rent functions. The equilibrium rent function, $r(x)$, is determined by two conditions:

1. Rent at the city boundary, x^*, equals the agricultural rent, r_A.
2. All household must be accommodated.

Therefore, the equilibrium land rent at any location is the maximum of the bid-rents at that location:

$$r(x) = max[b(x, u^e), r_A] = \begin{cases} b(x, u^e)\ for\ x \leq x^* \\ r_A\ for\ x > x^* \end{cases}$$

In the simple monocentric model, each parcel of land goes to its highest-bidding use and the land use is efficient.

Edwin S. Mills also incorporated transportation costs and housing into his model, which he developed to help explain the "size and structure" of urban areas [18]. Mills assumed competitive markets and considered three activities: goods production, intracity transportation, and housing. In the model, land in the CBD is used to produce goods and transportation, while suburban land is used to produce housing and transportation. To increase land input for CBD uses, the land must be bid away from suburban uses. Mills assumed that per-worker housing consumption is independent of distance, u, from the CBD, and that a fraction, ρ, of workers living at each u are employed in housing and transportation, and a fraction, $(1\text{-}\rho)$, commute to the CBD for employment.

Because a significant cost of intracity travel is the opportunity cost of the time spent traveling, and travel is slower in more dense areas, a worker currently at u could decrease his or her transportation costs by moving toward the CBD. Equilibrium in housing location would require that no such move provides additional benefits, i.e., a change in transportation cost is just offset by a change in the cost of housing [18].

Wheaton's 1974 paper provided a general comparative statics analysis on two equilibrium models of density and urban land rent, based on Alonso's work [19]. He used a utility function that differed from previous work in that its only requirement is that both goods in the function are normal with positive income effects. As with other previous work, the choice of location is a result of trading off land and travel. Both "closed" and "open" city models are presented by Wheaton. In the closed city case, the size of the population is exogenous while utility, or welfare, is determined within the model. The closed city most closely represents the situation of developed societies where most of the population lives in urban areas. In the case of the open city, utility is exogenous while the population size is endogenous, allowing for migration from rural to urban areas until the benefits of urban living begin to decline. Only the closed city case is discussed herein.

From the closed city case, Wheaton showed that increasing population size expands the city border and reduces utility while higher prices for rural land lower welfare by reducing the city size. These results impact rent and density gradients; effects of higher population or rural competition increase rents and densities for every location. Further, Wheaton showed that increasing marginal travel cost lowers utility, contracts the city boundary, and increases both rents and density at the city center. Finally, for the closed city case, Wheaton found that increasing incomes (per capita) expands the city border and lowers rents and density at the city center, thus flattening the rent and density gradients [19].

To bring the theory up to date, Brueckner presented a unified approach to the Alonso-Muth-Mills urban model. His treatment is based on the work of Alonso, Mills, and Muth, and the major result that differences in commuting (travel) costs within an urban area must be just offset by differences in the price of housing [20]. Alonso examined this observation within a framework where individuals directly consumed land, while Muth and Mills presented a more realistic model where land is an input to the production of housing, a final consumption good.

In Brueckner's intracity analysis, workers commute to their jobs in the CBD along a radial road network. Round-trip commuting cost per mile is represented by t, such that the commuting cost from a residence x radial miles from the CBD is tx per period. The CBD is a point where $x = 0$. All households have the same income, y, per period, and preferences are assumed to be identical. The utility function is denoted by $v(c, q)$, where c is a composite non-housing, non-transportation good and q is the consumption of housing, measured as square feet of floor space. Relevant to the work contained in this document, Brueckner notes that, in reality, housing is characterized by a vector of various attributes, but his analysis focuses instead on the single and important attribute of interior living space. Hedonic price analysis, as it relates to housing prices, is further discussed in Chapter Five (Empirical Methodology) of this document. The price of the composite commodity, c, is assumed to be the same everywhere in the urban area whose price is normalized to unity. Rental price per square foot of floor space varies with location and is represented by p.

In this analysis, because consumers have identical preferences, the urban equilibrium must result in the same utility for everyone. It is the spatial variation in p that provides the equal utilities throughout the urban area (some constant u). The budget constraint is $c + pq = y - tx$. As shown in Brueckner, the requirement that maximized utility must equal u is shown by:

$$\max_q v(y - tx - pq, q) = u \tag{1}$$

Brueckner reduces equation (1) above to two equations [20]. First, because consumers optimally choose q based on p, the first-order condition must hold (subscripts denote partial derivatives with respect to q and c):

$$\frac{v_2(y - tx - pq, q)}{v_1(y - tx - pq, q)} = p \tag{2}$$

Second, the selected consumption bundle must generate utility u:

$$v(y - tx - pq, q) = u \tag{3}$$

Brueckner notes that this determination is the reverse of the usual consumer optimization, because utility is fixed and then a price is determined (not vice versa). How p and q depend on the parameters x, y, t, and u can be derived mathematically by totally differentiating equations (2) and (3). As shown in Brueckner [20], totally differentiating equation (3) with respect to x yields the following:

$$-v_1\left(t + \frac{\partial p}{\partial x} q + p \frac{\partial q}{\partial x}\right) + v_2 \frac{\partial q}{\partial x} = 0 \tag{4}$$

From equation (2), $v_2 = pv_1$, and so equation (4) yields:

$$\frac{\partial p}{\partial x} = \frac{-t}{q} < 0 \tag{5}$$

Equation (5) is a key result that is directly relevant to the work in this dissertation, and shows that the price per square foot of housing is a decreasing function of distance (x) of the residence from the CBD. An additional result is that q is an increasing function of x, meaning that housing consumption increases with distance from the CBD:

$$\frac{\partial q}{\partial x} = \eta \frac{\partial p}{\partial x} > 0, \tag{6}$$

where $\eta < 0$ is the slope of the relevant income-compensated demand curve.

The relationship between changes in distance, x, and the behavior of p and q is intuitive. Those who live farther from the CBD must be compensated in some way for the relatively longer and costlier commutes; otherwise, no one would choose to live at such distances. In this model, compensation is in the form of lower housing prices relative to locations closer to the CBD. With constant utility, as distance, x, increases and the price of housing decreases, consumers substitute toward more housing.

The complete model contains more equations. A comprehensive summary of the comparative static analysis, as shown in Brueckner's closed city model, is presented in Table 2.1, where S is structural density, r is urban land rent, $\overline{x}$ is the city boundary, P is population size, r_A is rural land rent, and x' is the distance where S, r, and x pivot due to changes in income and transportation cost [21]. All other variables are as defined previously.

Table 2.1. Brueckner's Closed City Model: Comparative Statics (1987)

Exogenous Variables	Endogenous Variables					
	q	p	S	r	$\bar{x}$	u
x	+	-	-	-	No change	No change
P	-	+	+	+	+	-
r_A	-	+	+	+	-	-
y $x < x'$	+	-	-	-	+	+
y $x > x'$	?	+	+	+	+	+
t $x < x'$	?	+	+	+	-	-
t $x > x'$	+	-	-	-	-	-

Source: Geshkov and DeSalvo [21].

Another piece of work relevant to this dissertation research is a 1961 paper by Herbert Mohring that investigated land values and the benefits of investments in highways and other transportation facilities. Mohring notes that some of the benefits of transportation investments are those that are believed to accrue to owners of property. He notes some other characteristics of these benefits [22]:

- The benefits are non-user benefits, i.e., they are not related to the extent to which the impacted property owners actually use the facilities involved.
- The benefits are beyond those that accrue to the users of the facilities.
- The benefits are not net benefits, but reflect the transfer of benefits from one group to another.

- There is no guarantee that the net effect of the investment (or improvement) is an increase in land values as a whole.

The above is particularly relevant to the work presented in this document, as the increases in property values that are expected to be realized with access to BRT (thus improving access to the CBD and/or other locations with goods, services, and employment) would accrue to all property owners in the affected area, whether or not they use the BRT system. Further, if the BRT stations increase only those property values in proximity to the stations, holding all else constant, it is not necessarily implied that the property values will increase as a whole.

Mohring assumed a community with four general characteristics:

1. All workers commute to and from the CBD, which is a point.
2. Residences are all single-family and on identical lot sizes
3. Household size is the same throughout the city and households have identical incomes and preferences.
4. Cost of a trip is proportional to the time needed to make it.

In equilibrium, no household could gain by moving. As shown by Mohring, annual rents, R, differ between two properties, i and j, by:

$$R_i - R_j = 2N(T_j - T_i)V_T$$

where T = travel time, V_T = value of travel time, and N = number of trips to the CBD annually. The maximum travel cost is equal to $2NT_{max}V_T$, and rent is equal to zero at the city limit, where T_{max} is travel time to the urban boundary. Thus, the equilibrium annual rent on any property, i, is shown by:

$$R_i = 2N(T_{max} - T_i)V_T$$

As a household moves closer to the CBD, travel time costs decrease and rent increases just enough to offset the change in travel time costs. At the CBD, travel costs would be zero, and rent is equal to $2NT_{max}V_T$.

Finally, Andersson and Samartin extended Mohring's analysis by relaxing some of his assumptions to make the model more realistic. Their assumptions include [23]:

1. There can be several workplaces in a city (workers not just commuting to and from the CBD).
2. Residences can be multi-family as well as single-family, and the size of the home and lot can vary.
3. Household incomes can vary.
4. Commuting costs also depend on the layout of the transportation system, commute mode, congestion levels, and parking costs, etc.

Andersson and Samartin restate Mohring's equilibrium condition as:

$$cc_n + ar_n = cc_m + ar_m$$

where cc = commute costs from locations n and m, and ar = apartment rents at locations n and m. As before, an increase or decrease in cc is just offset by an increase or decrease in ar in equilibrium. While Mohring implied that $cc_n + ar_n$ was equal to a constant, Andersson and Samartin extended the equation to $cc_n + ar_n = ic^*$, where ic^* is a constant for a given income class. Perhaps most important for this dissertation work, Andersson and Samartin indicate that $cc_n + ar_n = ic^*$ can be expanded to include factors such as environmental benefits at location n: $-eb_n + cc_n + ar_n = ic^*$. They note that, in his empirical function, Mohring included the possible benefits of properties located near a lake [23].

Anas, et al., do summarize some criticisms of the monocentric model [15], including that people have different preferences for particular locations or types of locations. For example, some people like to have a large yard to take care of, while others prefer a small yard that requires less work to maintain. Others enjoy or do not mind long drives to work while others prefer shorter commutes; Anas, et al., contend that commuting, overall, is not well-explained by the simple monocentric model (or even polycentric models) [15]. Further, the monocentric model assumes only one worker in each household commutes to the CBD; however, two-worker households may have to compromise on residential location based on the location of the two jobs. Finally, job changes can be frequent and moving can be costly, which further impacts the residential location decision.

This chapter has provided a brief description of the urban economic models and basic theoretical framework that are relevant to this dissertation research. From the key results shown in the above exposition that housing prices are expected to increase closer to the CBD while transportation or commuting costs fall (and vice versa) comes the underlying hypothesis of this work. It is expected that proximity to BRT stations provides access to goods, services, and employment (not necessarily at the CBD), and lowers the transportation costs of traveling to those locations; as such, residential property values are expected to be higher closer to the stations and are likewise expected to fall off as distance from the stations increases. The next chapter provides a summary of the empirical literature on the impacts of transit access on property values.

CHAPTER THREE:

PREVIOUS EMPIRICAL LITERATURE ON THE EFFECT OF TRANSIT ON HOUSING PRICES

This research examines the extent to which access to BRT services in the U.S. are considered in the residential location decision. Indeed, there exists a large amount of qualitative and anecdotal evidence that the implementation of BRT services in the U.S. can lead to economic development and increased land values [24] [25]. As will be discussed in this chapter, there have been recent studies conducted on BRT systems outside the U.S. While these international studies are useful to examine from a methodological perspective, it is argued within this chapter that the international results are not applicable to U.S. experiences with BRT; thus the need for additional research in this area. This work goes beyond qualitative evidence and the international results by attempting to find a positive, statistically-significant impact on property values from proximity to BRT stations in a U.S. city.

Comparing Estimated Impacts from BRT

Until recently, no quantitative modeling studies on the property value impacts of BRT for systems operating in the U.S. were conducted. In 1990, one study analyzed some "busways" (including in Houston and Pittsburgh), but did not find any positive impacts [26]. In that study, the term "busway" referred to an exclusive right-of-way on which any number of regular local or express bus routes operates (not necessarily defined as BRT is known today, as described in Chapter One). Perk and Catalá published a study on the Pittsburgh Martin Luther King, Jr. East Busway in 2009. While several different routes operate on this busway, many of them exhibit the characteristics of modern BRT. The findings from the 2009 study showed that proximity to stations along the East Busway resulted in a positive and statistically significant impact on the assessed values of single-family homes along the East Busway corridor [11]. In 2013, a second study, on the Boston Silver Line Washington Street Corridor was published by Perk, et al. [12]. The focus was changes in actual market prices of nearby condominium units and, again, a positive, statistically significant impact was found on sale prices of condo units. Interestingly, when sales were examined two years prior to the implementation of the Silver Line service, no impact was found on sale prices based on the distance of condo units to the Washington Street corridor, where regular local bus service, stopping every block, had operated previously. In both the 2009 and 2013 studies, hedonic regression analysis was used. The 2013 study in Boston used actual market transactions rather than assessed values, as was used in the 2009 Pittsburgh study. In addition, the 2013 Boston study used the network distance from the property to the nearest station, rather than the straight-line distances used in the Pittsburgh study.

In a 2012 research paper, Nelson, et al., studied whether the EmX BRT system in Eugene, Oregon had attracted new employment using a shift-share analysis technique, and found that some additional jobs, particularly public sector jobs, had located near the BRT corridor [27]. However, no other work on property value impacts has been published for U.S. BRT systems.

Recent studies have been conducted on the BRT systems operating in other countries including in Bogotá, Colombia; Seoul, South Korea; and Sydney, Australia. Because of difficulties accessing data on sales transactions in Bogotá, researchers generally relied upon asking prices instead of actual final prices. In a Bogotá study, Rodriguez and Targa used asking prices for properties and found a premium of 6.8 to 9.3 percent for every 5 minutes of walking time closer to a BRT station [28]. In another Bogotá study, Muñoz-Raskin used asking prices for properties and found that properties within a five-minute walk to the BRT lines were more highly valued than those with a five- to ten-minute walk [29]. In Seoul, Cervero and Kang used assessed values and found premiums of 10 percent for those residences location within 300 meters of BRT stations [30]. In Sydney, Mulley found that prices were mainly determined by the features of the properties and the neighborhood characteristics;

however, small effects were found for decreased access times to the BRT transitway [31].

While this international work is useful to examine for methodological applications, the results may not reflect expectations for impacts in U.S. cities. The different political, cultural, and social environments in South America, Asia, and Australia may render the results of those studies incomparable to the U.S. experience. In Bogotá, the TransMilenio is known globally as one of the largest-scale BRT systems in the world, generating 2.2 million trips per day, many times more than any U.S. BRT system. TransMilenio comprises 70 miles and 114 stations and operates with more than 1,000 buses [32]. Service is very frequent, at between two and five minutes between bus arrivals. In Seoul, South Korea, the BRT system consists of 8 corridors covering 27 miles and serving 400,000 trips per day [32]. In Sydney, a 20-mile alignment with 35 stations provides 10,680 trips per day with buses arriving every 3 to 10 minutes [32]. Cleveland's BRT system, the case study for this work, is more similar to the one in Sydney in terms of operating characteristics and ridership. The Cleveland HealthLine operates along a 7-mile corridor with 40 stations and 5-minute peak headways. The HealthLine system currently generates approximately 15,800 trips per day [32].

Transit usage and attitudes toward transit and other alternative modes of transportation in other countries tend to be different from those in the U.S. The bus/BRT mode share for all trip purposes is a very large 62 percent in Bogotá, 28 percent in Seoul, and 6 percent in Sydney [33]. In Cleveland, transit's mode share for only work trips (including rail transit) is seven percent [34]. These differing transit mode shares represent how transit is used more intensively in other countries than in the U.S., particularly in Bogotá and Seoul. Further, the share of trips taken by private vehicles is quite different among these cities. Just 15 percent of trips in Bogotá are taken by private vehicle, 26 percent in Seoul, and 69 percent in Sydney, compared to 88 percent in Cleveland [32] [33] [34]. Sydney, in particular, has relatively larger percentages of people who travel by bicycling and walking, thus resulting in the lower percentage of trips taken by private vehicle than in Cleveland. Generally in the U.S., transit mode share is typically low, while private auto use is most prevalent. Recent data show that more than 86 percent of workers in the U.S. commute in private vehicles, while only approximately 5 percent of workers nationwide use public transportation [35]. It is because of these different levels of transit and private vehicle usage, as well as the different levels of transit/BRT investment (particularly in Bogotá and Seoul), that the impacts of BRT on housing prices may be very different in those places compared to the U.S. It is for these reasons that this additional research on U.S. BRT experiences is needed.

Comparing Estimated Impacts from Rail Modes

In the U.S., research on the impacts of transit on property or land values has focused on rail modes of transit, which might be expected to have a larger impact than

BRT. These studies generally attempted to isolate the effect of distance from rail transit (the rail stations, the rail right-of-way, or both) on property or land values. Most of the studies did find positive impacts on property values from proximity to rail transit; typically, however, the magnitudes are relatively small. A relatively small marginal impact from access to transit would likely be expected considering the myriad factors that influence property and land prices.

When looking at these studies, the relevant research generally acknowledges that two sets of factors can impact property values. The first is a set that can lead to increases in property values (amenities) and the second is a set that can negatively impact those values (disamenities, or nuisance effects). If the transit service is perceived well, surrounding property values could be positively affected. In addition, those who find the service to be comfortable, reliable, and have favorable travel times may be willing to pay more for housing within walking distance. Conversely, locations in very close proximity to rail lines or other fixed guideways may also introduce nuisance effects (real or perceived) such as noise, pollution, and crime, which could negatively impact discourage people from living there.

Most of the studies use hedonic regression techniques, though functional forms vary. The typical variables include various property and neighborhood characteristics, although there is variation in the use of these variables among the studies due to the nature and availability of data in different cities. Data from the U.S. Census are standardized, and most city or county property assessor's offices provide similar data on property characteristics. Distance variables can be readily computed using GIS software. However, some information, such as crime data, can be collected and reported quite differently among areas and is not always reported at a fine enough geographic level to be useful in this type of analysis [11].

Tables 3.1 and 3.2 summarize a selection of research on this topic. It should be noted that some of the information in Tables 3.1 and 3.2 was taken from a comprehensive literature review found in Perk, et al. [12]. Specifically, Table 3.1 includes studies of light rail transit's impacts on residential property values, while the studies listed in Table 3.2 focus on commercial property values. Lastly, Table 3.3 provides the brief results of a set of studies that were not individually reviewed for this effort but were included in the Transit Cooperative Research Program (TCRP) Report 118: Bus Rapid Transit Practitioner's Guide [36].

The results from TCRP Report 118 were intended to provide a sense of expected impacts from transit on property values for those who were planning BRT systems in their communities. At the time of its publication in 2007, only a few BRT systems were in operation in the U.S. The current level of research on this topic for U.S. BRT systems continues to be scant. Given that many more cities are exploring transit investments, often choosing between bus and rail modes, additional research on this topic for U.S. BRT systems is needed.

Table 3.1. Summary of Literature Estimating Impacts of LRT on Residential Property Values

Study Authors and Year	Study Information	Key Findings
Gatzlaff and Smith, 1993 [37]	Miami-Dade County Property Tax Records data on sales for a pooled sample of properties surrounding Miami Metrorail stations.	No significant change in sales index of homes before and after establishing Metrorail. Overall, weak evidence of positive residential property impacts, with high-income households accruing greater net benefits than low-income households.
Gruen, Gruen and Associates, 1997 [38]	Data on sales price of single-family homes, structural data, social data, station and transportation access data for Chicago Transit Authority.	Home prices decrease as distance from a rail station increases, for both low and high income neighborhoods.
Chen, et al., 1998 [39]	Prices of single-family homes sold from 1992 to 1994 in Portland, Oregon.	As distance to a MAX light rail station increases, housing price decreases, but at a decreasing rate.
Baum-Snow and Kahn, 2000 [40]	1980 and 1990 U.S. Census tract-level data for rail transit in Boston, Atlanta, Chicago, Portland (OR), and Washington, D.C.	Decreasing transit distance from 3 to 1 km (9843 to 3281 ft) increased monthly rents by $33 and home values by $8,557 (2011 $)
Bowes and Ihlanfeldt, 2001 [41]	Atlanta sales of single-family homes and crime density of the census tract from 1991 to 1994.	Proximity to MARTA rail stations has a positive effect on the value of single-family homes.
Garrett, 2004 [42]	1,516 single-family homes in St. Louis County (Missouri) within one mile of a Metrolink light rail station, sold from 1998-2001.	Home values increase an average of $185.63 (2011 $) for every 10 feet closer to a station, starting at 1,460 feet. The "nuisance" effect associated with the Metrolink is weak.
Hess and Almeida, 2007 [43]	City of Buffalo 2002 assessed value of single-family properties, 1990 and 2000 U.S. Census.	A property increases $1.24-2.89 (2011 $) for every foot closer to a light rail station.
Kent and Parilla, 2008 [44]	Used a repeat-sales approach but with assessed market values of single-family homes for two time periods, 1997-2000 and 2003-2006, representing before and after the Hiawatha line opened in Minneapolis.	Within a half-mile buffer of the stations, it was found that proximity to the stations resulted in a $18,723 (2011 $) increase in assessed values.
Yan, Delmelle, and Duncan, 2012 [45]	Applied hedonic regression using single-family home sale prices in Charlotte, NC to four time periods: pre-planning (1997-1998), planning (1999-2005), construction (2005-2007), and operation (2007-2008).	Using a one-mile buffer around stations, a positive relationship between distance and sale price was found in all four time periods. However, the effect was smallest in the operation period, suggesting that the light rail system was beginning to influence sale prices.

Note: Dollars adjusted to 2011 dollars using the Consumer Price Index for all urban consumers (CPI-U), U.S. city average, housing index, from the Bureau of Labor Statistics (BLS.gov).

Table 3.2. Summary of Literature Estimating Impacts of Rail Transit on Commercial Property Values

Study Authors and Year	Study Information	Key Findings
Cervero and Landis, 1995 [46]	On-line database of property tax records (TRW-REDI) and U.S. Census data for population and employment statistics.	No major commercial price or rent premiums associated with proximity to San Francisco BART rail stations.
Cervero 1994 [47]	Pooled data for five rail station areas, with large commercial development from 1978 to 1989 in Washington, D.C. and Atlanta.	Overall, the evidence supports a measurable land value benefit from rail transit investments and joint development projects. Vacancy rates are 11% lower in station areas with joint development projects.
Weinberger 2001 [48]	Santa Clara County lease transactions from 1984 to 2000 collected from a large brokerage firm.	Rental premium exists on office properties located within one half-mile of light rail stations.
Cervero and Duncan, 2002b [49]	1998 and 1999 Santa Clara County commercial property data.	Being near rail transit increased commercial land values. Land parcels within a quarter mile of a rail station in a business district were worth $33.75 (2011 $) per square foot more than comparable properties away from stations.
Cervero and Duncan, 2002a [50]	San Diego County sale prices from Metroscan database (maintained by First American Real Estate Solutions), 2000 U.S. Census, GIS.	Greatest amenity and disamenity factors for commercial properties, claim rents to be an inaccurate way to measure benefits.

Note: Dollars adjusted to 2011 dollars using the Consumer Price Index for all urban consumers (CPI-U), U.S. city average, housing index, from the Bureau of Labor Statistics (BLS.gov).

A majority of the studies reviewed in this chapter did find positive, yet relatively small, effects of transit on property values and of some factors related to economic development (such as employment, etc.). While nearly all of these studies analyzed impacts of rail transit on property values, they still provide a useful framework of reference for the research in this dissertation which focuses on the impacts of a bus mode, BRT, on property values.

Table 3.3. Summary of Other Literature Estimating Impacts of LRT on Property Values

Study Authors and Year	Study Information	Key Findings
Dueker and Bianco, 1999	Population Census' median house value in Portland, Oregon between 1980 and 1990.	Premium of $4,720 (2011$) for properties within 0.06 km (197 ft) of a MAX station.
Lewis-Workmann and Brod, 1997	Cadastral information for all properties (4,170) within 1.7 km (5577.43 ft) of three MAX stations in Portland, Oregon.	Premium of $107.52 (2011 $) per 0.03 km (98.43 ft) closer to a station.
Forrest et al., 1995	795 house sales in Manchester (UK) during 1990.	Premium ranging from 2.1- 8.1% depending on distance from a station.
Landis et al., 1995	134 single-family sales in San Diego during 1990.	Premium of $468 (2011 $) for every 0.1 km (328 ft) closer to a station.
Dabinett, 1998	Sheffield (UK) Supertram.	No evidence of any appreciable effects.
Al-Mosaind et al., 1993	235 single-family home sales in Portland, Oregon during 1988.	Premium of $1,261 (2011 $) per 0.03 km (98.43 ft) closer to a station.

Source: Transit Cooperative Research Program (TCRP) Report 118 [36] *Note: Dollars adjusted to 2011 dollars using the Consumer Price Index for all urban consumers (CPI-U), U.S. city average, housing index, from the Bureau of Labor Statistics (BLS.gov).*

CHAPTER FOUR:

CASE STUDY OF THE CLEVELAND HEALTHLINE BRT

As discussed previously, only two studies have been published to date on the property value impacts of access to BRT stations in U.S. cities [11] [12]. The two studies were conducted on Pittsburgh's Martin Luther King, Jr. East Busway and Boston's Silver Line Washington Street Corridor. This dissertation research will add a third U.S. study on this topic, focusing on the HealthLine BRT system, which began operating in Cleveland, Ohio at the end of October 2008. This chapter describes the case study site in Cleveland, located in Cuyahoga County, and the data used in the analysis.

History and Design of the Cleveland HealthLine

In the late 1800s, Cleveland's Euclid Avenue was known as Millionaire's Row, a place where wealthy individuals such as John D. Rockefeller had built several ornate mansions. However, by the 1960s, the area had experienced significant decline and was characterized by deteriorated housing, abandoned buildings, and empty lots [51]. As early as the 1950s, the "Dual Hub" transit project began, which aimed to construct a subway line between Downtown Cleveland and the University Circle area to the east. The Dual Hub project never gained traction and, by the 1980s, the Euclid Corridor

Transportation Project envisioned LRT service along this corridor between the region's two largest employment centers. By 1995, what eventually evolved into the BRT project operating today was selected as the locally preferred transit alternative, with preliminary engineering beginning in 1997. Originally dubbed the Euclid Corridor Silver Line, ground was broken for construction in 2004. The name was eventually changed to the HealthLine after the Cleveland Clinic and University Hospitals purchased the naming rights [52].

The project involved more than the construction of the basic transit infrastructure. Essentially, a 6.8-mile stretch of Euclid Avenue was demolished and completely rebuilt, with upgraded utilities, a new road surface, new curbs and sidewalks, bike lanes, and a host of aesthetically-pleasing amenities such as planters, lighting, landscaping, and public art. Figure 4.1 shows periods of the construction phase, while Figure 4.2 shows the completed project.

When it opened in October 2008, the HealthLine replaced the regular local Route 6, which had been the most heavily traveled route in the Greater Cleveland Regional Transit Authority's (GCRTA) bus system. However, ridership on the new HealthLine increased 60 percent over the Route 6 ridership, with a significant number of riders being new to transit [52]. The service is 7.1 miles in length, with nearly 16,800 daily passengers [52]. Characteristics of the service include an exclusive median guideway for most of the alignment, signal priority at intersections, off-board fare collection similar to rail systems, level boarding on elevated platforms (another rail-like characteristic), real-time passenger information, and a high frequency of five minutes in the peak periods. Figure 4.3 illustrates the HealthLine alignment along Euclid Avenue, which extends east from the CBD.

Photo Credits: Victoria A. Perk (see Appendix C)

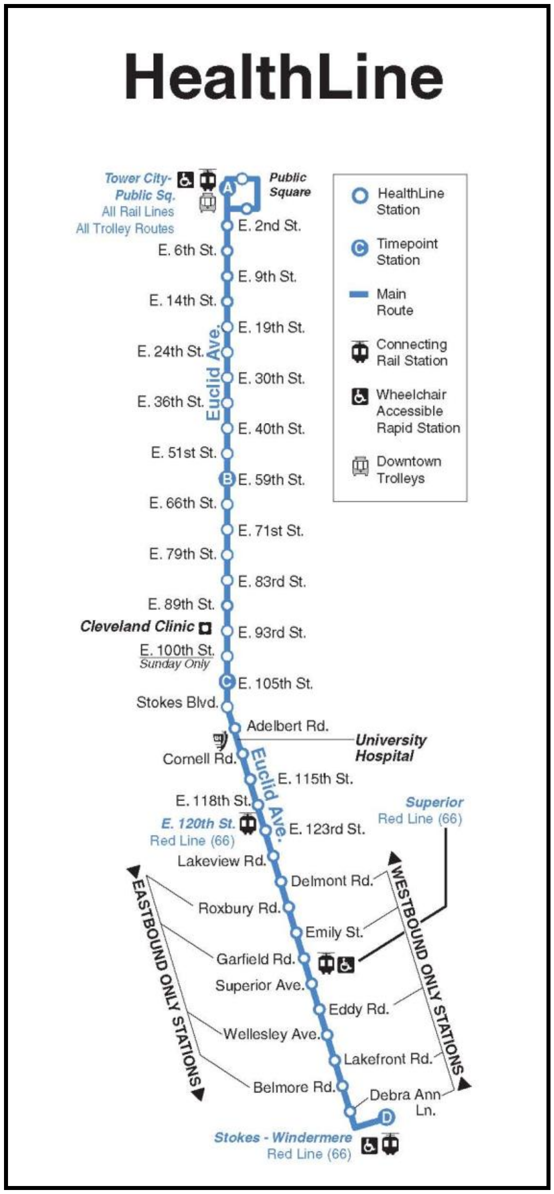
Figure 4.1. Cleveland HealthLine Construction along Euclid Avenue, 2007

Data Used in the HealthLine Analysis

The College of Urban Affairs at Cleveland State University (CSU) has tracked housing data in Cuyahoga County for many years. A substantial amount of data was available for this effort. The data are available county-wide with sales transactions

Photo Credits: Victoria A. Perk (see Appendix C)

Figure 4.2. Completed Cleveland HealthLine Euclid Avenue Corridor



Source: Greater Cleveland Regional Transit Authority (GCRTA).
Used with permission from Stephen Bitto, GCRTA (sbitto@gcrta.org).
(See Appendix C)

Figure 4.3. Cleveland HealthLine Stations

going back to the 1970s through 2011, with all the typical physical property characteristics that would be expected.

From the original data received from CSU, a data set of single-family homes was created with sales transactions going back to 2000. Several buffers were constructed around the corridor at two miles, one-and-a-half miles, one mile, and one-half mile. Data from the U.S. Census is also used for the analysis. The full data set contains 12,015 single-family homes within two miles of the Cleveland HealthLine BRT corridor, although in the final models data were only used up to one-and-a-half miles from the corridor (7,457 single-family homes).

Table 4.1 shows the distribution of sale prices from 2000 to 2011 for single-family homes for the full data set within two miles of the Euclid corridor, adjusted to constant 2011 dollars. The impact of the housing crisis and recession that began in 2007 is evident in Table 4.1, particularly in the category of homes selling for less than $20,000. Beginning in 2007, the number of homes selling for less than $20,000 increased from approximately 16 percent to nearly 32 percent, further climbing to nearly 61 percent in 2008. By 2011, the number of homes selling for less than $20,000 was approximately 55 percent. As discussed in later chapters of this dissertation, the economic conditions in the area during this time affected the results of this work even though homes selling for very low prices were excluded from the analysis (homes selling for less than $30,000 were not included in the final data sets).

Geographic Information Systems (GIS) software was used to compute the network distances from each property to the nearest BRT station which represents the

Table 4.1. Distribution of Sale Prices of Single-Family Homes (2011 $)

	< $20,000	$20,000 -- $49,999	$50,000 -- $74,999	$75,000 -- $99,999	$100,000 -- $199,999	$200,000 -- $499,999	>= $500,000
2000	12.4%	27.8%	16.3%	20.5%	16.9%	5.7%	0.3%
2001	14.1%	25.2%	11.5%	21.1%	23.3%	4.4%	0.3%
2002	12.2%	28.7%	10.7%	15.3%	29.7%	3.3%	0.0%
2003	13.2%	23.6%	6.6%	17.3%	31.4%	7.6%	0.2%
2004	13.3%	26.1%	6.6%	22.4%	23.4%	7.5%	0.8%
2005	9.4%	26.7%	8.8%	22.4%	25.4%	6.6%	0.8%
2006	16.4%	28.8%	8.3%	20.3%	21.1%	4.8%	0.4%
2007	31.8%	36.7%	8.2%	9.8%	9.0%	4.5%	0.0%
2008	60.6%	16.4%	5.3%	7.6%	7.5%	2.1%	0.5%
2009	58.8%	18.2%	5.2%	6.8%	7.4%	3.7%	0.0%
2010	56.7%	15.3%	5.3%	3.0%	7.0%	1.0%	0.0%
2011	55.3%	23.0%	5.3%	4.9%	8.8%	2.7%	0.0%

Note: Dollars adjusted to 2011 dollars using the Consumer Price Index for all urban consumers (CPI-U), U.S. city average, housing index, from the Bureau of Labor Statistics (BLS.gov).

key independent variable of interest in the research (DIST). The network distance measures the route along the street network from the property to the nearest transit station. This method produces an estimation of actual walking (or, perhaps, biking) distance. Network distances more accurately represent actual walking or biking distance than Euclidean, or straight-line, distances because the network distances account for the fact that people must generally stay on roads or adjacent sidewalks and cannot travel through other homes, buildings, other private property, or other geographic barriers to get to their destinations. In Tables 4.2, 4.3, 4.4, and 4.5, the distances are shown in miles; however, for the regression models, the miles were converted to feet in anticipation of relatively small marginal effects. Because the theory predicts non-linearity (i.e., the price function should decrease at a decreasing numerical rate with distance from the BRT station), a variable for the squared distance (DISTSQ) is included to control for possible non-linearities in the effect of distance on sale price. If the effect of distance on sale price is linear, each foot of distance from the parcel would be estimated to change the property value by the same marginal amount. However, if the effect is non-linear, the marginal effect can vary at an increasing or decreasing rate. Following the hypothesis of this research, the anticipated sign of the distance coefficient is negative; i.e., holding all other factors constant, the sale price of a property decreases as distance from the nearest BRT station increases. It should be noted that network distances have only been computed for the full data set including properties within two miles of the BRT corridor; however, for the models only data up to one-and-a-half miles were used (and only up to one mile for some of the models). However, two different specifications were estimated for each year; one in which the individual network distance for each property is used and one in which dummy variables are used to specify ranges of network distance (e.g., within one-quarter mile, between one-quarter and one-half-mile, between half-mile and one mile, etc.). This latter approach will provide an alternate measure of the effect of distance on sale prices.

Table 4.2 illustrates the distribution of network distances from the properties to the nearest BRT station for the homes sold in years 2004, 2008–2009 combined, 2010–2011 combined, and all single-family homes in the data set. Home sales for 2008–2009 and 2010–2011 were combined because of the relatively low number of homes sold in each of those individual years. In the few cases where a home sold in both 2008 and 2009, or 2010 and 2011, the latter year was used (i.e., 2009 or 2011). The Chi-Square Goodness of Fit test was used to compare the distributions of properties sold in each cross-section group with the total stock of all single-family homes within two miles of the corridor. To reduce the incidence of sample selection bias, the distributions of homes sold in each cross-section group should be similar to the distribution of all homes in the study area. In this case, it is desirable to accept the null hypothesis that the distributions are equal ($p > 0.05$). Therefore, in Table 4.2, the Chi-Square Goodness of Fit statistics are shown, along with the relevant p values. When the p value is greater than 0.05, it can be said that the distribution of single-family homes sold in a particular cross-section year is very similar to the distribution of all single-family homes in the study area (at the five percent level of significance). While there are some differences among the distributions, it appears that the distribution of single-family homes within two miles of the BRT corridor sold in each cross-section year is relatively similar to the distribution for the stock of all single-family homes within two miles of the corridor.

Table 4.3 shows the mean sale price, in nominal terms, within each range of distance ultimately used in the models. In 2004 and 2008–2009, all single-family homes within one mile were used. In 2010–2011, due to the low number of homes sold, data within one-and-a-half miles were used. The data in Table 4.3 exclude foreclosures and other very low sale prices (including “love and affection” transfers between family members where homes may have a price of as little as $1.00).

Property characteristics in the data set include:

- Square feet of lot size
- Square feet of living area (total usable area)
- Condition of the property (likert scale ranging from Very Poor to Excellent)
- Age of the structure
- Style of the residential structure (colonial, bungalow, townhouse, etc.)
- Number of bedrooms
- Number of bathrooms, number of half-bathrooms

Table 4.2. Distribution of Network Distances from Properties to Nearest BRT Station

Distance Category	**Sold in 2004**	**Sold in 2008–2009**	**Sold in 2010–2011**	**All Homes**
Less than 0.25 mile	2.6%	2.2%	1.8%	2.0%
0.25 – 0.499 mile	9.8%	7.4%	6.9%	8.3%
0.50 – 0.749 mile	14.2%	10.7%	8.8%	12.3%
0.75 – 0.999 mile	11.8%	12.0%	9.7%	11.5%
1.00 – 1.499 miles	24.4%	27.7%	28.8%	28.0%
1.50 – 2.00 miles	37.2%	40.0%	44.0%	38.0%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	9.221 $p = 0.101$	5.788 $p = 0.327$	21.941 $p = 0.001$	

Table 4.3. **Mean Sale Price by Distance Category (Nominal Values)**

Distance Category	**Sold in 2004**	**Sold in 2008–2009**	**Sold in 2010–2011**
Less than 0.5 mile	$131,822	$149,469	$102,727
0.5 – 0.99 mile	$106,653	$87,552	$75,555
1.00 – 1.5 miles	n/a	n/a	$101,510

The square footage of the lot area (LOTSIZE) and the square footage of the available living area (LIVAREA) are commonly included in hedonic housing price regression analysis. Additional common variables include those indicating the number of bedrooms (BEDRMS), bathrooms (BATHRMS), and half-bathrooms (HALFBTH) within a home. A variable interacting the number of bedrooms and square feet of living area (BED*LIVA) is used to allow the effect from living area to vary with the number of bedrooms, although it is not used in the final models. A likert-scale variable indicating the condition of the home (ranging from Very Poor to Excellent) as assessed by Cuyahoga County (COND) and a variable for the year the property was built (YRBUILT) are used to further describe the homes. Finally, a variable indicating the style of the single-family home is included. Appendix A presents the distributions of these variables along with the Chi-Square Goodness of Fit results comparing the distributions with the overall housing stock.

To describe additional characteristics of the communities along the BRT corridor, certain data available from the American Community Survey's (ACS) five-year estimates were provided at the block-group level were considered for this analysis (see Appendix A). These include:

- Median household incomes (MDHHINC)
- Percent of owner-occupied housing units (OWNOCC)
- Proportion of housing units built before 1940 (B1940)

If the BRT service provides favorable travel times to destinations of interest, sale prices of homes with proximity to the stations may be positively affected. It is not feasible to compute transit travel times from each parcel to various destinations accessible via the BRT service (which would include walking time, waiting time, and in-vehicle travel time). This is further discussed in Chapter Five (Empirical Methodology). However, for the purposes of this research, in-vehicle travel times have been calculated from each station to two major stations of interest: Tower City (in Public Square, which represents the CBD) and the main access point for the Cleveland Clinic campus, which is a major area employer and medical facility. The travel times, in minutes, were computed based on the schedules for the HealthLine weekday peak service. The weekday peak schedules have remained stable since the implementation of the service, with very frequent five-minute peak headways (i.e., time between buses).

End-to-end travel time along the entire HealthLine corridor is approximately 43 minutes, which represents the maximum travel time along the corridor from the eastern terminus station (Windermere) to the western terminus station (Tower City). The sprawling campus of the Cleveland Clinic Foundation spans three or even four HealthLine stations. However, the station that provides access to the main entrance is at East 93rd Street, located near the center of the HealthLine corridor. As such, the maximum distance from any HealthLine station to the East 93rd Street Station is approximately 24 minutes.

These two station locations, Tower City and the Cleveland Clinic, were chosen because it is likely that they attract a significant amount of trips. Both are major employment centers and also attract other types of trips besides the work commute. For Tower City, there are shopping and dining opportunities as well as connections to the rest of the GCRTA system, including bus and rail modes. The Cleveland Clinic also attracts a large number of medical trips. Appendix A includes the distributions for these two travel times for the relevant years in the data set.

Given that neighborhoods or other geographically-defined areas often have varying characteristics that may influence real estate prices, dummy variables are included to control for the location of the homes in the data set. For each location or neighborhood, the dummy variable takes the value of one (1) if the property is within the particular location or neighborhood and takes the value of zero (0) otherwise:

- Dummy variables to indicate particular neighborhoods
- Dummy variables to indicate Cleveland city wards
- Dummy variables to indicate the location of specific areas along the corridor

As indicated above, there are three ways available to control for the location of the properties: defined neighborhoods, city wards, and other specific divided areas along the corridor. GIS applications were used to place the observations into the appropriate categories for these three variables. In specifying the models, close attention is paid to these three ways of classifying the location of the properties; the neighborhood dummy variables performed best in the models and thus were the only set of location dummy variables used in the final models. First, there are several official bounded "neighborhoods" in the City of Cleveland, each with varying characteristics and differences that may be reflected in the sale prices of homes within them (NBRHD). Cleveland is also divided into wards, which are political boundaries that often span neighborhoods (WARD). Finally, the corridor itself is divided into a few distinct areas based on the characteristics and activities in those areas (BOUND). Referring back to Figure 4.3, the eastern part of the corridor, between Public Square and E. 24th Street, is considered to be a part of the CBD. Continuing east to E. 89th Street is the Midtown section, which has been blighted in the past but is undergoing redevelopment with housing and restaurant activity. The area between E. 89th Street and E. 123rd Street is known as the University Circle area, containing Case Western Reserve University, Severance Hall (symphony hall), museums of art and natural history, and the large campus of the Cleveland Clinic, a globally-recognized hospital and medical center. Beyond E. 123rd Street are Cleveland Heights and the City of East Cleveland. In addition to these four sections, there are areas to the north and south of Midtown and the University Circle area that include residential properties. Tables are available in Appendix A that provide data frequencies for these three variables for homes that sold in 2004, 2008–2009, and 2010–2011, as well as all single-family homes in the data set located within two miles of the BRT corridor.

Tables 4.4 and 4.5 provide some additional data descriptives for the variables available in the data set. Specifically, Tables 4.4 and 4.5 provide the minimum, maximum, mean, and standard deviation for the continuous variables for single-family homes sold in 2004, 2008–2009, and 2010–2011, as well as all single-family homes within either one mile (for 2004 and 2008–2009) or one-and-a-half miles (for 2010–20011) of the BRT corridor. The sold homes included in Tables 4.4 and 4.5 exclude foreclosures and other very low priced transactions. The next chapter, Chapter Five, introduces the empirical methodology for the dissertation research.

Table 4.4. Additional Data Descriptives for Homes Sold in 2004, 2008–2009, and All Homes

Variable Name	Description	Sold in 2004 (n=192)				Sold in 2008–2009 (n=127)				All Single-Family Homes (n=4,096)			
		Min.	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
PRICE (2011 $)	Sale Price of Home in Dollars	$40,028	$352,165	$133,064.07	$75,129.64	$35,145	$398,354	$114,689.00	$78,653.31	--	--	--	--
LOTSIZE	Sq. ft. of lot	1,285	57,408	5,585.06	4,728.68	1,568	36,547	5,931.36	4,555.58	1,080	62,441	5,161.62	3,467.39
LIVAREA	Sq. ft. living area	780	5,109	1,687.12	610.39	641	4,941	1,823.02	721.85	588	6,933	1,602.82	548.17
YRBUILT	Year built	1870	2004	1944.27	48.21	1870	2009	1952.20	48.41	1850	2009	1931.75	42.65
BEDRMS	# of bedrooms	2	9	3.48	1.09	1	7	3.62	1.01	1	9	3.49	0.96
BATHRMS	# full bathrooms	1	3	1.42	0.57	1	4	1.60	0.65	1	6	1.35	0.57
HALFBTH	# of half-baths	0	3	0.43	0.56	0	2	0.53	0.56	0	10	0.28	0.50
DIST	Distance (miles) to nearest BRT station	0.08	1.00	0.60	0.24	0.06	1.00	0.58	0.27	0.04	1.00	0.63	0.23
MDHHINC (2011 $)	Median household income for census block	$5,200	$40,714	$23,318.93	$9,057.50	$7,530	$40,375	$23,846.45	$8,950.16	$5,200	$46,771	$23,376.22	$8,681.61
OWNOCC	Percent of owner-occupied homes in census block	0%	73.33%	41.28%	18.68%	0%	71.43%	42.09%	18.50%	0	73.33%	41.90%	17.97%
B1940	Percent of homes in census block built before 1940	11.84%	82.24%	53.05%	18.48%	9.81%	83.89%	54.50%	18.35%	9.81%	83.89%	56.57%	18.57%
CBDTT	Travel time (min) from nearest BRT station to Public Square (CBD)	n/a	n/a	n/a	n/a	11.10	41.85	24.82	8.05	0	41.85	25.34	8.20
CCTT	Travel time (min) from nearest BRT station to Cleveland Clinic	n/a	n/a	n/a	n/a	1.10	17.90	6.86	4.26	0	23.95	6.86	4.71

Table 4.5. Additional Data Descriptives for Homes Sold in 2010–2011, and All Homes

Variable Name	Description	Sold in 2010–2011 (n=140)				All Single-Family Homes (n=7,457)			
		Min.	Max.	Mean	Std. Dev.	Min.	Max.	Mean	Std. Dev.
PRICE (2011 $)	Sale Price of Home in Dollars	$30,300	$412,588	$93,665.63	$68,873.59	--	--	--	--
LOTSIZE	Sq. ft. of lot	1,500	34,325	5,450.33	3,857.80	1,080	62,441	4,828.22	2,982.17
LIVAREA	Sq. ft. living area	887	5,580	1,775.17	751.65	480	6,933	1,545.71	483.20
YRBUILT	Year built	1870	2009	1940.89	46.11	1824	2009	1927.78	39.30
BEDRMS	# of bedrooms	2	7	3.46	1.0	1	9	3.45	0.93
BATHRMS	# full bathrooms	1	5	1.54	0.71	1	6	1.27	0.52
HALFBTH	# of half-baths	0	2	0.41	0.56	0	10	0.24	0.47
DIST	Distance (miles) to nearest BRT station	0.07	1.49	0.87	0.43	0.05	1.50	0.92	0.37
MDHHINC (2011 $)	Median household income for census block	$8,367	$93,292	$26,849.26	$15,287.74	$8,367	$93,292	$22,959.36	$9,070.81
OWNOCC	Percent of owner-occupied homes in census block	0	74.21%	40.78%	18.60%	0	87.06%	43.02%	17.46%
B1940	Percent of homes in census block built before 1940	9.81%	88.52%	58.51%	19.64%	9.81%	88.52%	59.62%	18.13%
CBDTT	Travel time (min) from nearest BRT station to Public Square (CBD)	0	43.45	23.37	11.13	0	43.45	26.06	9.36
CCTT	Travel time (min) from nearest BRT station to Cleveland Clinic	0	23.95	8.67	6.99	0	23.95	7.65	5.85

CHAPTER FIVE:

EMPIRICAL METHODOLOGY

This research applies hedonic regression analysis to estimate impacts of access to BRT stations on residential properties surrounding the Cleveland HealthLine BRT system. A brief discussion of hedonic price analysis and its theoretical basis is appropriate here. Housing is the largest asset in the country and the importance of housing prices within the overall economy and the importance of estimating them correctly must be emphasized [53]. Each house is unique – housing can be thought of as an extremely differentiated product! Hedonic methods allow house prices to be expressed as a vector of housing characteristics, and distinctions can be made between physical and locational characteristics. A few examples of physical characteristics are the number of bedrooms, lot size, number of stories, etc., while locational characteristics can include the exact location of a house within a neighborhood or school district, and distance to amenities such as a body of water, shopping areas, parks, or public transit, among others. Some characteristics may be considered undesirable and could be expected to have a negative effect on housing prices, such as a house in poor condition (physical characteristics), or a house located in a high-crime area, near an industrial area, or an interstate right-of-way, for example (locational characteristics). However, the first applications of these hedonic methods were not in housing markets.

History of the Hedonic Method

According to Hill, the hedonic approach dates back to "at least" 1928, but Goodman dates the first work and the coining of the term "hedonic" back to Andrew Court in a 1939 paper [53] [54]. Goodman's history of hedonic analysis notes that, after Court's work, very little similar work was done until nearly 1960, when Zvi Griliches used his first hedonic regression in a 1958 article on the demand for fertilizer. After that, the method gained more attention and became more widely used. It is likely that the gap from 1939 to 1960 was due to a lack of modern computing technology and the fact that, during that time, econometric methods were most often used with aggregate macroeconomic data. It is also interesting to note that this early work by Court and Griliches was not found in typical economics publications.

Goodman describes Court's work from 1939 (and also replicates and extends the analysis). Andrew Court was an economist for the Automobile Manufacturer's Association and later General Motors. Court had compiled several spreadsheets of auto model data and felt that the methods commonly used for constructing price indexes at the time were "wanting" and too simple [54]. He discussed the weighting of the relative importance of different product attributes in constructing a price index. From the idea of utilitarianism or, "the greatest happiness," he used the term "hedonic" believing that such analysis recognizes the "potential contribution" of any commodity to the "welfare and happiness of its purchasers and the community" [54]. Court's equation for a three-period model of auto prices is as follows:

$$p = a + b_g g + b_f f + b_h h + b_1 t_1 + b_2 t_2$$

where a = constant
p = auto price
g = dry weight
f = wheelbase
h = advertised horsepower
t_1, t_2 = time period variables

The equation above resembles the typical basic hedonic equation used today. Court did acknowledge that other factors influenced the price of automobiles, but chose to focus on the three major variables above.

Regarding their use in housing markets today, Hill outlines six major ways in which hedonic methods are used and also provides references to relevant literature [53]:

1. Construction of quality-adjusted price indexes (for housing or any differentiated product).
2. Provision of general appraisals of properties.
3. Explanation of variations in housing prices and a determination of the impact of various characteristics on housing prices.
4. Two-stage demand studies for non-market services.
5. Testing for market segmentation.
6. Evaluation of the effectiveness of government policies.

Early applications to housing generally addressed the third item in the list above. Hill noted a study by Ridker and Henning (1967) which may have been the first done for housing, and which focused on air pollution [53]. Work in the field began to increase significantly throughout the 1970s. The work completed for this dissertation also addresses the third item in the list above, with a goal of estimating the impacts on housing prices from the locational characteristic of distance to a BRT station.

Certainly location is a very important determination of housing prices, and so it follows that such prices are spatially dependent. With advances in geographic information systems (GIS), it has become easier to use geospatial data, particularly with exact latitude and longitude information. One way to address spatial dependence is to incorporate dummy variables for properties within certain neighborhoods, districts, or other relevant areas. Computing the distance to amenities (or disamenities) is another way to account for spatial dependence, and is important in this research. It should also be noted that the impact on prices of distance to an amenity may be nonmonotonic, i.e., some may want to live "not too close" to an amenity yet "not too far" [53]. Regarding this research on the impacts of distance to BRT stations, this may be quite relevant. For example, people may wish to live within relatively close walking distance to a station, but may not want to live right next door to such a station.

For the work herein, it is important to consider the possible effects of spatial autocorrelation and spatial heterogeneity in regression analysis. If, as is very likely, the sale price of a home is dependent on the sale prices of other homes in the area, then spatial correlation exists, which violates an assumption of the classical linear regression model. Specifically, it violates the assumption that the disturbance terms are not correlated. As a result, ordinary least squares estimators (OLS) are unbiased, but they are no longer efficient (i.e., minimum variance); therefore, the results of the typical *t* and *F* tests of significance will not be valid. The use of robust standard errors is one way to correct OLS estimators in the presence of autocorrelation, as well as heteroskedasticity. However, spatial regression models can incorporate spatial dependence either by using a spatially-lagged dependent variable as an additional regressor, or in the structure of the error term [55]. The former, referred to as a spatial-lag model, is used when the nature of the spatial dependence is considered "substantive," and the researcher is interested primarily in the strength of spatial interaction. The latter, known as a spatial error model, is used when the nature of the spatial dependence is a "nuisance" and the researcher is mostly interested in correcting for the influence of the spatial autocorrelation. This research does not apply any spatial regression techniques; however, such techniques may be used in subsequent research based on the findings in this dissertation.

A researcher applying a hedonic method must consider many factors while building the model and has several decisions to make. Hill summarizes these decisions well, and they include [53]:

- Selection of an appropriate functional form.
- Selection of explanatory variables, including any interactions or variable transformations.

- Inclusion of discrete variables as standard variables or dummy variables (e.g., number of bedrooms).
- Expectation of the signs of the coefficients.
- Inclusion of any locational distances.
- Determination of what to do with outliers.

Regarding the chosen explanatory variables, Hill also lists the nine physical attributes that appear most often in hedonic housing price regression analyses [53]:

1. Floor area
2. Land area
3. Age of structure
4. Number of bedrooms
5. Number of bathrooms
6. Garage
7. Swimming pool
8. Fireplace
9. Air-conditioning

Hedonic analysis is not without some drawbacks. Again, turning to Hill, he briefly describes some criticisms [53]. First, one major concern in such analyses is the problem of omitted variable bias. At times, the researcher may be constrained by data availability, and/or have difficulty quantifying other variables that may be important. Another issue is that the theory alone does not provide insight into the appropriate functional form, and so functional form misspecification can be a problem. Sample selection bias may also be an issue, particularly if the population of sales transactions is not representative of the entire population of housing. Finally, data mining and/or lack of reproducibility can affect the validity and robustness of hedonic analysis. For example, Hill notes that any two researchers given the same data set can likely generate two very different results.

Theoretical Basis for Hedonic Regression

This section includes a brief discussion of the theoretical basis for hedonic price regression, as outlined by Rosen. As mentioned previously, housing is an example of a highly differentiated product, and Rosen's analysis presents a "model of product differentiation based on the hypothesis that goods are valued for their utility-bearing" attributes or characteristics [56]. Rosen assumes that z represents the characteristics of a good, such that $z = (z_1, z_2, \dots, z_n)$ and z is continuous. Prices and characteristics are related according to $p(z) = p(z_1, z_2, \dots, z_n)$.

The consumption decision involves maximizing the utility function, $U(x, z_1, z_2, \dots, z_n)$ subject to the budget constraint, $y = x + p(z)$, where x represents expenditures on all other goods consumed. The amount a consumer is willing to pay for various values of z at a fixed level of utility can be represented by a bid function, $\theta(z; u, y)$, which determines a family of indifference curves relating the z_i to money, with x foregone. Rosen shows that the bid function is increasing in z_i at a decreasing rate [56].

Assuming competitive markets, the optimal production decision requires profit-maximization where marginal revenue equals marginal cost, $p(z) = C_M(M, z_1, z_2, \dots, z_n)$, where M equals the number of units produced, C represents total costs, and C_M represents marginal cost with respect to M. Similar to the consumption decision, an offer function, $\phi(z_1, z_2, \ldots , z_n; \pi, \beta)$, where β represents any factor that shifts the cost curves. The offer function shows the prices the firm is willing to accept keeping profit constant and determines a family of indifference curves (or isoquants) which are increasing in z_i at an increasing rate [56].

In equilibrium, the consumer's bid function and the producer's offer function are tangent to each other, with the "common gradient" at that point being equal to the market-clearing implicit price function, $p(z)$. Rosen shows that the observations of $p(z)$ represent a joint envelope of the families of bid functions and offer functions, and represent the equivalent of a hedonic price regression [56].

Method for Research

The conceptual hedonic model is:

$$P = f\,(\boldsymbol{D}, \boldsymbol{H}, \boldsymbol{L}, \boldsymbol{N})$$

where the dependent variable, P, representing the property value, is a function of four vectors of independent variables. The four vectors are $\boldsymbol{D}$, a vector of variables that measures the distance of parcels to transit stations (and to any other locations of interest); $\boldsymbol{H}$, a vector of variables that describes housing characteristics; $\boldsymbol{L}$, a vector of variables that describes locational amenities; and $\boldsymbol{N}$, a vector of variables that describes neighborhood characteristics.

Because theory does not dictate a pre-determined functional form for the hedonic price equations, it is critical to select an estimation strategy that is appropriate [56] [57].

While criticized by some [58], the comparison of goodness-of-fit among alternative functional forms is a common approach in the current relevant literature [59] [30] [60] [61] [62] [29].

The Box-Cox transformation is a widely used nonparametric method [63]. However, Cassel and Mendelsohn indicate that a Box-Cox application may not always be best suited for hedonic analysis [64]. When the purpose of the research is to estimate the price of a particular characteristic, as it is in the case of this work, the Box-Cox does not necessarily result in better estimates of the characteristics' prices. Further, if the key independent variable of interest has a relatively small role in explaining the variation in the dependent variable (in this case, sale price), then it follows that such a variable also has a relatively small role in the determination of the appropriate functional form. Based on the results of relevant literature, it is anticipated that the distance of a property to the nearest transit station plays a statistically significant, yet relatively small role in explaining the sale prices [40] [41] [39] [43] [12]. As such, this work will use specific functional forms.

Based on previous relevant literature on this topic, this work compares the results from three alternative functional forms: levels, log-level (semi-log), and log-log (log-linear). In the levels case, the estimated coefficients represent the marginal characteristic prices for unit changes in the independent variables. With the log-level form, only the dependent variable is transformed, and the interpretations of the coefficients represent percent changes in the dependent variable (sale price) for unit changes in the independent variables. Finally, in log-log form, both the dependent and independent variables are transformed, as appropriate (dummy variables and other categorical variables are not transformed) [65]. In the log-log case, the estimated coefficients indicate the percent changes in the dependent variable with a one percent change in the associated independent variables.

Hedonic regression analysis is prone to issues such as multicollinearity, spatial dependence, and heteroskedasticity [57] [59]. It is important to recognize these potential pitfalls and, as necessary, make adjustments or corrections to the models to reduce or eliminate their effects on the results. The use of White-Huber robust standard errors is one way to adjust for some of these effects, particularly heteroskedasticity [65].

As described previously in Chapter Four, sales data are available for the years 2000 through 2011. An interesting aspect of this research is to see whether the coefficient on the key variable, distance to the nearest BRT station, would be stable over time or how it might be expected to change over time. Because this research covers periods before and after the implementation of the BRT service, it might be expected that the coefficient on the key variable would change over time to reflect the changes on the corridor. Ideally, a panel data set would be employed for this analysis, making it possible to track the sale prices of individual homes over time. However, constructing a useful panel data set for this research would have serious shortcomings. An individual home does not typically sell every year, or even every few years. In the data set of single-family homes available for this research, there are no homes with a recorded sale in every year of the data (2000 to 2011). To investigate whether a sufficient number of individual homes sold in multiple years, the data were filtered in several combinations of year groupings to see how many of the same homes sold in each year of those groupings. In one example, the data were filtered to isolate individual homes that sold three times, once each in 2007, 2009, and 2011. In the year grouping of 2007, 2009, and 2011, it was found that only 12 of the same homes sold once in 2007, once in 2009, and once in 2011. Table 5.1 presents the total number of individual homes sold (price greater than $0) in the database of single-family homes within two miles of the BRT corridor for select years groupings. Since construction of the BRT system was already underway in 2005 (and thus station locations were known), 2005 is one of the years selected for some of the year groupings shown in Table 5.1. Construction was also continuing in 2007, and 2009, 2010, and 2011 represent one, two, and three years after the opening of the system, respectively, and so these years were also included in some of the year groupings shown in Table 5.1. When looking at year groupings comprising just two years, the numbers increase. For example, in Table 5.1, the pair of years in which the most homes sold twice, once in each year, is 2007 and 2011 with 74 homes. Additional year groupings could have been presented in Table 5.1; however, the objective of the information that is presented in the table is to emphasize the lack of sufficient data available to construct a balanced panel data set.

Table 5.1. Number of Same Single-Family Homes Sold in Each Year of Select Year Groupings

Years	Number of the Same Homes Sold in Each Year
2005, 2007, 2009, 2011	4
2007, 2009, 2011	12
2009, 2010, 2011	6
2005, 2010	69
2005, 2011	38
2007, 2009	27
2007, 2011	74
2009, 2011	44
2010, 2011	64

For additional reference, Table 5.2 shows the number of all single-family homes sold in each of the years available in the data set. The table shows all of the homes sold at prices greater than $0, and also at prices greater than $20,000. The data in Table 5.2 represent the number of observations available for individual cross-section analyses for the years shown. The data in the table reflect the relatively higher number of homes sold at lower prices as the housing market passed its peak after 2006 and in the aftermath of the subsequent recession.

Table 5.2. Number of Single-Family Homes Sold in Each Year

Years	Number of All Homes Sold	Number of Homes Sold at $20,000+
2000	577	480
2001	572	484
2002	590	499
2003	702	604
2004	769	663
2005	896	811
2006	881	742
2007	1044	717
2008	944	373
2009	557	217
2010	502	204
2011	440	196

Due to the nature of the available data, a set of separate cross-section regression models for different time periods are estimated and compared. Construction of the system began in 2004, and thus the station locations were known at that point. Additional years in the cross-section analysis include: a combination of sales from 2008 and 2009 (BRT service began in October 2008) and a combination of sales from 2010 and 2011 (two and three years, respectively, after implementation). Because of the relatively smaller number of single-family homes sold near the corridor during these years, and because it was determined that analyzing sales within one mile of the corridor would produce better models, sales were combined for the years 2008 and 2009, as well as 2010 and 2011 (although to increase sample size, the dataset for 2010 and 2011 extends to 1.5 miles beyond the corridor). In the very few cases where a home sold in both years (2008 and 2009; 2010 and 2011), the later sale was selected (i.e., 2009; 2011). Beyond one or one-and-a-half miles from the Euclid corridor it is very likely that there are other unmeasurable factors that influence sale prices, and beyond that distance there are also other barriers to access such as the Cuyahoga River, bridges, and heavy industrial areas, particularly toward the western part of the corridor. After eliminating foreclosures and other abnormally low sales, the 2004 data set represents 192 single-family home sales of at least $30,000 that occurred within one mile of the corridor. For the combined data set representing sales in 2008 and 2009, there are 127 observations. Finally, 140 sales are included in the data set for 2010 and 2011, with the buffer expanded to one-and-a-half miles due to the low number of sales in those years.

Looking beyond the data sufficiency problems confronting the construction of a panel data set is perhaps a more serious econometric problem. Specifically, in this type of empirical work, researchers must be concerned about unobserved heterogeneity, or unobserved variation across individual units of observations. In the empirical study completed herein, it is reasonable to believe that there may be other variables (macroeconomic or other kinds of variables) that may impact housing prices over time. Omitted variables lead to unobserved heterogeneity: for example, if X1 has any degree of correlation with X2, but X2 is omitted from the model, then X1 will be correlated with the error term resulting in biased and inconsistent estimates for X1. It is important to address this issue to the extent possible, and implications for the results are discussed in Chapter Six.

The use of instrumental variables is one way to correct for unobserved heterogeneity when omitted variables cannot be observed and included in the model. Successful implementation of this technique requires finding a variable that is observable, correlated with an omitted explanatory variable, and also uncorrelated with the error term. Finding suitable instrumental variables can be quite difficult and was not feasible for this research. Nonetheless, the use of additional variables to reduce any unobserved heterogeneity was considered, particularly relating to the recession and the housing market irregularities during the time period of this research.

Another method for addressing possible unobserved heterogeneity would be to estimate models using subsets of the sample data. For example, the sample could be restricted based on the type of single-family home (size or age), the location of the home (in particular neighborhoods or areas of the corridor that are expected to be different, such as Midtown and University Circle), or by condition of the home ("average" or better, or "good" or better, perhaps). At least one of these cases could represent another model specification to estimate; however, sample sizes are already relatively small.

An idea for future research would be to investigate whether the distance to the corridor impacted sale prices prior to the construction and implementation of the BRT system. There are four earlier years of data available (2000–2003) that could be used in "before" analyses; however, the only distances computed in the data are from the properties to the nearest BRT station. Because regular bus services operated along Euclid Avenue before the BRT, the current distances to the BRT stations would not be the best measure of accessibility to transit because the BRT stations are spaced more widely apart than typical local bus service stops. In future work, the network distance of the properties to the corridor itself, where bus stops had been placed at least every other block, can be calculated and used to estimate whether, prior to the BRT system, there were any marginal effects on sale prices based on the distance to the corridor itself. It should be noted that, in the early 2000s as the BRT system was planned, possible station locations may not have been widely known, but some home buyers may have been aware that such a system was in the planning stages and were likely aware of the plan to completely reconstruct Euclid Avenue. Raw data from the 1990s are available and could be incorporated into the existing data set for future research.

In addition, future research could pursue the acquisition of additional data for the years beyond 2011, as the local economy in Cleveland has improved more recently and the housing market may be more stable. This is discussed further in Chapter Six.

The models include relevant interaction terms, as well as any variable transformations that prove to be helpful. Previous work has used the square of the distance variable coupled with various interactions such as the number of rooms or bedrooms with the square feet of living area [11]. The former is used to control for possible non-linearities in the effect of distance from the BRT stations on sale price. In the case of the latter, such an interaction is used to allow the effect of square feet of living area on sale price to vary with the number of rooms or bedrooms. In addition, it is interesting to see how the distance coefficient may vary with the neighborhood characteristics. To explore this, a set of interaction terms between distance to the nearest BRT station and select neighborhood characteristics is included in one of the model specifications. There are several neighborhoods and city wards along the corridor, so it is more feasible to apply these interactions to only a few select areas. For example, it would be of interest to compare the distance effect in two very different areas such as the redeveloping Midtown section and the more established University Circle section (the latter of which contains important cultural and arts establishments).

It would also be interesting to investigate how the distance coefficient might be affected by the distances from the BRT stations to various destinations to which the system provides access. From the home-buyer's perspective, it is logical when one is consciously considering BRT access as a factor in the purchasing decision that one will also be accounting for the location of any preferred destinations. For example, might the coefficient be different depending on how far along the corridor one travels and then how far one must travel from the corridor to a destination? However, from the data available for research, it is not possible to know to which destinations particular individuals travel or would like to travel. Transit on-board surveys are typically conducted by transit agencies on a regular basis and usually collect information on passengers' origins and destinations. However, such information is useful to the agency in aggregate form for planning purposes, and consequently no identifying information is ever collected from the survey respondents (except for perhaps the passenger's home zip code). Most surveys of any type have an important shortcoming: either they do not collect identifying information on respondents or such information is not made available to researchers due to privacy concerns. Thus, it is not possible to know which home-owners even utilize the BRT system. In this research, it is hypothesized that proximity to BRT is a locational amenity, similar to being located near a park [59] or highly-ranked public schools, for example. Those who live near a nice park or in a good school district may be paying a housing premium for doing so, regardless of whether they use the park or have children who attend the good schools.

While not always feasible for some transit services, the design of the Cleveland BRT corridor is such that it provides walking access to many major destinations within the CBD as well as Cleveland State University, Case Western Reserve University, the Cleveland Clinic campus and other hospitals/medical offices, art/history/children's museums, a symphony hall, the Playhouse Square theater district, retail, other offices, parks, and housing. There are major employment centers at the western terminus of the corridor and in the eastern portion (the CBD and University Circle, respectively), so commutes occur in both directions. The HealthLine BRT also connects to the Red Line heavy rail metro system in two locations and to other bus routes in the GCRTA system. Therefore, it is not possible in this research to control for the distances from stations to particular destinations in the area. However, it is feasible to compute the travel time from each station to every other station along the line. For this research, as discussed previously, the two stations of particular interest are Tower City, which represents the CBD, and East 93rd Street, which represents the main access to the Cleveland Clinic, a major employer and trip generator.

This chapter has outlined the methodology used in this dissertation research. The results are presented and discussed in the next chapter, Chapter Six.

CHAPTER SIX:

RESULTS AND DISCUSSION

This chapter presents the results of the dissertation research. Results of 18 ordinary least squares (OLS) regression models are summarized in Tables 6.1 through 6.6. Discussion is provided, particularly with respect to the key variables relating to the distance to the BRT corridor. Then, a set of three models is selected, one from each time period, to be more closely examined, with interpretations of all coefficients. Finally, all relevant results are summarized, and future research ideas are proposed.

Regression Results

Overall, there were a few variables that were expected to positively impact sales prices but did not have any significant effect in any of the models. These included the lot size, number of bathrooms, and number of half-bathrooms. Much of the sample comprises relatively smaller homes on relatively smaller lots. Further, there is not much variability among the sample regarding the number of bathrooms and half-bathrooms, and this may explain the lack of a significant effect in the models. Therefore, these three variables were not included in any of the models shown in Tables 6.1 through 6.6.

During the course of the research, some of the initial variables were altered. For example, the variable representing the year the home was built was converted to the age of the structure in the year of sale (i.e., in the 2004 data set, this variable represented the age of the structure in 2004, etc.). This was originally done to better use the age variable in some interaction terms; however, those interaction terms were never used and the age variable remained in the models. As shown in Tables 6.1 through 6.6, the age variable is one of the only variables to be statistically significant at the five percent level of significance or better in nearly every model tested (the variable AGE was found to be statistically significant at the eight percent level in the 2004 levels model with distance as a continuous variable and statistically significant at the ten percent level of significance in the 2004 model with distance as a dummy variable).

In Chapter Five, dummy variables to control for the location of the single-family homes in the data set were discussed. There were three options: City of Cleveland Wards, Cleveland neighborhoods, and a set of development areas selected just for this research. After the initial models were estimated, it became clear that the specially drawn development areas were highly insignificant in explaining sale prices in all models. This may have been expected, given that these areas were each quite large compared to the city wards and neighborhoods. Indeed, the city wards and neighborhoods represent a much finer geographic level. Also, the neighborhoods tend to be smaller than the city wards. While both the neighborhoods and city wards performed similarly in the models, ultimately the neighborhoods were selected for the models shown in Tables 6.1 through 6.6. This is because the neighborhoods represent a smaller geography and also because the robust *F* statistic (Wald statistic) was stronger for most of the models estimated when the neighborhood dummies were used.

First, Tables 6.1 and 6.2 present the results for the 2004 cross-section of sales. Table 6.1 shows the results for the models using distance as a continuous variable, while Table 6.2 includes distance as a dummy variable. In 2004, the Euclid corridor was under construction and the BRT service had not yet been implemented. However, the station locations had been announced, and so these models test if there is any effect from the distance at that time. Alternately, it might be expected that the distance variables would show that property values are lower nearer to the corridor and/or station locations because not all buyers may have been aware of the station locations and there was no BRT service in operation, only typical regular local bus services. Table 6.1 shows no significance for the distance coefficients for any of the three models, although the expected signs of each indicate that property values increase as distance increases. For example, even though the estimated coefficient for DIST is negative for the levels model, the DIST-SQ coefficient makes the overall effect positive. See Appendix A for a discussion of how the effect of distance on sale price is obtained from the estimated regression coefficients.

Evaluated at a half-mile (2,640 feet) distance, the estimated effect is –6.18247 + 2(0.00203)(2,460) = \$3.80. This result would mean that, at a distance of 2,460 feet, a one foot increase in distance from the nearest BRT station would increase estimated sale price by \$3.80. Similarly, for the log-level model in Table 6.1, the estimated increase in sale price for a one foot increase in distance from the nearest station at 2,460 feet is 0.002 percent. This percentage change in price is equal to 100[–0.000077 + 2(1.99e-08)(2,460)]. In the log-log model, the estimated coefficient on the log of the total feet from the nearest station, LN_DIST, represents the elasticity of sale price with respect to distance from the nearest station. In this case, the estimated change in price for a one percent increase in distance would be 0.07 percent (or \$85.13, based on the mean sale price in 2004). It should be noted, as already discussed, that none of these coefficients is statistically significant; the interpretations are shown to demonstrate that the signs of the coefficients indicate that, for the 2004 data, sale prices increase as distance from a station increases. The remainder of the coefficients in Table 6.1 generally have the expected signs and most are statistically significant at the five percent level of significance. In addition, for all three models shown in Table 6.1, the neighborhood dummy variables are jointly significant using the heteroskedastic-robust *F* statistic.

Table 6.2 shows the models with distance as a dummy variable, with the distance dummy variable taking the value of one if the home is one half-mile (2,640 feet) or less from its nearest BRT station. The dummy variable takes a value of zero if the home is greater than one-half mile away up to one mile. In each of the three models, the distance dummy (DIST1_dum) has a negative sign, implying that homes within a half-mile of the nearest BRT station have lower sale prices than homes farther away. The table indicates that only the coefficient in the log-log model is statistically significant, and only at the ten percent level of significance (the exact level of significance is 8.5 percent). However, in the levels model, the coefficient on the distance dummy has an exact level of significance of 11.8 percent. In the log-level model, the exact level of significance of this coefficient is 11.3 percent. In the levels model, the coefficient on DIST1_dum indicates that a home one half-mile or less from the nearest BRT station would sell for $12,716.36 less than a home farther than one-half mile away. In the log-level model, the exact expression for the percentage change in sale price is $100(e^{-0.098021} - 1) = 100(0.906630 - 1) = 100(-0.09337) = -9.337$ percent. This means that a home one half-mile or less from the nearest BRT station would have a sale price approximately 9.3 percent less than a home farther away (or $11,310 less, based on the mean sale price in 2004). Similarly, in the log-log model, a home one half-mile or less from the nearest BRT station would be estimated to sell for approximately 10.2 percent less than a home farther away, as shown by: $100(e^{-0.108801} - 1) = 100(0.897627 - 1) = 100(-0.102373) = -10.237$ percent (or $12,374 based on the mean sale price in 2004). As expected, in Table 6.2, the magnitudes of the coefficients in the log-level model and the log-log model are very similar. Similar to Table 6.1, most of the estimated coefficients listed in Table 6.2 have the expected signs and nearly all are statistically significant. Also, as with the models in Table 6.1, the models in Table 6.2 have neighborhood dummy variables that are jointly significant using the heteroskedastic-robust *F* statistic. The log-log model shown in Table 6.2 is discussed in more detail later in this chapter.

In October 2008, construction was complete along the Euclid corridor and the HealthLine BRT service began operating. However, also during this time, the 2007–2009 recession and subsequent housing market decline began to affect sale prices of homes not just in Cleveland, but in many parts of the country. Between 2005 and 2009, the median sale price of a single-family home in the city of Cleveland fell from $91,200 to $73,400, a decline of approximately 20 percent [66]. Fewer homes were sold; therefore, to increase the sample size for this time period, home sales were taken from both 2008 and 2009. Tables 6.3 and 6.4 show results for the sample of single-family homes sold during 2008 and 2009 (there were just a handful of homes that sold in both years; in those cases, the 2009 sale was used in the analysis).

In Table 6.3, showing distance as a continuous variable, the coefficients on DIST and DIST_SQ have the expected negative sign and all are statistically significant at at least the five percent level of significance. These results would be expected if, after implementation of the BRT service, proximity to the stations is positively impacting sale prices. Evaluated at a distance of 2,640 feet (one-half mile), the levels model indicates that as distance from the nearest station increases one more foot, sale price is estimated

Table 6.1. 2004 Cross-Section with Distance as a Continuous Variable

Variable	Description	2004 (n=192) Level Coefficient *Robust Std Error*	2004 (n=192) Log-Level Coefficient *Robust Std Error*	2004 (n=192) Log-Log† Coefficient *Robust Std Error*
PRICE	Dependent variable: sale price of home	PRICE	LN_PRICE	LN_PRICE
Constant	Constant term in regression equation	159,587.9 * *(77,474.87)*	11.77161 * *(0.345070)*	5.117631 * *(1.324164)*
DIST (LN_DIST†)	**Distance (in feet) of home to nearest BRT station**	**–6.182469** ***(10.90342)***	**–0.000077** ***(0.0000757)***	**0.072538** ***(0.047239)***
DIST_SQ	Distance (in feet) of home to nearest BRT station squared	0.002029 *(0.001770)*	1.99e-08 *(1.27e-08)*	n/a
AREA (LN_AREA†)	Size of home's living area in square feet	46.45356 * *(12.86471)*	–0.000201 * *(0.000060)*	0.436165 * *(0.118828)*
BEDROOMS	Number of bedrooms	–8,096.733 ** *(4811.846)*	–0.049149 ** *(0.027815)*	–0.058972 ** *(0.030434)*
AGE	Age of the home in years	–296.8507 ** *(171.6148)*	–0.004588 * *(0.001055)*	–0.004322 * *(0.001160)*
COND	Condition of the home; likert scale	13,692.19 * *(4,095.951)*	0.068648 * *(0.030164)*	0.078595 * *(0.031242)*
MDHHINC (LN_MDHHINC†)	Median household income for census block group that includes	2.859797 * *(0.918804)*	0.000018 * *(4.85e-06)*	0.384538 * *(0.083577)*
B1940	Percent of homes built before 1940 in the census tract that includes the property	–1,307.083 * *(382.8362)*	–0.006786 * *(0.002043)*	–0.008613 * *(0.002076)*
CENTRAL	Dummy variables: Take value of 1 if property is located in the listed city neighborhood; 0 otherwise (base case is outside of these neighborhoods)	–128,278.3 ** *(75,616.77)*	–0.400443 *(0.271472)*	–0.599862 * *(0.241912*
FAIRFAX		–100,583.5 *(73,871.83)*	–0.273953 *(0.259136)*	–0.424428 ** *(0.245544)*
GLENVILLE		–131,326.0 ** *(73,522.69)*	–0.490599 ** *(0.254193)*	–0.61073 * *(0.238202)*
GOODRICH		–128,902.70 ** *(73,964.56)*	–0.533010 * *(0.263834)*	–0.720299 * *(0.248223)*
HOUGH		–152,812.9 ** *(77,675.03)*	–0.533925 * *(0.265448)*	–0.695322 * *(0.249264)*
UNIVERSITY		–69,220.36 *(77,077.34)*	0.036431 *(0.283179)*	–0.107361 *(0.254801)*

*Significant at the 5-percent level of significance. **Significant at the 10-percent level of significance.

† These variables are only used in the log-log specification. All others are entered as levels variables unless otherwise noted.

<u>Levels model:</u> adjusted $R^2 = 0.712$, F (Wald statistic) = 32.14 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (6, 177) = 4.42 (prob > F = 0.0003).

<u>Log-Level model:</u> adjusted $R^2 = 0.674$, F (Wald statistic) = 40.49 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (6, 177) = 4.55 (prob > F = 0.0003).

<u>Log-Log model:</u> adjusted $R^2 = 0.664$, F (Wald statistic) = 38.20 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (6, 178) = 6.27 (prob > F = 0.0000).

Table 6.2. 2004 Cross-Section with Distance as a Dummy Variable

Variable	Description	2004 (n=192)		
		Level	Log-Level	Log-Log†
		Coefficient *Robust Std Error*	Coefficient *Robust Std Error*	Coefficient *Robust Std Error*
PRICE	Dependent variable: sale price of home	PRICE	LN_PRICE	LN_PRICE
Constant	Constant term in regression equation	174,529.6 * *(79,233.55)*	11.85904 * *(0.345305)*	5.742406 * *(1.126172)*
DIST1_dum	**Distance from home to nearest BRT station between 0 and 2,640 feet (0.5 mile); base case is between 2,640.01 and 5,280 feet (0.5 mile to 1**	**–12,716.36** ***(8,105.654)***	**–0.098021** ***(0.061472)***	**–0.108001 **** ***(0.062404)***
AREA (LN_AREA†)	Size of home's living area in square feet	46.63263 * *(12.79539)*	-0.000204 * *(0.000060)*	0.431117 * *(0.118107)*
BEDROOMS	Number of bedrooms	–7,883.225 *(4,956.769)*	–0.0485556 ** *(0.029006)*	–0.052638 ** *(0.031515)*
AGE	Age of the home in years	–292.2287 ** *(175.4353)*	–0.004581 * *(0.001066)*	–0.0046 * *(0.0011)*
COND	Condition of the home; likert scale	13,598.63 * *(4,276.826)*	0.066815 * *(0.031005)*	0.074933 * *(0.031926)*
MDHHINC (LN_MDHHINC†)	Median household income for census block group that includes	2.794496* *(0.882806)*	0.000018 * *(4.72e-06)*	0.382661 * *(0.078968)*
B1940	Percent of homes built before 1940 in the census tract that includes the property	–1,320.434 * *(365.3235)*	–0.007062 * *(0.001925)*	–0.008194 * *(0.001928)*
CENTRAL	Dummy variables: Take value of 1 if property is located in the listed city neighborhood; 0 otherwise (base case is outside of these neighborhoods)	–135,715.0 ** *(77,711.85)*	–0.466565 ** *(0.288165)*	–0.567455 * *(0.256262)*
FAIRFAX		–106,381.3 *(76,986.21)*	–0.313849 *(0.283470)*	–0.423603 ** *(0.258221)*
GLENVILLE		–134,665.7 ** *(76,147.59)*	–0.518093 ** *(0.279018)*	–0.604837 * *(0.252551)*
GOODRICH		–135,843.0 ** *(76,514.58)*	–0.591701 * *(0.285053)*	–0.703082 * *(0.261537)*
HOUGH		–158,159.2 * *(80,174.49)*	–0.582725 * *(0.287664)*	–0.672981 * *(0.262348)*
UNIVERSITY		–76,657.32 *(80,034.19)*	–0.014391 *(0.300913)*	–0.090528 *(0.272340)*

*Significant at the 5-percent level of significance. **Significant at the 10-percent level of significance.

† These variables are only used in the log-log specification. All others are entered as levels variables unless otherwise noted.

Levels model: adjusted $R^2 = 0.711$, F (Wald statistic) = 33.18 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (6, 178) = 6.10 (prob > F = 0.0000).

Log-Level model: adjusted $R^2 = 0.672$, F (Wald statistic) = 42.65 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (6, 178) = 5.59 (prob > F = 0.0000).

Log-Log model: adjusted $R^2 = 0.667$, F (Wald statistic) = 38.86 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (6, 178) = 6.58 (prob > F = 0.0000).

to decrease approximately \$23.20. This is determined by $-72.1864 + 2(0.00928)(2{,}640)$. In the log-level model, the percentage change in price from an increase in distance of one foot at the one-half mile from the nearest station is $100[-0.000612 + 2(8.15\text{e-}08)(2{,}460)] = -0.02$ percent (or \$22.48 based on the mean sale price in 2008 and 2009). Lastly, the log-log model indicates that a one percent increase in distance from the nearest BRT station decreases sale price by approximately 0.28 percent (or \$314.76 based on the mean sale price in 2008 and 2009).

Because during this time period, 2008 to 2009, the BRT service was implemented, the BRT schedules and travel times between stations were determined and published. As discussed in Chapter Five, the models for this time period include the travel times from each station to the main station for accessing the Cleveland Clinic, and from each station to the major CBD station at Tower City. For individual homes in this data set, the CCTT variable represents the travel time from the nearest BRT station to the East 93rd Street station (the main station for accessing the Cleveland Clinic campus) according to the published HealthLine schedule for weekday peak service. Similarly, the CBDTT variable represents the travel time from the nearest BRT station to the Tower City station according to the published HealthLine schedule for weekday peak service. It might be expected that the signs of both of these coefficients would be negative, indicating that an increase in travel time to these stations (and area destinations) would decrease sale prices, all else constant. For the travel time to the

Cleveland Clinic campus, this expectation holds for the models in Table 6.3. However, none of the coefficients is considered statistically significant, although the coefficient in the log-log model has an exact level of confidence of 86.6 percent. For the travel time to the CBD, the signs on the coefficients are all positive, with the coefficient in the log-log model statistically significant at the five percent level of significance. In this log-log case, a one minute increase in travel time to the CBD via the BRT system is estimated to increase sale price by 1.3 percent [100(1.01327 – 1) = 100(0.01327) = 1.33 percent] or $1,461.41 based on the mean sale price in 2008 and 2009. It may be that single-family homes closer to the CBD are somewhat less desirable, in general, thus explaining this result.

In Table 6.4, the three models are shown with distance as a dummy variable. As expected, the signs on the three DIST1_dum coefficients are now positive, indicating that homes closer to the stations (within one-half mile) have higher sale prices than those homes farther away. Each is statistically significant at at least the five percent level of significance. Similar to the interpretations in Table 6.2, for the levels model, the coefficient on DIST1_dum indicates that a home located within one half-mile from its nearest BRT station would sell for $44,817.52 more, on average, than a home located between one half-mile and one mile from the nearest BRT station, holding all else constant. The exact interpretation for the log-level model is: $100(e^{0.361712} - 1) = 100(1.43579 - 1) = 100(0.43579) = 43.6$ percent, meaning that a home within one-half mile of the nearest BRT station is estimated to have a sale price 43.6 percent higher than a home farther than one-half mile away (or approximately $49,013 based on the mean sale price in 2008 and 2009). Lastly, the coefficient for DIST1_dum in the log-log model is equal to 0.344018, meaning that a home within the one-half mile distance from a station would sell for 41.1 percent more than one farther away [$100(e^{0.344018} – 1) = 100(1.41060 – 1) = 100(0.41060) = 41.1$ percent], or approximately $46,203 based on the mean sale price in 2008 and 2009).

As in Table 6.3, the travel time variables, CCTT and CBDTT enter into the models in Table 6.4. The coefficients on CCTT have the expected negative signs, but none is statistically significant. For the coefficients on CBDTT, they are significant in the log-level only (9.6 percent exact level of significance in the log-level model). For the log-level model, a one-minute increase in travel time to the CBD station at Tower City would be estimated to increase sale price 0.97 percent, all else constant [$100(e^{0.009675} – 1) = 100(1.00972 – 1) = 100(0.00972) = 0.97$ percent], or approximately $1,090 based on the mean sale price in 2008 and 2009.

Overall, the results of the models in Tables 6.3 and 6.4 may be showing the effects of the 2007–2009 recession and resulting irregular housing market. Two of the variables, AREA and AGE, do not have the expected signs in the log-level models. Also, the variable representing household median income has the unexpected negative sign in all of the models. Further, none of the coefficients for median income is statistically significant. One explanation for this unanticipated result may be that the median household income variable used for the 2008 and 2009 data did not represent actual incomes at that time. The reported median income information used in the analysis is based on five-year estimates from the American Community Survey (ACS), and actual incomes may have been quite different as the effects of the recession were felt among households in Cleveland. The 2004 median incomes are also based on five-year ACS estimates; however, 2004 was before the 2007–2009 recession. In addition, the five-year ACS median income data used for the 2010 and 2011 data may be more accurately reflecting actual household incomes in the study area (Tables 6.5 and 6.6 show the median income variable returning to statistical significance in those models).

Most other variables in Tables 6.3 and 6.4 do have expected signs, and in all models the neighborhood dummy variables are jointly significant using the heteroskedastic-robust *F* statistic. Overall, the results shown in Tables 6.3 and 6.4 support the hypothesis of this research that proximity to BRT stations is associated with relatively higher sale prices for single-family homes in the study area. The log-log model with distance as a dummy variable, shown in Table 6.4, is further discussed later in this chapter.

Tables 6.5 and 6.6 present the results for the most recent data available on single-family home sales along the BRT corridor. While the hypothesis of this work would

Table 6.3. 2008–2009 Cross-Section with Distance as a Continuous Variable

Variable	Description	2008–2009 (n=127) Level Coefficient *Robust Std Error*	Log-Level Coefficient *Robust Std Error*	Log-Log† Coefficient *Robust Std Error*
PRICE	Dependent variable: sale price of home	PRICE	LN_PRICE	LN_PRICE
Constant	Constant term in regression equation	216,362.5 * *(44,561.74)*	12.12051 * *(0.366363)*	8.924299 * *(1.507681)*
DIST (LN_DIST†)	**Distance (in feet) of home to nearest BRT station**	**–72.18642 *** ***(20.11583)***	**–0.000612 *** ***(0.000163)***	**–0.278642 *** ***(0.073095)***
DIST_SQ	Distance (in feet) of home to nearest BRT station squared	0.009278 * *(0.002907)*	8.15e-08 * *(2.56e-08)*	n/a
AREA (LN_AREA†)	Size of home's living area in square feet	48.64595 * *(10.31747)*	–0.000382 * *(0.000065)*	0.814910 * *(0.160015)*
BEDROOMS	Number of bedrooms	–10,180.48 *(5,661.671)*	–0.079868 ** *(0.044295)*	–0.089027 ** *(0.050350)*
AGE	Age of the home in years	–502.4693 * *(172.6184)*	0.003456 * *(0.001342)*	–0.180550 * *(0.046183)*
COND	Condition of the home; likert scale	13,953.62 * *(5,540.688)*	0.129479 * *(0.044651)*	0.105056 * *(0.042960)*
MDHHINC (LN_MDHHINC†)	Median household income for census tract that includes the	–0.488564 *(0.664679)*	–2.73e-06 *(4.90e-06)*	–0.037424 *(0.105752)*
CCTT	Travel time in minutes from the nearest BRT station to the Cleveland Clinic	–802.7411 *(1,497.437)*	–0.011012 *(0.014069)*	–0.020689 *(0.013695)*
CBDTT	Travel time in minutes from the nearest BRT station to Tower City	382.1515 *(582.1572)*	0.007425 *(0.005848)*	0.013181 * *(0.005734)*
B1940	Percent of homes built before 1940 in the census tract that includes the property	–698.0498 ** *(366.2902)*	–0.006421 ** *(0.003327)*	–0.008277 * *(0.003316)*
CENTRAL	Dummy variables: Take value of 1 if property is located in the listed city neighborhood; 0 otherwise (base case is outside of these neighborhoods)	–48,937.08 * *(22,519.79)*	–0.228355 *(0.188781)*	–0.286038 *(0.3630)*
FAIRFAX		–59,418.10 * *(21,304.49)*	–0.425675 * *(0.168705)*	–0.467314 * *(0.172243)*
HOUGH		–45,144.99 * *(19,177.44)*	–0.254944 *(0.167930)*	–0.392972 * *(0.153648)*

*Significant at the 5 percent level of significance. **Significant at the 10 percent level of significance.

† These variables are only used in the log-log specification. All others are entered as levels variables unless otherwise noted.

<u>Levels model:</u> adjusted $R^2 = 0.612$, F (Wald statistic) = 20.81 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (3, 113) = 2.90 (prob > F = 0.0380).

<u>Log-Level model:</u> adjusted $R^2 = 0.609$, F (Wald statistic) = 24.61 (prob > F = 0.0000). Neighborhood dummies are jointly significant (at the 90% level of significance) using the heteroskedastic-robust F statistic, F (3, 113) = 2.14 (prob > F = 0.0990).

<u>Log-Log model:</u> adjusted $R^2 = 0.607$, F (Wald statistic) = 31.65 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (3, 114) = 2.92 (prob > F = 0.0370).

Table 6.4. 2008–2009 Cross-Section with Distance as a Dummy Variable

Variable	Description	2008–2009 (n=127)		
		Level	Log-Level	Log-Log†
		Coefficient *Robust Std Error*	Coefficient *Robust Std Error*	Coefficient *Robust Std Error*
PRICE	Dependent variable: sale price of home	PRICE	LN_PRICE	LN_PRICE
Constant	Constant term in regression equation	127,924.1 * *(41,835.35)*	11.40716 * *(0.385024)*	5.492629 * *(1.312385)*
DIST1_dum	**Distance from home to nearest BRT station between 0 and 2,640 feet (0.5 mile); base case is between 2,640.01 and 5,280 feet (0.5 mile to 1 mile)**	**44,817.52 *** ***(13,432.16)***	**0.361712 *** ***(1.025221)***	**0.344018 *** ***(0.102827)***
AREA (LN_AREA†)	Size of home's living area in square feet	54.75167 * *(9.529861)*	–0.000426 * *(0.000063)*	0.925687 * *(0.161668)*
BEDROOMS	Number of bedrooms	-8,619.304 *(5,849.518)*	–0.080868 ** *(0.045967)*	–0.106635 * *(0.052764)*
AGE	Age of the home in years	–671.2398 * *(196.9275)*	0.004762 * *(0.001399)*	–0.004238 * *(0.001410)*
COND	Condition of the home; likert scale	9,689.966 ** *(5,342.956)*	0.09733 * *(0.043465)*	0.109153 * *(0.044538)*
MDHHINC (LN_MDHHINC†)	Median household income for census tract that includes the	–0.177847 *(0.673103)*	–3.08e-07 *(5.14e-06)*	–0.014574 *(0.113329)*
CCTT	Travel time in minutes from the nearest BRT station to the Cleveland Clinic	–1,209.471 *(1,523.107)*	–0.015599 *(0.013567)*	–0.014634 *(0.014351)*
CBDTT	Travel time in minutes from the nearest BRT station to Tower City	592.6985 *(595.0017)*	0.009675 ** *(0.005764)*	0.008987 *(0.006513)*
B1940	Percent of homes built before 1940 in the census tract that includes the property	–1,233.306 * *(330.0276)*	–0.011031 * *(0.003311)*	–0.011387 * *(0.003312)*
CENTRAL	Dummy variables: Take value of 1 if property is located in the listed city neighborhood; 0 otherwise (base case is outside of these neighborhoods)	–77,099.96 * *(26,598.47)*	–0.465107 * *(0.199452)*	–0.455854 * *(0.201378)*
FAIRFAX		–59,596.43 * *(23,074.33)*	–0.443575 * *(0.183248)*	–0.485879 * *(0.187107)*
HOUGH		–73,072.41 * *(20,488.0)*	–0.495959 * *(0.164522)*	–0.516659 * *(0.168461)*

*Significant at the 5 percent level of significance. **Significant at the 10 percent level of significance.

† These variables are only used in the log-log specification. All others are entered as levels variables unless otherwise noted.

Levels model: adjusted $R^2 = 0.594$, F (Wald statistic) = 22.94 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (3, 114) = 4.67 (prob > F = 0.0041).

Log-Level model: adjusted $R^2 = 0.587$, F (Wald statistic) = 25.79 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (3, 114) = 3.21 (prob > F = 0.0258).

Log-Log model: adjusted $R^2 = 0.578$, F (Wald statistic) = 23.11 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (3, 114) = 3.28 (prob > F = 0.0235).

suggest that the results found for the 2008–2009 models would persist in 2010 and 2011, the results summarized in Tables 6.5 and 6.6 do not show this to be true.

Certainly, the northeast Ohio region was impacted greatly by the recession of 2007–2009. Much of the area surrounding the BRT corridor is characterized by older neighborhoods with relatively smaller, aging homes. And, excluding foreclosures and other far below-market sales, very few homes were sold during this period, and very few were located closer to the corridor. As with the 2008–2009 data set, because of the small number of homes sold in each year, 2010 and 2011, the two years were combined to increase the sample size. In addition, because the sample size remained low even after the years were combined, the distance range was extended by an additional half-mile, up to 1.5 miles (7,920 feet). This change may partially explain some of the differences in the 2010–2011 models from the others in Tables 6.1 through 6.4. However, it is more likely that the state of the economy in Cleveland and the continued irregularities in the housing market resulted in models that are relatively less strong than for the earlier years in this research. The median sale price for single-family homes in the city of Cleveland continued to fall during this time, from $73,400 in December 2009 to a low of $56,500 in December 2011, a drop of 23 percent in two years [66]. However, the median sale price of single-family homes began increasing by December 2012 to $61,300 and, after a small dip in 2013, has grown to $76,000 as of August 2015 and is forecast to increase by another 1.8 percent by September 2016 [66].

In addition to a stabilizing of the housing market very recently, the areas surrounding the Euclid corridor have seen additional redevelopment and recovery since 2011. Ideally, a future research endeavor would involve acquiring more recent sales data to better understand the impacts of the BRT corridor on home sales without the distortion of the recent recession and housing crisis. For this research, it was difficult to find measurable variables that could help account for the macroeconomic situation at the time data were available.

In Table 6.5, showing the models with distance as a continuous variable, none of the distance variables is significant. While all appear to have the expected negative sign, upon closer inspection the levels model actually shows a positive impact on sale price with increasing distance from the nearest BRT station, i.e., at a distance of a half-mile (2,640 feet), an additional foot of distance away from the station would increase sale price by approximately \$0.15 [$-1.72576 + 2(0.00036)(2{,}640)$]. The other two coefficients, from the log-level and log-log models, show very slight negative impacts. Again, none of these results is statistically significant, but they are discussed and the interpretation of the levels model coefficient is provided to show the resulting sign.

Table 6.5 also shows the coefficients for the two travel time variables, CCTT and CBDTT. The sign of CCTT, the travel time in minutes from the nearest BRT station to the main Cleveland Clinic station, is negative in all three models, but only statistically significant in the levels model. For its interpretation, a one minute increase in travel time to the main Cleveland Clinic station (East 93rd Street) is estimated to decrease sale price by $2,354.53, all else constant. The CBDTT variable is highly insignificant and it moves from positive in the levels model to negative in the log-level and log-log models.

Table 6.6 presents the models with distance as a dummy variable for the 2010–2011 data set. In all three of these models, the coefficient on the DIST1_dum variable has the expected positive sign, indicating that sale prices are higher within one-half mile of properties' nearest BRT stations than they are beyond that distance. However, as with the results in Table 6.5, none is statistically significant.

Again, the two travel time variables are included in the models in Table 6.6. Similar to the results in Table 6.5, the signs on the estimated coefficients for the variable representing travel time from the nearest BRT station to the main Cleveland Clinic station (CCTT) are negative. The only statistically significant result is for the levels model, where a one minute increase in this travel time is estimated to decrease the sale price of a home by $1,916.88, all else constant. The coefficient for the CBDTT variable, representing the travel time from the nearest BRT station to the Tower City station in the CBD, has negative signs in all three models and is not statistically significant in any of them.

As for the remainder of the variables included in the models shown in Tables 6-5 and 6-6, they have the expected signs, although not all are statistically significant. The variables representing square feet of living area, the home's age, and the household median income all have the expected signs on their estimated coefficients and are statistically significant. The coefficients on the variables indicating the number of bedrooms in the home are not statistically significant in any of the models presented in Tables 6.5 and 6.6. In addition the variable representing the condition of the home, which was statistically significant at at least the 10-percent level of significance in the models for 2004 and 2008–2009, is not statistically significant at that level for the 2010–2011 data set. For the coefficients on the condition variable in the 2010–2011 data set, the exact level of significance ranges from 15.6 percent (in the levels model using distance measured as a continuous variable) to 11.6 percent (in the log-level model using distance as a dummy variable). In addition, for the 2010–2011 data set, the variable representing the percent of homes within a census tract built before 1940 became highly insignificant, after being statistically significant in all previous models shown for the 2004 and 2008–2009 data sets in Tables 6.1 through 6.4. Instead, a variable that had been highly insignificant in those previous models became statistically significant in the 2010–2011 models, the percentage of homes within a census tract that are owner-occupied. In all six models summarized in Tables 6.5 and 6.6, this variable has a negative sign and is statistically significant, indicating that as the percentage of owner-occupied homes increases, sale prices fall, all else constant. This result may be due to the changing characteristics of the housing market in the study area during this time, and also perhaps to the updated information from the 2010 U.S. Census. Finally, it must be noted that for all models shown in Tables 6.5 and 6.6, the neighborhood dummy variables are statistically significant using the heteroskedastic-robust *F* statistic. In the next section, the log-log model using distance as a dummy variable is discussed in more detail.

The Log-Log Models with Distance as a Dummy Variable

After analyzing all of the models presented in Tables 6.1 through 6.6, one particular functional form and specification appeared to emerge as a bit stronger or robust than the others. From Tables 6.2, 6.4, and 6.6, the log-log models with distance entered as a dummy variable are chosen for additional interpretation in this section. The variable DIST1_dum was used to denote homes in the data set that were one half-mile or less from the nearest BRT station (2,640 feet or less). The base case was homes outside of this distance, and up to one mile (5,280 feet) for the 2004 and 2008–2009 data sets and up to one-and-a-half miles (7,920 feet) for the 2010–2011 data set. For the 2004 data set, representing the beginning of the corridor construction and four years before the BRT service was actually implemented, there was not necessarily an *a priori* expectation for the sign of the distance dummy coefficient. First, it might be expected that the coefficient would be negative, indicating that homes closer to the BRT stations would sell for less than those farther away. This might be expected because there was only regular local bus service along the corridor at that time, and several parts of the corridor were blighted. Second, it might be that the coefficient would be positive,

Table 6.5. 2010–2011 Cross-Section with Distance as a Continuous Variable

Variable	Description	2010–2011 (n=140)		
		Level	**Log-Level**	**Log-Log†**
		Coefficient *Robust Std Error*	Coefficient *Robust Std Error*	Coefficient *Robust Std Error*
PRICE	Dependent variable: sale price of home	PRICE	LN_PRICE	LN_PRICE
Constant	Constant term in regression equation	99,715.27 * *(38,345.99)*	11.24524 * *(0.408939)*	4.848422 * *(1.64418)*
DIST (LN_DIST†)	**Distance (in feet) of home to nearest BRT station**	**–1.725756** ***(7.597802)***	**–0.000031** ***(0.000089)***	**–0.013241** ***(0.062528)***
DIST_SQ	Distance (in feet) of parcel to nearest BRT station squared	0.000357 *(0.000916)*	3.17e-09 *(1.06e-08)*	n/a
AREA (LN_AREA†)	Size of home's living area in square feet	19.8517 * *(9.061425)*	0.000183 ** *(0.000098)*	0.564306 * *(0.8989)*
BEDROOMS	Number of bedrooms	–5,102.099 *(6096.526)*	–0.027948 *(0.059500)*	–0.064242 *(0.060515)*
AGE	Age of the home in years	–543.9546 * *(163.4414)*	–0.005252 * *(0.001529)*	–0.005317 * *(0.8468)*
COND	Condition of the home; likert scale	7,891.275 *(5,525.891)*	–0.088625 *(0.057317)*	–0.084944 *(0.055264)*
MDHHINC (LN_MDHHINC†)	Median household income for census tract that includes the	2.471283 * *(0.501261)*	0.000015 * *(4.06e-06)*	0.321776 * *(0.129101)*
CCTT	Travel time in minutes from the nearest BRT station to the Cleveland Clinic	–2,354.528 * *(1,075.317)*	–0.01067 *(0.011054)*	–0.007225 *(0.010362)*
CBDTT	Travel time in minutes from the nearest BRT station to Tower City	13.49361 *(502.7104)*	–0.001751 *(0.004830)*	–0.005836 *(0.004673)*
OWNOCC	Percent of owner-occupied homes in the census tract that includes the property	–1,013.104 * *(330.9354)*	–0.007037 * *(0.002937)*	–0.007174 * *(0.003472)*
CENTRAL	Dummy variables: Take value of 1 if property is located in the listed city neighborhood; 0 otherwise (base case is outside of these neighborhoods)	–75,982.62 * *(18,769.44)*	–0.777960 * *(0.194394)*	–0.831071 * *(0.172063)*
FAIRFAX		–29,321.43 *(18,632.21)*	–0.128136 *(0.189931)*	–0.189547 *(0.195411)*
HOUGH		–43,795.94 * *(15,414.09)*	–0.309637 * *(0.155037)*	–0.338367 * *(0.150414)*
SHAKER		–73,276.38 * *(25,345.29)*	–0.394678 ** *(0.237182)*	–0.354204 *(0.244025)*
WOODHILL		–22,138.21 *(16,883.96)*	0.12293 *(0.180969)*	–0.035860 *(0.166632)*

*Significant at the 5 percent level of significance. **Significant at the 10 percent level of significance.

† These variables are only used in the log-log specification. All others are entered as levels variables unless otherwise noted.

Levels model: adjusted $R^2 = 0.624$, F (Wald statistic) = 16.29 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (5, 124) = 6.03 (prob > F = 0.0000).

Log-Level model: adjusted $R^2 = 0.517$, F (Wald statistic) = 19.34 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (5, 124) = 4.57 (prob > F = 0.0007).

Log-Log model: adjusted $R^2 = 0.521$, F (Wald statistic) = 20.22 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (5, 125) = 6.29 (prob > F = 0.0000).

Table 6.6. 2010–2011 Cross-Section with Distance as a Dummy Variable

Variable	Description	2010–2011 (n=140)		
		Level	Log-Level	Log-Log†
		Coefficient *Robust Std Error*	Coefficient *Robust Std Error*	Coefficient *Robust Std Error*
PRICE	Dependent variable: sale price of home	PRICE	LN_PRICE	LN_PRICE
Constant	Constant term in regression equation	94,449.0 * *(38,094.28)*	11.1665 * *(0.397403)*	4.74241 * *(1.521949)*
DIST1_dum	**Distance from home to nearest BRT station between 0 and 2,640 feet (0.5 mile); base case is between 2,640.01 and 7,290 feet (0.5 mile to 1.5 miles)**	**7,017.535** ***(9,598.617)***	**0.061648** ***(0.098190)***	**0.036498** ***(0.100530)***
AREA (LN_AREA†)	Size of home's living area in square feet	19.79256 * *(9.350814)*	0.000182 ** *(0.000099)*	0.564676 * *(0.8747)*
BEDROOMS	Number of bedrooms	–5,485.929 *(6,346.215)*	–0.030207 *(0.059762)*	–0.065912 *(0.061699)*
AGE	Age of the home in years	–533.1638 * *(157.0175)*	–0.005207 * *(0.001501)*	–0.005281 * *(0.001431)*
COND	Condition of the home; likert scale	8,496.831 *(5,649.106)*	0.091283 *(0.057664)*	0.085887 *(0.055831)*
MDHHINC (LN_MDHHINC†)	Median household income for census tract that includes the	2.397745 * *(0.480043)*	0.000015 * *(3.91e-06)*	0.319551 * *(0.126416)*
CCTT	Travel time in minutes from the nearest BRT station to the Cleveland Clinic	–1,916.875 * *(968.3005)*	–0.008822 *(0.009477)*	–0.007138 *(0.009664)*
CBDTT	Travel time in minutes from the nearest BRT station to Tower City	–72.86952 *(498.5112)*	–0.002094 *(0.004823)*	–0.005769 *(0.004595)*
OWNOCC	Percent of owner-occupied homes in the census tract that includes the property	–944.0878 * *(300.4223)*	–0.006796 * *(0.002841)*	–0.007064 * *(0.003393)*
CENTRAL	Dummy variables: Take value of 1 if property is located in the listed city neighborhood; 0 otherwise (base case is outside of these neighborhoods)	–78,420.16 * *(15,891.61)*	–0.790334 * *(0.175612)*	–0.827966 * *(0.172671)*
FAIRFAX		–30,165.49 *(18,548.62)*	–0.130179 *(0.188755)*	–0.189623 *(0.194947)*
HOUGH		–44,703.26 * *(14,820.16)*	–0.319379 * *(0.150760)*	–0.343644 * *(0.147448)*
SHAKER		–64,639.4 * *(22,233.7)*	–0.359811 ** *(0.217722)*	–0.351131 *(0.233458)*
WOODHILL		–13,056.19 *(15,783.66)*	0.051951 *(0.158928)*	–0.033969 *(0.1562703)*

*Significant at the 5 percent level of significance. **Significant at the 10 percent level of significance.

† These variables are only used in the log-log specification. All others are entered as levels variables unless otherwise noted.

<u>Levels model:</u> adjusted $R^2 = 0.627$, F (Wald statistic) = 16.35 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (5, 125) = 10.51 (prob > F = 0.0000).

<u>Log-Level model:</u> adjusted $R^2 = 0.521$, F (Wald statistic) = 21.37 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (5, 125) = 7.56 (prob > F = 0.0000).

<u>Log-Log model:</u> adjusted $R^2 = 0.521$, F (Wald statistic) = 20.41 (prob > F = 0.0000). Neighborhood dummies are jointly significant using the heteroskedastic-robust F statistic, F (5, 125) = 6.47 (prob > F = 0.0000).

indicating that the homes closer to BRT stations would have sale prices greater than those farther away. The positive result might be expected if the station locations, which had been announced, were known to home buyers and considered by them in the area to be an amenity. For a negative coefficient in 2004, it would be expected that the coefficient would turn positive in 2008–2009 after the BRT service began operating, and would persist in 2010–2011. For a positive coefficient in 2004, it would be expected that the magnitude might increase after the BRT began operating in 2008 and persist through the 2008–2009 and 2010–2011 models.

As the information in Table 6.2 shows, the coefficient on DIST1_dum for the log-log model is negative and significant at the 8.5 level of significance. As shown previously in this chapter, the exact interpretation of this coefficient, –0.108001, would be as follows: the percent change is sale price is equal to $100(e^{-0.108001} - 1) = 100(0.897627 - 1) = 100(-0.102373) = -10.237$ percent. By this interpretation, a home one half-mile or less from the nearest BRT station would be estimated to sell for approximately 10.2 percent less than a home more than a half-mile away, or approximately $12,450 less based on the mean sale price in 2004. As expected, in the 2008–2009 log-log model with distance as a dummy variable shown in Table 6.4, the coefficient on DIST1_dum has turned positive. For this time period and these data, the coefficient for DIST1_dum in the log-log model is equal to 0.344018, which means that a home a half-mile or less from the nearest station would have a sale price 41.1 percent more than one farther away

[100($e^{0.344018}$ – 1) = 100(1.41060 – 1) = 100(0.41060) = 41.1 percent], or approximately $46,203 based on the mean sale price in 2008 and 2009). This result is consistent with the hypothesis of this research. However, for the 2010–2011 time period, the coefficient remains positive, but is smaller in magnitude and not statistically significant, as shown in Table 6.6. The coefficient of 0.03650 would be interpreted as follows: a home located a half-mile or less from the nearest BRT station would have an estimated sale price 3.7 percent more than a home outside of that distance [100($e^{0.03650}$ – 1) = 100(1.03717 – 1) = 100(0.03717) = 3.7 percent], or approximately $3,436 less based on the mean sale price in 2010 and 2011. This result was unexpected, but may be due to the continuing and deepening irregularities in the area housing market at the time. As mentioned previously, the median sale price of single-family homes in the city of Cleveland fell 23 percent from $73,400 in 2009 to $56,500 in 2011. The median sale price in 2011 represented the lowest point since prices began falling after 2005. Median sale prices for single-family homes in Cleveland began rising again in 2012 and, after a small decline in 2013, began increasing again to $76,000 as of August 2015 [66].

Returning to the log-log model with distance as a dummy variable in Table 6.2, all of the remaining estimated coefficients in this 2004 model are statistically significant, except for one of the neighborhood dummy variables (UNIVERSITY). However, as discussed previously, all six neighborhood dummy variables were found to be jointly significant using the heteroskedastic-robust *F* statistic [F(6, 178) = 6.58, prob > F =

0.0000]. Using the estimated coefficients on the neighborhood dummy variables, the following interpretations were developed (note that the mean sale price in 2004 is $133,064 in 2011 dollars):

- Homes selling in the Central neighborhood have estimated sale prices 43.3 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.567455} - 1) = 100(0.56697 - 1) = 100(-0.43303) = -43.3$ percent].
- Homes selling in the Fairfax neighborhood have estimated sale prices 34.5 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.423603} - 1) = 100(0.65468 - 1) = 100(-0.34532) = -34.5$ percent].
- Homes selling in the Glenville neighborhood have estimated sale prices 45.4 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.604837} - 1) = 100(0.54616 - 1) = 100(-0.45384) = -45.4$ percent].
- Homes selling in the Goodrich neighborhood have estimated sale prices 50.5 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.703082} - 1) = 100(0.49506 - 1) = 100(-0.50494) = -50.5$ percent].
- Homes selling in the Hough neighborhood have estimated sale prices 49.0 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.672981} - 1) = 100(0.51019 - 1) = 100(-0.48981) = -49.0$ percent].

- Homes selling in the University neighborhood have estimated sale prices 8.7 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.090528} - 1) = 100(0.91345 - 1) = 100(-0.08655) = -8.7$ percent].

Regarding the other variables in this model, interpretations are as follows:

- An additional bedroom decreases the sale price of a home, all else constant, by 5.1 percent [$100(e^{-052638} - 1) = 100(0.94872 - 1) = 100(-0.05128) = -5.1$ percent]. This interpretation is intuitive because as the number of bedrooms increases, holding all other variables constant, including the square feet of living area, the rooms must become smaller. The smaller rooms result in a slightly lower sale price. It should be noted that an interaction term between the number of bedrooms and living area was not found to be significant in any of the models shown in Tables 6.1 through 6.6.
- A one year increase in the age of the home reduces the estimated sale price by 0.46 percent [$100(e^{-0.0046} - 1) = 100(0.99541 - 1) = 100(-0.00459) = -0.46$ percent].
- An improvement in the condition of the home by one level on the likert scale increases the estimated sale price by 7.8 percent [$100(e^{0.074933} - 1) = 100(1.07781 - 1) = 100(0.07781) = 7.8$ percent].
- A one percentage point increase in the number of homes within a census tract that were built before 1940 reduces the estimated sale price by 0.82 percent [$100(e^{-0.0081940} - 1) = 100(0.99184 - 1) = 100(-0.00816) = -0.82$ percent].

- A one percent increase in the square feet of living area increases the estimated sale price by 0.43 percent (0.431117).
- A one percent increase in the median household income of the census tract containing a home increases the estimated sale price by 0.38 percent (0.382661).

Overall, the 2004 log-log model with distance as a dummy variable has an adjusted R^2 value of 0.667, which is somewhat useful as a measure of goodness-of-fit. Further, the robust Wald statistic of 38.86 is significant (prob > F = 0.0000).

For the log-log model shown in Table 6.4, the 2008–2009 cross-section, the only two insignificant variables are median household income, the travel time from the nearest station to the Cleveland Clinic main campus at the East 93rd Station, and the travel time from the nearest station to the Tower City station in the CBD. Each of these insignificant variables has the expected signs, however. Interpretations of the remaining significant variables are as follows:

With a somewhat smaller sample size in 2008–2009 than in 2004 (127 and 192, respectively), there were fewer dummy variables needed. As discussed earlier in this chapter, all three neighborhood dummy variables were found to be jointly significant using the heteroskedastic-robust *F* statistic at the 2.35 percent exact level of significance [F(3, 114) = 3.28, prob > F = 0.0235]. Using the estimated coefficients on the neighborhood dummy variables, the following interpretations are presented below (note that the mean sale price in 2008 and 2009 is $114,689 adjusted to 2011 dollars) :

- Homes selling in the Central neighborhood have estimated sale prices 36.6 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.455854} - 1) = 100(0.63391 - 1) = 100(-0.36609) = -36.6$ percent].
- Homes selling in the Fairfax neighborhood have estimated sale prices 38.5 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.485879} - 1) = 100(0.61516 - 1) = 100(-0.38484) = -38.5$ percent].
- Homes selling in the Hough neighborhood have estimated sale prices 40.3 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.516659} - 1) = 100(0.59651 - 1) = 100(-0.40349) = -40.3$ percent].

For the other variables in this 2008–2009 model, interpretations are as follows:

- An additional bedroom decreases the sale price of a home, all else constant, by 10.1 percent [$100(e^{-0.106635} - 1) = 100(0.89885 - 1) = 100(-0.10115) = -10.1$ percent]. As discussed above, this interpretation is intuitive because, as the number of bedrooms increases, holding all other variables constant (including the square feet of living area), the rooms therefore become smaller. Smaller rooms result in a lower sale price, all else constant. It should be noted again that an interaction term between the number of bedrooms and living area was not found to be significant in any of the models shown in Tables 6.1 through 6.6.

- A one year increase in the age of the home reduces the estimated sale price by 0.42 percent [$100(e^{-0.004238} - 1) = 100(0.99577 - 1) = 100(-0.00423) = -0.42$ percent].
- An improvement in the condition of the home by one level on the likert scale increases the estimated sale price by 11.5 percent [$100(e^{0.109153} - 1) = 100(1.11533 - 1) = 100(0.11533) = 11.5$ percent].
- A one percentage point increase in the number of homes within a census tract that were built before 1940 reduces the estimated sale price by 1.1 percent [$100(e^{-0.011387} - 1) = 100(0.98868 - 1) = 100(-0.01132) = -1.1$ percent].
- A one percent increase in the square feet of living area increases the estimated sale price by 0.93 percent (0.925687).

Overall, the 2008–2009 log-log model with distance as a dummy variable has an adjusted R^2 value of 0.578, which is somewhat useful as a measure of goodness-of-fit. In addition, the robust Wald statistic of 23.11 is statistically significant (prob > F = 0.0000).

The 2010–2011 log-log model shown in Table 6.6 represents the final model to be discussed in this section. As shown in Table 6.6, there are several insignificant variables in this model. As discussed earlier, it is likely that the unfavorable macroeconomic situation and housing market in the study area during this time impacted these results in ways that could not easily be measured. Insignificant variables include the DIST1_dum, the number of bedrooms, the condition of a home, and the two travel time variables, CCTT and CBDTT. The condition variable is significant at the 12.6 level of significance, however. In addition, two of the neighborhood dummy variables were not significant, although all five were found to be jointly significant using the heteroskedastic-robust *F* statistic [$F(5, 125) = 6.47$, prob > F = 0.0000].

The 2010–2011 data set had an even smaller sample size than the other two time periods when staying within one mile of BRT stations. Therefore, the distance was expanded to 1.5 miles from the stations (7,920 feet) to increase the sample size to more than 100 homes (the final sample yielded 140 observations). Using the estimated coefficients on the neighborhood dummy variables, the following interpretations are summarized below (note that the mean sale price in 2010 and 2011 is $93,666 adjusted to 2011 dollars):

- Homes selling in the Central neighborhood have estimated sale prices 56.3 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.827966} - 1) = 100(0.43694 - 1) = 100(-0.56306) = -56.3$ percent].
- Homes selling in the Fairfax neighborhood have estimated sale prices 17.3 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.189623} - 1) = 100(0.82727 - 1) = 100(-0.17273) = -17.3$ percent].

- Homes selling in the Hough neighborhood have estimated sale prices 29.1 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.343644} - 1) = 100(0.70918 - 1) = 100(-0.29082) = -29.1$ percent].
- Homes selling in the Shaker neighborhood have estimated sale prices 29.6 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.351131} - 1) = 100(0.70389 - 1) = 100(-0.29611) = -29.6$ percent].
- Homes selling in the Woodhill neighborhood have estimated sale prices 3.3 percent lower than homes outside the neighborhoods included in this model [$100(e^{-0.033969} - 1) = 100(0.96660 - 1) = 100(-0.03340) = -3.3$ percent].

For the other significant variables in the 2010–2011 model, interpretations are as shown below:

- A one year increase in the age of the home reduces the estimated sale price by 0.53 percent [$100(e^{-0.005281} - 1) = 100(0.99473 - 1) = 100(-0.00527) = -0.53$ percent].
- An improvement in the condition of the home by one level on the likert scale increases the estimated sale price by 9.0 percent [$100(e^{0.085887} - 1) = 100(1.08968 - 1) = 100(0.08968) = 9.0$ percent]. This variable is included in this list because its coefficient is significant at the 12.6 level of significance.
- A one percentage point increase in the number of homes within a census tract that are owner-occupied reduces the estimated sale price by 0.70 percent

$[100(e^{-0.007064} - 1) = 100(0.99296 - 1) = 100(-0.00704) = -0.70$ percent]. Perhaps revealing new characteristics about the housing market in 2010 and 2011, the variable for the percentage of homes within a census tract built before 1940 was highly insignificant in all 2010-2011 models. Instead, the variable representing the percentage of homes within a census tract that are occupied by their owners became significant in all 2010-2011 models tested.

- A one percent increase in the square feet of living area increases the estimated sale price by 0.56 percent (0.564676).

The 2010–2011 log-log model with distance as a dummy variable has an adjusted R^2 value of 0.521, which is somewhat useful as a measure of goodness-of-fit. Also, the robust Wald statistic of 20.41 is statistically significant (prob > F = 0.0000).

Discussion and Conclusion

The hedonic regression models analyzed for this dissertation seek to paint a picture of sale prices of single-family homes along the Euclid Avenue BRT corridor in Cleveland, Ohio and how the determinants of those prices have changed over time. In particular, three time periods were examined: 2004, the year construction began; 2008–2009, after the HealthLine BRT service began operations; and 2010–2011, the latest year for which sales data are available. Variables such as the square feet of living area, the home's age, and the home's condition were reliable determinants of the sale prices of single-family homes within 1.0 and 1.5 miles of the BRT stations for all time periods included. The variables of interest were related to the distance of the homes to the nearest BRT station along the corridor. The primary purpose of this research was to find that the presence of the BRT service and access to the stations impacted the sale prices of single-family homes in a statistically significant, positive way. Despite the documented decline in median sale prices of single-family homes in the city of Cleveland from 2005 to 2011, this work investigated whether a portion of sale price could be explained by proximity to a BRT station and how the effect may be expected to change over time.

Beginning in 2004, the evidence presented in this dissertation found that home prices were lower for homes within one half-mile of the nearest proposed BRT station than for homes farther away (recall that the station locations were known at the time construction began). Even though station locations were made available, this result was not entirely unexpected, given that many home buyers may not necessarily have been aware of the scope of the Euclid Avenue project and the BRT services (which were still four years in the future). As expected according to the hypothesis of this work, the results changed in the 2008–2009 models after the BRT service opened in 2008. In the 2008–2009 analysis, now homes within one half-mile of the station had estimated sale prices greater than those farther away, and the result is statistically significant. The result is promising, although due to the recession of 2007–2009 the models were not quite as strong as the 2004 models. Further, the number of home sales in the study area dropped, and the income variable, the coefficient of which should have a positive sign in a model of housing price, suddenly became highly insignificant in all models. Clearly, the impacts of the recession and subsequent housing crisis were impacting sale prices. Within the city of Cleveland, the median sale price for single-family homes fell 20 percent, a significant amount, from 2005 to 2009. It is therefore likely that the typical variables used to determine housing values were not sufficient to fully explain what was happening to sale prices. This issue became exacerbated in the 2010–2011 analysis. In just two more years, from 2009 to 2011, the median single-family home sale price in Cleveland fell an additional 23 percent, to a low of $56,500 [66]. Still fewer homes were sold, resulting in a smaller sample size for the two years of 2010 and 2011 (to address this, the distance was increased from 1.0 mile to 1.5 miles away from the nearest BRT stations). It was also not expected that the effect of distance in the models would essentially disappear. While the results of 2008–2009 analysis were promising regarding the impact of proximity to the BRT stations, this result did not persist in the 2010–2011 analysis.

There are some shortcomings to the analysis conducted for this dissertation. Significant among them is the relatively small number of homes sold during the time periods of study. In addition, it was not feasible to properly account for the impacts of the recession and housing crisis, which impacted northeast Ohio significantly, in the models, and so there are likely some omitted variables. Ideally, such a study would occur outside a time period of such irregularities. However, after a massive investment and major reconstruction of the Euclid Avenue corridor, the HealthLine BRT service began operating in October 2008. The service was very well-received and has been highly popular from the start. Developers and others in the private sector took notice of the positive changes along the corridor and the question arose as to how soon the impacts might begin to be capitalized into home prices in the surrounding communities.

The research presented in this dissertation provides an excellent starting point for future work. As median sale prices for single-family homes in Cleveland have been rising steadily since 2013 (reaching $76,000 in August 2015 [66]), the logical next step in this work would be to acquire the sales data through at least 2015 (or 2016 and perhaps 2017) for the single-family homes in the study area and determine, within a stable housing market, whether distance to a BRT station can still explain a portion of sale prices in a statistically significant, positive way. Also, additional variables can be collected and added to the models, such as additional distance variables, housing characteristics, and even other variables expected to negatively affect sale prices such as crime statistics or other nuisance effects (crime statistics in a usable format were not available for this research). Different analytical techniques could be employed in future studies, such as spatial regression analysis, which is becoming easier to handle with the latest GIS software and the availability of the parcel data representing the properties.

Finally, other types of housing can be examined in future work, such as multi-family units and apartments (rents could be used instead of sale prices), several of which have been recently constructed along the corridor (since 2011). Commercial property data could be analyzed, as well. Finally, future research could examine other measures of economic activity along the Euclid corridor, such as changes in property taxes, for example.

It is clear that, while this dissertation contributes to the still small body of literature on the impacts of BRT services that operate in the U.S., there is more work that can be undertaken. As communities in the U.S. continue to explore various public transit investments and modes, it is important that the best and most up-to-date information is available to aid in the decision-making process; this research contributes to that end.

REFERENCES

[1] U.S. Government Accountability Office, "Highway and Transit Investments: Options for Improving Information on Projects' Benefits and Costs and Increasing Accountability for Results," Rep. GAO-05-172, January 2005. [Online]. Available: http://www.gao.gov/new.items/d05172.pdf.

[2] U.S. Government Accountability Office, "Bus Rapid Transit: Projects Improve Transit Service and Can Contribute to Economic Development," Rep. GAO-12-811, July 2012. [Online]. Available: http://www.gao.gov/assets/600/592973.pdf.

[3] T. M. Ben-Akiva Moshe, "Comparing Ridership Attraction of Rail and Bus," *Transport Policy,* vol. 19, no. 2, pp. 107-116, 2002.

[4] K. D. Milena Scherer, "Bus or Rail: An Approach to Explain the Psychological Rail Factor," *Journal of Public Transportation,* vol. 15, no. 1, 2012.

[5] "Smart Growth America," 22 August 2013. [Online]. Available: http://www.smartgrowthamerica.org/tag/light-rail/.

[6] Federal Transit Administration, U.S. Department of Transportation, "Capital Investment Program: New Starts, Small Starts and Core Capacity Improvements," 2014. [Online]. Available: http://www.fta.dot.gov/documents/MAP-21_Fact_Sheet_-_Fixed_Guideway_Capital_Investment_Grants.pdf.

[7] National Bus Rapid Transit Institute, "National BRT Systems," [Online]. Available: http://nbrti.org/systems.html.

[8] Federal Transit Administration, U.S. Department of Transportation, "NTD Glossary," [Online]. Available: http://www.ntdprogram.gov/ntdprogram/Glossary.htm.

[9] Federal Transit Administration, U.S. Department of Transportation, "Bus Rapid Transit," [Online]. Available: http://www.fta.dot.gov/12351_4240.html.

[10] D. Hinebaugh and R. Diaz, "Characteristics of Bus Rapid Transit for Decision-Making," Federal Transit Administration, U.S. Department of Transportation, Washington, D.C., Rep. FTA-FL-26-7109.2009.1, 2009.

[11] V. Perk and M. Catalá, "Land Use Impacts of Bus Rapid Transit: Effects of Station Proximity on Property Values of Single Family Homes Along Pittsburgh's Martin Luther King, Jr., East Busway," Federal Transit Administration, U.S. Department of Transportation, Washington, D.C., Rep. FTA-FL-26-7109.2009.6, 2009.

[12] V. Perk, S. Bovino, M. Catalá, S. Reader, and S. Ulloa, "Silver Line Bus Rapid Transit in Boston, Massachusettes: Impacts on Sale Prices of Condominiums Along Washington Street," *Transportation Research Record: Journal of the Transportation Research Board, No. 2350,* pp. 72-79, 2013.

[13] C. Cochran, "Overview of Greenlight Pinellas," presented at the Institute of Transportation Engineers Student Chapter Seminar at the University of South Florida, Tampa, 2014.

[14] R. Trigaux, "Separating Fact from Fantasy: Tampa Bay's Economy Will Suffer without Strong Mass Transit," *Tampa Bay Times*, October 3, 2014.

[15] A. Anas, R. Arnott, and K. A. Small, "Urban Spatial Structure," *Journal of Economic Literature,* vol. 36, no. 3, pp. 1426-1464, 1998.

[16] R. F. Muth, "Economic Change and Rural-Urban Land Conversions," *Econometrica,* vol. 29, no. 1, pp. 1-23, January 1961.

[17] W. Alonso, *Location and Land Use*, Cambridge: Harvard University Press, 1964.

[18] E. S. Mills, "An Aggregative Model of Resource Allocation in a Metropolitan Area," in *The American Economic Review, Papers and Proceedings of the Seventy-ninth Annual Meeting of the American Economic Association*, 1967.

[19] W. C. Wheaton, "A Comparative Static Analysis of Urban Spatial Structure," *Journal of Economic Theory,* pp. 223-237, 1974.

[20] J. K. Brueckner, "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model," in *Handbook of Regional and Urban Economics, Volume II*, E. S. Mills, Ed., Elsevier Science Publishers B.V., 1987.

[21] M. Geshkov and J. S. DeSalvo, "The Effect of Land-Use Controls on the Spatial Size of U.S. Urbanized Areas," *Journal of Regional Science,* vol. 52, no. 4, pp. 648-675, 2012.

[22] H. Mohring, "Land Values and the Measurement of Highway Benefits," *Journal of Political Economy,* vol. 69, no. 3, pp. 236-249, June 1961.

[23] R. Andersson and A. Samartin, "An Extension of Mohring's Model for Land Rent Distrbution," *Journal of Urban Economics,* vol. 18, pp. 143-160, 1985.

[24] W. Vincent and L. Callaghan Jerram, "Bus Rapid Transit and Transit Oriented Development," Breakthrough Technologies Institute, 2008.

[25] A. Stokenberga, "Does Bus Rapid Transit Influence Urban Land Development and Property Values: A Review of the Literature," *Transport Reviews: A Transnational Transdisciplinary Journal,* vol. 34, no. 3, pp. 279-296, April 2014.

[26] J. A. Mullins III , E. J. Washington, and R. B. Stokes, "Land Use Impacts of the Houston Transitway System," *Transportation Research Record 1237,* 1990.

[27] A. C. Nelson, S. Kannan, B. Appleyard, M. Miller, G. Meakins and R. Ewing, "Bus Rapid Transit and Economic Development," in *Transportation Research Board 2012 Annual Meeting Compendium of Papers*, Washington, D.C., Conference Proceedings, 2012.

[28] D. A. Rodriguez and F. Targa, "Value of Accessibility to Bogotá's Bus Rapid Transit System," *Transport Reviews,* pp. 587-610, 2004.

[29] R. Muñoz-Raskin, "Walking Accessibility to Bus Rapid Transit: Does it Affect Property Values? The Case of Bogota, Columbia," *Transport Policy,* pp. 72-84, 2010.

[30] R. Cervero and C. D. Kang, "Bus Rapid Transit Impacts on Land Uses and Land Values in Seoul, Korea," *Transport Policy* , pp. 102-116, 2011.

[31] C. Mulley, "Accessibility and Residential Land Value Uplift: Identifying Spatial Variations in the Accessibility Impacts of a Bus Transitway," *Urban Studies,* August 2013.

[32] EMBARQ, Global BRT Data, 2015. [Online]. Available: http://brtdata.org/. [Accessed 19 June 2015].

[33] Land Transport Authority, "Passenger Transport Mode Shares in World Cities," 2011. [Online]. Available: http://www.lta.gov.sg/ltaacademy/doc/J11Nov-p60PassengerTransportModeShares.pdf. [Accessed 19 June 2015].

[34] Northeast Ohio Areawide Coordinating Agency, "NOACA 2014 Annual Report -- Planning with Vision," 2014. [Online]. Available: http://noaca.org/modules/showdocument.aspx?documentid=6177. [Accessed 17 June 2015].

[35] American Association of State Highway and Transportation Officials, "Commuting in America 2013, Brief 10: Commuting Mode Choice," AASHTO, 2013.

[36] Kittelson and Associates, Inc., Herbert S. Levinson Transportation Consultants, DMJM+Harris, "Transit Cooperative Research Program - TCRP Report 118: Bus Rapid Transit Practitioner's Guide," Transportation Research Board of the National Academies, Washington, D.C., 2007.

[37] D. H. Gatzlaff and M. T. Smith, "The Impact of the Miami Metrorail on the Value of Residences Near Station Locations," *Land Economics,* pp. 54-66, 1993.

[38] Gruen, Gruen and Associates, "The Effect of CTA and Metra Stations on Residential Property Values," Regional Transportation Authority/ JHK and Associates, 1997.

[39] H. Chen, A. Rufolo and K. J. Dueker, "Measuring the Impact of Light Rail Systems on Single-Family Home Values: A Hedonic Approach with Geographic Information System Application," *Transportation Research Board* , pp. 38-43, 1998.

[40] N. Baum-Snow and M. E. Kahn, "The Effects of New Public Projects to Expand Urban Rail Transit," *Journal of Public Economics,* pp. 241-263, 2000.

[41] D. R. Bowes and K. R. Ihlanfeldt, "Identifying the Impacts of Rail Transit Stations on Residential Property Values," *Journal of Urban Economics* , pp. 1-25, 2001.

[42] T. Garrett, "Light-Rail Transit in America," Federal Reserve Bank of St. Louis, 2004. [Online]. Available: http://www.cfte.org/news/garrett.pdf.

[43] D. B. Hess and T. M. Almeida, "Impact of Proximity to Light Rail Rapid Transit on Station-area Property Values in Buffalo, New York," *Urban Studies,* pp. 1041-1068, 2007.

[44] A. Kent and J. Parilla, "Did the Hiawatha Light Rail Line Increase Single-Family Residential Property Values?" 2008. [Online]. Available: http://www.reconnectingamerica.org/assets/Uploads/20080610parilla.pdf. [Accessed 18 September 2015].

[45] S. Yan, E. Delmelle, and M. Duncan, "The Impact of a New Light Rail System on Single-Family Property Values in Charlotte, NC," *The Journal of Transport and Land Use,* vol. 5, no. 2, pp. 60-67, 2012.

[46] R. Cervero and J. Landis, "Twenty Years of the Bay Area Rapid Transit System: Land Use and Development Impacts," Department of City and Regional Planning, Institute of Urban and Regional Development, University of California, Berkeley, 1995.

[47] R. Cervero, "Rail Transit and Joint Development: Land Market Impacts in Washington, D.C. and Atlanta," *Journal of the American Planning Association,* vol. 60, no. 1, pp. 83-94, Winter 1994.

[48] R. R. Weinberger, "Light Rail Proximity: Benefit or Detriment in the Case of Santa Clara County, California?," *Transportation Research Record of the Transportation Research Board,* vol. 1747, pp. 104-113, 2001.

[49] R. Cervero and M. Duncan, "Transit's Value-Added: Effects of Light and Commuter Rail Services on Commercial Land Values," in *81st Annual Meeting of the Transportation Research Board Compendium of Papers*, Washington, D.C., 2002.

[50] R. Cervero and M. Duncan, "Land Value Impacts of Rail Transit Services in San Diego County," National Association of Realtors, 2002.

[51] V. Mays, "Health on Wheels," *Landscape Architecture Magazine,* vol. 101, no. 6, 2011.

[52] Greater Cleveland Regional Transit Authority, "Greater Cleveland Regional Transit Authority Euclid Corridor Transportation Project - Bus Rapid Transit Before and After Study," 2012.

[53] R. J. Hill, "Hedonic Price Indexes for Residential Housing: A Survey, Evaluation and Taxonomy," *Journal of Economic Surveys,* vol. 27, no. 5, pp. 879-914, 2013.

[54] A. C. Goodman, "Andrew Court and the Invention of Hedonic Price Analysis," *Journal of Urban Economics,* vol. 44, pp. 291-298, 1998.

[55] L. Anselin, "Spatial Econometrics," in *A Companion to Theoretical Econometrics*, B. H. Baltagi, Ed., Blackwell Publishing, Ltd., 2003.

[56] S. Rosen, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy,* vol. 82, no. 1, pp. 34-55, 1974.

[57] A. Can, "Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban Economics,* vol. 22, pp. 453-474, 1992.

[58] R. Halvorsen and H. O. Pollakowski, "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics,* vol. 10, pp. 37-49, 1981.

[59] E. G. Irwin, "The Effects of Open Space on Residential Property Values," *Land Economics,* pp. 465-480, 2002.

[60] N. C. Poudyal, D. G. Hodges and C. D. Merrett, "A Hedonic Analysis of the Demand for and Benefits of Urban Recreation Parks," *Land Use Policy,* pp. 975-983, 2009.

[61] C. A. Agostini and G. A. Palmucci, "The Anticipated Capitalisation Effect of a New Metro Line on Housing Prices," *Fiscal Studies,* pp. 233-256, 2008.

[62] J. G. Xin, H. MacDonald, and S. Ghosh, "Assessing the Impact of Rail Investment on Housing Prices in North-West Sydney," in *18th Annual Pacific-Rim Real Estate Society Conference Proceedings*, Adelaide, 2012.

[63] R. M. Sakia, "The Box-Cox Transformation Technique: A Review," *The Statistician,* vol. 41, pp. 169-178, 1992.

[64] E. Cassel and R. Mendelsohn, "The Choice of Functional Forms for Hedonic Price Equations: Comment," *Journal of Urban Economics,* vol. 18, pp. 135-142, 1985.

[65] J. M. Wooldridge, *Introductory Econometrics: A Modern Approach*. 2nd ed., Ohio: Thomson South-Western, 2003.

[66] "Cleveland Home Prices and Values," [Online]. Available: http://www.zillow.com/cleveland-oh/home-values/. [Accessed 27 October 2015].

APPENDIX A:

DATA FREQUENCIES

This Appendix contains the distributions for the variables used in this research for the relevant years of home sales, including 2004, 2008–2009, and 2010–2011. As described in Chapter Four, home sales for 2008–2009 and 2010–2011 were combined because of the relatively low number of homes sold in each of those individual years. In the few cases where a home sold in both 2008 and 2009, or 2010 and 2011, the latter year was used (i.e., 2009 or 2011). The data shown are for two miles within the BRT corridor on Euclid Avenue.

The Chi-Square Goodness of Fit test was used to compare the distributions of properties sold in each cross-section group with the total stock of all single-family homes within two miles of the corridor. To reduce the incidence of sample selection bias, the distributions of homes sold in each cross-section group should be similar to the distribution of all homes in the study area. In this case, it is desirable to accept the null hypothesis that the distributions are equal ($p > 0.05$). Therefore, in Tables A.1 through A.14, the Chi-Square Goodness of Fit statistics are shown, along with the relevant p values. When the p value is greater than 0.05, it can be said that the distribution of single-family homes sold in a particular cross-section year is very similar to the distribution of all single-family homes in the study area (at the five percent level of significance). While there are some differences among the distributions, overall it appears that the distribution of single-family homes within two miles of the BRT corridor sold in each cross-section year is relatively similar to the distribution for the stock of all single-family homes within two miles of the corridor.

Tables A.15 through A.17 show the distributions of Cleveland city wards, named neighborhoods, and other bounded areas for the homes sold in 2004, 2008–2009, and 2010–2011.

Table A.1. Distribution of Lot Sizes (Square Feet)

Lot Size Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Less than 2,000 sq. ft.	3.4%	4.2%	7.51%	3.7%
2,000 – 4,999 sq. ft.	65.0%	72.3%	69.6%	65.6%
5,000 – 9,999 sq. ft.	28.0%	21.5%	21.2%	27.5%
10,000 – 19,999 sq. ft.	2.6%	1.7%	1.5%	2.8%
20,000 sq. ft. or greater	1.0%	0.3%	0.2%	0.4%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	0.454 $p = 0.797$	31.525 $p = 0.000$	54.137 $p = 0.000$	

Table A.2. Distribution of Living Area Sizes (Square Feet)

Area Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Less than 1,000 sq. ft.	6.2%	6.8%	5.4%	6.0%
1,000 – 1,499 sq. ft.	49.7%	54.5%	57.1%	53.6%
1,500 – 1,999 sq. ft.	28.3%	28.3%	28.8%	30.0%
2,000 – 2,999 sq. ft.	13.5%	9.0%	7.4%	9.0%
3,000 sq. ft. or greater	2.3%	2.0%	1.3%	1.4%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	24.028 $p = 0.000$	3.905 $p = 0.272$	5.229 $p = 0.156$	

Table A.3. Distribution of Number of Bedrooms

Bedroom Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
1	0.1%	0.2%	0.5%	0.5%
2	10.7%	11.8%	10.0%	10.2%
3	48.8%	48.8%	53.9%	51.6%
4	30.2%	29.0%	27.6%	28.6%
5 or more	10.2%	10.2%	8.1%	9.1%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	3.245 $p = 0.355$	5.604 $p = 0.133$	2.107 $p = 0.550$	

Table A.4. Distribution of Number of Full Bathrooms

Full Bathroom Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
1	70.5%	79.5%	81.3%	78.4%
2	27.3%	18.7%	16.4%	19.6%
3 or more	2.2%	1.8%	2.3%	2.0%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	29.633 $p = 0.000$	6.259 $p = 0.044$	5.879 $p = 0.053$	

Table A.5. Distribution of Number of Half Bathrooms

Half Bathroom Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
0	76.2%	83.3%	83.5%	78.9%
1	22.8%	16.5%	15.7%	20.3%
2 or more	1.0%	0.2%	0.8%	0.8%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	3.360 $p = 0.067$	14.829 $p = 0.000$	11.088 $p = 0.001$	

Table A.6. Distribution of Home Condition

Condition Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Excellent	0.0%	0.0%	0.0%	0.0%
Very Good	2.5%	1.2%	1.9%	1.2%
Good	19.6%	12.3%	10.5%	13.0%
Average	34.3%	31.5%	35.8%	35.7%
Fair	25.9%	32.7%	31.1%	31.2%
Poor	13.3%	16.0%	15.7%	13.9%
Very Poor	4.4%	6.3%	5.0%	5.0%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	41.841 $p = 0.000$	81.351 $p = 0.000$	3.529 $p = 0.317$	

Table A.7. Distribution of Home Age (Year Built)

Age Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
1899 or earlier	9.2%	10.3%	12.6%	10.3%
1900 – 1939	71.1%	77.5%	72.9%	71.2%
1940 – 1959	1.0%	1.2%	0.8%	1.5%
1960 – 1979	0.6%	1.8%	4.4%	0.9%
1980 – 1989	0.2%	0.2%	0.1%	0.4%
1990 – 1999	2.6%	2.5%	4.0%	6.5%
2000 – 2010	15.3%	6.5%	5.1%	9.2%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	54.041 $p = 0.000$	46.851 $p = 0.000$	60.340 $p = 0.000$	

Table A.8. Distribution of Home Style

Style Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Bungalow	1.8%	2.3%	3.0%	2.5%
Cape Cod	13.3%	14.1%	11.4%	12.8%
Colonial	78.7%	76.7%	77.9%	79.5%
Contemporary	1.6%	0.3%	1.0%	0.5%
Ranch	2.7%	3.0%	1.5%	2.9%
Split Level	0.0%	0.0%	0.0%	0.1%
Townhouse	1.9%	3.6%	5.2%	1.7%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	3.713 $p = 0.446$	29.702 $p = 0.000$	70.529 $p = 0.000$	

Table A.9. Distribution of Median Household Income (by Census Block Group)

Income Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Less than $20,000	41.0%	36.2%	34.4%	36.7%
$20,000 -- $29,999	39.3%	46.7%	47.2%	44.2%
$30,000 -- $39,999	10.0%	9.3%	9.2%	9.9%
$40,000 -- $49,999	5.7%	4.6%	4.1%	5.3%
$50,000 or greater	4.0%	3.2%	5.1%	3.9%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	8.334 $p = 0.080$	4.834 $p = 0.305$	9.169 $p = 0.057$	

Table A.10. Distribution of Percentage of Owner-Occupied Homes (by Census Block Group)

Percentage Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
0 – 24.99%	15.8%	12.4%	13.2%	14.2%
25 – 49.99%	58.8%	59.4%	59.0%	59.5%
50 – 74.99%	23.8%	25.2%	22.1%	23.9.%
75 – 100%	1.6%	3.0%	5.7%	2.4%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	3.830 $p = 0.280$	5.718 $p = 0.126$	41.550 $p = 0.000$	

Table A.11. Distribution of Percentage of Homes Built Before 1940 (by Census Block Group)

Percentage Category	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
9.8 – 24.99%	5.2%	5.7%	8.0%	5.8%
25 – 49.99%	20.0%	15.6%	15.8%	16.1%
50 – 74.99%	54.9%	58.0%	54.7%	55.6%
75 – 89.8%	19.9%	20.7%	21.5%	22.5%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	10.227 $p = 0.017$	3.589 $p = 0.309$	7.692 $p = 0.053$	

Table A.12. Distribution of Travel Time (Minutes) from Nearest BRT Station to Public Square (Tower City – CBD)

Travel Time (minutes)	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
0 – 9.99	11.2%	9.4%	14.2%	10.5%
10 – 19.99	18.6%	14.4%	11.7%	16.0%
20 – 29.99	36.3%	33.6%	35.5%	35.5%
30 – 43.5	33.9%	42.6%	38.6%	38.0%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	7.048 $p = 0.070$	12.071 $p = 0.007$	22.304 $p = 0.000$	

Table A.13. Distribution of Travel Time (Minutes) from Nearest BRT Station to Cleveland Clinic (Major Area Employer/Medical Facility)

Travel Time (minutes)	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
0 – 9.99	59.7%	53.5%	52.2%	56.1%
10 – 19.99	25.3%	32.2%	28.4%	29.0%
20 – 30	15.0%	14.3%	19.4%	14.9%
TOTAL	100.0%	100.0%	100.0%	100.0%
Chi-Square Goodness of Fit Compared to All Homes	5.284 $p = 0.071$	6.374 $p = 0.041$	14.241 $p = 0.001$	

Table A.14. Distribution of Homes by Cleveland City Ward

Cleveland City Ward	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Not within City of Cleveland	0.1%	0.5%	0.8%	0.6%
Ward 3	10.4%	9.1%	13.7%	9.8%
Ward 4	3.1%	3.7%	3.8%	3.8%
Ward 5	7.2%	4.8%	8.4%	6.9%
Ward 6	17.3%	15.4%	12.5%	15.4%
Ward 7	19.4%	16.9%	15.6%	18.3%
Ward 8	2.1%	2.0%	2.3%	2.2%
Ward 9	25.1%	29.3%	26.2%	26.7%
Ward 10	14.6%	18.0%	16.3%	15.5%
Ward 15	0.7%	0.3%	0.5%	0.8%
TOTAL	100.0%	100.0%	100.0%	100.0%

Table A.15. Distribution of Homes by Other Areas along Euclid Corridor

Area	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Not within any Area	25.0%	23.8%	26.4%	24.9%
East Cleveland	0.4%	0.6%	0.7%	0.8%
Midtown	2.0%	0.9%	0.8%	1.3%
North Area 1	21.1%	19.7%	18.4%	19.9%
North Area 2	25.8%	30.9%	25.9%	27.9%
North Area 3	8.2%	11.0%	11.5%	9.8%
South Area	13.3%	9.7%	12.3%	11.5%
University Circle	4.2%	3.4%	4.0%	3.9%
TOTAL	100.0%	100.0%	100.0%	100.0%

Table A.16. Distribution of Homes by Cleveland Neighborhood

Cleveland Neighborhood	Sold in 2004	Sold in 2008–2009	Sold in 2010–2011	All Homes
Not within City of Cleveland	0.1%	0.5%	0.8%	0.6%
Broadway – Slavic Village	0.3%	0.7%	0.3%	0.4%
Buckeye – Shaker Square	5.7%	5.5%	4.1%	5.4%
Buckeye – Woodhill	4.6%	5.4%	5.4%	5.6%
Central	5.2%	3.7%	6.7%	4.3%
Clark – Fulton	0.0%	0.1%	0.1%	0.2%
Collinwood – Nottingham	2.3%	3.1%	4.1%	2.9%
Detroit Shoreway	0.7%	0.3%	0.5%	0.8%
Euclid – Green	1.6%	1.8%	1.0%	1.5%
Fairfax	10.0%	7.5%	6.3%	8.3%
Glenville	32.5%	39.0%	33.5%	35.1%
Goodrich – Kirtland Park	2.3%	1.6%	1.7%	2.6%
Hough	13.4	11.2%	9.8%	11.6%
Kinsman	0.5%	0.4%	0.5%	0.9%
Mount Pleasant	0.4%	0.2%	0.5%	0.2%
Ohio City	4.2%	4.2%	7.9%	5.5%
St. Clair – Superior	8.7%	8.8%	9.7%	8.6%
Tremont	6.2%	4.8%	5.7%	4.1%
University	1.3%	1.2%	1.4%	1.4%
TOTAL	100.0%	100.0%	100.0%	100.0%

APPENDIX B:

HOW THE EFFECT OF DISTANCE ON SALES PRICE IS OBTAINED FROM ESTIMATED REGRESSION COEFFICIENTS

I. Regressions with Distance as a Continuous Variable

A. Definitions

1. p is the sale price of a house
2. d is distance between the house and a BRT station
3. x represents all other variables in the regression
4. β_i are regression coefficients

B. Level Regression (with distance and distance squared)

1. Estimated equation: $p = \beta_0 + \beta_1 d + \beta_2 d^2 + \beta_3 x$
2. Interpretation of the effect of distance on price: $\frac{\partial p}{\partial d} = \beta_1 + 2\beta_2 d$

C. Log-Level Regression (with distance and distance squared)

1. Estimated equation: $\ln p = \beta_0 + \beta_1 d + \beta_2 d^2 + \beta_3 x$
2. Interpretation of the effect of distance on price: $\frac{\partial \ln p}{\partial d} = \frac{1}{p}\frac{\partial p}{\partial d} = \beta_1 + 2\beta_2 d$

   a. $\beta_1 + 2\beta_2 d$ is the proportionate change in p per unit change in d

b. So the percentage change is $\%\Delta p = 100\left(\beta_1 + 2\beta_2 d\right)$

D. Log-Log Regression

1. Estimated equation: $\ln p = \beta_0 + \beta_1 \ln d + \beta_2 \ln x$

2. Interpretation of the effect of distance on price: $\dfrac{\partial \ln p}{\partial \ln d} = \dfrac{d}{p}\dfrac{\partial p}{\partial d} = \beta_1$, which is the percentage change in *p* per 1-percent change in *d*, or the elasticity of *p* with respect to *d*

II. Regressions with Distance as a Dummy Variable

A. Definitions

1. *p*, *x*, and β_i as above

2. $d = \begin{cases} 1 & \text{if } 0 < d \le 2{,}640 \\ 0 & \text{if } 2{,}640 < d \le 5{,}280 \end{cases}$

B. Level Regression

1. Estimated equation: $p = \beta_0 + \beta_1 d + \beta_2 x$

2. Interpretation of the effect of distance on price

a. When $d = 1$, $p_1 = \beta_0 + \beta_1 + \beta_2 x$

b. When $d = 0$, $p_0 = \beta_0 + \beta_2 x$

c. Then, $p_1 - p_0 = \beta_1$

d. So β_1 represents the change in price for houses located between zero and one-half mile away from the BRT station relative to those located between one-half and one mile from the BRT station

B. Log-Level Regression

1. Estimated equation: $\ln p = \beta_0 + \beta_1 d + \beta_2 x$

2. Interpretation of the effect of distance on price

   a. $\dfrac{\Delta p}{p} = e^{\beta_1} - 1$ is the proportionate change in price

   b. $\%\Delta p = 100\left(e^{\beta_1} - 1\right)$ is the percentage change in price (This interpretation holds when the independent variable is not a dummy variable and is not log-transformed, and the dependent variable is logged.)

3. Proof

   a. Note that $p = e^{\beta_0 + \beta_1 d + \beta_2 x}$

   b. When $d = 1$, we have $p_1 = e^{\beta_0 + \beta_1 + \beta_2 x}$

   c. When $d = 0$, we have $p_0 = e^{\beta_0 + \beta_2 x}$

   d. Hence $\Delta p = p_1 - p_0 = e^{\beta_0 + \beta_1 + \beta_2 x} - e^{\beta_0 + \beta_2 x} = e^{\beta_0} e^{\beta_2 x}\left(e^{\beta_1} - 1\right)$

   e. Therefore, the proportionate change in p is

   $$\frac{\Delta p}{p} = \frac{p_1 - p_0}{p_0} = e^{\beta_1} - 1$$

   f. The percentage change in p is $\%\Delta p = 100\left(e^{\beta_1} - 1\right)$

C. Log-Log Regression

1. Estimated equation: $\ln p = \beta_0 + \beta_1 \ln d + \beta_2 \ln x$

2. Interpretation of effect of distance on price: This has the same interpretation as in the log-level regression because in both regressions the dependent variable is logged while the dummy variable is not. This interpretation holds when the independent variable is not a dummy variable and is not log-transformed, and the dependent variable is logged.

APPENDIX C:

IMAGE CREDITS

This Appendix contains the credits for the images used in this document.

Figure 1.1. Examples of Commuter Rail Systems (page 5)
Virginia Railway Express photo: By John from Southern Maryland, USA (VRE V09(RP39-2C)) via Wikimedia Commons [CC BY 2.0 (http://creativecommons.org/licenses/by/2.0)].
Caltrain photo: By Lucius Kwok from New York, NY, United States (caltrain-927-2) via Wikimedia Commons [CC BY-SA 2.0 (http://creativecommons.org/licenses/by-sa/2.0)].
SunRail photo: By SunRail [Public domain], via Wikimedia Commons.
Tri-Rail photo: By Phillip Pessar (Flickr: Tri Rail Miami Airport Station) via Wikimedia Commons [CC BY 2.0 (http://creativecommons.org/licenses/by/2.0)].

Figure 1.2. Examples of Heavy Rail Systems (page 6)
New York photo: By Aude (Own work) via Wikimedia Commons [CC BY-SA 2.5 (http://creativecommons.org/licenses/by-sa/2.5)].
Washington, D.C. photo: By Andrew Bossi via Wikimedia Commons [CC BY-SA 2.0 (http://creativecommons.org/licenses/by-sa/2.0)].
Chicago photo: By Rene Schwietzke from Jena, Germany (Chicago 'L') via Wikimedia Commons [CC BY 2.0 (http://creativecommons.org/licenses/by/2.0)].
Miami photo: By DearEdward from New York, NY, USA (Miami Metrorail Uploaded by russavia) via Wikimedia Commons [CC BY 2.0 (http://creativecommons.org/licenses/by/2.0)].

Figure 1.3. Examples of Light Rail/Streetcar Systems (page 6)
San Diego photo: By Josh Truelson (San Diego Trolley) via Wikimedia Commons [CC BY-SA 2.0 (http://creativecommons.org/licenses/by-sa/2.0)].
Portland photo: By Tim Adams (Own work) via Wikimedia Commons [CC BY 3.0 (http://creativecommons.org/licenses/by/3.0)].
Charlotte photo: By James Willamor via Wikimedia Commons [CC BY-SA 2.0 (http://creativecommons.org/licenses/by-sa/2.0)].
Tampa photo: TheCustomOfLife [Public domain], via Wikimedia Commons.

Figure 1.4. Select BRT Systems Operating in the U.S. (page 10)
L.A. Orange Line photo: By Carren Jao via Zocalo Public Square – July 2, 2012 http://zocalopublicsquare.org/wp-content/ uploads/ 2011/05/orangeline_ myclockworkorange.jpg
Kansas City MAX photo: By Victoria A. Perk, author of dissertation.
L.A. Metro Rapid photo: By Mariordo Mario Roberto Duran Ortiz (Own work) via Wikimedia Commons [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)].
Las Vegas MAX photo: By Cello06 at English Wikipedia (Transferred from en.wikipedia to Commons.) [Public domain], via Wikimedia Commons.
Cleveland photo: By Victoria A. Perk, author of dissertation.
Pittsburgh photo: By Dllu (Own work) via Wikimedia Commons [CC BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0)].
Boston photo: By Victoria A. Perk, author of dissertation.
Eugene photo: By Victoria A. Perk, author of dissertation.

Figure 4.1. Cleveland HealthLine Construction along Euclid Avenue, 2007 (page 40)
All photos by Victoria A. Perk, author of dissertation.

Figure 4.2. Completed Cleveland HealthLine Euclid Avenue Corridor (page 41)
All photos by Victoria A. Perk, author of dissertation.

Figure 4.3. Cleveland HealthLine Stations (page 42)
Source: Greater Cleveland Regional Transit Authority (GCRTA). Used with permission from Stephen Bitto, GCRTA (sbitto@gcrta.org).